В.В. ЛИТВИНОВ*, Н. СТОЯНОВ**, І.С. СКІТЕР***, О.В. ТРУНОВА*,
А.Г. ГРЕБЕННИК***

ВИКОРИСТАННЯ МЕТОДІВ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ПРИ ПОШУКУ ДЖЕРЕЛ АТАК НА КОМП'ЮТЕРНІ МЕРЕЖІ В УМОВАХ НЕВИЗНАЧЕНОСТІ

*Чернігівський національний технологічний університет, м. Чернігів, Україна

**Болгарський інститут оборони імені Цветана Лазарова, м. Софія, Болгарія

***Інститут проблем математичних машин і систем НАН України, м. Київ, Україна

Анотація. Запропоновані теоретичний і, частково, методологічний підходи до створення спеціальних систем підтримки прийняття рішень (СППР) для пошуку найімовірніших джерел атак. Систематизовані, узагальнені і розвинені уявлення про методи підтримки прийняття рішень при пошуку джерел атак на комп'ютерні мережі. Запропонована схема структури процесу пошуку джерел атак на базі методології підтримки прийняття рішень, основна мета якої максимально скоротити число висунутих варіантів рішень, щоб на наступних етапах детальній оцінці підлягала обмежена кількість альтернатив. Коротко охарактеризований кожен етап окремо, а саме: виявлення проблемної ситуації; формування та узгодження цільової ієрархії; формування альтернатив (варіантів рішення); аналіз альтернатив; формування критеріїв відбору; оптимальний вибір; рішення проблемної ситуації. Зокрема, розглянута задача багатокритеріального вибору та її системну модель, узагальнені методи розв'язання. Запропоновано рішення проблемної ситуації з залученням методу визначення нечітких суджень ОПР, що поєднує класичні підходи при нечіткому автоматичному управлінні на основі експертних суджень та ідею розбиття простору критеріїв на області в комбінованому методі підтримки прийняття рішень. Запропоновано структуру СППР, за допомогою якої здійснюється вибір альтернативи (рішення), під яким розуміється визначення найбільш імовірного джерела атак, що дозволяє, використовуючи інформацію з сайтів стакерів, оцінити загальний рівень небезпеки кіберпростору для даної мережі, який визначається як кількість невідомих атак в одиницю часу; здійснити налаштування адаптивної СЗА; розробити нове покоління мережесих екранів, які базуються не тільки на протоколах, а й на адресній інформації з мережесих пакетів.

Ключові слова: СППР, джерела атак, комп'ютерні мережі.

Аннотация. Предложены теоретический и, частично, методологический подходы к созданию специальных систем поддержки принятия решений (СППР) для поиска наиболее вероятных источников атак. Систематизированы, обобщены и развиты представления о методах поддержки принятия решений при поиске источников атак на компьютерные сети. Предложена схема структуры процесса поиска источников атак на базе методологии поддержки принятия решений, основная цель которой максимально сократит число выдвинутых вариантов решений, чтобы на следующих этапах детальной оценке подлежало ограниченное количество альтернатив. Коротко охарактеризован каждый этап отдельно, а именно: выявление проблемной ситуации; формирование и согласование целевой иерархии; формирование альтернатив (вариантов решения); анализ альтернатив; формирование критериев отбора; оптимальный выбор; решение проблемной ситуации. В частности, рассматриваются задача многокритериального выбора и системная модель, обобщенные методы решения. Предложено решение проблемной ситуации с привлечением метода определения нечетких суждений ЛПР, сочетающее классические подходы при нечетком автоматическом управлении на основе экспертных суждений и идею разбиения пространства критериев на области, в комбинированном методе поддержки принятия решений. Предложена структура СППР, с помощью которой осуществляется выбор альтернативы (решения), под которым понимается определение наиболее вероятного источника атак, что позволяет, используя информацию с сайтов атакеров, оценить общий уровень опасности киберпространства для данной сети, который определяется как количество неизвестных атак в единицу времени; осуществить

настройку адаптивной СЗА; разработать новое поколение сетевых экранов, которые базируются не только на протоколах, но и на адресной информации из сетевых пакетов.

Ключевые слова: СППР, источники атак, компьютерные сети.

Abstract. It was proposed theoretical and, in part, methodological approach to the creation of special decision support systems (DSS) to search the most likely sources of attacks. Systematic, generalized and developed ideas about methods decision support searching the sources of attacks on computer networks. It was proposed structure scheme of the process of finding sources of attacks based on decision-making methodology, the main goal of which is to minimize the number of options put forward so that at the next stages a limited number of alternatives should be evaluated in detail. Each stage was briefly characterized separately, namely: identifying a problem situation, formation and coordination of the target hierarchy, formation of alternatives (solutions), analysis of alternatives, and formation of selection criteria, optimal choice, and solution of a problem situation. In particular, the problem of multi-criteria choice and its system model, generalized methods of solution are considered. The proposed solution to the problem situation involving the method of determining fuzzy judgments DM, which combining classical approaches with fuzzy automatic control based on expert judgments and the idea of splitting the criterion space into domains, the combined method of decision support. It was proposed structure of DSP, with the help of which the choice of alternatives (solutions), which is the definition of the most likely source of attacks, allows to use general information about the threat of cyberspace for a given network, which is defined as the number of unknown attacks per unit of time, set up adaptive SDA. It also allows developing a new generation of firewalls that are based not only on protocols, but on address information from network packets as well.

