**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**

**КАФЕДРА** **КОМП’ЮТЕРИЗОВАНИХ СИСТЕМ ЗАХИСТУ ІНФОРМАЦІЇ**

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ С.В. Казмірчук

«\_\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_ 20\_\_ р.

На правах рукопису

УДК 004.056.5:510.(043.3)

**МАГІСТЕРСЬКА АТЕСТАЦІЙНА РОБОТА**

**ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ**

**«МАГІСТР»**

**Тема**:Система виявлення вторгнень на об’єктах критичної інфраструктури

|  |  |
| --- | --- |
| **Автор:** | А. Есипенко |
| **Науковий керівник: д**.т.н., професор | С. Толюпа |
| **Нормоконтролер:** |  |

**Київ 2020**

**НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ**

**Факультет:** Кібербезпеки, комп’ютерної та програмної інженерії

**Кафедра:** Компютеризованих систем захисту інформації

**Освітній ступінь:** Магістр

**Спеціальність:** 125 «Кібербезпека»

**Освітньо-професійна програма**: «Безпека інформаційних і комунікаційних систем»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_\_\_\_\_\_\_\_\_ С.В. Казмірчук

«\_\_» \_\_\_\_\_\_\_\_\_ 2020 р.

**ЗАВДАННЯ**

**на виконання магістерської атестаційної роботи**

**магістранта Олексія Єсипенка**

1. Тема: Система виявлення вторгнень на об’єктах критичної інфраструктури

затверджена наказом ректора від «\_\_» \_\_\_\_\_\_\_\_\_\_ 20\_\_ № \_\_\_\_\_/ст*.*

1. Термін виконання з \_\_.\_\_.20\_\_р. по \_\_.\_\_.20\_\_р*.*
2. Вихідні дані: проаналізувати існуючі системи та методики аналізу виявлення вторгнень на об’єкти критичної інфраструктури; на основі аналізу бази даних кібератак - DARPA і KDD виділити вхідні і вихідні параметри, завдяки яким можливо побудувати ефективну систему виявлення вторгнень; проаналізувати методи інтелектуального аналізу даних в виділити серед них найбільш ефективні для нашої мети; запропонувати методи та програмне забезпечення.
3. Зміст пояснювальної записки: аналіз існуючих систем та методик аналізу і оцінки ризиків інформаційної безпеки; розробка методики системи аналізу та оцінки ризиків на основі нечіткої логіки; розробка програмного забезпечення запропонованої системи, верифікація отриманих результатів.

**КАЛЕНДАРНИЙ ПЛАН**

**виконання магістерської роботи**

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Етапи виконання магістерської роботи** | **Термін виконання етапів** | **Примітка** |
|  | Уточнення постановки задачі |  | *Виконано* |
|  | Аналіз літературних джерел |  |  |
|  | Обґрунтування вибору рішення |  |  |
|  | Збір інформації |  |  |
|  | Аналіз проблеми державного управління у сфері кібербезпеки |  |  |
|  | Особливості захисту об’єктів критичної інфраструктури від кібернетичного впливу |  |  |
|  | Теоретичні та практичні основи виявлення мережевих атак на об’єктах критичної інфраструктури |  |  |
|  | Розробка моделі системи виявлення мережевих атак на об’єктах критичної інфраструктури |  |  |
|  | Апробація роботи |  |  |
|  | Перевірка на антиплагіат |  |  |
|  | Оформлення і друк пояснювальної записки |  |  |
|  | Оформлення презентації |  |  |
|  | Отримання рецензій від рецензента |  |  |
|  | Захист в ЕК |  |  |

Магістрант О. Есипенко

(підпис, дата)

Науковий керівник С. Толюпа

(підпис, дата)

**РЕФЕРАТ**

Магістерська атестаційна робота складається зі вступу, трьох розділів, загальних висновків, списку використаних джерел, додатків і має 122 сторінки основного тексту, 42 рисунка, 25 таблиць, 13 сторінок додатків. Список використаних джерел містить 74 найменування і займає 6 сторінок. Загальний обсяг роботи 135 сторінок.

Метою роботи є побудова систем виявлення вторгнень на об’єктах критичної інфраструктури.

В роботі вирішено задачу побудови систем виявлення вторгнень на об’єкти критичної інфраструктури на основі застосування методів інтелектуального аналізу даних.

В роботі розроблено програмне забезпечення для аналізу та оцінки впливу кібератак на обєкти критичної інфраструктури.

Впроваджені методи та програмне забезпечення відносяться до галузі інформаційної безпеки і можуть бути використані для підвищення рівня захищеності.

*Ключові слова:* кіберпростір, інформаційна безпека, кібербезпека, національна безпека, об’єкти критичної інфраструктури, загрози кібернетичній безпеці.

**ЗМІСТ**

|  |  |
| --- | --- |
| ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ .......................................................... | 7 |
| ВСТУП ............................................................................................................ | 8 |
| РОЗДІЛ 1. ПРОБЛЕМИ ДЕРЖАВНОГО УПРАВЛІННЯ У СФЕРІ КІБЕРБЕЗПЕКИ …......................................................................................... | 11 |
| 1.1. Теоретичні аспекти державного регулювання кібернетичної безпеки ……………………………………………………….. | 11 |
| 1.2 Аналітичний огляд стратегічних пріоритетів та обґрунтування актуальних проблем й можливих методів удосконалення державного управління у галузі кібербезпеки ………………….. | 17 |
| РОЗДІЛ 2. ОСОБЛИВОСТІ ЗАХИСТУ ОБ’ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ ВІД КІБЕРНЕТИЧНОГО ВПЛИВУ …………………. | 22 |
| 2.1. Міжнародний досвід формування й реалізації системи захисту критичної інфраструктури держави ……………………… | 22 |
| 2.2. Сучасний стан та проблеми захисту критичної інфраструктури в Україні …………………………………………………….. | 24 |
| 2.3. Ідентифікація структурних елементів державної критичної інфраструктури …………………………………………………… | 28 |
| 2.4. Способи та шляхи досягнення цілей стратегії захисту критичної інфраструктури …………………………………………… | 34 |
| РОЗДІЛ 3. ТЕОРЕТИЧНІ ТА ПРАКТИЧНІ ОСНОВИ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АТАК НА ОБ’ЄКТАХ КРИТИЧНОЇ ІНФРАСТРУКТУРИ ………………………………………………………………………... | 38 |
| 3.1. Віддалені мережеві атаки. Поняття віддаленої мережевої атаки ……………………………………………………………… | 38 |
| 3.2. Виявлення мережевих атак. Методи виявлення мережевих атак ……………………………………………………………….. | 43 |
| РОЗДІЛ 4. РОЗРОБКА МОДЕЛІ СИСТЕМИ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АТАК НА ОБ’ЄКТАХ КРИТИЧНОЇ ІНФРАСТРУКТУРИ ……… | 63 |
| 4.1. Застосування методів інтелектуального аналізу даних в завданні виявлення мережевих атак ………………………………. | 63 |
| 4.2. Основи обраних методів інтелектуального аналізу даних .. | 67 |
| 4.3. Формалізація моделі системи виявлення мережевих атак .. | 77 |
| 4.4. Формування методики застосування методів інтелектуального аналізу даних в задачі виявлення мережевих атак ………. | 85 |
| ВИСНОВОК ……………………………………………………………….. | 92 |
| СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ………………………………….. | 93 |
| Додаток А Результати експериментів по виявленню атак ……………… | 98 |
| Додаток В Програмна реалізація експерементальної частини …………. | 102 |
| Додаток Д Слайди ………………………………………………………… | 110 |

**ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ**

|  |  |  |
| --- | --- | --- |
| АС | − | автоматизована система; |
| БД | − | база даних; |
| ВД | − | вихідні дані; |
| ІБ | − | інформаційна безпека; |
| ІР | − | інформаційні ресурси; |
| ІКС | − | інформаційно-комунікаційна система; |
| ІТ | − | інформаційні технології; |
| ІТС | − | інформаційні телекомунікаційні системи; |
| КМУ | − | Кабінет Міністрів України |
| НБ | − | національна безпека |
| РНБО | − | Рада національної безпеки й оборони |
| ПЗ | − | програмне забезпечення; |
| ISO | − | International Organization for Standardization. |
| МГК | − | метод головних компонент |
| ФА | − | факторний аналіз |
| ММП | − | метод максимальної правдоподібності |
| МК | − | методи кластеризації |
| МПНП | − | метод пошуку найкращої проекції |
| НМГК | − | нелінійний метод головних компонент |
| МСР | − | методи скорочення розмірності |
| МГК | − | метод головних кривих |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**ВСТУП**

**Актуальність**. Безпечне існування громадянського суспільства і держави безпосередньо залежить від діяльності та умов функціонування об’єктів її інфраструктури, а також різноманітних систем та мереж, що складають інформаційно-комунікаційне середовище. Вагома частина систем, об’єктів і ресурсів складають таку важливу для функціонування економіки, суспільства і держави нішу, що навіть, здавалося, незначне їх пошкодження може призвести до негативних наслідків на загальнонаціональному рівні. Перераховане вище відносять до критичної інфраструктури, і саме їх захист має бути першочерговим завданням державного управління для сучасної країни.

**Відомі підходи до вирішення поставленої задачі.** На сьогоднішній день існує велика кількість різноманітних методів та засобів захисту від кібератак на об’єктах інформаційної діяльності. Нажаль на теперішній час не існує абсолютно універсального методу протидії кібератакам і тому виникає необхідність комплексного підходу до вирішення даної задачі. Застосування методів інтелектуального аналізу даних та штучного інтелекту дає можливість підвищити ефективність протидії вторгненням і захистити об’єкти від потенційних порушників. Часто перед керівництвом об’єкту для підвищення ефективності вирішення завдань захисту виникає питання про вибір відповідного методу та засобу, що задовольняє поточним вимогам інформаційної безпеки. Існуючі на даний момент методи не задовільняють вимогам сьогодення і тому виникає необхідність постійного удосконалення систем захисту. Вирішенням проблем кібербезпеки та протидії загрозам займаються такі вітчизняні та закордонні вчені: Г. Вільський, Ф. Каргл, В. Лахно, Дж. Петі, зокрема питанням захисту протидії інцидентам та загрозам державним інформаційним ресурсам присвячені роботи Корченка О.Г., Гнатюка С.О., Бурячка В.Л., Юдіна О.К., Бучика С.С., Дубова Д.В., Горбенко І.Д., Кузнецова О.О. та ін..

**Метою роботи є** побудова ефективної системи виявлення вторгнень на об’єкти критичної інфраструктури на основі застосування методів інтелектуального аналізу даних.

Для досягнення поставленої мети вирішуються такі **задачі:**

* дослідження відомих методів протидії вторгненням на об’єкти критичної інфраструктури;
* впровадження методів інтелектуального аналізу даних та їх комбінаторіка для ефективної протидії атакам;
* розробка програмного забезпечення на основі впровадження запропонованих методів.

**Галузь застосування**. Запропоновані методи та програмне забезпечення відносяться до галузі інформаційної безпеки і можуть бути використані для підвищення рівня захищеності ОКІ за рахунок використання методів інтелектуального аналізу даних.

**Об’єктом дослідження** є процес захисту об’єктів критичної інфраструктури.

**Предметом дослідження** є застосування методів інтелектуального аналізу даних для побудови ефективної системи протидії вторгненням на ОКІ.

**Методи дослідження** базуються на основі методів штучного інтелекту й методів інтелектуального аналізу даних та об’єктно-орієнтованого програмування (для програмної реалізації запропонованого методу).

**Новизна одержаних результатів полягає в наступному:**

* на основі застосування методів штучного інтелекту та методів інтелектуального аналізу даних, розроблена ефективна система протидії вторгненням на об’єкти критичної інфраструктури.

**Практичне значення отриманих результатів:**

* розроблено інструмент для аналізу і оцінки кібернетичних впливів на об’єкти критичної інфраструктури відповідно до вітчизняних та міжнародних нормативно правових актів;
* розроблена система протидії вторгненням на ОКІ з використанням розробленого програмного продукту.

**Апробація.** Основні положення роботи доповідалися та обговорювалися на таких конференціях:

С. Толюпа, Берестовенко Г., Єсипенко О. Системи виявлення вторгнень в інформаційні системи. Матеріали ІІ Міжнародної науково-практичної конференції “Проблеми кібербезпеки інформаційно- телекомунікаційних систем” (PCSIТS). – Київ. - 2019. – с. 167-168.

Толюпа С., Єсипенко О., Маляренко Д. Застосування математичної теорії ігор при побудові систем управління інформаційною безпекою. Матеріали доповідей. VІ Міжнародна науково-практична конференція «Інформаційні технології та взаємодії» (IT&I 2019). – с. 313-317.

**РОЗДІЛ 1****.** **ПРОБЛЕМИ ДЕРЖАВНОГО УПРАВЛІННЯ У СФЕРІ КІБЕРБЕЗПЕКИ**

**1.1. Теоретичні аспекти державного регулювання кібернетичної безпеки**

Наукові дослідження з питань забезпечення національної безпеки, надважливою складовою якої є кібербезпека держави показують, що саме вона є тією комплексною проблемою, яка передусім стосується економіки країни, розвитку інфраструктури електронних комунікацій, технологій захисту державних інформаційних ресурсів, а також об’єктів критичної інфраструктури. Оскільки термін «кібербезпека» є багатогранним, його розглядають з різних аспектів – політичного, технологічного, економічного, правового, соціального тощо. Захист державних інтересів, у політичному контексті, є першоосновою забезпечення національної безпеки країни, що і пояснює потребу створення та невпинного розвитку потужної кібернетичної безпеки.

В Україні кібербезпека забезпечується з урахуванням положень, що закладені в таких документах нормативно-правової бази у даній сфері: Конституція України, Закон України “Про основи національної безпеки”, Стратегія національної безпеки України, Стратегія кібербезпеки України, Рішення Ради національної безпеки і оборони України від 28 квітня 2014 р. «Про заходи щодо вдосконалення формування та реалізації державної політики у сфері інформаційної безпеки України», що введений в дію Указом Президента України від 1 травня 2014 р. № 449, а також Конвенція «Про кіберзлочинність», ратифікована Законом України від 7 вересня 2005 р., Постанова КМУ №518 від 19 червня 2019 року. Нормативно-правова база з у галузі державного регулювання з питань кібернетичної безпеки детальніше описана на рисунку 1.1.

З урахуванням вищезазначеного, основним пріоритетним завданням кібербезпеки, як складової національної безпеки, за Стратегією кібербезпеки України, являється забезпечення національних інтересів України, що порушуються в результаті сучасних кіберзагроз. За Законом України «Про основи національної безпеки», такими інтересами є:



Рис. 1.1. Базові документи з кібернетичної безпеки України

* гарантування конституційних прав і свобод людини і громадянина;
* розвиток громадянського суспільства, його демократичних інститутів; захист державного суверенітету, територіальної цілісності та недоторканності державних кордонів, недопущення втручання у внутрішні справи України;
* зміцнення політичної і соціальної стабільності в суспільстві;
* забезпечення розвитку і функціонування української мови як державної в усіх сферах суспільного життя на всій території України, гарантування вільного розвитку, використання і захисту російської, інших мов національних меншин України; створення конкурентоспроможної, соціально орієнтованої ринкової економіки та забезпечення постійного зростання рівня життя і добробуту населення;
* збереження та зміцнення науково-технологічного потенціалу, утвердження інноваційної моделі розвитку;
* забезпечення екологічно та техногенно безпечних умов життєдіяльності громадян і суспільства, збереження навколишнього природного середовища та раціональне використання природних ресурсів; розвиток духовності, моральних засад, інтелектуального потенціалу українського народу, зміцнення фізичного здоров'я нації, створення умов для розширеного відтворення населення, збереження та зміцнення інституту сім’ї; створення умов для повноцінного життя і розвитку дитини, її безпеки, благополуччя, зростання у безпечному сімейному оточенні;
* інтеграція України в європейський політичний, економічний, правовий простір з метою набуття членства в Європейському Союзі та в євроатлантичний безпековий простір з метою набуття членства в Організації Північноатлантичного договору; розвиток рівноправних взаємовигідних відносин з іншими державами світу в інтересах України [1].

У зв’язку з тими умовами, що склалися в Україні в 2014 - 2018 років, державі довелось набути певного досвіду протиборства з Російською Федерацією на теренах кіберпростору, що використовувався країною-агресором для досягнення власних цілей у “гібридній війні” проти України. Хоч Україна і відмовилася від політики позаблоковості та багато спрямованості на шляху до формування політики приєднання до НАТО та прозахідного геополітичного й цивілізаційного вибору, нашій державі просто необхідно формувати пріоритети своєї політики щодо міжнародного та національного кіберпростору [2].

Від так кібернетична безпека стала однією з найвагоміших складових національної безпеки України, підтримання оптимального стану якої неможливе без розроблення загальнодержавної системи, що буде ґрунтуватися на толерантного ставлення до норм і принципів міжнародного права, захисті першочергових цінностей, визначених чинним законодавством України, а також забезпеченні національних інтересів у кіберпросторі.

Покладаючись на досвід науковців, неможливо не погодитись з тим, що національна система кібернетичної безпеки повинна бути рушієм взаємодії її суб’єктів у цілях об’єднання спецслужб, державних та правоохоронних органів, що здійснюють регулювання у сфері телекомунікацій та захисту інформації (рис. 1.2).

Рис. 1.2. Національна система кібернетичної безпеки

А головною метою державного управління у даній сфері в свою чергу, має бути розробка та застосування всіх можливих методів своєчасного виявлення, припинення та запобігання загрозам кібернетичного характеру, ліквідація передумов до їх настання, щоб мінімізувати негативні наслідки у результаті їх настання. Це дійсно важливо, адже на сучасному етапі існує низка реальних та потенційних загроз національній безпеці України як у зовнішньополітичній та внутрішньополітичній сфері, так і у сфері державної безпеки, воєнній, економічній, науково-технологічній, соціальній, гуманітарній, екологічній та сфері цивільного захисту [3].

В [4] відмічено, що в найповнішому розумінні середовище кіберпростору утворене в результаті взаємодії інформаційних, автоматизованих, телекомунікаційних системам. При цьому ж кібернетикний простір ще й надає можливості здійснення комунікацій та реалізовує розвиток суспільних відносин, що стає можливим внаслідок функціонування спільних (з’єднаних) комунікаційних систем та забезпечення електронних комунікацій із використанням глобальної мережі передачі даних. Відповідно до Стратегії кібербезпеки України забезпечення кібернетичної безпеки є станом захищеності життєво важливих інтересів людини і громадянина, суспільства та держави в кіберпросторі, що досягається комплексним застосуванням сукупності правових, організаційних, інформаційних заходів. Даний стан досягається за допомогою сукупності активних захисних та розвідувальних дій, що у процесі інформаційного протистояння діяльністю поодиноких інсайдерів чи організованих кібернетичних угруповань розгортаються навколо ІР, ІКТ і ІТС (рис. 1.3).

Вище викладений механізм створення процесів впливу націлений на досягнення та утримання потенційними протиборчими сторонами переваги у протидії новим загрозам безпеці для власних об’єктів особливо важливої фізичної, інформаційної та кіберінфраструктури. Під котроль останньої складові фізичної інфраструктури держави переходять принципом моделі переходу (рис. 1.4).



Рис. 1.3. Механізм та цілі здійснення кібернетичних впливів

З огляду на вищевикладене, можна констатувати той факт, що термін «кібербезпека» є дуже неоднозначним та багатогранним. Таким чином, його політичний контекст являє собою складову національної безпеки держави виходячи із викликів та загроз сучасності. З технічної точки зору, кібербезпека є захистом кібернетичного простору від реальних та потенційних кіберзагроз та кібератак, що можуть порушити основні властивості інформації, що циркулює в ньому. Завдяки потужним можливостям кібербезпеки з погляду економічної сфери, від її стану залежить розвиток економічного потенціалу країни, особливо у сфері передових цифрових та IT–технологій.

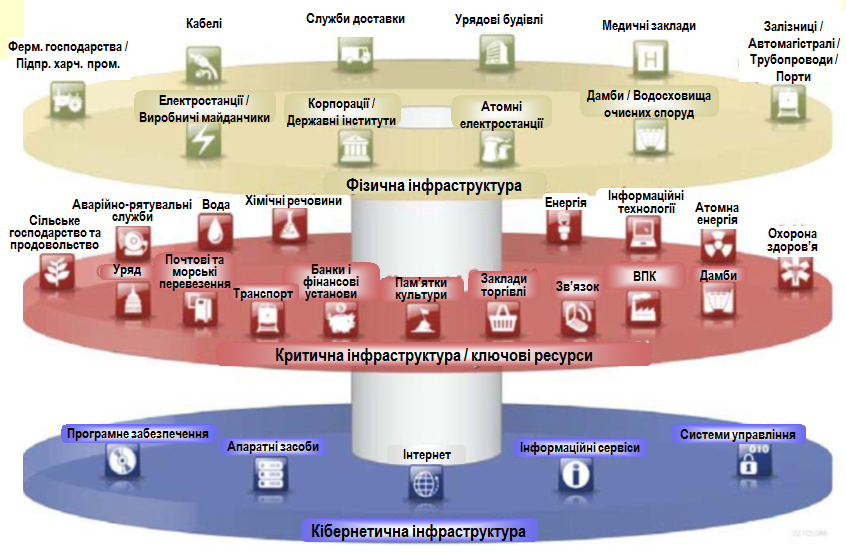


Рис. 1.4. Перехід складових фізичної інфраструктури під контроль

кібернетичної безпеки

Підсумовуючи термінологічні особливості терміну «кібербезпека», можна стверджувати, що це поняття не зводиться тільки до комп’ютерних та інформаційно-комунікаційних систем. Кібербезпека являється інформаційним, організаційно-правовим, науково-технічним забезпечення інформаційно-комунікаційної безпеки за кількома напрямами одночасно: адміністративна безпека інформаційно-телекомунікаційних систем; технічне забезпечення систем безпеки, системне та програмне забезпечення, криптографічний захист інформації тощо, які в сукупності формують оптимальний захист кібернетичного простору в цілому.

**1.2 Аналітичний огляд стратегічних пріоритетів та обґрунтування актуальних проблем й можливих методів удосконалення державного управління у галузі кібербезпеки**

Введення в дію рішення Ради національної безпеки і оборони України «Про Стратегію кібербезпеки України» у 2016 році стало поштовхом до обґрунтування основної мети Стратегії – створення умов для безпечного функціонування кіберпростору, його використання в інтересах особи, суспільства і держави. Досягнення поставлених цілей передбачає посилення здатності суб'єктів в галузі безпеки і оборони ведення ефективної боротьби з різноманітними кіберзагрозами воєнного характеру, кібертероризмом та кіберзлочинністю, що неможливе без тісної міжнародної співпраці. Стратегія кібербезпеки України переслідує мету створити умови кібернетичного захисту державних електронних інформаційних ресурсів та загалом – всієї інформаційної інфраструктури, порушення нормального функціонування якої створить негативний вплив на стан національної безпеки і оборони України.

Надзвичайно важливим чинником усвідомлення суті державного управління є розуміння його принципів. Вищезгадані фундаментальні принципи, що лягли в основу формування Стратегії є тими засадами державного управління сектору кібернетичної безпеки України, що являють собою певні закономірності та взаємовідносини і взаємозв’язки, які формують здійснення цього управління. Значення принципів політичної спрямованості полягає в тому, що вони завдяки здійснюваному свідомо узагальненню, у стислому вигляді відображають найсуттєвіші характеристики вказаної галузі права. З огляду на те, що формулювання принципів є результатом творчої діяльність практиків та фахівців у сфері захисту національних інтересів держави, саме вони є віддзеркаленням першочергових цінностей, що дозволяють побачити нинішній стан справ та точки вимірювання прогресу з часом. Виділення і характеристика принципів державного управ­ління вимагають глибоких і цілеспрямованих досліджень. Запропоновані принципи базуються на об'єктивних закономірностях державного управ­ління та отримали практичне застосування в управлінській діяльності. Таким чином, серед загальних принципів можна виділити такі: законність, демократизм, публічність, єдиноначальність і колегіальність, централізація і децентралізація, плановість, ефективність [5].

Результати досліджень фахівців у сфері кібербезпеки показують, що в умовах постійно зростаючого розвитку інформаційних та комп’ютерних систем, у більшості провідних країн світу стійка посилюється імпульс розширення діапазону кібератак з метою порушення конфіденційності, цілісності і доступності державних інформаційних ресурсів. При цьому під ризиковий стан підпадає інформація, що циркулює на об’єктах критичної інфраструктури. Зазіхання реалізовуються через інформаційно-телекомунікаційні системи, а також автоматизовані системи керування, широко використовувані в роботі структур економіки, органів державної влади та повсякденному житті громадян України.

Згідно з даними, зібраними у результаті проведення аналітичних досліджень проблем і ризиків у кібернетичному просторі, економічні втрати від усіх глобальних кібернападів сягають $53 млрд, в тому числі $850 млн від недавнього нападу вірусу "NotPetya", що уразив у липні український фінансовий і державний сектор та створив проблеми в інших країнах. При цьому зауважується, що кіберзагрози для фінансових установ зростають, масштаби та частота нападів збільшуються, а природа атак тільки вдосконалюється. Кіберзлочинці маскують їх під легальне ПЗ у банківській сфері, що спричиняє чималих фінансових втрат.

Лихо не оминає і Україну, про що свідчать неодноразові спроби реалізацій кібератак в інформаційному просторі нашої держави. Тільки 2016 році в результаті кібератаки за допомогою вірусу Petya, що торкнулася України, ряду країни Європи та Азії, постраждали кілька приватних українських та польських компаній, які працювали в умовах партнерства. За інформацією українських засобів масової інформації, хакерських атак 27 червня зазнали банки "ТАС", Ощадбанк, "Південний", ОТП, Укрзалізниця, аеропорт Бориспіль, мережа "Епіцентр", "Нова пошта", ДТЕК, "Укренерго", "Київенерго", "Укрпошта", "Укртелеком", мережа заправок ТНК, ПБГ "Ковальська", ТРК "Люкс", "Київводоканал", видання "Кореспондент", сайт Кабміну і ряду міністерств, Пенсійний фонд, Київська міська державна адміністрація, а також "Запоріжжяобленерго", "Дніпроенерго", Дніпровська електроенергетична система та багато інших важливих структур.

Причиною беззаперечної актуальності кіберзагроз, що є причиною здійснення численних кібератак на інформаційні ресурси зокрема нашої країни, є певніфактори, що впливають на стан захищеності загальноважливих для країни ресурсів, а саме:

* недостатній рівень захищеності критичної інформаційної інфраструктури, державних електронних інформаційних ресурсів;
* невідповідність інфраструктури електронних комунікацій держави, рівня її розвитку та захищеності сучасним вимогам; безсистемність заходів кіберзахисту критичної інформаційної інфраструктури;
* недостатня ефективність суб’єктів сектору безпеки і оборони України у протидії кіберзагрозам воєнного, кримінального, терористичного та іншого характеру;
* недостатній розвиток організаційно-технічної інфраструктури забезпечення кібербезпеки та кіберзахисту критичної інформаційної інфраструктури та державних електронних інформаційних ресурсів;
* недостатній рівень координації, взаємодії та інформаційного обміну між суб’єктами забезпечення кібербезпеки тощо.

З огляду на подане вище, є всі підстави стверджувати, що попри прагнення української держави побудувати комплексну систему опору внутрішнім і зовнішнім загрозам власному інформаційному просторі, станом на зараз існує ряд проблем, що перешкоджають їй у цьому. До таких проблем слід віднести:

* уразливість інформаційної сфери України через надмірно широке впровадження до неї західних програмних продуктів (зокрема фірми Microsoft) та використання матеріально-технічних засобів іноземного виробництва;
* можливість недержавних суб’єктів та неавторизованих (індивідуальних) користувачів виступити у ролі гравців в кіберпросторі та проблематичність щодо їх виявлення;
* деградацію науково-технічного потенціалу України, нерозвиненість національної інноваційної системи в інфосфері та низький рівень конкурентоспроможності в ній;
* непрозорість розподілу обов’язків між певними відомствами, правоохоронними органами і силовими структурами України, що спеціалізуються на проблемах кіберзахисту та їх незадовільне кадрове забезпечення кваліфікованими фахівцями з цих питань;
* можливість скритного (прихованого) проведення протиборчими сторонами кібератак та кібероперацій у кіберпросторах один одного;
* складність структури ІКТ та національного кіберпростору;
* ускладненість щодо розмежування воєнних і цивільних об’єктів критичної інфраструктури держави у кіберпросторі [5].

Очевидно, що подібний стан речей тільки заохочує зловмисників та кіберзлочинці до несанкціонованих втручань в інформаційний простір нашої держави.

Можливими варіантами методів та дій для поліпшення ситуації, що склалася може бути:

* формування власних систем захисту національної інфосфери від стороннього кібервпливу та вдосконалення існуючої нормативно-правової бази;
* удосконалення методів, засобів та способів отримання суспільно значущої інформації з відкритих, відносно відкритих і закритих електронних джерел, оцінювання рівня захищеності власних ІТС і мереж від впливу внутрішніх та зовнішніх кібернетичних втручань і загроз, а також злому систем захисту протиборчих;
* забезпечення автоматизації процесів пошуку та збору інформації (ІРМ, відомостей, даних та знань) у відкритих і відносно відкритих та її добування із закритих електронних джерел, а також накопичення і оброблення такої інформації та обміну нею.

Без вдосконалення міжнародних правил поведінки держав у кібернетичному просторі та відповідного нормативно-правової бази, а також впровадження програм підтримки вітчизняної інноваційної продукції, створення національної системи кібербезпеки залишатиметься на рівні ідеї без реального подальшого її втілення у життя.

**Висновки до першого розділу**

Беручи до уваги переконання авторитетних фахівців, кібернетичну безпеку слід виділяти як самостійну складову національної безпеки України, що у той самий час тісно пов’язана з іншими елементами її системи, зокрема з інформаційною безпекою. Формування та функціонування дієздатної національної системи кібербезпеки передбачає спільну участь як військових, так і цивільних державних інститутів, що здійснюється на планових засадах, у рамках виконання положень національних стратегічних документів, що регулюють проблематику безпечного кіберпростору, забезпечення кіберзахисту об’єктів критичної інфраструктури, державних інформаційних ресурсів, боротьби з кіберзлочинністю тощо.

Основою національної системи кібербезпеки є її організаційно-технічна модель, яка призначена для забезпечення безперебійного функціонування автоматизованих систем органів військового та державного управління і включає такі складові: постійний моніторинг кіберпростору, комплексний захист інформації, оперативне оповіщення про кібератаки або кіберзагрози та протистояння їм; стандарти управління кібербезпекою. Впровадження сучасної організаційнотехнічної моделі кібербезпеки дасть змогу покращити координацію діяльності сектору безпеки і оборони України, підвищити його техніко-технологічні можливості в рамках цілісної управлінської системи з метою боротьби з кіберзагрозами незалежно від способу, мети та суб’єкта їх реалізації, цілеспрямовано орієнтувати науковий та людський потенціал державних органів і приватного сектору на забезпечення кібербезпеки.

**РОЗДІЛ 2****.** **ОСОБЛИВОСТІ ЗАХИСТУ ОБ’ЄКТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ ВІД КІБЕРНЕТИЧНОГО ВПЛИВУ**

**2.1. Міжнародний досвід формування й реалізації системи захисту критичної інфраструктури держави**

Згідно даних, що викладені в численних наукових публікаціях та аналітичних матеріалах, можна дійти до висновку, що організація діяльності стосовно захисту критичної інфраструктури у різних країнах реалізовується по-різному. Така діяльність іноді має несистематизований характер, коли заходи здійснюються у неофіційному режимі, тоді як в деяких державах – організаційна модель є визначеною та складає певну структуру, а заходи – цілеспрямованими та системними.

Як показує світова практика, процес становлення нормативно-правової бази в галузі захисту критичної інфраструктури є достатньо трудомістким і тривалим. При цьому ті чи інші законодавчі положення різних держав із питань захисту критичних інфраструктур є неузгодженими, до того ж спостерігається недосконалість механізмів віднесення об’єктів до критичної інфраструктури та інші проблеми. Тому кожна держава визначає свою критичну інфраструктуру, зважаючи на критичність окремих секторів та важливість певних послуг для економіки держави та безпеки її суспільства. Таким чином, поняття «критична інфраструктура» не має сталого тлумачення і кожна країна вкладає у нього власний зміст та специфіку. Це зумовлено тим, що національні потреби й проблеми дуже відрізняються залежно від територіального розташування, рівня розвитку держави та інших специфічних факторів, що складають основну перешкоду на шляху до міжнародної стандартизації у сфері захисту об’єктів критичної інфраструктури. Попри всі ці відмінності, все-ж-таки простежується спільна риса критичної інфраструктури різних держав світу, а саме: її беззаперечне значення для безпеки громадян, суспільства й держави [6].

Досвід США, Великої Британії, Ізраїлю, Німеччини свідчить про створення системи кібербезпеки з єдиним координуючим органом, який протягом короткого проміжку часу буде здатен акумулювати сили й засоби різних державних і недержавних органів, з метою протидії та нейтралізації кібератак.

Аналіз міжнародного досвіду свідчить, що основою забезпечення захищеності й безпеки критичної інфраструктури є вирішення низки питань, з-поміж яких основними є такі:

• координація і взаємодія органів державної влади та обмін інформацією про загрози;

• організація державно-приватного партнерства у сфері безпеки;

• використання ризик-орієнтованого підходу при попередженні загроз критичній інфраструктурі.

Варто зазначити, що провідні країни світу здійснюють захист своїх національних інтересів, не обмежуючись національними кордонами. Крім національних критичних інфраструктур, розглядаються зарубіжні об’єкти, безпека яких має важливе значення для тієї чи іншої держави. На підтвердження цього можна згадати, зокрема, той факт, що на веб-сайті *WikiLeaks* був оприлюднений список компаній та установ, нібито створений дипломатичними місіями на запит Державного департаменту США у 2009 р. Тому природно (з огляду на сучасні геополітичні реалії), що, наприклад, газотранспортна система України може розглядатися європейськими і трансатлантичними партнерами як елемент критичної інфраструктури, який має загальноєвропейське значення. Цей аспект, безперечно, заслуговує на увагу під час розгляду проблем захисту національної критичної інфраструктури, вирішення питання власності на об’єкти національної газотранспортної системи, що матиме безпосередній вплив на процес формування ціни на газ для України. Для нашої держави може бути корисним досвід імплементації концепції захисту критичної інфраструктури в законодавствах деяких східноєвропейських країн [5].

**2.2. Сучасний стан та проблеми захисту критичної інфраструктури в Україні**

У наш час високий рівень впровадження новітніх технологій у світі являється визначальним фактором економічної конкурентоспроможності будь-якої держави. Звідси, розвиток саме технологічний розвиток є умовою для досягнення цілей, які визначаються національними інтересами. Разом з численними перевагами технологічного процесу, він також утворив стан певної залежності як суспільства, так і окремих людей від систем, що постачають різні види послуг: транспортні, енергетичні, фінансові, інформаційні, комунікаційні тощо. Руйнуванням саме таких систем на сьогодні може спричинити найбільш небезпечні безпекові сценарії провідним державам світу, зокрема країнам ЄС. Отже, зважаючи на обмеженість ресурсів, об’єктивну неможливість забезпечити абсолютний захист і безпеку всіх інфраструктурних систем, в багатьох країнах світу впроваджується концепція КІ, що концентрує увагу на системах, мережах та окремих об’єктах, знищення або порушення роботи яких матиме найсерйозніші негативні наслідки для національної безпеки цих країн.

У Європейському Союзі створення правових та організаційних механізмів захисту КІ було ініційовано ще у 2004р. у зверненні Європейської ради до Європейської комісії, в якому ЄК доручалося підготувати загальну стратегію захисту КІ. В жовтні 2004 р. [7]. ЄК оприлюднила офіційне повідомлення, яке містило рекомендації стосовно вжиття таких заходів усім країнам Європейського Союзу:

* розробити національну програму захисту КІ;
* забезпечувати такий рівень охорони здоров’я, технологічної безпеки, соціально-економічного благополуччя, який би гарантував «стійкість» нації до загроз;
* уніфікувати зусилля, спрямовані на захист КІ, надавши єдиному державному органу, що звітує з цього питання, функції координації дій державних органів влади, що відповідають за окремі галузі промисловості, до яких належать об’єкти КІ;
* визначити органи державної влади, відповідальні за сектори КІ, та відповідні приватні компанії;
* створити умови для ефективної взаємодії та обміну інформацією, даними і досвідом між країнами-членами ЄС, урядовими структурами та приватним сектором;
* зробити внесок у створення гармонізованої методології аналізу ризиків.

Визначення критеріїв та показників, за якими певні інфраструктури чи їх елементи можна віднести до КІ, є окремим питанням, що подається до розгляду. Пропозиції стосовно процедури та критеріїв визначення об’єктів КІ на загальноєвропейському рівні були представлені в Зеленій книзі (2005р.). В ній розглядалося 11 секторів КІ, в які було включено 37 підсекторів. Надалі, під час підготовки проекту директиви, були внесені цих 11 секторів із 29 підсекторами, а вже в ухваленій директиві ЄК згадується тільки два сектори енергетика та транспорт.

Згідно з Постановою № 563 Кабінету Міністрів України від 23 серпня 2016 р., було зроблено вагомий крок у напрямі підвищення рівня захисту інформації, що обробляється в інформаційно-телекомунікаційних системах об'єктів критичної інфраструктури держави, а саме визначено механізм, за яким відбуватиметься формування переліку таких систем. Передбачалось, що завдяки цьому механізму в державі буде затверджено перелік інформаційно-телекомунікаційних систем, що експлуатуються найважливішими для економіки та промисловості, функціонування суспільства та безпеки населення підприємствами та установами. Включені до цього переліку системи першочергово захищатимуться від кібератак [9].

Таким чином, у грудні 2017 Кабінет міністрів України схвалив концепцію системи захисту критичної інфраструктури. Згідно потребам НБ створення критичної інфраструктури, тобто сукупності об'єктів, які є стратегічно важливими для економіки і безпеки держави, суспільства, населення і порушення функціонування яких може завдати шкоди життєво важливим національним інтересам України, є беззаперечно необхідним кроком для нашої держави. Згідно з розпорядженням, метою концепції є визначення основних напрямів, механізмів та строків правового врегулювання питання захисту КІ. Наголошується, що створення системи спрямоване на забезпечення стійкості критичної інфраструктури до загроз всіх видів, включаючи загрози техногенного природного характеру, загрози, викликані протиправними діями, і інші загрози [10]. Концепцію передбачається реалізувати протягом 2017-2027 років. Міністерству економічного розвитку і торгівлі спільно із зацікавленими центральними органами виконавчої влади, іншими державними органами та установами доручається у двомісячний термін розробити і подати на розгляд Кабміну проект закону України "Про критичну інфраструктуру та її захист".

За Концепцією, основними проблемами, які потребують розв’язання, є:

* відсутність єдиної загальнодержавної системи захисту КІ;
* недостатність та неузгодженість нормативно-правового регулювання з питань захисту систем і ОКІ, зокрема відсутність спеціального закону про критичну інфраструктуру та її захист;
* відсутність державного органу, відповідального за координацію дій у сфері захисту критичної інфраструктури;
* невизначеність повноважень, завдань і відповідальності центральних органів виконавчої влади та інших державних органів у сфері захисту критичної інфраструктури, а також прав, обов’язків та відповідальності власників (розпорядників) ОКІ;
* відсутність єдиних критеріїв та методології віднесення об’єктів інфраструктури до КІ, порядку їх паспортизації та категоризації;
* відсутність єдиної методології проведення оцінки загроз КІ, а також відсутність спеціального правоохоронного органу, відповідального за проведення аналізу та оцінки загроз критичній інфраструктурі внаслідок проведення іноземними державами економічної експансії та дискримінаційної політики, недопущення заподіяння шкоди економічному і науково-технічному потенціалу держави, а також організацію та вжиття відповідних заходів протидії;
* нерозвиненість державно-приватного партнерства у сфері захисту КІ та невизначеність джерел фінансування заходів із захисту критичної інфраструктури;
* недостатній рівень міжнародного співробітництва у сфері захисту критичної інфраструктури [10].

Створення державної системи захисту критичної інфраструктури (як комплексу організаційних, нормативно-правових, інженерно-технічних, наукових та інших заходів, спрямованих на забезпечення безпеки та стійкості критичної інфраструктури) потребує нормативно-правового врегулювання основоположних принципів її функціонування, запровадження єдиних підходів до організації управління об’єктами системи на державному та місцевому рівні, визначення засад взаємодії залучених до захисту критичної інфраструктури державних органів та суб’єктів господарювання, суспільства та громадян.

Гострота проблеми захисту критичної інфраструктури полягає у тому,що жодна з систем не призначена для реагування на усі види загроз, що обумовлює відсутність системного підходу на національному рівні до захисту критичної інфраструктури, який мав би враховувати численні взаємозв’язки її елементів. При цьому жоден орган державної влади не опікується проблемами захисту критичної інфраструктури у комплексі. З огляду на це, виходячи із потреб національної безпеки і необхідності запровадження системного підходу до розв’язання проблеми на національному рівні, створення системи захисту критичної інфраструктури є одним із пріоритетів у реформуванні сектору безпеки і оборони України на сучасному етапі.

Визначені Концепцією створення державної системи захисту критичної інфраструктури актуальні проблеми у даному секторі безпеки можна доповнити численними загрозами, з-поміж якихмають місце епідемії, промислові аварії, терористичну та злочинну діяльності, кібератаки, стихійні лиха тощо. Зважаючи на це, держава має забезпечувати захист об’єктів критичної інфраструктури від усіх суттєвих загроз, техногенного та природного характеру, а також соціально-політичних загроз.

У підходах до захисту ОКІ поступово відбувалися зміни під впливом тих чи інших подій, що становлять їм суттєву загрозу. З-поміж техногенних загроз особлива увага приділяється спробам втручання в роботу автоматизованих систем управління технологічним процесом на підприємствах та об’єктах інфраструктури. Хоча досі не повідомлялося про факти руйнувань унаслідок кібератак на КІ, однак було зафіксовано численні випадки зараження автоматизованих систем управління технологічним процесом відноманітними вірусами, а також спроби втручання в роботу АС управління об’єктів газотранспортних систем (США).

Зважаючи на зростання негативних наслідків для держави, які завдаються кібератаками на інформаційну інфраструктуру органів державної влади, та на небезпеки, пов’язані з можливими атаками на промислові об’єкти, дедалі частіше лунають заклики посилити відповідальність за вчинення кіберзлочинів. До безпеки таких об’єктів, як метрополітен, стадіони, виставкові центри, вищі навчальні заклади, висуваються підвищені вимоги з безпеки, проте забезпечити їх стовідсоткову захищеність від терористичних загроз неможливо. Варто зазначити, що суттєвою частиною аналізу й оцінки загроз є вивчення взаємозв’язку між елементами КІ всередині секторів, і між різними секторами.

Зволікання з розв’язанням зазначених проблем буде не тільки гальмувати створення системи захисту критичної інфраструктури, але й створювати перешкоди для реформування та подальшого розвитку суміжних безпекових напрямів забезпечення національної безпеки.

**2.3. Ідентифікація структурних елементів державної критичної інфраструктури**

Об’єднання економік країн ЄС, їх взаємозалежність, але і необхідність протистояти сумісним або подібним погрозам, відбилися в ухваленні документа Critical Ibfrastructure protection in the fight against terrorism. У зазначеному документі критична інфраструктура визначається як фізичні засоби виробництва, інформаційні технології, мережі (транспортні, енергетичні тощо), служби і інші активи, розлад або руйнування яких мав би серйозні наслідки на здоров’я, охорону, надійність або життєвий рівень громадян або на штатне функціонування урядів в цих державах.

За таким тлумаченням до сектору КІ входять:

* державне управління, зокрема критичні служби і установи, інформаційні мережі, важливі економічні об’єкти, стратегічні об’єкти, а також культурні пам’ятники;
* енергетичні об’єкти і мережі, наприклад, електричні розподільні мережі, газопроводи, нафтопроводи, збірки пального тощо.;
* транспорт, особливо авіаційний, шосейний, залізничний, комбінований, комунікаційні вузли, а також системи управління транспортом;
* комунікаційні і інформаційні технології (наприклад, телекомунікації, радіомовні і телевізійні передавачі і мережі, інтернет);фінансова система(банкова справа, капіталові ринки, інвестування);
* виробництво, зберігання і транспорт небезпечних товарів, особливо хімічних, біологічних, радіологічних ядерних матеріалів;
* охорона здоров’я, особливо лікарні, поліклініки, установи постачання крові, лабораторії, сантехнічна рятувальна служби;
* харчова промисловість, сільське господарство, торгівля і постачання продовольством;
* вода, особливо греблі, гідроресурси [3].

До того ж, щоб визначити елемент інфраструктури держави як критичний, потрібно співвіднести його за набором критеріїв, а саме:

* територіальна досяжність негативних результатів (транснаціональний, народний, регіональний, локальний і т.д.);
* велика кількість наслідків (гуманітарних, матеріальних, економічних, політичних), збитків і втрат по відношенню до навколишнього середовища;
* часовий діапазон настання наслідків, особливо негативних (наприклад: негайно, за 24 год.) і як довго вони можуть продовжуватися (наприклад: до 24 годин, до 3 днів і т.д.).

Розробка методології ідентифікації ОКІ певної держави потребує окремого дослідження. Досвід інших країн свідчить, що завдання сформувати єдину методологію ідентифікації об’єктів, систем та сервісів, критичних на державному рівні, а також створення глобальної (на національному рівні) бази даних для ведення їх реєстру, являється досить нелегким. Складність полягає у тому, що самі об’єктів та систем є дуже неоднорідними та належать до різних секторів критичної інфраструктури. Більше того, кількість таких об’єктів надзвичайно велика, так само як і їх характеристики з огляду на всі типи можливих загроз [3].