Keywords: DSS, source of the attack, computer networks.

DOI: 10.34121/1028-9763-2019-4-38-51

1. Вступ

Ефективність залучення інформації глобальної мережі для виявлення нестандартної поведінки і комп'ютерних атак у мережі залежить від двох чинників:

- ступеня налаштованості СЗА на протидію найбільш імовірним вторгненням, включаючи і загрози невідомого типу;
- можливості відшукування найбільш імовірних джерел атак.

Якщо перший чинник вимагає високого ступеня адаптивності СЗА до множини різновидів вторгнень, що, як правило, досягається використанням методів штучного інтелекту, які мають механізми навчання, то інформація про джерело загрози до певної міри дозволяє оцінити ризики проведення атак різних видів від груп найбільш імовірних атакерів для комп'ютерних мереж.

Хоча інформація про джерело часто вже присутня в пакетах протоколів обміну, проте необхідно пам'ятати, що атака передбачає три складових:

- розвідку;
- саму атаку;
- приховування наслідків.

Остання складова атаки істотно знижує ступінь достовірності інформації пакета про джерело. Крім того, атаки можуть вестися не безпосередньо з комп'ютера атакера, а через зомбовані комп'ютери.

Разом з тим, наявність інформації про джерело атаки надає можливість організації активного захисту мережі, що дозволяє заздалегідь адаптуватися СЗА до ймовірних різновидів атак і у принципі здійснювати цілеспрямований вплив на атакерів.

Тим самим, другий чинник також набуває важливого значення для захисту мереж.

2. Постановка завдання

Актуальність даної роботи обумовлена недостатнім освітленням у літературі принципів пошуку джерел атак на базі методології підтримки прийняття рішень. У відомих публіка-

ціях [1–5, 13, 14] в основному висвітлюються питання, пов'язані з побудовою вузькоспеціалізованих експертних систем щодо забезпечення безпеки комп'ютерних мереж. Однак, на даний момент актуальною є побудова більш універсальних систем підтримки прийняття рішень, здатних аналізувати і охоплювати широке коло питань аналізу функціонування мереж.

Мета даної роботи – запропонувати теоретичний і, частково, методологічний підхід до створення спеціальних систем підтримки прийняття рішень (СППР) для пошуку найбільш імовірних джерел атак.

Глобальна мережа, в кіберпросторі якої можуть відбуватися «бойові дії», може розглядатися як складна соціоекономічна система (СЕС) [6, 12], в якій здійснення атаки на комп'ютерну мережу трактується як виникнення проблемної ситуації, а процес пошуку можливих джерел атак як процес розробки найкращого варіанта вирішення проблемної ситуації. Останній можна розглядати як процес прийняття управлінського рішення при вирішенні слабкоструктурованих проблем.

Як правило, рішення доводиться приймати в умовах інформаційної невизначеності та багатокритеріальності.

Це складні задачі, які потребують проведення їх дослідження на різних рівнях абстрактного опису: лінгвістичного, логіко-математичного, інформаційного, евристичного.

Більш того, пошук джерел атак – це процес колективної роботи висококваліфікованих спеціалістів.

Основними етапами абстрактного рівня опису складних задач прийняття рішень можна вважати:

- формулювання основної проблемної ситуації, яка потребує вирішення у вигляді задачі прийняття рішень;
- складання морфології проблемної ситуації та відповідної задачі. Під морфологією розуміють перелік підпроблем (підзадач), необхідних для вирішення основної проблемної ситуації;
- опис основних параметрів: цілей, обмежень, вхідних і вихідних параметрів, які характеризують як виконання підзадач, так і задачу в цілому;
- вибір критеріїв, які характеризують різні властивості цілей;
- побудову моделей для задач прийняття рішень, критеріїв, обмежень тощо;
- формування принципу вибору ефективного рішення підзадач;
- визначення принципу узгодження рішень підзадач для вирішення основної задачі.

3. Виклад основного матеріалу

Структуризація процесу пошуку джерел атак на базі методології підтримки прийняття рішень представлена на рис. 1.

Основна мета даної схеми – це максимально скоротити число висунутих варіантів рішень, щоб на наступних етапах детальній оцінці підлягала обмежена кількість альтернатив. Коротко охарактеризуємо кожен етап процесу пошуку окремо.

Виявлення проблемної ситуації

На цьому етапі завдання виявлення проблемної ситуації відповідають завданням виявлення атак або завданням виявлення нестандартної поведінки мережі, які вирішуються засобами СЗА.

Оскільки до СЗА в комплексі оперативного управління висуваються вимоги до часу обчислень вхідної інформації, то для вирішення завдання управління в умовах неповноти, суперечливості та невизначеності даних про стан інформаційного середовища пропонуємо використовувати механізм нечіткого логічного висновку. Інформацією, яка надходить на вхід системи нечіткого логічного висновку, є вхідні змінні – число ознак аномальних по-

дій. Ці змінні відповідають реальним процесам у мережі. Інформація, яка формується на виході системи нечіткого логічного висновку, відповідає вихідній змінній, яка є ймовірністю того, що сукупність аномальних подій у мережі є атакою (ймовірність атаки).