Таким чином, в Україні визначено 18 секторів: аграрний і продовольчий, банківська система й фінанси, хімічна промисловість, комерційні об’єкти (музеї, виставки й інші місця масово го скупчення людей), критичне виробництво, дамби, оборонно-промисловий комплекс, сервіси допомоги (пожежна, медична швидка тощо), енергетика, урядові об’єкти, охорона здоров’я, інформаційні технології, зв’язок, національні пам’ятки й символи, пошта й доставка, транспортні системи, водопостачання. У кожної країни існують власні підходи до ідентифікації об’єктів критичної інфраструктури та кількості їх секторів, тоді як, до прикладу, в Канаді неї віднесено десять: харчова промисловість, фінанси, промисловість, безпека, енергетика, уряд, охорона здоров’я, інформаційні технології та зв’язок, транспорт, водопостачання. Проте в найближчій перспективі Директивою ЄС лише три сектори було визнано пріоритетними: енергетика, телекомунікації і транспорт. До сектору енергетики було включено такі системи та об’єкти: електромережі та об’єкти з генерування й передачі електроенергії; нафтовидобувна та нафтопереробна промисловість, нафтопроводи і сховища; газовидобувна промисловість, газопроводи, термінали зрідженого газу (рис. 2.1-2.2).

До сектору транспорту належать такі його види та об’єкти: автодорожній, залізничний та авіаційний транспорт; річковий флот, океанічний і морський флот; порти (рис. 2.3).



Рис. 2.1. Сектор енергетики та хімічної промисловості: нафтовидобувна та нафтопереробна промисловість, нафтопроводи і сховища; газовидобувна промисловість, газопроводи, термінали зрідженого газу



Рис. 2.2. Сектор енергетики: електромережі та об’єкти з генерування й передачі електроенергії



Рис. 2.3. Сектор транспорту: автодорожній, залізничний та авіаційний транспорт; річковий флот, океанічний і морський флот, порти

До сектору телекомунікацій долучились такі системи та об’єкти: інформаційні технології та зв’язок, телекомунікаційні системи та мережі державного управління (рис. 2.4).



Рис. 2.4. Сектор телекомунікацій: інформаційні технології та зв’язок, телекомунікаційні системи та мережі державного управління

У процесі визначення елементів критичної інфраструктури в нашій державі будується ієрархія критеріїв, що охоплює такі основні групи:

* економічна безпека (значна частка продукції на ринку, велика кількість зайнятих співробітників, крупний платник податків);
* безпека життєдіяльності та здоров’я населення (забезпечення роботи аварійно-рятувальних служб, екстреної допомоги населенню;
* недо пущення техногенних аварій регіонального або національного масштабів);
* державна безпека та оборона (недопущення порушень в управлінні державою, зниження боєздатності збройних сил, розголошення таємної інформації);
* національна самоповага та імідж держави (збереження культурних цінностей, авторитету держави).

Крім того, під час визначення структурних елементів критичної інфраструктури беруться до уваги певні чинники та характеристики, а саме:

* масштаб (географічне охоплення території, для якої втрата елемента критичної інфраструктури викликає значну шкоду) – міжнародний, національний, регіональний або територіальний;
* тяжкість можливих наслідків за такими показниками:

– вплив на населення (кількість постраждалих, загиблих, осіб, які отримали серйозні травми, а також чисельність евакуйованого населення);

– економічна шкода (вплив на ВВП, розмір економічних втрат – і прямих, і непрямих);

– екологічна шкода (вплив на населення і навколишнє середовище);

– взаємозв’язок з іншими елементами критичної інфраструктури;

– політичний ефект (втрата впевненості в дієздатності влади);

– тривалість впливу (як саме й коли виявлятимуться збитки, по в’язані із втратою чи відмовою об’єктів критичної інфраструктури) [9].

Здійснення ідентифікації ОКІ дозволяє встановити диференційовані вимоги до забезпечення безпеки цих об’єктів, враховуючи ступінь потенційної небезпеки здійснення акту незаконного втручання або теракту з усіма ймовірними наслідками.

Порядок проведення ідентифікації потенційно небезпечних об’єктів встановлено відповідною методикою. Одним з етапів процедури ідентифікації є виявлення за результатами аналізу джерел небезпеки, які за певних умов (аварії, порушення режиму експлуатації, виникнення природних небезпечних явищ тощо) можуть спричинити надзвичайну ситуацію, а також оцінка можливих наслідків надзвичайної си туації для кожного із джерел небезпеки (кількість загиблих, постраждалих, тих, яким порушено умови життєдіяльності, матеріальні збитки) з використанням відповідної методики.

При встановленні рівня можливих надзвичайних ситуацій визначається таке:

* територіальне поширення ймовірних надзвичайних ситуацій;
* кількість осіб, умови життєдіяльн ості яких можуть бути порушен і в результаті можливої аварії на об’єкті;
* збитки від наслідків можливих надзвичайних ситуацій.

Таким чином, основною характеристикою, що враховується під час визначення та віднесення об’єктів до категорії «потенційно небезпечні об’єкти» є розмір наслідків можливої аварії на об’єкті. Зважаючи на загальноприйняте визначення критичної інфраструктури, частина переліку потенційно небезпечних об’єктів може увійти й до переліку ОКІ [3].

**2.4. Способи та шляхи досягнення цілей стратегії захисту критичної інфраструктури**

Згідно [10] Концепції створення державної системи захисту критичної інфраструктури, основною метою державного управління у цій сфері повинно бути визначення основних напрямів, механізмів і строків комплексного правового врегулювання питання захисту критичної інфраструктури на шляху до створення національної системи захисту критичної інфраструктури.

Безумовно, щоб для досягнення мети Концепції необхідно слідувати наступним принципам:

* суб’єкти державної системи захисту критичної інфраструктури мають безперешкодно та оперативно взаємодіяти між собою;
* рівень власних спроможностей громадян, суб’єктів господарювання, установ та організацій, що є вразливими до припинення або погіршення функціонування критичної інфраструктури необхідно суттєво підвищити;
* будівництво та експлуатація об’єктів критичної інфраструктури має відбуватись з дотриманням інженерно-технічних, організаційних, експлуатаційних вимог тощо, у різних режимах функціонування критичної інфраструктури;
* необхідно забезпечити обмін інформацією між суб’єктами державної системи захисту критичної інфраструктури про загрози критичній інфраструктурі;
* здійснення державно-приватного партнерства у сфері захисту критичної інфраструктури має бути обов’язковим;
* потрібно сприяти міжнародному співробітництву у сфері захисту критичної інфраструктури з урахуванням глобальних та регіональних безпекових процесів;
* необхідно розроблювати та удосконалювати законодавство у сфері захисту критичної інфраструктури, враховуючи норми та стандарти НАТО, а також положення річних національних програм під егідою Комісії Україна - НАТО на відповідні роки.

Врегулювання наявних проблем у галузі забезпечення захисту КІ потребує тільки комплексного вдосконалення правової основи діяльності органів державної влади, суб’єктів господарювання та суспільства загалом. Разом з тим, їх вирішення практично неможливе без створення організаційно-інституційної структури функціонування державної системи управління безпекою критичної інфраструктури, а також без визначення відповідальності, завдань та повноважень зацікавлених суб’єктів державної системи захисту критичної інфраструктури. Зокрема, вдосконалення правової основи захисту об’єктів критичної інфраструктури потребує визначення базових принципів її функціонування, запровадження єдиних підходів до управління елементами системи на державному, місцевому та корпоративному рівнях, визначення засад взаємодії залучених державних органів та приватного сектору, суспільства та громадян, а також формування єдиної методологічнотермінологічної бази усіма суб’єктами державної системи захисту критичної інфраструктури [4].

Невідкладними заходами на шляху до формування структури системи захисту критичної інфраструктури на загальнодержавному рівні є так:

* визначення органу, відповідального за формування та реалізацію державної політики у сфері захисту критичної інфраструктури;
* затвердження переліку секторів критичної інфраструктури з визначенням відповідальних за їх захист державних органів, а також переліку об’єктів критичної інфраструктури;
* визначення повноважень, завдань та відповідальності державних органів у сфері захисту критичної інфраструктури, а також прав, обов’язків та відповідальності власників (розпорядників) об’єктів критичної інфраструктури;
* організацію взаємодії суб’єктів державної системи захисту критичної інфраструктури, обміну інформацією між ними про загрози критичній інфраструктурі, створення мережі ситуаційних центрів;
* розроблення і затвердження єдиної методології проведення оцінки загроз критичній інфраструктурі та визначення критеріїв віднесення до неї об’єктів інфраструктури;
* створення системи підготовки та перепідготовки кадрів у сфері захисту критичної інфраструктури [3].

Визначення повноважень, завдань та відповідальності суб’єктів державної системи захисту критичної інфраструктури повинно здійснюватись шляхом визначення державних органів, відповідальних за забезпечення здійснення заходів щодо пріоритетних секторів критичної інфраструктури, зокрема паливно-енергетичного комплексу, транспорту, мережі життєзабезпечення, телекомунікації та зв’язку, хімічної промисловості, харчової промисловості та агропромислового комплексу.

**Висновки до другого розділу**

Не так давно ухвалений парламентом [закон](http://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=55657) “Про основні засади забезпечення кібербезпеки в Україні” вводить та визначає різні терміни та поняття, пов’язані з кібербезпекою. До цього часу поняття кібербезпеки та механізми її забезпечення не були визначені в українському законодавстві. Даний заакон визначає відповідальних за кібербезпеку та дає розуміння того, які та у якій послідовності закони та нормативно-правові акти мають бути прийняті у цій сфері.

Основними суб'єктами Національної системи кібербезпеки є: Державна служба спецзв'язку і захисту інформації, Національна поліція, Служба безпеки України, Міністерство оборони, Генеральний штаб ЗСУ, розвідувальні органи, Національний банк України.

Основоположники прийнятого Закону про захист ОКІ і значна частина експертів вважають, що реалізація закону допоможе під егідою держави і в тісній співпраці з приватним сектором і громадянським суспільством запровадити комплексний підхід до визначення основних принципів формування держполітики у сфері кібербезпеки і створити умови для забезпечення кіберзахисту інформаційної інфраструктури країни.

Отже, закон встановлює правила гри на полі кіберзахисту і містить дуже багато планів та намірів. Експерти загалом схвально відгукуються про його ухвалення – як базового документу, хоч у ньому й є ціла низка недопрацювань. Водночас на думку аналітиків, ще зарано говорити про ефективність цього закону для підвищення рівня кібербезпеки, адже офіційно в дію він буде впроваджений згодом, а до того часу потрібно ще придбати відповідне обладнання, сформувати відповідні установи та найняти фахівців високого класу для забезпечення описаних норм закону.

Одна з основних категорій об’єктів, визначена в національному законодавстві й така, що може бути використана при визначенні критично важливих об’єктів та інфраструктури в Україні, – це потенційно небезпечні об’єкти. У межах Урядової інформаційно-аналітичної системи з питань надзвичайних ситуацій створено Державний реєстр потенційно небезпечних об’єктів. Реєстрації на безоплатній основі підлягають всі (незалежно від форми власності) розташовані на території України небезпечні об’єкти, на яких існує загроза виникнення надзвичайних ситуацій. Метою створення реєстру було введення державно го обліку потенційно небезпечних об’єктів та інформаційного забезпечення процесів підготовки управлінських рішень і виконання зобов’язань України згідно з міжнародними договорами щодо запобігання та ліквідації наслідків надзвичайних ситуацій, у т.ч. транскордонного характеру, пов’язаних із функціонуванням небезпечних об’єктів.

**РОЗДІЛ 3.** **ТЕОРЕТИЧНІ ТА ПРАКТИЧНІ ОСНОВИ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АТАК НА ОБ’ЄКТАХ КРИТИЧНОЇ ІНФРАСТРУКТУРИ**

**3.1. Віддалені мережеві атаки.** **Поняття віддаленої мережевої атаки**

Поняття інформаційної безпеки грунтується на збереженні трьох основних властивостей інформації - конфіденційності, цілісності та доступності.

Конфіденційність інформації - це характеристика інформації, яка вказує на обмеження кола суб'єктів, що мають певний доступ до даної інформації, і забезпечує вимога збереження зазначеної інформації в таємниці від суб'єктів, які не мають доступу до неї.

Цілісність інформації - це властивість інформації, яка полягає в її збереження в неспотвореному вигляді.

Доступність інформації - це здатність системи забезпечувати своєчасний доступ певного безлічі суб'єктів до інформації або до містить її ресурсу.

Загроза безпеки обчислювальної системи - потенційно можлива подія, процес, дію, явище, яке може привести до нанесення збитку ресурсів обчислювальної системи. Загроза безпеки вимагає реалізації уразливості об'єкта для нанесення шкоди [11].

Уразливість обчислювальної системи - це умова чи безліч умов, які можуть дозволити загрозу впливати на об'єкт або з більшою частотою, або з великим впливом, або спільно.

Атака на обчислювальну систему - це дії, що робляться порушником, що полягає в пошуку тієї або іншої уразливості і її використанні.

* + Залежно від взаємного розташування порушника і його цілі можна виділити наступні види атак:
  + мережеві віддалені атаки;
  + локальні атаки.

Основною особливістю будь-якої мережевої системи є розподіленість її компонентів в просторі і організація зв'язків між ними засобами мережевих з'єднань за допомогою повідомлень. Всі дані і сигнали управління передаються між вузлами розподіленої обчислювальної мережі в формі мережевих пакетів.

Віддалена мережева атака - це інформаційне деструктивний вплив на розподілену обчислювальну систему, що здійснюється по каналах зв'язку. Це визначення охоплює обидві особливості мережевих систем - розподіленість обчислювальних вузлів і розподіленість самої інформації.

На відміну від локальних обчислювальних систем, де найбільш частими є погрози конфіденційності і цілісності, в мережевих системах найбільш поширена загроза відмови в обслуговуванні.

Вторгнення складається з однієї або декількох мережевих атак. Для більшості вторгнень можна виділити три основні етапи:

* збір інформації;
* виконання атаки;
* консолідація.

Труднощі виявлення проведеної віддаленої мережевої атаки і відносна простота її проведення робить даний вид впливу одним з найнебезпечніших і складним для швидкого реагування на здійснену загрозу [12-13].

**Підходи до класифікації атак.** Існує безліч технологій, що застосовуються порушниками для здійснення віддалених мережних атак. Мережеві атаки рідко використовуються окремо і в більшості випадків застосовуються в сукупності для досягнення атакуючим своїх цілей.

* таблиці 3.1 наведено список найбільш поширених класифікацій віддалених атак.

**Тренувальні бази даних з мережевими атаками.** При розробці і проведенні досліджень систем виявлення вторгнень однією з ключових завдань є вибір масивів даних, на яких буде проводитися тестування. Великі компанії-розробники в першу чергу орієнтуються на власні бази даних, спеціалізовані під конкретні завдання і область застосування.

Таблиця 3.1

Класифікації віддалених мережних атак

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Критерій** | **Класи** | | | |
| **За характером впливу** | активні | | пасивні | |
| **За метою впливу** | порушення конфіденціальності | Порушення цілосності | | Порушення доступності |
| **За наявністю каналу зворотнього зв'язку** | Наявний зворотній звязок | | Відсутній зворотній звязок | |
| **За умовою запуску атаки** | За запитом від атакованого об’єкта | За настанням події на атакованому об’єкті | | Безумовна атака |
| **За розташування атакуючого відносно досягнення мети** | внутрішньосегментного | | міжсегментного | |

При розробці систем виявлення мережевих атак важливо мати можливість зіставляти результати тестування з іншими розробками і дослідженнями, що призводить до необхідності використання відомих загальнодоступних баз даних.

На сьогоднішній день можна виділити дві найбільш поширені тренувальні бази даних з відомими атаками - DARPA і KDD [12].

Тренувальна база даних DARPA (Defense Advanced Research Project Agency) була сформована в рамках досліджень лабораторії Лінкольна Массачусетського технологічного інституту (MIT Lincoln Laboratory) в 1998-1999 роках в рамках дослідження можливостей різних систем виявлення вторгнень. Під час цього дослідження використовувалися дані мережевого трафіку і відомості від файлової системи для можливості ідентифікації змодельованих вторгнень, проведених фахівцями під час запису мережевих дампов. Тренувальні дані містять як реальний потік мережевого трафіку, так і спеціально змодельований фоновий трафік. Всі атаки були спрямовані на реальні обчислювальні машини.

* 1999 році, після проведення повторного дослідження роботи різних систем виявлення вторгнень, DARPA надало збережені тренувальні дані у вільному доступі. В даний час ці тренувальні дампи доступні всім дослідникам, тому значна частина публікацій у науковій літературі, пов'язаних з пропозицією нових методів і підходів з виявлення мережевих атак або аномалій, спираються на ці тестові дані. Використання даної бази даних дозволяє дослідникам порівняти основні характеристики якості виявлення: ймовірності помилок пропуску (false negative) і помилкового спрацьовування (false positive) [14].

Загальна кількість типів атак, включених в тестові дані DARPA 1998 року, склало 32 атаки. З точки зору атакуючого ці атаки можна розділити на чотири категорії:

* + атаки відмови в обслуговуванні (Denial of Service, DoS);
  + атаки переходу від віддаленого використання до локального (Remote to Local, R2L);
  + атаки отримання користувачами прав суперкористувача (User to Root,U2R);
  + атаки сканування або проб (Probing / surveillance, PRB).

У підсумкових тренувальних записах, скоригованих і доповнених у 1999 році, поміщені записи мережевого трафіку і аудиту, тривалістю шість тижнів для тренувальних даних і два тижні для тестових даних. Всього в цих дампи присутні 65 різних атак [19].

Всі атаки, використані в дослідженнях DARPA 1998 і 1999 років, перераховані в таблиці 3.2.

Інформація про атаки DARPA зберігається у вигляді текстового опису, в якому вказується час початку атаки, тривалість, адреса жертви, назва атаки, категорія атаки та інші параметри.

На відміну від тренувальних даних DARPA, база даних KDD містить не дампи мережевого трафіку, а оброблені відомості у вигляді масивів з 42 ключових значень. Дана база успішно застосовується багатьма дослідниками для аналізу застосовності різних математичних методів в завданню виявлення мережевих атак, в основному через можливість використання масивів даних з більшості програмних засобів без виконання додаткової обробки.

Таблиця 3.2

Атаки, присутні в тренувальних мережевих дампи DARPA

|  |  |  |  |
| --- | --- | --- | --- |
| **Denial of Service** | **Remote to Local** | **User to Root** | **Probing / surveillance** |
|  |  |  |  |
| Neptune | Dict | Eject | Portsweep |
|  |  |  |  |
| pod | Ftpwrite | Fdformat | Queso |
|  |  |  |  |
| Processtable | Guest | Ftbconfig | Ntinfoscan |
|  |  |  |  |
| Selfping | guesstelnet | Ps | Ls\_domain |
|  |  |  |  |
| Smurf | guessftp | Casesen | Mscan |
|  |  |  |  |
| Syslogd | guesspop | Sechole | satan |
|  |  |  |  |
| Tcpreset | Httptunnel | Yaga | lllegalsniffer |
|  |  |  |  |
| warez | Xlock | ntfsdos | Ipsweep |
|  |  |  |  |
| warezclient | Xsnoop | Anypw | portsweep |
|  |  |  |  |
| sshproctable | Framespoofer | Loadmodule |  |
|  |  |  |  |
| Arppoison | Netbus | Perl |  |
|  |  |  |  |
| Crashiis | ppmacro | Sqlattack |  |
|  |  |  |  |
| Dosnuke | Netcat | Xterm |  |
|  |  |  |  |
| Land | Netcat\_set | secret |  |
|  |  |  |  |
| Mailbomb | Netcat\_brea |  |  |
|  |  |  |  |
| Apache2 | Telnet |  |  |
|  |  |  |  |
| Back | Imap |  |  |
|  |  |  |  |
| Teardrop | Named |  |  |
|  |  |  |  |
| udpstorm | Ncftp |  |  |
|  |  |  |  |
|  | Phf |  |  |
|  |  |  |  |

Склад 42 параметрів, що розглядаються в базі даних KDD, був обгрунтований декількома науковими працями [21-23], присвяченими виявлення аномалій в мережевому трафіку. Однак при дослідженні можливостей по виявленню конкретних мережевих атак виявляється недостатньо аналізувати тільки представлені параметри, але також необхідно розглядати корисне навантаження мережевих пакетів - вищі рівні стека протоколів TCP / IP.

Крім позначених тренувальних баз даних існує безліч більш вузько спеціалізованих, але вони не набули такого широкого поширення в науковому середовищі.

**Опис поширених атак.** У таблиці 3.3 наведено опис деяких поширених атак, що зустрічаються в розглянутих тренувальних базах.

Наведений набір атак показує різноманіття існуючих на сьогоднішній момент методів і засобів проведення віддалених атак.

Для проведення атак використовуються різні протоколи стека TCP / IP. Таким атакам як відмова в обслуговуванні (Denial of Service), сканування портів або прослуховування каналу, можливо перешкоджати за допомогою міжмережевих екранів і використовуючи додаткове мережеве обладнання. Для виявлення атак отримання прав доступу або підвищення привілеїв в локальній мережі організації необхідні більш складні програмні або програмно-апаратні засоби. Опис подібних систем наводиться в наступному розділі.

**3.2 Виявлення мережевих атак.** **Методи виявлення мережевих атак**

Виявлення вторгнень є областю активних досліджень вже кілька десятиліть. Існує велика кількість різних методів і підходів для виявлення віддалених мережних атак [25-26]. Для захисту обчислювальної мережі використовуються такі найпоширеніші засоби і методи:

* політика безпеки корпоративної мережі;
* міжмережеві екрани;
* захист на рівні маршрутизаторів;
* мережевий аудит;
* системи виявлення вторгнень;
* регламент реагування на виявлені атаки;

Таблиця 3.3

Опис найбільш поширених атак

|  |  |
| --- | --- |
| **Найменування** | **ОПИС** |
| **DoS-атаки** | Метою атаки є порушення можливості доступу до обчислювальному середовищі легітимними користувачами. При цьому обмін даними між користувачами атакованого комп'ютера або сегмента мережі неможливий або утруднений в прийнятному режимі. |
| **Сканування портів** | Метою сканування портів є збір відомостей про вузол обчислювальної мережі. У процесі сканування портів визначається безліч відкритих портів вузла, а інформація про сукупності відкритих портів допомагає виявити потенційно вразливі місця всієїобчислювальної мережі. |
| **Атаки по протоколу ICMP** | Застосування протоколу ICMP не супроводжується механізмом аутентифікації, що відкриває порушнику певні можливості. В першу чергу за протоколом ICMP організовуються атаки відмови в обслуговуванні. вузли обчислювальної мережі, які піддалися атаці по протоколу ICMP, відчувають серйозні труднощі при підключенні до мережевих ресурсів |
| **Нав'язування хосту помилкового маршруту** | Ідея атаки полягає в посилці помилкового redirect-повідомлення від імені маршрутизатора на комп'ютер-жертву. При цьому на комп'ютері-жертві зміниться таблиця маршрутизації, і трафік даного хоста буде проходити через комп'ютер-відправник, відіслав redirect-повідомлення. |
| **Ping flooding** | За замовчуванням, багато операційні системи, отримуючи ICMP-запит, генерують відповідний пакет. При відправці дуже великого числа таких запитів порушник може домогтися помітною перевантаження каналу зв'язку. |
| **Smurf** | Ідея даної атаки полягає в широкомовної відправці широкомовних ICMP-запитів від імені комп'ютера-жертви. В результаті атаки комп'ютери, які взяли дані пакети, практично одночасно відповідають компьютеру-жертві, що призводить до перевантаження каналу зв'язку, а іноді до повної ізоляції мережі. При своїй простоті дана атака є вкрай ефективною |
| **Фрагментовані пакети** | У структуру протоколу мережевого рівня IP закладена можливість поділу переданої інформації на частини -фрагменти. На машині-одержувачі інформація відновлюється з цих фрагментів. порушник може змоделювати і відправити спеціальну послідовність фрагментів, що може привести до переповнення виділеної пам'яті або навіть до аварійного завершення роботи системи. |
| **Мережеві**  **черв'яки** | Основна мета мережевих черв'яків - це найбільше поширення шкідливого коду по мережі. черви легко поширюються по обчислювальної мережі, використовуючи мережеві уразливі місця різних програм. Чи по кишені мережі Інтернет мережеві черв'яки можуть поширюватися по мільйонам обчислювальних машин за лічені хвилини. |
| **Sniffing - про-**  **слуховування**  **каналу** | Будь-який комп'ютер, що знаходиться в одному сегменті мережі з комп'ютером-жертвою може перемкнути мережеву карту в прослуховує режим і перехоплювати весь трафік, що надходить на комп'ютер-жертву. |
| **IP spoofing** | Ідея даної атаки полягає в передачі на комп'ютер-жертву повідомлення від імені іншого комп'ютера мережі. шляхом передачі помилкових повідомлень можна замести сліди протиправних дій чи нав'язати інформацію. |
| **DNS spoofing** | Метою даної атаки є підміна IP-адрес в кеші DNS сервера. Підміна здійснюється посилкою великого кількості DNS-пакетів з одним і тим же ім'ям. В разі успіху, всі користувачі DNS-сервера отримають недостовірну інформацію про IP-адреси і доменні імена. |
| **WinNuke** | У старих версіях операційних систем сімейства Windows існувала наступна вразливість: при посилці невеликого пакету з нульовим значенням поля Urgent Pointer на 139 порт (Використовується протоколом NetBIOS), то служба NetBIOS аварійно завершить роботу операційної системи |

Міжмережеві екрани організовують фільтрацію вхідного і вихідного, відсікаючи всі заборонені дані і пропускаючи дозволені. Міжмережеві екрани працюють на різних рівнях стека протоколів TCP / IP. Реалізований контроль доступу з зовнішньої мережі до захищеної внутрішньої сильно залежить від використовуваного рівня аналізованих протоколів. Ступінь захисту підвищується зі збільшенням рівня, але разом з нею підвищується обчислювальна навантаження проведеного аналізу і складність адміністрування. Міжмережевий екран (МСЕ) - це пристрій забезпечення безпеки мережі, яке здійснює моніторинг вхідного і вихідного мережевого трафіку і на підставі встановленого набору правил безпеки приймає рішення, пропустити або блокувати конкретний трафік [27].

Розрізняють декілька типів міжмережевих екранів в залежності від наступних характеристик:

• забезпечує екран з'єднання між одним вузлом і мережею або між двома або більше різними мережами;

• відбувається контроль потоку даних на мережевому рівні або більш високих рівнях моделі OSI;

• чи аналізуються стану активних сполук чи ні.

Залежно від охоплення контрольованих потоків даних міжмережеві екрани поділяються на:

• традиційний мережевий (або міжмережевий) екран - програма (або невід'ємна частина операційної системи) на шлюзі (пристрої, що трафік між мережами) або апаратне рішення, контролюючі вхідні та вихідні потоки даних між підключеними мережами (об'єктами розподіленої мережі);

• персональний міжмережевий екран - програма, встановлена ​​на користування земельними діл-тельских комп'ютері і призначена для захисту від несанкціонованого доступу тільки цього комп'ютера.

Залежно від рівня OSI, на якому відбувається контроль доступу, мережеві екрани можуть працювати на:

• мережевому рівні, коли фільтрація відбувається на основі адрес відправника і одержувача пакетів, номерів портів транспортного рівня моделі OSI і статичних правил, заданих адміністратором;

• сеансовому рівні (також відомі, як stateful), коли відслідковуються сеанси між додатками і не пропускаються пакети, що порушують специфікації TCP / IP, часто використовувані в зловмисних операціях - сканування ресурсів, зломи через неправильні реалізації TCP / IP, обрив / уповільнення з'єднань, ін'єкція даних;

• прикладному рівні (або рівні додатків), коли фільтрація проводиться на підставі аналізу даних програми, що передаються всередині пакету. Такі типи екранів дозволяють блокувати передачу небажаної і потенційно небезпечної інформації на підставі політик і налаштувань [28].

З основних недоліків міжмережевих екранів можна виділити неможливість детального контролю активності авторизованих користувачів. На думку експертів переважна більшість порушень інформаційної безпеки відбувається в зв'язку з діяльністю легітимних користувачів, що призводить до необхідності контролю призначених для користувача дій на більш високому рівні.

Одним з варіантів розв'язання проблеми є системи виявлення вторгнень. У своїй роботі системи виявлення вторгнень керуються не тільки мережевим трафіком і безліччю правил, а й аудитом системи, різними журналами, показниками роботи операційної системи і т.д. Також існують системи запобігання вторгнень, що дозволяють не тільки виявити факт відбувся вторгнення в систему, але і мінімізувати наслідки, розірвавши мережеве з'єднання, заблокувавши підозрілу активність користувача або навіть адміністратора.

Найбільш ефективним способом запобігання несанкціонованому використанню комп'ютерних систем і мережевих ресурсів є підтримка багаторівневого захисту, коли спільно використовуються міжмережеві екрани, системи виявлення вторгнень, системи аудиту, політика безпеки і інші засоби захисту.

**Системи виявлення вторгнень.** Система виявлення вторгнень (СВВ, англ. Intrusion Detection System, IDS) - це програма або програмно-апаратний засіб, призначений для виявлення фактів несанкціонованого доступу (вторгнення або мережевої атаки) в обчислювальну систему або мережу [13].

Системи виявлення атак (CВA) являють собою окремий клас програмних засобiв (ПЗ), пiд яким розумiють програми, процедури, правила, а також, якщо передбачено, супутнiх їм документацiї та даних, що вiдносяться до функцiонування системи обробки iнформацiї. Повна назва CВA – це системи виявлення i запобiгання атак, так як саме в можливостi автоматизованої протидiї атакам полягає одна з основних переваг таких систем, у порiвняннi, наприклад, iз засобами, заснованими на людському факторi. Проте надалi буде використовуватися найбiльш усталена назва - система виявлення атак. Використання СВА дозволяє вирiшити цiлий ряд завдань, що забезпечують досягнення цiлей iнформацiйної безпеки:

• розпiзнавання вiдомих i, по можливостi, невiдомих атак та попередження персоналу, що вiдповiдає за забезпечення iнформацiйної безпеки (IБ);

• статистичний аналiз шаблонiв аномальних дiй;

• монiторинг i аналiз користувацької, мережевої та системної активностi;

• контроль цiлiсностi файлiв та iнших ресурсiв iнформацiйної системи (IС);

• аудит системної конфiгурацiї i виявлення вразливостей;

• iнсталяцiя i пiдтримка роботи серверiв-пасток для запису iнформацiї про порушникiв;

• зниження навантаження на персонал (або звiльнення вiд нього), що вiдповiдає за IБ, вiд поточних рутинних операцiй з контролю за користувачами, системами i мережами, якi є компонентами IС;

• надання можливостi управлiння функцiями захисту не спецiалiстам в областi iнформацiйної безпеки.

Перший, пiдготовчий, етап полягає в пошуку передумов для здiйснення тiєї чи iншої атаки. На даному етапi шукають вразливостi, використання яких робить можливим в принципi реалiзацiю атаки, яка i складає другий етап. На третьому етапi атака завершується. При цьому перший i третiй етапи самi по собi можуть бути атаками. Наприклад, пошук порушником вразливостей за допомогою сканерiв безпеки вже вважається атакою. Технологiї виявлення атак постiйно розвиваються i удосконалюються, i ця область постiйно залучає нових виробникiв i розробникiв. Незважаючи на брак теоретичних основ технологiї виявлення атак, iснують досить ефективнi методи, що використовують на сьогоднi. Механізми виявлення атак, застосовувані в сучасних системах виявлення атак IDS (IntrusionDetectionSystem), засновані на декількох загальних методах. Слід зазначити, що ці методи не є взаємовиключними. У багатьох системах використається комбінація декількох методів [28].

Класифікація систем виявлення атак може бути виконана по декількох ознаках:

* по способу реагування;
* по способу виявлення атаки;
* по способу збору інформації про атаку.

По способу реагування розрізняють пасивні й активні IDS. Пасивні IDS просто фіксують факт атаки, записують дані у файл журналу та видають попередження. Активні IDS намагаються протидіяти атаці, наприклад, шляхом реконфігурації міжмережевого екрану або генерації списків доступу маршрутизатора [31].

По способу реагування розрізняють пасивні й активні IDS. Пасивні IDS просто фіксують факт атаки, записують дані у файл журналу та видають попередження. Активні IDS намагаються протидіяти атаці, наприклад, шляхом реконфігурації міжмережевого екрану або генерації списків доступу маршрутизатора. По способу виявлення атаки системи IDS прийнято ділити на дві категорії:

* виявлення аномального поводження (anomaly-based);
* виявленнязловживань (misuse detection або signature-based).

Технологія виявлення атак шляхом ідентифікації аномального поводження заснована на наступній гіпотезі. Аномальне поводження користувача (тобто атака або яка-небудь ворожа дія) проявляється як відхилення від нормального поводження. Прикладом аномального поводження може служити велика кількість з'єднань за короткий проміжок часу, високе завантаження центрального процесора й т.п.

Суть іншого підходу до виявлення атак - виявлення зловживань - полягає в описі атаки у вигляді сигнатури (signature) і пошуку даної сигнатури в контрольованому просторі (мережному трафике або журналі реєстрації). Як сигнатура атаки може виступати шаблон дій або рядок символів, що характеризують аномальну діяльність. Ці сигнатури зберігаються в базі даних, аналогічної тієї, котра використається в антивірусних системах.

На сьогодні найбільш популярною є класифікація по способу збору інформації про атаку, а саме:

* виявлення атак на рівні мережі (network-based);
* виявлення атак на рівні хоста (host-based);
* виявлення атак на рівні додатка (application-based).

Система першого типу (network-based) працює по типу сниффера, «прослуховуючи» трафік у мережі й визначаючи можливі дії зловмисників. Пошук атаки йде за принципом «від хоста до хоста». Системи, що входять у перший клас, аналізують мережний трафік, використовуючи, як правило, сигнатури атак й аналіз «на льоту». Метод аналізу «на льоту» полягає в моніторингу мережного трафіка в реальному або близькому до реального часу й використанні відповідних алгоритмів виявлення. Найчастіше використовується механізм пошуку в трафіку певних рядків, які можуть охарактеризувати несанкціоновану діяльність [32].

Системи другого типу (host-based) призначена для моніторингу, детектування та реагування на дії зловмисників на певному хості. Система перевіряє й виявляє спрямовані проти хоста дії. Ці системи аналізують реєстраційні журнали операційної системи або додатка. Аналіз журналів реєстрації є одним з найперших реалізованих методів виявлення атак. Він полягає в аналізі журналів реєстрації (log, audit trail), створюваних операційною системою, прикладним програмним забезпеченням, маршрутизаторами й т.д. Запис у журналі реєстрації аналізуються й інтерпретуються системою виявлення атак. До переваг цього методу можливо віднести простоту його реалізації.

Однак за цією простотою приховано ряд недоліків:

* для достовірного виявлення тієї або іншої підозрілої діяльності необхідна реєстрація в журналах великого обсягу даних, що негативно позначається на швидкості роботи контрольованої системи;
* при аналізі журналів реєстрації дуже важко обійтися без допомоги фахівців, що істотно знижує коло поширення цього методу;
* до дійсного моменту немає уніфікованого формату зберігання журналів;
* аналіз записів у журналах реєстрації здійснюється не в реальному режимі часу, тому цей метод не може бути застосований для раннього виявлення атак у процесі їхнього розвитку. Очевидним недоліком є те, що системи рівня хоста не можуть виявити атаки, які спрямовані на вузли, що не використовують журнали реєстрації або для яких не існує відповідної реалізації агента.

Як правило, аналіз журналів реєстрації є доповненням до інших методів виявлення атак, зокрема до виявлення атак «на льоту». Використання цього методу дозволяє проводити «розбір польотів» вже після того, як була зафіксована атака, для того щоб виробити ефективні заходи запобігання аналогічних атак у майбутньому [33-34].

Третій тип IDS (application-based) заснований на пошуку проблем у певному додатку.

Кожний із цих трьох типів систем виявлення атак (на рівні мережі, на рівні хоста й на рівні додатка) має свої переваги та недоліки. Необхідно відмітити, що лише деякі системи виявлення атак можуть бути однозначно віднесені до одного з названих класів. Гібридні IDS, що представляють собою комбінацію різних типів систем, як правило, містять у собі можливості декількох категорій.

* + таблиці 3.4 наведені основні підходи до класифікації систем виявлення вторгнень.

Таблиця 3.4

Класифікація систем виявлення вторгнень

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **критерій** |  |  | **класи** | | |  |  |
|  |  |  |  |  |  | |  |
| Захищається | Мережеві системи |  |  | хостової системи | | | |
| система |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Методологія | виявлення | виявлення | | |  | гібридне | |
| виявлення | сигнатур | аномалій | | |  | виявлення | |
|  |  |  |  |  |  |  |  |
| Реакція системи | пасивна |  |  | активна | |  |  |
|  |  |  |  |  | |  |  |
| Структура | централізована |  | децентралізована | | |  | агентські |
| системи |  |  |  |  |  |  | системи |
|  |  |  |  |  |  |  |  |
| Час аналізу | У реальному часі |  | | відкладений аналіз | | | |
|  |  |  | |  |  | | |
| Джерела даних | Мережеві пакети | записи аудиту | | | об'єкти системи | | |
|  |  |  | |  |  | | |

Найбільш загальна структура системи виявлення вторгнень, розроблена групою дослідників CIDF (Common Intrusion Detection Framework), представлена ​​на рис. 3.1.

Блок збору даних (сенсор, Event-box) - витягує дані для обробки та прийняття рішення аналізатором. У даних можуть міститися імена контрольованих параметрів, їх особливості та значення. Сенсор може виконувати перетворення даних для перекладу в необхідний формат або для скорочення обсягу переданих даних.

Блок аналізатора (Analyzer-box) - приймає рішення про наявність чи відсутність ознак атаки або аномалії на підставі даних від сенсорів. В рамках аналізу даний блок може виконувати функції фільтрації, нормалізації, перетворення і кореляції даних. При виявленні атаки блок аналізатора може додати до вихідних даних опис виявленої атаки. Блок аналізатора може мати багаторівневу систему.



Рис. 3.1 - Загальна структура системи виявлення вторгнень.

Блок бази даних (сховище даних, Database-box) - містить безлічі вирішальних правил і семантичний опис атак, а також накопичувальну інформацію від сенсорів. Дані можуть перебувати в текстових файлах, базі даних, і т.д.

Блок реакції (Response-box) - інформує адміністратора про зафіксовані атаки, а в разі систем запобігання вторгнень формує активну реакцію. Системи запобігання вторгнень відстежують активність в режимі реального часу і швидко реалізовують дії щодо запобігання атак. Можливі заходи - блокування потоків трафіку в мережі, скидання сполук, видача сигналів оператору. Також системи запобігання вторгнень можуть виконувати дефрагментацію пакетів, переупорядкування пакетів TCP для захисту від пакетів з зміненими номерами послідовності і підтвердження [34].

Системи виявлення мережевих атак збирають інформацію з пакетів мережевого трафіку, системних журналів і показників функціонування системи. Традиційні системи виявлення мережевих атак будуються на сигнатурному підході: за допомогою набору правил або сигнатур, що формуються експертами і розміщені в базу вирішальних правил, описуються всі можливі сценарії і особливості атак. У цього підходу існує безліч відомих недоліків. За допомогою аналізу сигнатур неможливо виявити нові види атак, тому що база вирішальних правил не містить інформації про відповідну атаці. Процес аналізу сигнатур для розподілених атак є вкрай складним завданням. Крім того, бази вирішальних правил популярних систем виявлення вторгнень практично є загальнодоступними, тому порушник може протестувати можливості приховування атаки.

Перераховані проблеми підходу пошуку сигнатур змушують фахівців шукати альтернативні шляхи для організації захисту від мережевих атак. Одним з популярних напрямків досліджень є застосування різних методів Інтелектуального аналізу даних(інтелектуального аналізу даних) в системах виявлення мережевих атак. В основі даних методів лежить припущення, що вся легітимна активність в системі може бути представлена ​​у вигляді математичної моделі. Застосовувані для виявлення мережевих атак методи Інтелектуального аналізу данихпереслідують одну з наступних цілей:

* виявлення порушень;
* виявлення аномалій.

Перші моделюють атаки і застосовують засоби класифікації, другі моделюють нормальну поведінку і виконують пошук винятків.

При використанні методів Інтелектуального аналізу данихдля виявлення мережевих атак можна виділити наступні проблеми:

* дані, аналізовані системами виявлення, мають високу розмірність і обсяг;
* вимога обробки даних в режимі реального часу;
* велика кількість шумів і викидів в оброблюваних даних,
* викликають неадекватну реакцію методів інтелектуального аналізу даних.

**Сигнатурні системи виявлення вторгнень.** Класичні системи виявлення вторгнень засновані на понятті сигнатура. Сигнатура - це безліч умов, при задоволенні яких настає подія, яке визначається як атака або вторгнення. Спочатку, під сигнатурою розумілася характерна рядок (сукупність електронних даних), притаманна тільки даної атаці.

Сигнатурні системи виявлення мають ряд переваг:

* швидкість роботи, так як повний аналіз пакету - трудомістке завдання;
* простота створення і налаштування правил;
* підтримка комп'ютерного співтовариства в швидкому виробництві сигнатур для виявлення нових небезпек;
* складні атаки використовують ряд попередніх дій, які легко розпізнати за допомогою сигнатур.

З іншого боку СВВ, що грунтуються тільки на аналізі сигнатур, мають певні недоліки:

* + згодом швидкість роботи системи сповільнюється, оскільки зростає число перевіряються сигнатур. Кожна знову придумана атака або засіб доповнює довгий перелік перевірених сигнатур;
  + такі СОВ дозволяють виявляти тільки вже відомі атаки, і непридатні для виявлення щойно з'явилися атак;
  + можливі помилкові спрацьовування прі не унікальності сигнатури;
  + можливі пропуски атаки при незначній модифікації;
* обов'язковість поновлення бази даних для отримання сигнатур нових атак.

Згідно зі статистикою близько 80% нових атак ґрунтується на давно відомих сценаріях. Наявність в системі виявлення переліку сигнатур відомих атак дозволяє виявити більшість вторгнень з високою ймовірністю [41].

**Недоліки систем виявлення вторгнень.** Існуючі системи виявлення мережевих атак будуються на сигнатурному підході і підході, заснованому на пошуку аномалій. Для сигнатурних систем основним недоліком є ​​принципова неможливість виявити нові типи атак і якісні варіації відомих атак.

Проведена СВВ робота є складною обчислювальної завданням: необхідно проаналізувати кожен вступник мережевий пакет в режимі реального часу. Одним з варіантів розв'язання проблеми для мережевих СВВ є апаратна реалізація, що значно підвищує продуктивність, але породжує певні складності з адмініструванням, зокрема оновленням баз даних. Найчастіше виробники сильно обмежують спектр виявляються атак для можливості оперативного аналізу трафіку.

Одним з головних недоліків роботи будь-якої системи виявлення вторгнення є неможливість аналізу зашифрованої інформації. Основний прояв даної проблеми - використання віртуальних приватних мереж (VPN).

Мережеві СВВ не здатні оцінити результати виявленої підозрілої активності - атаки, тому що не мають відомості хостового аудиту або станом операційної системи.

Для систем виявлення аномалій основними недоліками є складність формування профілю нормального поведінки і неможливість віднесення виявленої аномалії до певного класу атак.

Для систем виявлення аномалій властиво значно більше число помилкових спрацьовувань, ніж для систем виявлення атак. Головна причина цього явища - складність навчання системи виявлення аномалій. Для побудови якісного профілю нормального поведінки необхідно зібрати достатньо повну базу і при цьому в період формування даної бази необхідно виключити появу будь-яких аномалій. В іншому випадку порушник зможе обійти захист, присвоївши аномальному впливу статус легітимного. Системи виявлення аномалій сильно залежать від профілю нормального поведінки і при значних варіаціях в поведінці користувача, викликаних, наприклад, тимчасовим переключенням на виконання іншої роботи, система буде фіксувати аномальна поведінка.

Для ефективного захисту розподіленої обчислювальної мережі необхідно рознесення зазначених компонентів і їх частин з різних вузлів мережі. Більшість існуючих хостових і мережевих СОВ мають монолітну структуру, що не дозволяє ефективно розподілити функціонал по вузлах обчислювальної мережі.

**Виявлення атак за допомогою прихованої марковської моделі.** Прихована марковська модель являє собою статистичну модель, де система моделюється як процес Маркова з невідомими параметрами. Завдання методу полягає в оцінці прихованих параметрів, грунтуючись на спостережуваних параметрах. Послідовності подій, зібрані з нормальних операцій системи, використовуються в якості навчальних даних для оцінки параметрів прихованої марковської моделі. Після навчання прихованої марковської моделі імовірнісні оцінки використовуються в якості порогових значень для ідентифікації мережевих аномалій в тестових даних. Існують наступні проблеми використання прихованих марковських моделей для виявлення аномалій:

* проблема оцінки ймовірності того, що спостережувана послідовність була згенерована моделлю;
* проблема побудови з даних аудиту моделі або набору моделей,
* які б правильно описували спостерігається поведінка;
* проблема обчислення найбільш ймовірного набору прихованих станів,
* що призвели до отриманих спостереженнями в рамках прихованої марковської моделі.

При виявленні 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 [43]

В роботі прихована марківських модель для нормального поведінки будується з журналів і даних нормального поведінки мережевий системи. Спостерігається поведінка системи аналізується для оцінки ймовірності того, що побудована модель підтримує спостерігається поведінка. Низька вірогідність підтримки вказує на аномальну поведінку.

* роботі [43] описується застосування прихованої марковської моделі для виявлення аномалій з використанням профілю послідовностей системних викликів і shell-команд. При навчанні модель обчислює ймовірність вибірки спостерігається послідовності. Граничні значення ймовірності формуються на основі мінімальної ймовірності серед всіх послідовностей в учнів даних і використовуються для розмежування між нормальним і аномальним поведінкою. Головним обмеженням даного підходу є слабка узагальнення, що може привести до блокування користувачів, неоднозначно визначених у даній системі.

**Виявлення атак за допомогою байесовских мереж.** Байєсова мережа являє собою модель, яка кодує імовірнісні взаємозв'язку між змінними. При використанні в поєднанні зі статистичними методами, байєсовські мережі мають такі переваги для аналізу даних:

* байєсовські мережі застосовні для обробки даних з пропусками,
* оскільки розглядається тільки взаємозалежність між змінними;\
* байєсовські мережі можуть представляти причинно-наслідкові зв'язки і можуть застосовуватися для прогнозування наслідків дії;
* байєсовські мережі можуть бути використані для вирішення завдань моделювання, де є необхідність об'єднати попередні знання з даними.

Основний метод передбачає незалежність серед атрибутів. Кілька варіантів застосування байесовских мереж були запропоновані для виявлення мережевих аномалій. Більшість методів направлено на формування умовних залежностей між атрибутами з використанням складних мереж Байеса. Байєсовські методи часто використовуються в процедурі класифікації і придушення помилкових спрацьовувань.