Рисунок 1 – Структура процесу пошуку джерел атак на базі методології підтримки прийняття рішень

Формування та узгодження цільової ієрархії

Ціль у теорії прийняття рішень визначається досить різними способами. Аналізуючи різні джерела, наведемо деякі визначення цілі [10]:

- ціль – це результат, на досягнення якого направлені всі зусилля;
- ціль визначає критерії оцінки ефективності рішень;
- ціль – це бажаний стан або результат діяльності системи.

У нашому випадку загальною метою може служити виявлення ймовірного джерела атак. Однак можна висунути і приватні підцілі. Наприклад, знаходження джерела в середині певної хакерської групи або знаходження джерела в середині певної корпоративної структури, або знаходження джерела серед певної державної організації тощо. Для того, щоб сформулювати мету, необхідно знати характеристики об'єкта атаки, його місце на ринку, його потенційних конкурентів, структуру державних органів потенційної атакуючої сторони, тощо. Потенційно ця інформація є у глобальній мережі. Отже, основним завданням даного етапу є інформаційний пошук, параметри якого визначаються в основному підцілями. Однак це вимагає чималого часу і зусиль для систематизації інформації.

Формування альтернатив (варіантів рішення)

Це найбільш невизначений етап у даній схемі. Його призначення - формування множини потенційних джерел атак.

З одного боку, формування цієї множини може йти за рахунок обміну інформацією про атаки з державними центрами безпеки, спеціалізованими компаніями, які займаються проблемами інформаційної безпеки, з другого боку, шляхом аналізу трафіка у своїй корпоративній мережі, з третього – шляхом інформаційного пошуку в соціальних мережах, корпоративних сайтах.

Найважчою формою пошуку джерела є контент-аналіз потоків, які виходять із потенційних джерел або приймаються корпоративною мережею з задіянням завдань класифікації контенту.

Аналіз альтернатив

На даному етапі, шляхом задіяння інформації з сайтів потенційних джерел атак, форумів і соціальних мереж, визначаються наслідки від прийняття тієї чи іншої альтернативи. Якщо прийнята альтернатива дійсно визначає правильне джерело втручання, то його вплив на корпоративну мережу може бути нівельовано. В іншому випадку втрати від неправильного визначення обумовлюються видом атаки. Таким чином, можна визначити ризик від кожної альтернативи:

$$R_i = (1 - p_i) * C_i,$$

де p_i – імовірність правильного визначення джерела;

C_i – втрати від неправильного визначення джерела для i -ї альтернативи.

Ризик є не єдиною характеристикою, яка описує якість альтернативи. Найчастіше альтернатива A_i описується вектором характеристик-наслідків $X_i = \langle x_{i1}, x_{i2}, \dots, x_{in} \rangle$ і відповідних функцій їх розрахунку f_{ij} , що дозволяють оцінити наслідки альтернативи з різних сторін. Іноді ці характеристики-наслідки розглядаються як критерії, на основі яких можна здійснити вибір найкращої альтернативи, яка в нашому випадку розглядається як найбільш правдоподібне джерело атак.

Формування критеріїв відбору

Критерій – це математична модель, яка відображає суть, мотив, стимул, бажання, відчуття та інші змінні (прості та узагальнені) стану об'єкта.

Припустимо, що відомі допустима множина альтернативних рішень A і довільна непорожня множина $K = \{K_1, K_2, \dots, K_m\}$, елементи якої ми будемо називати критеріями, за якими можна отримати оцінку будь-якої альтернативи $a \in A$. Під критерієм можна розуміти функцію, яка вимірює цінність або якість деяких об'єктів.

Виходячи з вищесказаного, оцінка досягнення цілі може бути представлена у вигляді трирівневого критеріального дерева.

Глобальний критерій – це головний критерій, вершина критеріального дерева, який підпорядковує собі всі інші критерії. Глобальний критерій – це фактор, який дозволяє визначити успіх вирішення проблеми.

Критерії-показники – це величини, які можуть вимірюватись якісно або кількісно. Критерії-індикатори є більш конкретними оцінками, наприклад, розмір, якість, об'єм, кількість тощо. Критерії-показники і підкритерії-індикатори можуть самі служити глобальним критерієм.

Ступінь або рівень вимірності критерію залежить, у значній мірі, від вимірності цілі, яку він описує, або вимірності результату, який очікується.

Існує кілька різновидів шкал вимірності:

- номінальна шкала – це класифікація з фіксованим номіналом родової якості;
- кардинальна шкала характеризує кумулятивність і адитивність вимірювання;
- ординальна шкала характеризує властивості через інтенсивності.

Кардинальний і ординальний класи шкал є числовими.

У відповідності з логічною структурою вимірювання важливості використовує три групи аксіом: тотожність, ранговий порядок і адитивність. У залежності від того, які з цих трьох властивостей присутні або відсутні і в яких комбінаціях зустрічаються для вимірювання критеріальних оцінок, розрізняють чотири основних типи шкал: назв, порядку, інтервалів (різниць), відношень [10]. Дамо їх коротку характеристику.