Для виявлення вторгнень або прогнозування поведінки порушника байєсовські мережі можуть бути ефективними в деяких випадках, але в загальному випадку точність цього методу залежить від припущень, пов'язаних з поведінкової моделлю цільової системи. Таким чином, будь-яка значна відхилення від припущень призведе до зменшення точності виявлення [43].

**Виявлення атак за допомогою методів кластеризації.** Методи кластеризації групують дані в кластери на підставі схожості об'єктів. Більшість методів кластеризації починається з вибору центральної точки для кожного кластера, а безліч елементів розподіляються по кластерам. Після цього центри коригуються, а елементи перерозподіляються. Кластеризація дозволяє вивчити і виявити аномалії, не вимагаючи явного безлічі класів або типів аномалій, тобто для виявлення аномалій за допомогою методів кластеризації може не турбуватися навчальну множину. Кластеризація досить широко застосовується для виявлення мережевих аномалій.

В дослідженні [18] після проведення кластеризації на навчальній множині на основі зумовлених критеріїв кластерам присвоюються мітки - нормальний трафік або відомі класи атак. Система формує профілі для цих двох класів. Відібрані точки кластерів використовуються для побудови класифікаторів. При реальній роботі системи виявлення аналізований фрагмент трафіку класифікується як нормальна або аномальний.

Виявлення невідомих мережевих атак найчастіше будується саме на методах кластеризації. Однорідні групи зі схожими характеристиками або кластери формуються шляхом розбиття набору елементів без будь-яких позначок.

В системі вкрай важливо правильно визначити кластери, щоб максимально віддалити їх від викидів. Кінцева мета даних методів полягає у визначенні ступеня відхилення викидів від кластерів. За допомогою простого порівняння з пороговим значенням викиди з високим ступенем відхилення від кластерів позначаються як аномалії.

**Виявлення атак за допомогою методу опорних векторів.** Метод опорних векторів є одним з найбільш популярних методів класифікації. Метод будує оптимальну гіперплощину в просторі характеристик: w · x - b = 0, що розділяє нормальні і аномальні елементи. У підсумку завдання можна звести до квадратичного програмування:

при

де - величина помилки на об'єктах .

* методу опорних векторів існують певні недоліки для застосування в задачі виявлення мережевих атак:
* висока залежність вирішальної функції f (x) від встановлюваного заздалегідь параметра С;
* висока чутливість до наявності шумів і викидів;
* потреба в масштабуванні тренувальних даних;
* відсутність можливості ранжувати важливість атак.
* роботі [34] розглядається система виявлення вторгнень, заснована на модифікації методу опорних векторів - додавання нечіткості. Нечіткість дозволяє скоротити число помилок першого і другого роду за рахунок більш гнучкої класифікації записів біля кордону між класами.

**Виявлення атак за допомогою нейронних мереж.** Інтерес до штучних нейронних мереж викликаний тим фактом, що людський мозок виробляє обчислювальні операції принципово іншим чином, ніж звичайна цифрова обчислювальна машина. Нейронні мережі являють безліч інструментів для самих різних застосувань: кластеризація даних, витяг ознак, скорочення розмірності і т.д. В деяких випадках застосовується модифікована версія навчання з підкріпленням для вивчення нових атак. При зустрічі нової атаки використовується зворотний зв'язок для оновлення сигнатур. Також використовуються ієрархії нейронних мереж для виявлення аномалій. Нейронні мережі навчаються з використанням даних, які охоплюють нормальний простір і здатні розпізнавати невідомі атаки.

Запропонована мережа має можливість кодування тимчасової інформації. Автори розробляють інкрементне ядро методу головних компонент для попередньої обробки даних, які надходять в нейронну мережу.

**Виявлення атак за допомогою правил нечіткої логіки.** Нечіткі системи виявлення мережевих вторгнень використовують безліч нечітких правил для визначення ймовірності конкретних або загальних мережевих атак. Нечітке безліч може бути сформовано для опису трафіку в конкретній мережі.

В роботі [46] описується метод для побудови класифікаторів, що використовують нечіткі асоціативні правила, що застосовуються для виявлення вторгнення в мережу. Нечіткі набори правил асоціації використовуються для опису нормальних і аномальних класів. Належність записи певного класу визначається за допомогою відповідної метрики.

Нечіткі асоціативні правила формуються на основі звичайних навчальних вибірок. Тестований зразок класифікується як нормальний, якщо згенерований сукупністю правил показник буде вищою певного порогового значення. Зразки з більш низьким показником вважаються аномальними. Крім того, автори пропонують метод для прискорення індукції правил шляхом зменшення числа елементів з добутих правил.

* роботі [34] представлений метод формування нечітких правил асоціації, заснований на генетичному мережевому програмуванні, для виявлення мережевих вторгнень. Генетичне мережеве програмування є еволюційним шляхом оптимізації, який використовує орієнтований граф замість рядків в стандартних генетичних алгоритмах, що істотно спрощує опису за рахунок багаторазового використання вузла в графі.

В роботі [47] запропонований метод неконтрольованої нечіткої кластеризації на основі клонального відбору для виявлення аномалій. Метод здатний отримувати глобальні оптимальні кластери швидше, ніж у конкуруючих алгоритмів.

Якщо зробити висновки з аналізу досліджень в області виявлення мережевих атак, заснованих на методах інтелектуального аналізу даних, можна зробити висновок про широкі можливості методів інтелектуального аналізу даних для вирішення поставлених завдань (табл. 3.5).

Найкращі показники роботи подібних систем були досягнуті при використанні методу опорних векторів для виявлення порушень, при цьому в навчальну множину входили як приклади нормальної активності, так і приклади атак.

**Висновки до 3 розділу**

Проблема організації захисту від віддалених мережних атак є актуальною і постійно розвивається. Серед всіх існуючих програмних і апаратних засобів за хисту немає універсального, що дозволяє організувати захист довільної обчислювальної машини або сегмента комп'ютерної мережі від усіх видів загроз і зокрема від віддалених мережних атак.

Таблиця 3.5

Показники виявлення атак в різних дослідженнях

|  |  |  |
| --- | --- | --- |
| **Метод інтелектуального аналізу даних** | **Показник**  **розпізнавання (%)** | **Помилкові**  **спрацьовування (%)** |
|  |
| C4.5 | 95 | 1 |
|  |  |  |
| Метод опорних векторів (SVM) | 95,5 | 1 |
|  |  |  |
| багатошаровий персептрон | 94,5 | 1 |
|  |  |  |
| Метод k-найближчих сусідів | 92 | 1 |
|  |  |  |
| Машина лінійного програмування | 94 | 1 |
|  |  |  |
| Упорядкований дискримінантний  аналіз | 94 | 1 |
|  |  |
|  | | |
| Дискримінантний аналіз Fisher | 89 | 1 |
|  |  |  |
| γ-алгоритм | 80 | 1 |
|  |  |  |
| k-means кластеризация | 65 | 1 |
|  |  |  |
| Single leakage кластеризация | 69 | 1 |
|  |  |  |
| Quarter-sphere SVM | 65 | 1 |
|  |  |  |
| Y-means кластеризация | 89,89 | 1 |
|  |  |  |
| Генетичне програмування для розподіленої системи виявлення вторгнень | 91 | 0,43 |
| SVM + Генетичні алгоритми | 99 | - |
|  |  |  |
| SVM + Нечітка логіка | 99,56 | 0,44 |
|  |  |  |
| Нейронні мережі + МГК | 92,22 | - |
|  |  |  |
| C4.5 + МГК | 92,16 | - |
|  |  |  |
| генетичні алгоритми | 97,47 | 0,69 |
|  |  |  |
| C4.5 + Гібридні нейронні мережі | 93,28 | 0,2 |
|  |  |  |
| Прихована марковська модель | 79 | - |
|  |  |  |

Системи виявлення вторгнень є одним з ключових компонентів комплексу засобів захисту інформації. Всі існуючі промислові СВВ і наукові розробки мають ті чи інші недоліки: обмежений спектр виявляються атак або підтримуваної програмно-апаратної середовища, складність адміністрування або створення профілю, висока обчислювальна складність.

Питання вибору тренувальної бази з атаками не має простого рішення, тому що широко поширені бази даних містять багато в чому застарілі типи атак, а більш сучасні бази мають специфічну структуру, що вимагає складної попередньої обробки, і використовуються тільки окремими дослідниками, що перешкоджає порівнянні якісних показників результатів роботи.

* + останні десятиліття методи Інтелектуального аналізу даних отримали широке застосування в багатьох наукових напрямах, і проблема виявлення мережевих атак не є винятком з цієї тенденції. Існують кілька сотень наукових досліджень щодо застосування різних методів Інтелектуального аналізу данихдля виявлення мережевих атак і для вирішення пов'язаних з виявленням подзадач.

Дослідниками пропонується безліч нестандартних рішень, що використовують конкретні особливості застосування даних методів.

Формування СВВ на основі методів Інтелектуального аналізу данихдозволяє позбутися від деяких відомих недоліків систем пошуку сигнатур і систем виявлення аномалій. Вибір конкретних методів і формування методик щодо застосування є складним завданням, що вимагає значних обсягів експериментів, і може сильно залежати від навчальної множини.

**РОЗДІЛ 4.** **РОЗРОБКА МОДЕЛІ СИСТЕМИ ВИЯВЛЕННЯ МЕРЕЖЕВИХ АТАК НА ОБ’ЄКТАХ КРИТИЧНОЇ ІНФРАСТРУКТУРИ**

**4.1. Застосування методів інтелектуального аналізу даних в завданні виявлення мережевих атак**

Функціональні компоненти системи виявлення мережевих атак. Традиційна модель системи виявлення вторгнень складається з чотирьох блоків: збору даних, бази вирішальних правил, аналізатора і блоку реакції, рис. 4.1. Метою дослідження є організація захисту розподілених обчислювальних мереж, тому на відміну від хостових і мережевих СВВ необхідна можливість розподілу компонентів системи виявлення мережевих атак і їх частин з різних вузлів мережі.

Більшість існуючих СВВ мають монолітну структуру, що не дозволяє ефективно розподілити обчислювальне навантаження системи. На рис. 4.1 виділені компоненти традиційної моделі СВВ, які можуть бути рознесені по вузлах обчислювальної мережі.

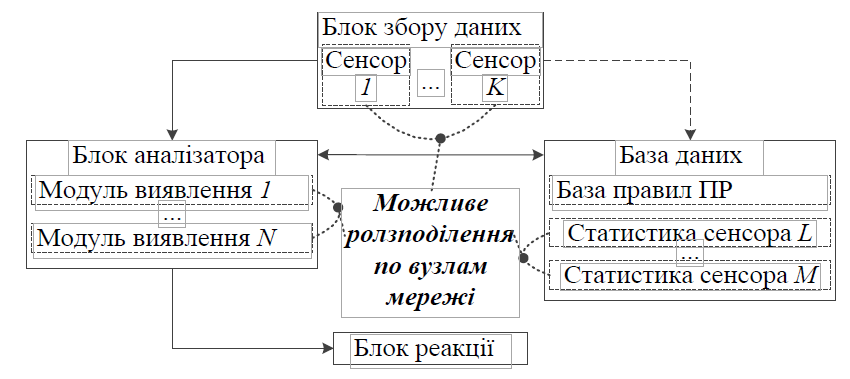


Рис. 4.1 - Можливість розподілу функціоналу СВВ

В залежності від програмної і апаратної складової окремих вузлів системи та мережевого обладнання можливе розміщення різного складу сенсорів, модулів виявлення і компонентів бази даних. Для ефективного функціонування СВВ на вузлах необхідно виконання наступних умов:

* склад модулів виявлення повинен відповідати безлічі потенційно можливих видів атак, що може бути визначено на основі оброблюваної на хості інформації та встановлене програмне забезпечення;
* склад сенсорів повинен відповідати аналізованих потокам інформації, повинні оброблятися тільки ті мережеві протоколи, за допомогою яких можливе здійснення атаки;
* внутрішня структура бази даних повинна відповідати беруть участь сенсорам і модулів виявлення.

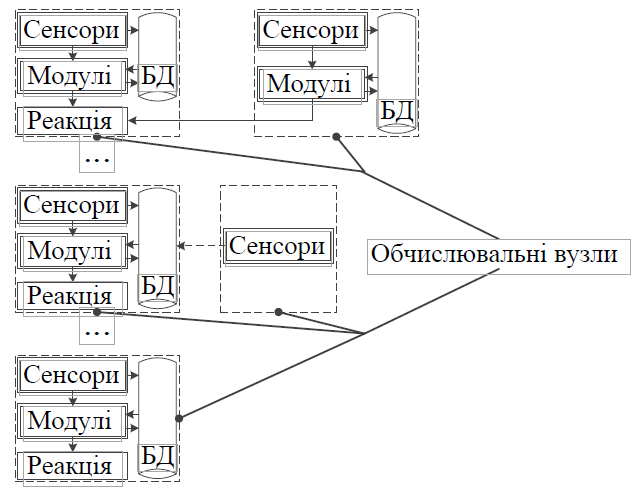
На рис. 4.2 зображені варіанти розміщення компонентів СВВ в розподіленої обчислювальної мережі. Обчислювальне навантаження СВВ на вузол також може варіюватися за рахунок створення єдиної бази даних для декількох вузлів і за рахунок розміщення на окремих вузлах тільки блоку сенсорів. При такому підході може значно збільшитися обсяг мережевого трафіку до вузла, який містить загальну базу даних, і збільшуються ризики проходження атаки відмови в обслуговуванні для вузлів, що не мають свого блоку аналізатора.

Рис. 4.2. Варіанти розміщення компонентів системи виявлення мережевих атак в розподіленої обчислювальної мережі

Грунтуючись на розглянутих дослідженнях по застосуванню методів Інтелектуального аналізу данихможна виділити наступні підзадачі процесу виявлення мережевих атак:

Класифікація елемента - віднесення аналізованого вектора до класу нормальних або аномальних.

Пошук шумів і викидів в тренувальних даних і очищення від них.

Розбиття тренувальних даних на частини з метою оптимізації процесу аналізу трафіку.

Визначення необхідного і достатнього безлічі видобутих сенсором параметрів для віднесення конкретного безлічі мережевих пакетів до класу атак.

Автоматичне формування модульної архітектури за коштами процедури кластеризації.

Введення додаткового рівня перевірки сигналів окремих модулів для скорочення числа помилкових спрацьовувань.

Віднесення виявленої аномалії до відомої атаці або класу атак.

Рішення деяких з перерахованих подзадач можливо за рахунок застосування різних методів інтелектуального аналізу даних. Рішення будь-якої з представлених подзадач веде до зміни методик застосування методів інтелектуального аналізу даних для вирішення інших завдань [34].

Групи методів інтелектуального аналізу даних для проектування системи виявлення мережевих атак. Згідно з поданими раніше підзадачами, повязаними з виявленням мережевих атак можна виділити кілька груп методів інтелектуального аналізу даних. Реалізовані функції представлені в таблиці 4.1.

В основі більшості СВВ лежить процес класифікації, яка формує висновок про фіксацію атаки або аномальної поведінки. В даний час проводиться безліч досліджень на тему виявлення мережевих атак. В основі цих досліджень лежать такі методики як нейронні мережі, дерева прийняття рішень, асоціативні правила, генетичні алгоритми та багато інших.

Таблиця 4.1

Методи і засоби інтелектуального аналізу даних для виявлення атак

|  |  |
| --- | --- |
| Найменування методу | Реалізовані функції в задачі виявлення мережевих атак |
| Методи класифікації | Віднесення аналізованих векторів до множинам нормальних і аномальних |
| Методи скорочення  розмірності | Збільшення швидкодії з допомогою формування оптимізованого простору ознак |
| Методи кластеризації | Побудова оптимізованого безлічі модулів виявлення |
| Нечітка логіка | Організація взаємодії модулів виявлення  створення надлишкової модульної архітектури СВВ |

За результатами аналізу безлічі досліджень в якості класифікатора обраний метод опорних векторів. Даний метод показує одні з кращих показників виявлення атак і має широкі можливості по внутрішньої налаштуванні.

Метод опорних векторів відшукує зразки, що знаходяться на кордонах між двома класами, які і називаються опорними векторами. Решта функціональні завдання, які вирішуються системами виявлення мережевих атак, в основному спрямовані на підвищення якості виявлення, швидкодії системи, уніфікації і виконання інших допоміжних завдань.

Навчання методу головних компонент є складною обчислювальної завданням. Крім того, якість класифікації істотно залежить від внутрішніх налаштувань даного методу, індивідуальних для різних видів атак і навчальних множин. У зв'язку з цим існує необхідність спрощення процедури підготовки моделі класифікації. Рішенням даної проблеми є скорочення розмірності, в тому числі відкидання шумів і викидів, і розбиття тренувального безлічі на частини - виконання процедури кластеризації.

Основна мета методів скорочення розмірності - пошук простору меншої розмірності, в якому зберігаються внутрішні властивості вихідних даних. Ці методи дозволяють визначити безліч найбільш важливих параметрів для виявлення конкретної атаки. Вибір конкретного методу скорочення розмірності сильно залежить від тренувальних даних. Для розв'язуваної задачі найбільш перспективними є метод головних компонент [46].

Мета кластерного аналізу - розбивка безлічі даних на групи таким чином, щоб мінімізувати відмінності між елементами однієї групи і максимізувати відмінності між елементами різних груп. Основні методи кластерного аналізу підрозділяються на ієрархічні і неієрархічні.

Ієрархічні методи дозволяють побудувати оптимальну структуру кластерів, але мають експонентну залежність від кількості записів. У зв'язку з великими розмірами масивів тренувальних даних застосування ієрархічної кластеризації до всього масиву даних неможливо. Але при розгляді фрагментів тренувальних даних, що містять записи окремих атак, обсяг інформації дозволяє застосувати ієрархічну кластеризацію.

З неієрархічних методів кластерного аналізу найбільшого поширення набули ітераційні методи. Ітераційні методи більш універсальні, ніж ієрархічні, але мають один серйозний недолік - необхідність апріорного знання про число кластерів, що істотно ускладнює їх автоматизацію.

Для завдання формування складу модулів виявлення обрані як ієрархічні методи, так і ітераційні. Ієрархічні застоСВВуються до подмножеству атак з навчальної множини, ітераційні дозволяють кластеризувати все навчальну множину, вихідне число кластерів і їх центри визначається на основі ієрархічної кластеризації підмножини атак.

Нечітка логіка - це узагальнення класичної формальної логіки і теорії множин. Замість значень брехня і істина розглядається функція приналежності елемента нечітких множин, значення якої знаходяться в інтервалі [0; 1]. Для нечіткої логіки є узагальнені логічні операції, засновані на операціях з нечіткими множинами. Сформована база нечітких правил дозволяє співвіднести виявлену аномалію з безліччю відомих атак. Модульна архітектура дозволяє побудувати надлишкову модель, збільшуючи ймовірність виявлення атак, а нечіткий висновок дозволяє скоротити число помилкових спрацьовувань [48].

Нечітка логіка дозволяє вирішити проблему наявності однакових записів c різними мітками в навчальних даних: «сумнівні» мережеві пакети відносяться до безлічі аномальних з певною ймовірністю.

Побічна застосування нечіткої логіки - це побудова пересічних кластерів і розширення методу опорних векторів.

**4.2. Основи обраних методів інтелектуального аналізу даних**

***Метод опорних векторів***. Метод опорних векторів (англ. Support Vector Machine, SVM) - це набір схожих алгоритмів категорії «навчання з учителем», застоСВВуваних у задачах класифікації та регресійного аналізу. Даний метод належить до сімейства лінійних класифікаторів. Характерною особливістю методу опорних векторів є постійне скорочення емпіричної помилки класифікації і збільшення зазору між класами. Тому даний метод часто називають методом класифікатора з максимальним зазором.

Метод відшукує елементи, що знаходяться на кордонах між двома класами, які і називаються опорними векторами.

Метод опорних векторів здійснює пошук лінійної функції, яка дозволяє віднести елементи набору даних до одного з двох класів. Завдання бінарної класифікації може бути сформульована як пошук лінійної функції *f(x)*, яка приймає значення менше нуля для елементів одного класу і більше нуля для елементів іншого.

Розподілена гіперплощина має наступний вигляд:

де *w* - вектор, перпендикулярний до розподіленої гіперплощини, параметр *b* визначає відстань гіперплощини від початку координат.

Гіперплощина, паралельні оптимальної гіперплощини і найближчі до опорним векторах двох класів, можуть бути описані наступними рівняннями:

Якщо навчальна множина даних лінійно нероздільні, то можна вибрати гіперплощини так, щоб в смугу між ними не потрапляла жодна точка навчальної вибірки і потім максимізувати відстань між гіперплоскостямі. Ширина смуги в цьому випадку дорівнює , тому слід мінімізувати ‖𝑤‖. Для виключення всіх точок з смуги, повинна виконуватися умова:

де - мітка класу, що приймає значення -1 і +1,

- вектор робочої вибірки з міткою класу.

Це завдання квадратичної оптимізації еквівалентна задачі пошуку сідлової точки функції Лагранжа:

де *L* - функція Лагранжа,

*λ*𝑖 - множники Лагранжа.

Щоб узагальнити SVM на випадок лінійної нероздільності, вводиться константа *С* - внутрішній параметр методу, що дозволяє регулювати відношення між максимізацією ширини розділової смуги і мінімізацією сумарної помилки.

Основною проблемою застосування методу опорних векторів в завданні бінарної класифікації є складність пошуку лінійної кордону між двома класами. У разі якщо таку кордон побудувати не вдається, одне з рішень - це збільшення розмірності (перенесення даних в інший простір, більш високої розмірності), де існує можливість побудови площині, що розділяє безліч елементів на два класи [47].

***Методи скорочення розмірності***. Методи скорочення розмірності - це широке коло різних алгоритмів, основним завданням яких є пошук простору меншої розмірності, в якому зберігаються внутрішні властивості вихідних даних. Скорочення розмірності може бути обумовлено рядом причин:

необхідність візуалізації вихідних даних;

спрощення обчислень і інтерпретація отриманих статистичних висновків;

необхідність стиснення обсягів збереженої статистичної інформації.

З математичної точки зору завдання скорочення розмірності можна представити в наступному вигляді: дана p-мірна змінна x = (x*1, x2, ..., xp*)T і необхідно знайти простір меншої розмірності, в якому змінна s =*(s1, s2, ..., sk)T*,

*K ≤ p* відображає зміст вихідних даних відповідно до деякого критерію. Компоненти s іноді називають прихованими (або латентними) компонентами. У різних випадках використовуються різні назви p-мірних змінних: термін «змінна» в основному використовується в області статистики, в той час як терміни «особливість» і «атрибут» широко використовуються в комп'ютерній науці і машинному навчанні [46].

Розрізняються два глобальних класу методів скорочення розмірності: лінійні методи і нелінійні методи. Для лінійних методів результатом кожної *k ≤ p* компоненти буде лінійна комбінація вихідних змінних:

*si = wi; 1x1 + ... + wi; pxp,*

де *i = 1, ..., k; s = Wx; Wkxp* - матриця ваги лінійних перетворень.

Цей же взаємозв'язок можна представити як *x = As,*

де *Apxk*- матриця, а *s* - це, так звані, приховані або латентні показники. Тоді в позначеннях спостережень *Xpхn*, Дані співвідношення виглядають наступним чином:

*si, j = wi, 1 x1, j + ... + wi, pxp, j, I = 1, ..., k, j = 1, ..., n,*

де *j* відповідає *j*-й вибірці.

*Skхn = WkхpXpхn,*

*Xpхn = Apxk Skxn,*

Лінійні методи більш наочні і значно простіше у використанні, ніж більш сучасні методи, засновані на нелінійних перетвореннях.

В окремий клас можна виділити методи, які обмежуються розглядом статистичних моментів другого порядку. Такі методи обчислювально прості і привертають лише класичні операції з матрицями, не вимагаючи розробки процедури пошуку в просторі параметрів перетворення. Це історично перші і найбільш розроблені методи, серед яких найбільш відомими є класичні методи: метод головних компонент (МГК) і факторний аналіз (ФА). Природно, для спільного вирішення завдання уявлення багатовимірних даних необхідно подолати два обмеження: обмеження на лінійність перетворення і обмеження на розгляд моментів другого порядку.

В даних методах пошук уявлення ведеться при використанні тільки інформації, що міститься в коваріційній матриці випадкового вектора, відповідного вихідними даними. Це є виправданим, якщо розподіл випадкового вектора містить в якості параметрів тільки моменти другого порядку. Типовим прикладом такого розподілу може служити нормальний розподіл. І дійсно, якщо розподіл описується гаусовською функцією, то воно визначається тільки інформацією другого порядку і залучення іншої інформації безглуздо.

Оскільки дані методи гуртуються на моментах другого порядку, то критерієм оптимальності в них буде виступати дисперсія по кожній з компонент вектора, що означає оптимізацію опису по його точності в сенсі середньоквадратичного відхилення. Відразу можна навести такий недолік такого підходу. Нехай дані висловлюють деякі фізичні величини, наприклад, масу і відстань. Тоді зміна одиниці виміру, скажімо, грама на кілограм, призведе до іншого результату. Тобто різні ознаки, складові вектор, повинні апріорі покладатися в деякому сенсі рівноправними [46].

До використання більш загальних, ніж нормальний розподіл, моделей вдаються рідко. Це пов'язано з тим, що визначення параметрів цих моделей вимагає істотно більшого обсягу обчислень, і з тим, що такі моделі залишаються досить приватними, чи не застосовуються в загальному випадку. Тому більш популярний інший підхід, що полягає у використанні деякого універсального критерію якості напрямків в просторі ознак, але на відміну від середньоквадратичного відхилення цей критерій повинен задіяти всю інформацію про функції розподілу.

Найчастіше розглядають два таких критеріїв. При першому вибираються такі компоненти, які дозволяють мінімізувати надмірність даних. Оскільки цей підхід часто аргументується на основі нейрофізіологічних даних, згідно з якими в природних нейронних мережах (виконують первинну обробку сенсорної інформації) проводиться зменшення надмірності, то і сам критерій зазвичай використовується при проектуванні нейронних мереж. При цьому під надмірністю зазвичай розуміється не надмірність в сенсі класичної теорії інформації, а то, наскільки часто відбувається спільна активація нейронів. Тобто перший шар персептрона розглядається в якості пристрою, що повертає систему координат в просторі первинних ознак, а ваги налаштовуються таким чином, щоб нейрони другого шару спільно активувалися як можна рідше.

При використанні другого критерію прагнуть виділити найбільш «цікаві» напрями. Під «цікавістю», як правило, розуміють те, наскільки по даному напрямку розподіл відрізняється від нормального розподілу. Можна показати, що при фіксованій коваріційній матриці максимум ентропії припадає саме на гаусовський розподіл. Для будь-якого іншого розподілу, ентропія строго менше.

***Метод головних компонент***. Метод головних компонент (англ. Principal components analysis, РСА) є одним з основних способів зменшення розмірності даних, при якому втрачається найменшу кількість інформації. Обчислення головних компонент здійснюється за рахунок обчислення власних векторів і значень ковариційної матриці навчальної множини. Метод головних компонент також відомий як перетворення Карунєна-Лоєва або перетворення Хотеллінга [48].

Завдання аналізу головних компонент має чотири базових версії:

* апроксимація даних лінійними різноманіття меншої розмірності;
* пошук підпростору меншою розмірності, в ортогональної проекції на яке середньоквадратичне відхилення від середнього значення (розкид) максимальний;
* пошук підпростору меншою розмірності, в ортогональної проекції на яке середньоквадратичне відстань між точками максимально;
* побудова такого ортогонального перетворення координат для вихідної багатовимірної випадкової величини, що в результаті кореляції між окремими координатами звернуться в нуль.

Ідея методу головних компонент полягає в зменшенні розмірності даних шляхом пошуку декількох ортогональних лінійних комбінацій (головних компонент) первинних змінних з найбільшою дисперсією. Перша головна компонента, s*1,* є лінійною комбінацією початкових величин і має найбільшу дисперсію:

*s1* = xTw1, де p-мірний вектор *w1* = (W*1,1, ..., w1, p)T* знаходиться як *w1 = argmax|| w = 1 ||Var {xTw}*

Друга головна компонента - це лінійна комбінація первинних змінних з другою за величиною дисперсією і ортогональна першій компоненті. Аналогічно можна визначити і всі інші головні компоненти. Число головних компонент збігається з числом початкових величин. Для багатьох типів даних кілька перших головних компонент відображають основні залежності даних від первинних величин, так що іншими компонентами можна знехтувати з мінімальними втратами інформації.

***Факторний аналіз***. Факторний аналіз (Factor analyze, FA), також як і метод головних компонент, є методом другого порядку, що працює з лінійними перетвореннями. У факторному аналізі використовується припущення, що ступінь відхилення змінних залежить від невідомих, і часто невимірних, загальних факторів. Метою методу факторного аналізу є виявлення таких залежностей, які можуть бути використані для скорочення розмірності даних.

Випадковий вектор *x* з нульовим середнім значенням задовольняє *k*-фактору моделі, якщо виконується така умова: *x = Λf +* u, де *Λp × k* - це матриця констант, *fk* і *up*- це, відповідно, приховані чинники і специфічні фактори. Крім того, необхідно, щоб всі приховані чинники не корелювали між собою і були стандартизованими таким чином, щоб їх дисперсія дорівнювала одиниці. За допомогою даного методу виявляються приховані залежності між вихідними величинами, в результаті чого можна виключити надлишкові дані. На практиці застосовуються два методи, засновані на факторному аналізі: метод головних факторів і метод максимальної правдоподібності [48].

***Метод пошуку найкращої проекції***. Пошук найкращої проекції - це також лінійний метод, але він, на відміну від методу головних компонент і факторного аналізу, ґрунтується на статистиках високих порядків, що корисно при невизначених даних. Цей метод вимагає більшого обсягу обчислень, ніж методи другого порядку.

Метод пошуку найкращої проекції заснований на понятті «цікавість» напряму проектування, яка визначається функцією якості (індексом проекції). Розподіл Гаусса є найменш цікавим розподілом, тому індекс проекції зазвичай вимірює деякі аспекти негаусовських розподілів. Якщо, проте, використовувати максимальну дисперсію другого порядку в якості індексу проекції за умови, що проекції будуть ортогональні, тоді метод пошуку найкращої проекції дає схожі з методом головних компонент результати [48].

***Нелінійний метод головних компонент***. Нелінійний метод головних компонент вводить нелінійність в цільову функцію, але, в результаті, компоненти як і раніше є лінійними комбінаціями початкових змінних. Цей метод показує сильний зв'язок між методом незалежних компонент і методом головних компонент. Вводячи нелінійність g, засновану на щільності ймовірностей незалежних компонентів в цільову функцію методу головних компонент, виходить модель незалежних компонент:

***Метод головних кривих***. Метод головних кривих - нелінійний метод скорочення розмірності, що продовжує концепцію методу головних компонент. Головні криві - це гладкі криві, що проходять через «середину» багатовимірного набору даних. Лінійні головні криві еквівалентні головних компонентів. Концепція головних кривих може бути розширена до головних поверхонь, але оцінна процедура стає ще більш складною.

***Генетичні і еволюційні алгоритми***. Генетичні і еволюційні алгоритми - це евристичні алгоритми пошуку, які використовуються для вирішення завдань оптимізації та моделювання шляхом послідовного підбору, комбінування і варіації шуканих параметрів з використанням механізмів, що нагадують біологічну еволюцію . Повний перебір кінцевого невеликого простору або будь-який ефективний алгоритм спуску буде завжди ефективніше, ніж генетичний алгоритм. Коли про завдання нічого апріорі невідомо, можна покладатися на результат роботи найпростішого генетичного алгоритму як якогось наближення.

Висновки по методам скорочення розмірності. В попередніх підрозділах наведені основні методи скорочення розмірності. Вибір конкретного методу може бути заснований на апріорному знанні ймовірнісної або геометричної природи оброблюваних даних або виходячи із специфіки розв'язуваної задачі. Для вирішення завдань візуалізації даних рекомендується використовувати методи, які використовують різні проекції, криві і багатовимірне шкалювання. Якщо візуалізації не потрібно, краще використовувати метод головних компонент або метод незалежних компонент. Альтернативним варіантом вирішення будь-яких зазначених завдань є використання нейронних мереж [46].

Для задач визначення необхідного та достатнього безлічі параметрів трафіку для виявлення конкретної атаки, а також для очищення тренувальних даних від шумів і викидів обраний метод головних компонент. Результатом його роботи, є побудова матриці ваг для обчислення нових параметрів. Для цього методу має значення вибір аналізованої матриці: матриця кореляції, ковариации або матриця суми квадратів і змішаних творів.

***Методи кластеризації***. Методи кластеризації - це безліч алгоритмів, метою яких є поділ безлічі об'єктів на групи таким чином, щоб схожі об'єкти потрапили в одну групу, а різні групи містили об'єкти з різнорідними характеристиками. Якість кластеризації характеризується високим схожістю об'єктів всередині кожної групи і високими відмінностями між групами.

Ключовими вимогами до вихідних даних для виконання процедури кластеризації є - однорідність і повнота даних. Однорідність вимагає, щоб всі кластерізуемие записи описувалися подібним набором характеристик.

Кластерний аналіз застосовується для вирішення наступних завдань:

розбиття об'єктів на групи з метою оптимізації подальших обчислень і аналізу;

скорочення обсягу даних шляхом аналізу лише окремих представників кластеру замість обробки всіх даних;

виділення нетипових об'єктів, яких неможливо віднести до відомих категоріям (класах).

Ієрархічні алгоритми кластеризації. Ієрархічні методи кластеризації будуються на послідовному поділі множини об'єктів на вкладені групи. Структура кластерів представляється у вигляді дерева, яке розщеплює набір даних рекурсивно на більш дрібні підмножини. Кожен вузол дерева, за винятком листя, є об'єднанням своїх подкластера, а корінь дерева є кластером, що містить всі об'єкти. Іноді лист дерева називають одноелементна кластером.

Обидві категорії ієрархічних алгоритмів мають наступні властивості: не вимагається апріорне знання про число кластерів; високу швидкодію; дерево допомагає візуалізувати кластери і асоціації між ними [48].

Головним недоліком цих методів є неможливість виконувати коригування після прийняття рішення про об'єднання або дробленні. Для великих наборів даних обчислювальна складність створення ієрархічних кластерів є дуже високою.

Ієрархічні методи кластеризації використовують різні критерії для визначення кластерів, які повинні бути об'єднані або розділені, при цьому використовуються різні метрики відстані між підмножинами:

оцінка відстані між двома найбільш схожими елементами двох кластерів. Дана метрика добре справляється з неелліптіческімі формами, але дуже чутлива до шумів і викидів;

оцінка відстані між двома найбільш віддаленими елементами двох кластерів. Дана метрика менш чутлива до шумів і викидів, але має проблеми з опуклими формами і може неадекватно розділяти великі кластери;

одночасна оцінка відстані між найбільш і найменш схожими елементами. Дана метрика є компромісом між двома попередніми.

**4.3. Формалізація моделі системи виявлення мережевих атак**

Формування архітектури системи виявлення мережевих атак. Як зазначалося раніше, система виявлення мережевих атак складається з ряду функціональних блоків і бази вирішальних правил. Особливості внутрішнього побудови блоку вилучення даних - сенсорів і блоку реакції виходять за рамки дослідження, тому що не мають принципової залежності від використовуваних засобів аналізу даних. Єдиною вимогою до даних блокам є можливість їх застосування в складі розподіленої обчислювальної мережі, тобто модульна структура блоку вилучення даних і можливість мережевої взаємодії основних компонентів системи з блоком реакції.

Для функціонування системи виявлення мережевих атак в розподіленої обчислювальної мережі також важливі елементи, що відповідають за взаємодію між компонентами системи виявлення.

В якості основного елемента системи виступає модуль виявлення - неподільна в рамках розподіленої обчислювальної мережі частина, що відповідає за виявлення певних атак або аномальних характеристик.

Обов'язковим елементом модуля виявлення є блок класифікації. З його допомогою аналізовані багаторозрядні вектора відзначаються як нормальні або аномальні, рис. 4.3.

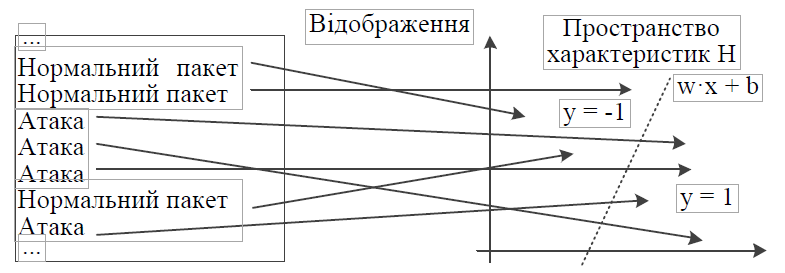


Рис. 4.3. Задача блока класифікації

Як класифікатор в роботі розглядається метод опорних векторів. Застосування даного методу сильно залежить від природи оброблюваних даних. Зокрема, існує ряд налаштувань, які необхідно виконати перед навчанням даного методу. У зв'язку з цим для побудови якісного класифікатора необхідно виконання не тільки навчання, а й тестування методу опорних векторів [48].

Для можливості системного формування модуля виявлення необхідна наявність блоку автоматичної настройки, аналізує побудовану SVM-модель (число опорних векторів) і результати тестування роботи методу опорних векторів (кількість правильно класифікованих пакетів, помилок першого і другого роду) і приймає рішення про зміну внутрішніх налаштувань методу опорних векторів.

Спрощена схема застосування методів опорних векторів представлена на рис. 4.4.

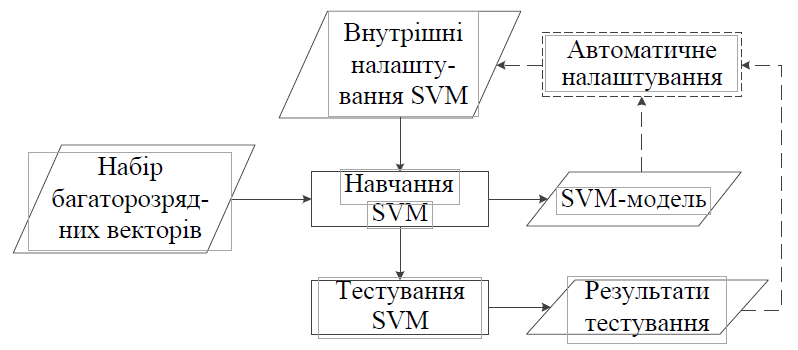


Рис. 4.4. Застосування методу опорних векторів (SVM)

Для можливості навчання методу опорних векторів існує ряд вимог до тренувальних даними. Дані, витягнуті з мережевого трафіку, складають масиви високої розмірності і досить значного обсягу, що містять безліч шумів і викидів. У зв'язку з цим необхідна наявність блоку попередньої обробки даних. З цим завданням якнайкраще справляються методи скорочення розмірності.

Блок скорочення розмірності (рис. 4.5) вирішує два основні завдання: визначає підмножина вихідних параметрів (назвемо їх базовими) і формує безліч параметрів в розрахунковому просторі (назвемо їх новими). Для методу головних компонент правило перекладу базових параметрів в нові являє собою лінійне перетворення.

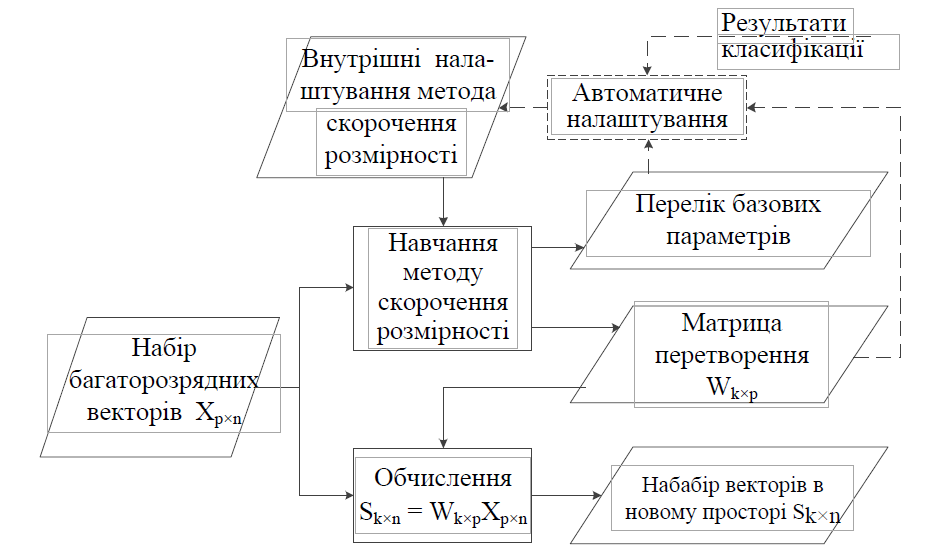


Рис. 4.5. Застосування методів скорочення розмірності

Аналогічно методу опорних векторів для блоку скорочення розмірності необхідно розміщення блоку автоматичного підбору параметрів.

В ході експериментального дослідження була виявлена наступна проблема: складні розподілені атаки складаються з безлічі мережевих пакетів, знаходяться на великій відстані один від одного в просторі базових параметрів.

При цьому безліч пакетів, що мають мітку нормального трафіку, знаходяться на малій відстані від пакетів атак. В результаті навчання методу опорних векторів на подібному навчальній множині за прийнятний час неможливо і скорочення розмірності ніяк не впливає на ситуацію [46].

У той же час за допомогою блоку візуалізації було виявлено, що для подібних за типом атак часто формуються подібні SVM-моделі.

У зв'язку з цими спостереженнями виникла необхідність перерозподілу навчальних пакетів між модулями виявлення: пакети складних атак розбиваються на кілька груп - кластерів та обробляються незалежно один від одного; аналогічні фрагменти схожих за типом атак поміщаються в єдині модулі виявлення.

Поділ тренувального безлічі на групи - це завдання методів кластеризації.

На рисунку 4.6 відображена схема застосування методів кластеризації.

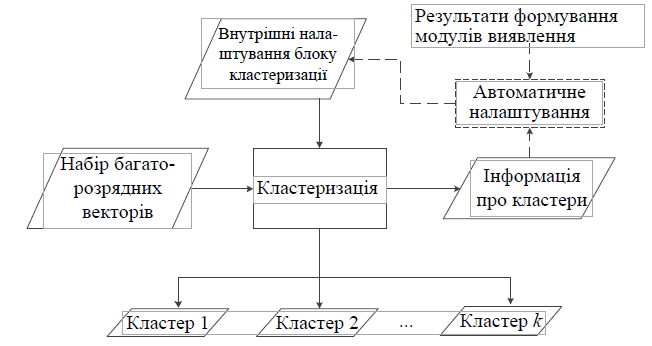


Рис. 4.6. Застосування методів кластеризації

Проведення кластеризації можливо як на всьому навчальній множині, так і на безлічі атак. У першому випадку можливе виділення підмножин, що складаються з одних атак або тільки з векторів нормального трафіку, що дозволяє виключити з навчальної множини окремо стоять кластери, які не потребують підготовки класифікатора. У другому випадку формується безліч кластерів, що описують центроїди для векторів з мітками «атака», що дозволяє побудувати відносно прості локальні SVM-моделі, що дозволяють класифікувати вектора, що знаходяться поблизу даних центрів, з мінімальною обчислювальною складністю.

При підготовці бази вирішальних правил (навчанні системи) блок кластеризації є першим функціональним компонентом в ланцюжку блоків системи. Тому час, необхідний на навчання системи виявлення, в першу чергу залежить від якості проведеної процедури кластеризації. В результаті аналізу методів кластеризації для розв'язуваної задачі були обрані метод k-середніх для кластеризації великих обсягів даних і агломеративного ієрархічний метод кластеризації для побудови оптимального безлічі кластерів на невеликих навчальних множинах, наприклад, на безлічі пакетів окремого виду атаки [48].

На рисунку 4.7. відображена схема формування модулів виявлення.

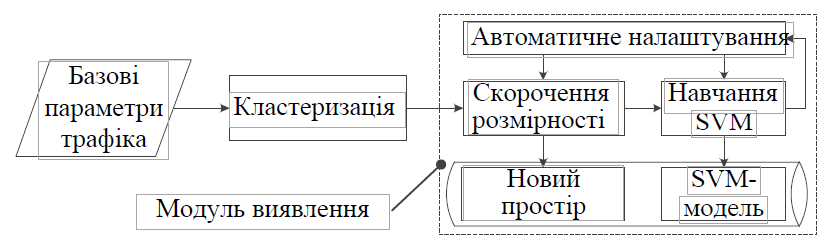


Рис. 4.7. Схема формування модулів виявлення

Застосування методів кластеризації дозволяє не тільки з високою ймовірністю виявити складні розподілені атаки, але і істотно підвищити продуктивність системи. За рахунок формування декількох простих модулів виявлення замість одного складного, що містить сотні опорних векторів, скорочується час навчання і підвищується швидкість аналізу трафіку.

При формуванні кожного модуля виявлення набір даних, складений з базових параметрів трафіку з мітками нормальний / аномальний, навчається на методах скорочення розмірності. В результаті відкидаються найменш значущі базові параметри і визначаються найбільш важливі параметри в новому просторі. У цьому новому просторі набір даних навчається на методі опорних векторів і формується розділяюча гіперплощина - SVM-модель. Блок автоматичного налаштування підбирає внутрішні налаштування інших блоків. Результатом створення модулів являються набори даних, які розміщуються в базі вирішуючих правил.

Схема аналізу мережевого пакета приведена на рисунку 4.8.

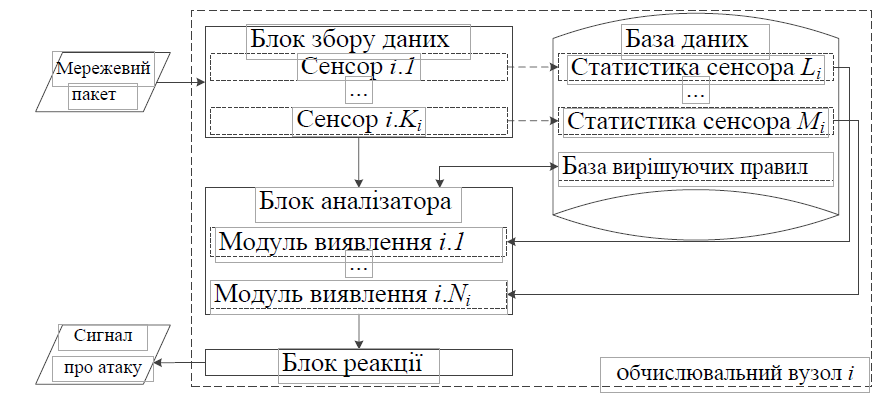


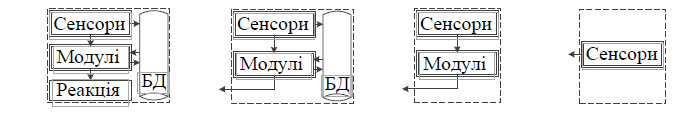
Рис. 4.8. Схема аналізу мережевого пакета

Після вилучення безліччю сенсорів загальний список базових параметрів передається на модулі виявлення. В процесі роботи системи виявлення в базі даних відбувається накопичення сигналів модулів про виявлення потенційно небезпечного трафіку. Рішення про вторгнення приймається за допомогою нечітких правил виводу, що застосовуються до поточних сигналам всіх модулів.