Шкала назв (ідентифікація альтернатив). Клас найпростіших шкал, які відображають якісні властивості. Їх елементи характеризуються тільки відношеннями еквівалентності (тотожності) і подібності проявлення властивості. Такі шкали не мають нуля і одиниць вимірювання, на них не можна проводити арифметичні операції. Вимірюваннями у шкалі назв є результати якісного аналізу.

Шкала порядку (визначення порядку переваг). Дана шкала задовольняє аксіомам рангового порядку. Елементи, які розміщені на шкалі, повинні бути порівнювані і транзитивні за деякою загальною ознакою. Порівняння елементів на цій шкалі відповідають відношенням «більше-менше», «краще-гірше», «сильніший-слабший», «складніший-простіший» тощо. Дані шкали є нелінійними і не мають одиниць виміру. Вимірювання у шкалі порядку є найнедосконалішими і найменш інформативними.

Для полегшення вимірювань на шкалі порядку на практиці застосовуються деякі опорні точки як «реперні». Такі шкали називаються реперними.

Шкала інтервалів показує ступінь близькості альтернатив у шкалі порядку. Якщо множина елементів впорядкована за допомогою множини дійсних чисел, то говорять, що виміри здійснені у шкалі інтервалів. Шкала інтервалів – це шкала без визначення точки відліку, тобто їй не притаманна властивість адитивності. Інтервали між кожними сусідніми елементами в даній шкалі показують інтенсивність (силу) переваги одного елемента в порівнянні з іншими і можуть бути як рівномірними, так і нерівномірними. У зв'язку з тим, що в даній шкалі не визначено початок відліку, для неї можливі адитивні операції (додавання і віднімання) і не визначені мультиплікативні (множення і ділення).

Шкала відношень. Якщо одна з реперних точок має значення вимірності, яке прирівняне до нуля, то в такій шкалі можна відрахувати абсолютне значення величини і визначити, у скільки разів одна величина більша за іншу. Шкала відношень є найдосконалішою і найбільш інформативною. В ній визначені всі арифметичні операції. Прикладами шкал відношення є температурна шкала, шкала відстаней, шкала ваг тощо.

Важливим компонентом задачі вибору рішень є простір критеріїв або критеріальний простір. Під простором критеріїв будемо розуміти сукупність критеріїв, за якими оцінюється ефективність альтернатив (наслідків), тобто критеріальний простір – це простір критеріальних оцінок альтернатив. Якщо даний простір складається з однієї скалярної характеристики, то задача вибору є тривіальною. Але, як правило, при вирішенні складних проблем, звичайно, доводиться розглядати декілька взаємозв'язаних показників ефективності, які у свою чергу виражені в кількісній або якісній формі, вимірюються в різних шкалах і мають різну вагомість. Це ускладнює проведення співставлення альтернатив у такому багатокритеріальному аморфному просторі. Вирішення даного питання полягає у нормуванні критеріального простору [6, 12].

Під нормалізацією критеріїв розуміють перехід до однаково направлених переваг, вираження їх значень в однакових абсолютних величинах або перехід до безрозмірних шкал. Приведемо декілька відомих видів нормалізації.

1. Нормалізація за заданим значенням: $\overline{K}_i = \frac{k_i}{K_i^*}$, де K_i^* – задана величина критерію (еталонна, ідеальна, задовільна) та її частинний випадок відносна нормалізація: $\overline{K}_i = \frac{k_i}{\max_{a \in A} k_i}$

або $\overline{K}_i = \frac{k_i}{\min_{a \in A} k_i}$.

2. Порівняльна нормалізація: $\overline{K}_i = K_i - \min_{a \in A} K_i$ або $\overline{K}_i = \max_{a \in A} K_i - K_i$, природна нормалізація: $\overline{K}_i = \frac{k_i}{\max_{a \in A} k_i - \min_{a \in A} k_i}$ та повна нормалізація: $\overline{K}_i = \frac{K_i - \min_{a \in A} K_i}{\max_{a \in A} K_i - \min_{a \in A} K_i}$.

3. Нормалізація Севіджа: $\overline{K}_i = \frac{\max_{a \in A} K_i - K_i}{\max_{a \in A} K_i - \min_{a \in A} K_i}$.

4. Інтегральна нормалізація: $\overline{K}_i = \frac{k_i}{\int_{a \in A} k_i dx}$. Часто використовується у задачах, коли множина A задана дискретно, тобто $A = \{a_1, a_2, \dots, a_n\}$, тоді $\overline{K}_i = \frac{k_i}{\sum_{a_j \in A} k_i(a_j)}$.