Метод адаптації системи виявлення мережевих атак під програмно-апаратну структуру розподіленої обчислювальної мережі. Вузли розподіленої обчислювальної мережі мають різну апаратну платформу, операційну систему, безліч встановлених сервісів і призначених для користувача програм, а також різне призначення в рамках всієї мережі. Від перерахованих чинників залежать потенційні цілі порушника, його можливості і засоби, які застосовуються. Так само критичним фактором може виступати продуктивність вузла, що вимагає мінімізувати будь-яку допоміжну навантаження, включаючи засоби забезпечення інформаційної безпеки [47].

В зв'язку з цим необхідна побудова гнучкої архітектури системи виявлення, здатної змінювати склад функціональних блоків і модулів виявлення. Обов'язковою функціональним блоком для контрольованого вузла є блок вилучення даних - безліч сенсорів. Блок реакції, блок аналізатора і блок бази даних можуть бути відсутні в залежності від вимог до роботи вузла.

Варіанти наповнення вузлів обчислювальної мережі функціональними блоками системи виявлення представлені на рис. 4.9.



а) б) г) д)

Рис. 4.9. Варіанти розміщення функціональних блоків системи виявлення мережевих атак в окремому вузлі

Формування загальної бази даних для декількох вузлів дозволяє організувати мережу з простих варіантів підсистеми виявлення, що складаються тільки з безлічі сенсорів.

Основна можливість адаптації системи виявлення організовується за рахунок модульної архітектури. Окремі модулі виявлення можуть бути сформовані для точного виявлення вузького кола атак або аномалій.

Всі можливості адаптації функціональних блоків системи виявлення представлені на рисунку 4.10.

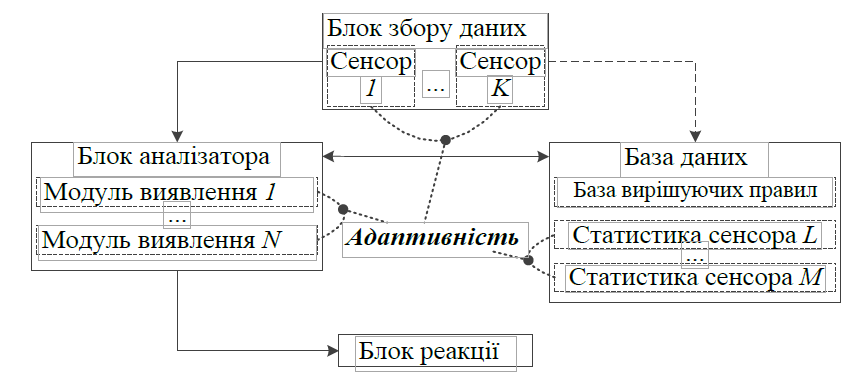


Рис. 4.10. Можливості адаптації функціональних блоків системи виявлення під програмно-апаратну структуру вузла

Кожному модулю виявлення ставиться у відповідність безліч класів атак, на які він реагує з певною ймовірністю. У навчальній вибірці для розглянутих атак заповнюються такі параметри як види потенційних цілей, категорія атаки і будь-які інші настроюються характеристики. Для динамічного зміни блоку збору даних з кожним модулем виявлення асоціюється список сенсорів, що забезпечують витяг необхідних даних з трафіку.

Залежність модулів виявлення з компонентами інших функціональних блоків і принципи оцінки необхідності їх використання зображені на рис 4.11.

Адаптивність системи реалізується можливістю автоматичної зміни безлічі сенсорів і модулів виявлення в залежності від структури програмно-апаратного середовища і множини потенційно можливих атак.

В якості висновку можна зтверджувати, що представлена модель системи виявлення мережевих атак в розподіленої обчислювальної мережі ґрунтується на безлічі різнопланових методів інтелектуального розподілу даних.

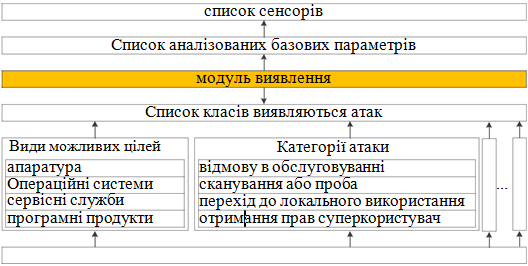


Рис. 4.11. Місце модуля виявлення в адаптивній системі

Для класичної моделі системи виявлення вторгнень виділені функціональні компоненти і можуть бути вирішені підзадачі, проаналізовані можливості застосування різних категорій методів інтелектуального аналізу даних і сформульована модель розподіленої системи.

В завданні виявлення мережевих атак методи інтелектуального аналізу даних дозволяють вирішити такі завдання:

формування ефективної модульної архітектури системи виявлення;

співвіднесення аналізованих векторів до безлічі атак або нормальних;

перетворення даних для оптимізації роботи блоку класифікації.

Представлена ​​модель системи дозволяє виявляти більшість мережевих атак і аномалій за рахунок широких можливостей щодо формування модулів виявлення.

Представлена ​​архітектура розподіленої системи і метод адаптації дозволяє підлаштовувати функціонал системи виявлення під фактичну програмно-апаратну структуру і призначення вузлів обчислювальної мережі.

**4.4. Формування методики застосування методів інтелектуального аналізу даних в задачі виявлення мережевих атак**

Основи застосування методу головних компонент. Попередня обробка тренувальних даних полягає в очищенні від викидів і шумів, масштабування, нормалізації і розбитті на фрагменти для можливості не тільки навчити метод головних компонент, а й протестувати результати побудованого класифікатора. У запропонованій моделі системи виявлення мережевих атак очищення від викидів і шумів проводиться за допомогою процедур кластеризації і скорочення розмірності. Методи кластеризації обмежують тренувальне підмножина, звільняючи його від викидів, а методи скорочення розмірності переводять дані в новий простір, виконуючи нормалізацію.

Для навчання і тестування методу головних компонент необхідно розділити набір тренувальних даних на дві частини, одна з яких вважається невідомою. Для першої частини проводиться навчання, для другої обчислюється точність передбачення. Найбільш об'єктивна оцінка ефективності класифікації здійснюється за рахунок виконання процедури перехресної перевірки (англ. Cross-validation) - узагальнення ідеї розбиття тренувального безлічі на дві частини. Набір даних розділяється на v підмножин однакового розміру. Послідовно виконується навчання методу головних компонент на одному з v підмножин і тестування на інших *v* - 1 підмножинах. Таким чином, виключається вплив шумів і скорочується проблема перенавчання [48].

Механізм ядер застосовується для побудови кордонів між класами в разі лінійної нероздільності вихідних даних. Природа розглянутих в задачі виявлення мережевих атак даних практично виключає лінійну роздільність між множинами векторів, витягнутих з нормального трафіку і мережевих пакетів, що містять атаки. Численні дослідження по застосуванню методу опорних векторів в самих різних областях доводять, що вибір конкретного ядра залежить від конкретних навчальних даних і не може бути визначено апріорі. Крім особливостей навчальної множини даних на вибір ядра впливає складність підбору внутрішніх параметрів ядра. У таблиці 4.1 перераховані параметри, що настроюються в методі опорних векторів.

Таблиця 4.1

Перелік налаштувань для методу опорних векторів

|  |  |
| --- | --- |
| Ядро SVM | Параметри |
| Радіально-базисне | *γ* і *C* |
| Сигмоїдальне | *γ, coef0* і *C* |
| Поліноміальне | *γ, coef0, degree* і *C* |

У більшості досліджень, пов'язаних з методом опорних векторів, розглядається тільки радіальне базисне ядро. Це викликано числом підбираються параметрів в цілому, хорошими результатами для більшості даних і можливістю двовимірної візуалізації результату.

Згідно з дослідженнями авторів бібліотеки SVM для визначення оптимальних параметрів необхідно побудувати сітку пошуку за параметрами *C* і *γ*, використовуючи експоненціально зростаючі послідовності *C* і *γ*. Вибирається пара (*C, γ*) з найкращим показником розпізнавання. Далі, при необхідності, обчислюються значення матриці для інших *i* і *j*. Після знаходження оптимальних значень матриці, виконується побудова такої ж матриці в області найкращих значень, тільки крок зміни *i* і *j* скорочується з 1 до 0,25.

Особливості даних розглянутої предметної області для застосування методів опорних векторів. Для дослідження роботи створеної системи виявлення вторгнень було проведено масштабне експериментальне дослідження щодо формування класифікаторів для виявлення різних мережевих атак за допомогою методу опорних векторів. Для розглянутих даних, що надходять в метод опорних векторів, були виявлені такі характерні особливості:

в досліджуваних даних число пакетів з міткою «атака» не перевищує декількох відсотків від загального числа пакетів, а в більшості досліджень щодо застосування методу опорних векторів з різними даними використовується порівнянне кількість векторів у двох класах;

відносно невисока розмірність досліджуваних даних - не більше трьох десятків параметрів проти кількох десятків або декількох сотень в інших відомих дослідженнях методу опорних векторів.

Проблема малого числа векторів з міткою атака була вирішена наступним чином: безліч векторів з мітками атака випадковим чином ділиться на 3 підмножини, безліч векторів з міткою норма - на 9. Потім формуються 9 підсумкових підмножин, що містять підмножина з векторів атак і підмножина з нормальних векторів. Для додаткової перевірки поділяються підмножини векторів з мітками атаки, витягнуті з різних дампів.

В увазі перерахованих особливостей був сформований наступний алгоритм підбору оптимальних параметрів методу опорних векторів з урахуванням застосування описаної модифікації процедури перехресної перевірки при кожному циклі навчання і тестування роботи блоку:

1. Обчислюється число опорних векторів і відсоток правильно класифікованих пакетів в інтервалі *i* = 4 ... 8 і кроком 1 при *j* = 25. На підставі цих 5 точок вибирається перспективний напрямок для подальших обчислень по параметру *i*.

Визначається точка (параметр *i*) з мінімальною кількістю опорних векторів в розрахунковому інтервалі і для цього параметра i обчислюються сусідні точки по параметру *j* з кроком 1. Таким чином, знаходиться найменше значення *j*, Відповідне найменшому числу опорних векторів. З усього набору обчислених значень знаходиться точка з мінімальним кількістю опорних векторів. Точка з мінімальним числом опорних векторів вважається кращим рішенням [49].

Завдяки проведеному аналізу та огляду методів виявлення атак ІБ в кіберпросторі, було визначено, що ці завдання є невід’ємною частиною процесу.

На підставі порівняльного аналізу методу визначення аномалій, а також сигнатурного методу та методу опорних векторів, було зроблено висновок про доцільність застосування комплексного підходу, що можна класифікувати як гібридний метод виявлення інцидентів, до рішення задач виявлення інцидентів, що включає метод опорних векторів, в доповненні до сигнатурних систем, що застосовуються на практиці (рис. 4.12).

З одного боку, сигнатурний метод дозволяє описати інцидент набором правил або за допомогою формальної моделі, в якості якої може застосовуватися символьний рядок, семантичне вираження на спеціальній мові тощо. Суть даного методу полягає в використанні спеціалізованої бази даних шаблонів (сигнатур) інцидентів для пошуку дій, що підпадають під визначення "інцидент".

З іншого боку, метод опорних векторів можна використовувати для класифікації нових даних. Користувач вводить вхідні дані, і на основі розробленої (за допомогою навчання) моделі отримує вихідні результати. Теоретично, число входів для методу опорних векторів лежить в діапазоні від одного до нескінченності. Показник розпізнавання в даному методі дорівнює 95,5%, а доля хибних спрацювань складає лише 1% .

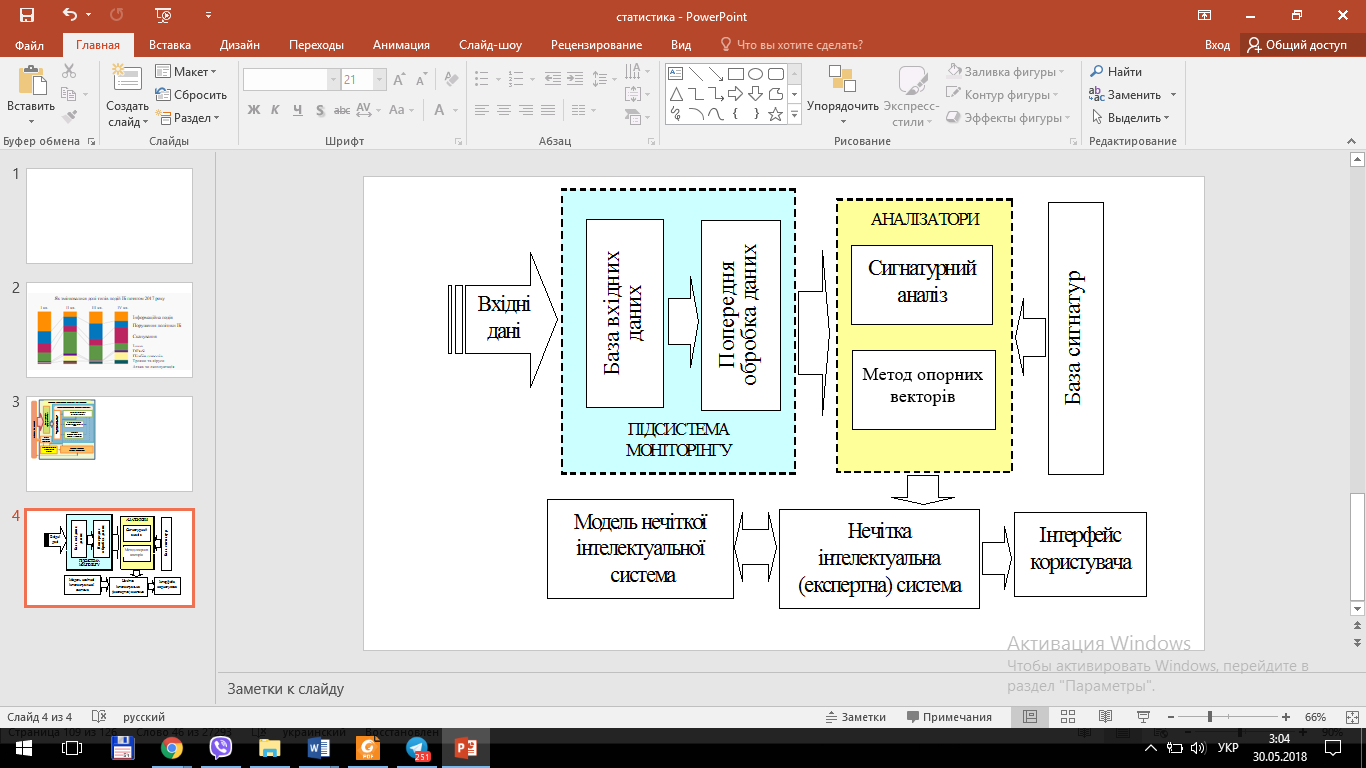


Рис. 4.12. Комплексний підхід до виявлення атак: гібридний метод сигнатурного

аналізу та методу опорних векторів для визначення вторгнень

Перша складова запропонованого комплексного підходу – універсальний сигнатурний аналізатор, структура якого зображена на рисунку 4.13. Механізм функціонування сигнатурного аналізатора включає два етапи: фільтрація і збірка фрагментів пакетів, розпізнавання вторгнень по сигнатурам.



Рисунок 4.13 - Структура сигнатурного аналізатора

Висновки по застосуванню методу опорних векторів. В результаті експериментального дослідження для всіх атак категорії U2R (відповідно до класифікації DARPA) і частини атак інших категорій з різними варіантами розрядності базових параметрів було сформовано загальний розподіл найкращих точок.

Алгоритм автоматического подбору параметрів методу опорних векторів виглядає наступним чином:

для параметра *C* = 225 обчислюються результати навчання і тестування методу опорних векторів при *γ* = 2, *γ* = 20 і *γ* = 21, якщо використовуються 8-бітові параметри; і при *γ* = 24, *γ* = 25 і *γ* = 26 для багаторозрядних параметрів;

на підставі обчислених точок визначається напрямок, в якому слід змінювати *γ*;

для зазначеного напрямку обчислюються результати роботи методу опорних векторів до отримання кращого значення;

на підставі всіх обчислених точок визначається діапазон значень *γ* з найкращими значеннями (найменшим числом опорних векторів);

для отриманого діапазону *γ* обчислюється «сітка пошуку» по параметру *C* як в меншу, так і в більшу сторону;

для найкращих точок і областей з сітки будуються нові сітки пошуки за параметрами *C* і *γ* з меншим кроком;

якщо серед точок з найбільшим відсотком розпізнавання число опорних векторів істотно більше, ніж в сусідніх точках сітки (зазвичай це точки під кривою 2), фіксується проблема перенавчання і змінюється навчальна множина (блок скорочення розмірності) з наступним запуском повного циклу підбору параметрів методу опорних векторів;

якщо проблеми перенавчання не зафіксовано, кращої вважається точка з найбільшим відсотком правильно розпізнаних векторів, найменшим числом опорних векторів і найменшими значеннями параметрів *C* і *γ*.

Зазначений алгоритм дозволяє автоматично визначити кращі параметри методу опорних векторів для роботи з даними заголовками мережевих пакетів. Після підбору параметрів порівнюються результати роботи класифікатора в звичайному і нечіткому режимі і вибирається найкращий варіант.

Методика експериментальних досліджень представлені в додатку А, В.

**Висновки до 4 розділу**

Описані в розділі методики дозволяють використовувати всі вибрані методи інтелектуального аналізу даних для вирішення відповідних підзадач, пов'язаних з виявленням мережевих атак.

У представлених методиках присутня значна кількість внутрішніх налаштувань. Для більшості налаштувань описані алгоритми по автоматичному вибору. Для інших видані рекомендації по ручному застосування при експериментальному побудові модулів виявлення для конкретних мережевих атак.

Метод опорних векторів в автоматичному режимі формує розділяє гіперплоскость між класами атак і нормальних векторів, при цьому застосовується радіально-базисна ядро і виконується підбір налаштувань і C для отримання SVM-моделі з мінімальним числом опорних векторів і показниками якості виявлення (максимальна кількість правильно класифікованих пакетів, мінімальна кількість помилок першого і другого роду).

Метод головний компонент застосовується в повністю автоматичному режимі і дозволяє сформувати простору ознак для виявлення заданого безлічі векторів (мережевих атак), а також скоротити обсяг інформації, що зберігається.

В базі вирішальних правил інформації, необхідної для класифікації мережевих пакетів, і підвищити швидкість формування модулів виявлення. Застосування двох порогових значень, коректованих блоком автоматичної настройки, дозволяє сформувати необхідне і достатнє підмножина параметрів нового простору для виявлення конкретної атаки.

Застосування нечіткої логіки дозволяє підвищити показники класифікаторів (розширення методу опорних векторів) і побудувати безліч пересічних кластерів для підвищення ймовірності виявлення мережевих атак. Нечіткі правила дозволяють віднести виявлену атаку до відомим при навчанні атакам.

**ВИСНОВОК**

В магістерській атестаційній роботі для захисту розподілених обчислювальних мереж від мережевих атак проведений аналіз існуючих засобів захисту та виявлено їх основні недоліки. Проведено аналіз застосовності різних методів і засобів Інтелектуального аналізу данихдля задач, пов'язаних з виявленням мережевих атак.

Побудована модель системи виявлення мережевих атак в розподілених обчислювальних мережах на основі різних груп методів інтелектуального розподілу даних.

Розроблено метод адаптації системи виявлення мережевих атак під програмно-апаратне середовище розподіленої обчислювальної мережі.

Розроблено архітектуру адаптивної системи виявлення мережевих атак в розподіленої обчислювальної мережі.

Розроблено методики застосування таких методів і засобів інтелектуального аналізу даних в завданню виявлення мережевих атак: метод головних компонент, метод опорних векторів, метод k-середніх, нечітка логіка.

Розроблено програмний прототип системи виявлення мережевих атак на основі вибраних методів інтелектуального аналізу даних і проведені експериментальні дослідження, що підтверджують ефективність створеної моделі виявлення для захисту розподіленої інформаційної мережі.

Проведені експерименти з програмним прототипом показали високу якість виявлення мережевих атак і довели правильність вибору методів інтелектуального аналізу даних і застосовність вироблених методик. Застосування різних методів, можливість настройки внутрішніх параметрів і порогових значень дозволяють домогтися оптимального співвідношення продуктивності системи і точності розпізнавання атак в розподіленої мережі.