Зазвичай як критерії використовуються такі індикатори (ознаки):

- місце реєстрації IP-адрес і доменів, які беруть участь в атаці або надають інфраструктуру для реалізації атаки;
- трасування атаки до її джерела або хоча б локалізація області, в якій джерело знаходиться;
- часові параметри;
- аналіз програмного коду, в якому можуть бути знайдені коментарі, посилання на сайти, домени, IP-адреси, які беруть участь в атаці;
- почерк програмістів і школа програмування;
- стильометрія, яка дозволяє визначити стилістику мови у коментарях або супутніх текстах;
- обманні системи або honeypot/honeynet;
- оперативна розробка;
- аналіз активності на форумах і в соціальних мережах;
- ідентифікація постфактум за діями.

Оптимальний вибір

Задача прийняття рішень буде вирішена, якщо будуть здійснені три процедури: – генерація допустимої множини альтернатив, багатокритеріальний аналіз і побудова правила вибору (принципу оптимальності).

Постановка задачі багатокритеріального вибору рішень. Нехай відома множина альтернативних рішень, яка може бути задана як у вигляді дискретної, так і неперервної множини. Відомі також критерії, за якими можуть бути оцінені наслідки будь-якої альтернативи із заданої множини. Потрібно здійснити вибір рішень на даній множині альтернатив у вигляді однієї із задач:

- відбір альтернатив (screening);
- класифікація (кластеризація) простору альтернатив (sorting methods);
- упорядкування альтернатив (ranking);
- вибір найкращої альтернативи (choice problem).

Системну модель такої задачі можна описати таким чином:

$$\{V, K, Q, A, P_V \mid R_V\}. \quad (1)$$

Відомі:

V – вид задачі вибору рішень;

K – множина критеріїв;

Q – шкали оцінок;

A – множина альтернатив;

P_V – правило вибору.

Невідомим є R_V – результат вибору. Результатом вибору може бути:

– множина кращих (допустимих) альтернатив

$$R_V \{a \mid a \in A \wedge K(a) \in X, X - \text{допустима множина оцінок}\};$$

– альтернативи, розбиті на класи,

$$R_V \{(a, q) \mid a \in A, q - \text{клас (назва)}\};$$

– визначений порядок альтернатив

$$R_V \{(a, q) \mid a \in A, p - \text{ранг (номер)}\};$$

– вибір однієї або декількох альтернатив

$$R_V = \{a^* = \arg \underset{a \in A}{\text{opt}} P_V(K(a))\}.$$

Загальна математична постановка задачі багатокритеріального вибору. Нехай A – множина допустимих альтернатив (рішень), а $C(A) \subseteq A$ – множина рішень, які вибираються. Визначити множину $C(A)$, для якої:

$$\Psi(K(C(A))) \rightarrow \text{opt}, \quad (2)$$

де K – множина критеріїв (показників) оцінювання альтернатив, Ψ – оператор, який реалізовує процедуру регуляризації значень оцінок ефективності рішень на основі визначеного принципу (правила) оптимальності.

Методи розв'язання багатокритеріальних задач можна поділити на три групи:

1. Метод головного показника: полягає у перетворенні вихідної багатокритеріальної оптимізаційної задачі у однокритеріальну з додатковими обмеженнями, що враховують вимоги, висунуті іншими критеріями.

2. Метод результуючого показника полягає у заміні вихідної багатокритеріальної задачі на скалярну через формування узагальненого критерію:

– адитивний: $R_V = \sum_{i=1}^n w_i \cdot \bar{K}_i$, $\sum_{i=1}^n w_i = 1$, $w_i > 0$, $i = \overline{1, n}$, де $\bar{K}_i$ – нормоване значення i -го критерію; w_i – ваговий коефіцієнт i -го критерію, який має тим більшу величину, чим більше він впливає на якість системи;

– мультиплікативний: $R_V = \prod_{i=1}^n \bar{K}_i^{w_i}$, де $\bar{K}_i$ – нормоване значення i -го критерію; w_i – ваговий коефіцієнт i -го критерію;

– максимінний: $\max_{P_V \in A} \min_{1 \leq j \leq m} \{\bar{K}_1(P_V), \dots, \bar{K}_j(P_V), \dots, \bar{K}_m(P_V)\}$, якщо вагові коефіцієнти час-тинних показників відсутні; $\max_{P_V \in A} \min_{1 \leq j \leq m} \{\bar{K}_1^{w_1}(P_V), \dots, \bar{K}_j^{w_j}(P_V), \dots, \bar{K}_m^m(P_V)\}$, якщо вагові коефіцієнти визначені.

3. Лексикографічний метод або метод послідовних поступок — це метод аналогічний до описаного вище методу головного показника, використовується за умови, упорядкованості критеріїв за важливістю, при цьому кількісні оцінки їх ваги залишаються невизначеними. Метод послідовних поступок полягає в тому, що за рахунок введення додаткових обмежень багатокритеріальна задача перетворюється на послідовність однокритеріальних, у якій враховуються вимоги до критеріїв області допустимих розв'язок, що призводять до її поступового звуження. При постановці задачі, відповідно до найважливішого критерію, робиться поступка. Її величина залежить від висунутих вимог й оптимального розв'язку за відповідним критерієм.

Рішення проблемної ситуації

Важливою особливістю процесів прийняття рішень при пошуку джерел атак на комп'ютерні мережі є необхідність урахування впливу невизначених факторів і розгляд усіх можливих наслідків альтернатив, пропонуваніх для вибору. Тому актуальною є проблема розробки моделей прийняття рішень саме в умовах невизначеності.

Ці моделі забезпечують обробку та структурування інформації, що сприяє принаймні частковому подоланню проблеми неповноти наявних вихідних даних. Отримані при цьому рекомендації щодо прийняття рішень використовуються тільки у випадках, коли припущення, на основі яких вони були побудовані, відповідають природі джерел невизначеності. Встановлювати цю відповідність необхідно тільки при розумінні сутності і різноманітності факторів невизначеності, що впливають на проведення атак на комп'ютерні мережі і пов'язаних з ними понять шансу, ризику і небезпеки.

Для розв'язання задачі виявлення найбільш імовірного джерела атак пропонуємо використовувати метод визначення нечітких суджень ОПР, що поєднує класичні підходи при нечіткому автоматичному управлінні на основі експертних суджень [7, 15], та ідею розбиття простору критеріїв на області в комбінованому методі підтримки прийняття рішень [8].

Нехай визначено n критеріїв $K = \{K_1, K_2, \dots, K_n\}$. Для кожного з них задана нечітка шкала Q_i , тобто множина можливих значень критерію, що розбита на окремі нечіткі градації, c_{ij} — нечітка j -та градація i -го критерію шкали критерію Q_i , де i — номер критерію ($i = 1, \dots, n$).

Число градацій для критеріїв може бути різним: $j = 1, \dots, q_i$, де q_i — число градацій шкали $Q_i = \{c_{i1}, c_{i2}, \dots, c_{iq_i}\}$ для i -го критерію.

Нечіткі градації задаються за допомогою функції належності $\mu_{ij}(x)$ — функція належності чітких значень i -го критерію j -ї градації. Функцію належності градації зручно задавати багатокутником. Зазвичай вони представлені у вигляді трикутника або трапеції (рис. 2, 3).

Нечіткі градації — це звичайні нечіткі обмежені числа. При побудові чіткої шкали критеріїв у якісних методах багатокритеріального аналізу потрібно, щоб градації не перетиналися [7, 15], однак по відношенню до нечіткої градації допустимий деякий ступінь перетину.

Ефективність розбиття визначається із співвідношень:

$$\forall i: \max_{j_1, j_2} \frac{\int_{x_1}^{x_2} \min \mu_{ij_1}(x), \mu_{ij_2}(x) dx}{\int_{x_1}^{x_2} \max \mu_{ij_1}(x), \mu_{ij_2}(x) dx} \leq \alpha, \inf_x (\max_j \mu_{ij}(x)) \geq \beta.$$

Конкретні значення констант α і β вибираються, виходячи з вимог до якості прийнятих рішень і «нечіткості» суджень ОПР. При роботі з чіткими критеріями $\alpha = 0$ і

$\beta = 1$, проте нечіткість не дозволяє досягти нам таких значень. У практичних задачах вдалим наближенням вважаються: $\alpha = 0,3$ і $\beta = 0,7$.

Рисунок 2 – Розбиття шкали на нечіткі трикутні градації

Для побудови функцій належності пропонуються такі алгоритми градацій:

– для трикутних градацій:

1) ОПР задає істотні для нього значення критеріїв (точки на шкалі);

2) відбувається автоматична заміна точок на трикутні функції належності. Так точка

x_k замінюється на нечітку градацію $c = \left\{ \frac{0}{x_{k-1}}, \frac{1}{x_k}, \frac{1}{x_{k+1}} \right\}$;

3) отримані градації пред'являються ОПР:

– для трапецієподібних градацій:

1) ОПР задає чіткі інтервали розбиття;

2) відбувається автоматичне «розмиття» границь інтервалів: інтервал (x_1, x_2) замінюється на нечітку градацію $c = \left\{ \frac{0}{x_1 - d}, \frac{1}{x_2 + d}, \frac{1}{x_2 - d}, \frac{0}{x_2 + d} \right\}$, де $d = (x_2 - x_1) \cdot 100\%$;

3) отримані градації пред'являються ОПР.

Рисунок 3 – Розбиття шкали на нечіткі трапецієподібні градації

У заданих шкалах ми можемо розбити критеріальний простір на області, які визначаються як комбінації значень градацій. Для лексичних критеріїв як такі градації виступають слова (терми). Для числових критеріїв градація – це інтервал значень. Інтервал вказується у вигляді правої і лівої границь.

За кожним критерієм переваги повинні бути монотонні або мати один екстремум (ідеальне значення, наприклад, час проведення атаки). Градації повинні бути відсортовані в порядку збільшення переваг:

$$\forall i, j: c_{ij+1} > c_{ij}.$$

Повний простір всіх можливих комбінацій значень градацій критеріїв задається як декартовий добуток Q_i :

$$A = \{c_{11}, c_{12}, \dots, c_{1q_1}\} \times \{c_{21}, c_{22}, \dots, c_{2q_2}\} \times \dots \times \{c_{n1}, c_{n2}, \dots, c_{nq_n}\}$$

і називається повною множиною нечітких альтернатив.

Потужність цієї множини обчислюється за формулою

$$|A| = \prod_{i=1}^n q_i.$$

В ідеальному випадку для визначення джерела атаки нам потрібно задати рівень переваг для всіх елементів повної множини нечітких альтернатив A . У нечіткій постановці градації можуть покривати чималі області критеріального простору з різним рівнем функції належності, тому тільки для деяких з цих областей ми задамо нечіткий рівень переваг. Введемо такі позначення:

s – номер області переваг;

S – число областей, на яких ОПР висунула свої переваги (зрозуміло, що виконується співвідношення $S \leq |A|$);

$M_s = (j_1, j_2, \dots, j_n)$ – кортеж номерів градацій, які беруть участь у визначенні області s (припускаємо, що повинні бути задані нечіткі області за всіма критеріями);

p_s – нечіткий рівень переваг області s ;

$p_s(y)$ – нечітка функція належності для області s .

Розглянемо задачу визначення рівня переваг альтернативи з конкретними значеннями критеріїв $(x_1, x_2, \dots, x_n)$. Залежно від типу значень критеріїв необхідна їх попередня обробка – зведення до нечіткого виду шляхом задання нечіткої функції належності конкретного значення критерію $\mu_{x_i}(x)$. Оскільки розглядається задача ранжування на скінченній множині альтернатив, то число таких функцій буде скінченним.

Якщо критерій приймає чіткі числові/лінгвістичні значення – функція належності набуде вигляду

$$\mu_{x_i}(x) = \begin{cases} 1, & x = x_i, \\ 0, & x \neq x_i. \end{cases}$$

На практиці навіть для об'єктивних результатів вимірювань значень критеріїв присутня деяка похибка, тому доцільно задавати функцію належності у трикутному вигляді з урахуванням похибки σ :

$$\mu_{x_i}(x) = \begin{cases} \frac{x - x_i}{\sigma} + 1, & x \in [x_i - \sigma, x_i], \\ \frac{x_i - x}{\sigma} + 1, & x \in (x_i, x_i + \sigma), \\ 0, & \text{в інших випадках.} \end{cases}$$

Визначення нечітких областей розглядається як нечітка імплікація виду: якщо нечіткі значення критеріїв дорівнюють нечітким градаціям області M_k , то перевага альтернативи дорівнює заданому для області нечіткому рівню переваг p_k . Дана імплікація записується у вигляді

$$p_s(y, x_1, x_2, \dots, x_n) = \min_i (\min_x [\sup_x (\min(\mu_{x_i}(x), \mu_{x_i p_s}(x)))] \cdot p_s(y)).$$

Об'єднання всіх нечітких значень переваг за всіма областями дасть підсумкову нечітку перевагу альтернативі:

$$p(y, x_1, x_2, \dots, x_n) = \max_s p_s(y, x_1, x_2, \dots, x_n).$$

Після побудови підсумкової функції належності може виявитися, що $p(y, x_1, x_2, \dots, x_n) = 0$ або наближається до 0. Це говорить про те, що є точки критеріального простору, не охоплені областями переваг. У цьому випадку слід провести зміну шкал і/або додати нові області (збільшити S).

Далі можна визначити нечітке відношення домінування альтернатив, що є достатньо трудомістким і може призводити до неоднозначного ранжування. У разі великої кількості порівнюваних альтернатив доцільно провести дефазифікацію переваг методом центра тяжіння:

$$p(x_1, x_2, \dots, x_n) = \frac{\int_{y_1}^{y_2} y p(y, x_1, x_2, \dots, x_n) dy}{\int_{y_1}^{y_2} p(y, x_1, x_2, \dots, x_n) dy}.$$

Якщо всі вхідні функції належності були задані у формі ламаної лінії, то отримана функція належності рівня переваг $p(y, x_1, x_2, \dots, x_n)$ також є ламана лінія. У цьому випадку інтегрування проводиться за формулою

$$p = \frac{\sum_{l=1}^{v-1} a_l \frac{y_{l+1}^3 - y_l^3}{3} + b_l \frac{y_{l+1}^2 - y_l^2}{2}}{\sum_{l=1}^{v-1} (y_{l+1} - y_l) (p(y_{l+1}) - p(y_l))},$$

де v – число вершин ламаної;

y_l – координата вершини ламаної;

$p(y_l)$ – значення функції належності в вершині ламаної;

a_l і b_l – коефіцієнти рівняння відрізка ламаної $p(y_l) = a_l y_l + b_l$.

Зауважимо, що p , a і b є функціями від $(x_1, x_2, \dots, x_n)$, a_l і b_l визначаються співвідношеннями

$$a_l = \frac{p(y_{l+1}) - p(y_l)}{y_{l+1} - y_l},$$

$$b_l = p(y_l) - a_l y_l.$$

Далі необхідно провести ранжування альтернатив з отриманими чіткими рангами p . Загальна схема алгоритму роботи методу показана на рис. 4.

У алгоритмі задіяні:

ОПР – відповідає за введення критеріїв, задає шкали, коригує шкали, вводить переваги;

експерт – може вводити оцінки альтернатив за первинними критеріями після того, як задані критерії та шкали;

СППР – система підтримки прийняття рішень реалізує необхідні математичні методи, враховуючи операції з нечіткими множинами і числами.

Рисунок 4 – Алгоритм процедури нечіткого ранжування альтернатив

У разі незадовільного результату ОПР повертається на етап введення і редагування шкал і переваг. Етап завдання переваг і введення нечітких оцінок альтернатив можуть проходити паралельно.

Якщо для деякої альтернативи не існує хоча б однієї точки з достатнім рівнем належності (більше 0,5), то потрібно вводити нову область переваг на градаціях, що забезпечують максимум належності заданої альтернативи. Якщо не існує такої комбінації градацій, при якій забезпечується належність альтернативи області на рівні не менше, ніж 0,5, то слід встановити нові градації.

4. Висновки

Запропонована структура СППР, за допомогою якої здійснюється вибір альтернативи (рішення), під яким розуміється визначення найбільш імовірного джерела атак, дозволяє:

- використовуючи інформацію з сайтів атакерів, оцінити загальний рівень небезпеки кіберпростору для даної мережі, що визначається як кількість невідомих СЗА атак в одиницю часу;
- здійснити налаштування адаптивної СЗА;
- розробляти нове покоління мережевих екранів, які базуються не тільки на протоколах, а й на адресній інформації з мережевих пакетів.

Подальший розвиток дослідження передбачає перехід від теоретичних засад запропонованої структури до практичної реалізації автоматизованого налаштування систем підтримки прийняття рішень для пошуку найімовірніших джерел атак у структурі СЗА корпоративних мереж.

Подяка

Робота проведена та фінансована в рамках проекту NATO CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation – CyRADARS) – grant agreement number: G5286.

СПИСОК ДЖЕРЕЛ

1. Baddar S., Merlo A., Migliardi M. Anomaly detection in computer networks: a state-of-the-art review. *Journal of Wireless Mobile Networks. Ubiquitous Computing, and Dependable Applications*. 2014. Vol. 5, Iss. 4. P. 29–64.
2. Guan Y., Ghorbani A. A., Belacel N. Y-means: a clustering method for intrusion detection. *CCECE 2003. Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane Technology (Cat. No.03CH37436)*. 2003. Vol. 2. P. 1083–1086.
3. Gyanchandani M., Rana J.L., Yadav R.N. Taxonomy of Anomaly Based Intrusion Detection System: A Review. *International Journal of Scientific and Research Publications*. 2012. Vol. 2, Iss. 12. P. 1–13.
4. Jyothisna V., Prasad Rama V.V. A review of anomaly based intrusion detection systems. *International Journal of Computer Applications*. 2011. Vol. 28, Iss. 7. P. 26–35.
5. Lahno V. Information security of critical application data processing systems. *TEKA. Commission of motorization and energetics in agriculture*. 2014. Vol. 14, Iss. 1. P. 134–143.
6. Malyar M., Polishchuk V. Choice and evaluation methodics of investment projects. *Košická bezpečnostná revue*. Košice, 2013. P. 117–126.
7. Mamdani E. Applications of Fuzzy Algorithms for Control of Simple Dynamic Plant. *Proc. of IEEE*. 1974. Vol. 121, N 12. P. 1585–1588.
8. Riadi Im., Istiyanto J., Ashari A. Log Analysis Techniques using Clustering in Network Forensics. *International Journal of Computer Science and Information Security*. 2013. Vol. 10, Iss. 7. P. 23.
9. Simon H., Kadane J. Optimal Problem-Solving Search: All-or-None Solutions. *Artificial Intelligence, Fall*. 1975. Vol. 6. P. 235–248.
10. Метрологія та стандартизація в теплоенергетиці: підручник для студ. спеціальності 144 «Теплоенергетика» / уклад. Л.О. Кесова, В.І. Промоскаль, В.В. Червоний. Київ: КПІ ім. Ігоря Сікорського, 2018. 451 с.
11. Котенко И.В. Модели и методы построения и поддержки функционирования интеллектуальных адаптивных систем защиты информации. *Математические методы распознавания образов: 13-я Всероссийская конф. (ММРО-13) (Зеленогорск, Ленинградская обл., 30 сентября – 6 октября 2007 г.)*: сб. докладов. М.: МАКС Пресс, 2007. С. 599–602.
12. Маляр М.М., Поліщук В.В., Шаркаді М.М. Використання динамічних критеріїв у моделях багатокритеріального вибору. *Комп'ютерна математика*. 2015. Вып. 1. С. 125–133.
13. Литвинов В.В., Казимир В.В., Стеценко І.В. [та ін.]. Методи аналізу та моделювання безпеки розподілених інформаційних систем: навч. посібн. Чернігів: Чернігівський національний технологічний університет, 2016. 254 с.
14. Литвинов В.В., Казимир В.В., Стеценко І.В. [та ін.]. Моделювання та аналіз безпеки розподілених інформаційних систем. Чернігів: Чернігівський національний технологічний університет, 2017. 206 с.
15. Осипов В.П., Судаков В.А. Многокритериальный анализ решений при нечетких областях предпочтений. *Препринты ИПМ им. М.В. Келдыша*, 2017. № 6. 16 с.

Стаття надійшла до редакції 05.07.2019