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

**СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ**

1. Стратегія кібербезпеки України (електорнний ресурс).
2. В. П. Горбулін. Світова гібридна війна: український фронт: монографія. – К.: НІСД, 2017. – 496 с.
3. Постанова КМУ №518 від 19.06.2019р. «Про затвердження Загальних вимог до кіберзахисту об’єктів критичної інфраструктури» (електронний ресурс).
4. В. Л. Бурячок, В. Б. Толубко, В. О. Хорошко, С. В. Толюпа Інформаційна та кібербезпека: соціотехнічний аспект: підручник. - К.: ДУТ, 2015.- 288 с.
5. Дубов Д.В. Геополітичне суперництво у кіберпросторі як чинник впливу на національну безпеку україни. Дисертація на здобуття наукового ступеня доктора політичних наук. с. 286.
6. Дубов Д. Проблеми нормативно-правового забезпечення інформаційного суверенітету в Україні : аналіт.зап./ Д.Дубов [Електронний ресурс].– Режим доступу: http://www.niss.gov.ua/ articles/1466/
7. Дубов Д. Стратегічні аспекти кібербезпеки України / Д. Дубов // Стратегічні пріоритети.– 2013.– № 4.– С.119–126.
8. Дубов Д. Сучасні тенденції забезпечення кібербезпеки на міжнародному рівні / Д. Дубов // Стратегічні пріоритети.– 2011.– № 4.– С.5–11.
9. Постанова № 563 Кабінету Міністрів України від 23 серпня 2016 (електронний ресурс).
10. Постанова Кабінету міністрів України. Концепція системи захисту критичної інфраструктури.- 2017р. (електронний ресурс)
11. Барсегян, А. А. Технологии анализа данных : Data Mining, Visual Mining, Text Mining, OLAP / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. — СПб.: БХВ-Петербург, 2007. — 384 с.
12. Бобров, А. Системы обнаружения вторжений. Intrusion Detection System IDS. [Электронный ресурс] / А. Бобров. — Режим доступа: win/lexion/ ids/ids. html (дата обращения: 20.05.2016)
13. Боршевников, А. Е. Сетевые атаки. Виды. Способы борьбы. / А. Е. Боршевников // Современные тенденции технических наук: материалы междунар. заоч. науч. конф. / Под общ. ред. Г. Д. Ахметовой. — Уфа: Лето, 2011. — С. 8-13.
14. Вапник, В. Н. Теория распознавания образов / В. Н. Вапник, А. Я. Червоненкис. — М.: «Наука», Главная редакция физико-математической литературы, 1974. — 416 с.
15. Вапник, В. Н. The Nature of Statistical Learning Theory / В. Н. Вапник. — 2 изд. — Springer, 2000. — 314 с.
16. Васильев, В. И. Применение нейронных сетей при обнаружении атак на компьютеры в сети Internet (на примере атаки SYNFLOOD) / В. И. Васильев, А. Ф. Хафизов // Нейрокомпьютеры: разработка и применение. — 2001. — №4-5. — с. 108-114.
17. Вятченин, Д. А. Нечёткие методы автоматической классификации / Д. А. Вятченин. — Минск: Технопринт, 2004. — 219 с.
18. Галатенко, В. А. Основы информационной безопасности, В. А. Галатенко. — М.: Интернет-Университет Информационных Технологий, 2004. — с. 343
19. Дюк, В. А. Data Mining: учебный курс / В. А. Дюк, А. Самойленко. —СПб.: Питер, 2001. — 368 с.
20. Жамбю, М. Иерархический кластер-анализ и соответствия / М. Жамбю.М.: Финансы и статистика, 1988. — 345 с.
21. Зегжда, Д. П. Основы безопасности информационных систем: учеб. пособие / Д. П. Зегжда, A. M. Ивашко. — М.: Телеком, 2014. — 451 с.
22. Зегжда, П. Д. Способы защиты информации / П. Д. Зегжда. — М.: Яхтсмен, 2009. — 234 с.
23. Зегжда, П. Д. Теория и практика обеспечения информационной безопасности. / П. Д. Зегжда. — М.: Яхтсмен, 2006. — 300 с.
24. Зима, В. М. Безопасность глобальных сетевых технологий / В. М. Зима, А. А. Молдовян, Н. А. Молдовян. — 2-е изд. — СПб.: БХВ-Петербург, 2014. — с. 234.
25. Ильин, А. А. Архитектура систем интеллектуального обнаружения атак / А. А. Ильин, А. А. Сыпин // Материалы межрегиональной научно-технической конференции «Интеллектуальные и информационные системы» (Интеллект-2004) Тула: ТулГУ, 2004. — С. 40-41.
26. Ильин, А. А. Системы искусственного интеллекта в задачах обнаружения и классификации объектов / А. А. Ильин, А. А. Сыпин // Материалы межрегиональной научно-технической конференции «Интеллектуальные и информационные системы» (Интеллект-2003) — Тула: ТулГУ, 2003. — С. 8-9.
27. Касперски, К. Техника сетевых атак. Приемы противодействия. Том 1 /К. Касперски — М.: "Солон-Р", 2001. — 400 с.
28. Корт, С. С. Теоретические основы защиты информации / С. С. Корт. —

М.: Гелиос-АРВ, 2004. — 240 с.

1. Круглов, В. В. Нечёткая логика и искусственные нейронные сети / В. В.

Круглов, М. И. Дли, Р. Ю. Голунов — М.: Физматлит, 2001. — 221 с.

1. Лукацкий, А. В. Краткий толковый словарь по информационной безопасности / А. В. Лукацкий. — М., 2000. — 177 с.
2. Лукацкий, А. В. Обнаружение атак. 2-е изд., перераб. и доп. / А. В. Лукацкий. — СПб.: БХВ-Петербург, 2001. — 624 с.
3. C. Макклури Секреты хакеров. Безопасность сетей – готовые решения /

С. Макклури, Д. Скамбрэй, Г. Куртц — М.: Вильямс, 2004. — 656 с.

1. Мандель, И. Д. Кластерный анализ / И. Д. Мандель. — М.: Финансы и статистика, 1999. — 176 с.
2. Медведовский, И. Д. Атака через Internet / И. Д. Медведовский, П. В.

Семьянов, В. В. Платонов. — СПб.: НПО "Мир и семья-95", 1997. — 328 с.

1. Новак, В. Математические принципы нечёткой логики / В. Новак, И. Мочкрож; под ред. А. Н. Аверкина. — М.: Физматлит, 2006. —с. 323.
2. Осовский, С. Нейронные сети для обработки информации / С. Осовский.

М.: Финансы и статистика, 2002. — 344 с.

1. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы : Пер. с польского И. Д. Рудинского / Д. Рутковская, М. Пилиньский, Л.

Рутковский. — М.: Горячая линия – Телеком, 2006. — 452 с.

1. Рутковский, Л. Искусственные нейронные сети. Теория и практика / Л. Рутковский. — М.: Горячая линия – Телеком, 2010. — 520 с.
2. Паклин, Н. Нечеткая логика — математические основы / Н. Паклин

[Электронный ресурс]—Режим доступа: https:// basegroup. ru/ com- munity /articles /fuzzylogic-math (дата обращения: 20.05.2019)

1. Петровский, М. И. Применение методов интеллектуального анализа данных в задачах выявления компьютерных вторжений / М. И. Петровский //Труды конф. Методы и средства обработки информации. — М.: МГУ им. М.В. Ломоносова, 2005. — С. 158–167.
2. Петрунин, Ю. Ю. Информационные технологии анализа данных. Data Analysis. / Ю. Ю. Петрунин. — М.: КДУ, 2010. — 292 с.
3. Платонов, В. В. Методы интеллектуального анализа данных для построения систем обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Сб. материалов 22 конференции «Методы и технические средства обеспечения безопасности информации». — СПб.: Изд-во Политехн. Ун-та, 2013 — С. 72-74.
4. Платонов, В. В. Методы кластерного анализа для формирования модулей обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Сб. материалов 23 конференции «Методы и технические средства обеспечения безопасности информации». — СПб.: Изд-во Политехн. Ун-та, 2014 — С. 80-81.
5. Платонов, В. В. Методы сокращения размерности в системах обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Проблемы информационной безопасности. Компьютерные системы. — 2012 — №3. — С. 40-45.
6. Платонов, В. В. Обзор методов сокращения размерности в задачах обнаружения сетевых атак и вторжений / В. В. Платонов, П. О. Семенов // Сб. материалов 19 конференции «Методы и технические средства обеспечения безопасности информации». — СПб.: Изд-во Политехн. Ун-та, 2010 — С. 46-48.
7. Платонов, В. В. Построение адаптивной системы обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Сб. материалов 24 конференции «Методы

и технические средства обеспечения безопасности информации». — СПб.: Изд-во Политехн. Ун-та, 2015 — С. 95-96.

1. Платонов, В. В. Применение методов Data mining в задаче обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Проблемы информационной безопасности. Компьютерные системы. — 2013 — №4. — С. 40-44.
2. Платонов, В. В. Применение модульного подхода к построению нейронных сетей для поиска аномалий. / В. В. Платонов, Е. В. Жульков // Проблемы информационной безопасности. Компьютерные системы. СПб.: —

2006. — №3. — С. 30-34

1. Платонов, В. В. Программно-аппаратные средства обеспечения информационной безопасности вычислительных сетей: учеб. пособие для студ. высш. учеб. заведений / В. В. Платонов. — М.: Издательский центр «Академия», 2006. — 240 с.
2. Платонов, В. В. Формирование признакового пространства для обнаружения сетевых атак / В. В. Платонов, П. О. Семенов // Проблемы информационной безопасности. Компьютерные системы. — 2011 — №2. — С. 33-39.
3. Сыпин, А. А. Идентификация сетевых атак / А. А. Сыпин // Известия ТулГУ. Серия Вычислительная техника. Информационные технологии. Системыуправления. Вып. 1. Вычислительная техника. — Тула: ТулГУ, 2005. —

с. 151-155.

53. Сыпин А. А. Использование нейронных сетей для обнаружения атак /А. Сыпин // Материалы межрегиональной научно-технической конференции «Интеллектуальные и информационные системы» (Интеллект-2005). — Тула: ТулГУ, 2005. — С. 53-54.

1. Типы удалённых атак [Электронный ресурс] — Режим доступа: <http://www.securrity.ru/terms/1254-tipy-udalennyh-atak.html> (дата обращения: 20.05.2019).
2. С. Толюпа, Берестовенко Г., Єсипенко О. Системи виявлення вторгнень в інформаційні системи. Матеріали ІІ Міжнародної науково-практичної конференції “Проблеми кібербезпеки інформаційно- телекомунікаційних систем” (PCSIТS). – Київ. - 2019. – с. 167-168.
3. Толюпа С., Єсипенко О., Маляренко Д. Застосування математичної теорії ігор при побудові систем управління інформаційною безпекою. Матеріали доповідей. VІ Міжнародна науково-практична конференція «Інформаційні технології та взаємодії» (IT&I 2019). – с. 313-317.

**Додаток А**

**Результати експериментів по виявленню атак**

**Методика проведення експериментального дослідження**. В рамках дослідження були сформовані модулі виявлення для окремо взятих атак категорій User-to-Root і Remote-to-Local з тренувальних баз даних DARPA, які є найбільш складними для виявлення. Для більшості атак був отриманий результат в 100% правильно класифікованих пакетів. Для подібних атак отримані однакові набори «базових» параметрів. При об'єднанні кількох атак одного типу в класи також досягається 100% розпізнавання, при цьому збільшується кількість опорних векторів. Процес тестування складався з п'яти етапів. У першій частині тестування використовувалися багаторозрядні параметри трафіку, які добувають із заголовків IP і TCP пакетів. Всього використовувалося 14 базових параметрів, 6 для IP і 8 для TCP. Для значної частини атак було досягнуто 100% розпізнавання. На другому етапі, шляхом поділу багаторозрядних параметрів на кілька частин, кратних 8 бітам, число базових параметрів було збільшено до 24. В результаті аналогічного тестування для більшого числа атак було досягнуто 100% розпізнавання. У порівнянні з багаторозрядних параметрами збільшилася кількість опорних векторів, і велику роль став грати вибір даної матриці в методі головних компонент.

На третьому етапі тестування в набір розглянутих базових параметрів були включені статистичні параметри TCP-сеансів: час з'єднання, число переданих і прийнятих пакетів, байт і число пакетів з різними мітками - всього 49 базових параметрів. Для всіх розглянутих атак істотно збільшилася кількість опорних векторів в SVM-моделях, що викликано збільшенням розрядності простору. Для кількох атак так і не було отримано 100% результат. Для деяких атак виявилося досить від 2 до 5 нових параметрів з 49 для досягнення 100% розпізнавання і незначного збільшення числа опорних векторів. На четвертому етапі для атак, які не вдавалося виявити на попередніх етапах, була проведена кластеризація тренувальних даних і проведені процедури навчання нових модулів виявлення. В результаті майже для всіх розглянутих атак були побудовані кілька простих SVM-моделей, які дозволили класифікувати пакети зі 100% вірогідністю. Атаки, для яких не вдалося побудувати SVM-моделі, були проаналізовані та виявлено, що в складі тренувальних даних були присутні однакові пакети з різними мітками, що призводило до неможливості побудови класифікатора. На п'ятому етапі було реалізовано розширення можливостей блоків кластеризації і класифікацій шляхом внесення нечіткості. В результаті побудовані пересічні кластери, однакові пакети з різними мітками були віднесені до класу атак з певною ймовірністю. Метод опорних векторів із застосуванням нечіткої логіки підвищив показники виявлення для окремих модулів. В результаті експериментального дослідження були отримані залежності числа опорних векторів від кількості нових параметрів для ряду атак. У всіх точках представлених залежностей досягнуто 100% розпізнавання. На на рис. 4.12 показані залежності при використанні тільки параметрів IP і TCP заголовків (всього 23 параметра).

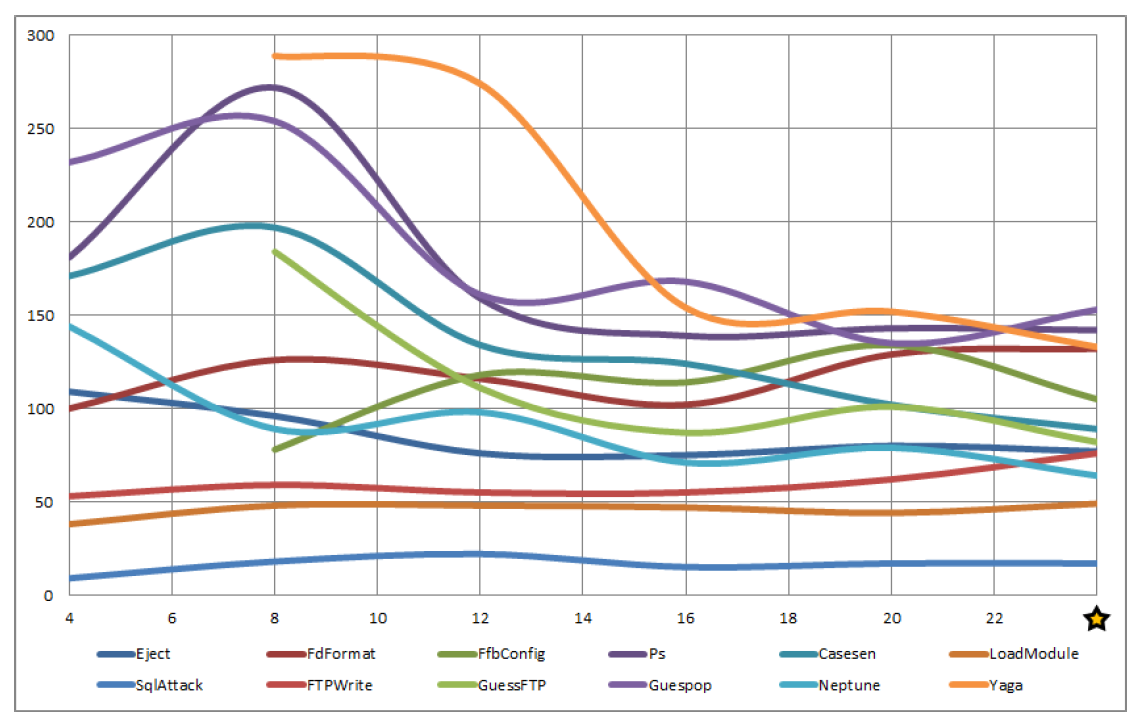
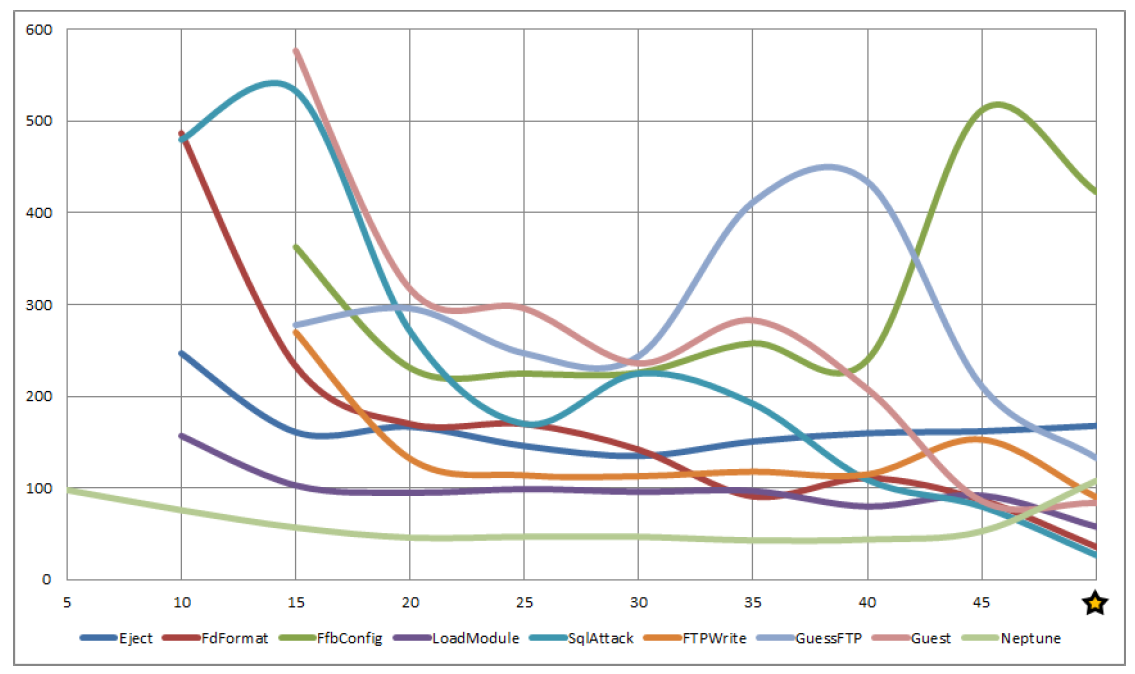


Рис. 4.12. Залежність числа опорних векторів від числа нових параметрів

На рис. 4.13 безліч параметрів доповнено параметрами TCP-сеансів (всього 49 параметрів). Зірочкою позначена робота без скорочення розмірності.



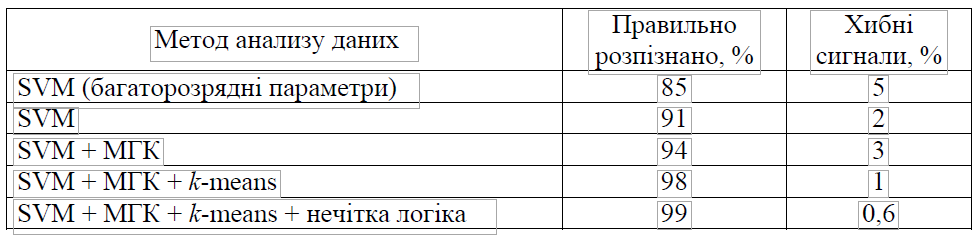
Для деяких атак, які не розпізнані програмним прототипом за допомогою одного модуля виявлення зі 100% ймовірністю, застосування декількох модулів виявлення з надмірною кількістю розглянутих «базових» і «нових» параметрів дозволяє скоротити число помилок другого роду до нуля. Проведені експерименти з окремими модулями виявлення показали хорошу працездатність системи і застосовність обраних інтелектуальних методів аналізу даних для поставленої мети. Метод опорних векторів дозволяє ідентифікувати значну частину розглянутих атак зі 100% ймовірністю, а в решті випадків помилка не перевищує декількох відсотків від числа всіх пакетів. Методи скорочення розмірності допомагають скоротити обсяг інформації, необхідної для класифікації мережевих пакетів і істотно підвищити продуктивність системи.

Проведене експериментальне дослідження підтвердило правильність запропонованої моделі та вибору безлічі методів інтелектуального розподілу даних, що лежать в її основі. Метод опорних векторів дозволив ідентифікувати більшість атак з результатом 98-100%. Метод головних компонент скоротив обсяг інформації, необхідної для класифікації мережевих пакетів, і підвищив швидкість формування модулів виявлення, але виявив проблему перенавчання. Методи кластеризації дозволили сформувати безліч модулів виявлення, виділивши типові фрагменти атак в окремі модулі виявлення і розбивши комплексні атаки на окремі модулі. Застосування нечіткої логіки підвищило результати роботи системи і дозволило класифікувати вектора, що мають різні мітки в навчальній вибірці.

На основі даного дослідження були детально опрацьовані сформульовані раніше методики по застосуванню методів інтелектуального розподілу даних по завданню виявлення мережевих атак. Результати етапів експериментального дослідження наведені в таблиці 1.

Таблиця 1.

Результати етапів експериментального дослідження



В таблиці 2 представлені результати досліджень в сфері застосування методів інтелектуального розподілу даних в задачах виявлення мережевих атак

Таблиця 2



**Додаток В**

**Програмна реалізація експерементальної частини**

#!/usr/bin/python3

# -\*- coding: utf-8 -\*-

f8 = "clearlyu 8"

f10 = "clearlyu 10"

def sizeMainWindow():

w = root.winfo\_screenwidth()

h = root.winfo\_screenheight()

w = w//2

h = h//2

w = w - 450

h = h - 325

root.geometry('900x650+{}+{}'.format(w, h))

def MainWindow():

global labelNameColumns

global boxNameColumns

global labelNameObject

global boxNameObject

global entNameObject

global ObjectOk

global ObjectPlot

global labelNameAttacks

global boxNameAttacks

global entAttack

global AttackOk

global AttackClear

global labelNameInt

global boxNameInt

global entNameInt

global IntOk

global IntPlot

global entNameInt2

global IntOk2

global IntPlot2

global labelNameFloat

global boxNameFloat

global entNameFloat

global FloatOk

global FloatPlot

global entNameFloat2

global FloatOk2

global FloatPlot2

global textAboutColumn

global ListOk

global QuitButton

########################### Первая колонка элементов #########################

labelNameColumns = Label(root, text="Атрібути KDD-99", font=f10, width=30, height=1)

labelNameColumns.place(x=5, y=5)

boxNameColumns = Listbox(root, selectmode=SINGLE, font=f8, width=30, height=42)

boxNameColumns.place(x=5, y=30)

########################## Вторая колонка элементов ##########################

labelNameObject = Label(root, text="Атрибути типу Object", font=f10, width=30, height=1)

labelNameObject.place(x=225, y=5)

boxNameObject = Listbox(root, selectmode=SINGLE, font=f8, width=30, height=4)

boxNameObject.place(x=225, y=30)

entNameObject = Entry(root, font=f8, width=30)

entNameObject.place (x=225, y=100)

ObjectOk = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok')

ObjectOk.place(x=225, y=135)

ObjectPlot = Button(root, width=8, height=1, bg="white", fg="blue", text='Plot')

ObjectPlot.place(x=349, y=135)

##################################################

labelNameAttacks = Label(root, text="Класи атак", font=f10, width=30, height=1)

labelNameAttacks.place(x=225, y=185)

boxNameAttacks = Listbox(root, selectmode=SINGLE, font=f8, width=30, height=23)

boxNameAttacks.place(x=225, y=220)

entAttack = Entry(root, font=f8, width=30)

entAttack.place (x=225, y=560)

AttackOk = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok')

AttackOk.place(x=225, y=590)

AttackClear = Button(root, width=8, height=1, bg="white", fg="blue", text='Clear')

AttackClear.place(x=349, y=590)

########################## Третья колонка элементов ##########################

labelNameInt = Label(root, text="Атрибути типу Integer", font=f10, width=30, height=1)

labelNameInt.place(x=445, y=5)

boxNameInt = Listbox(root, selectmode=SINGLE, font=f8, width=30, height=23)

boxNameInt.place(x=445, y=30)

entNameInt = Entry(root, font=f8, width=30)

entNameInt.place (x=445, y=368)

IntOk = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok')

IntOk.place(x=445, y=398)

IntPlot = Button(root, width=8, height=1, bg="white", fg="blue", text='Plot')

IntPlot.place(x=568, y=398)

entNameInt2 = Entry(root, font=f8, width=30)

entNameInt2.place (x=445, y=440)

IntOk2 = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok-2')

IntOk2.place(x=445, y=470)

IntPlot2 = Button(root, width=8, height=1, bg="white", fg="blue", text='Plot-2')

IntPlot2.place(x=568, y=470)

########################## Четвертая колонка элементов ##########################

labelNameFloat = Label(root, text="Атрибути типу Float", font=f10, width=30, height=1)

labelNameFloat.place(x=665, y=5)

boxNameFloat = Listbox(root, selectmode=SINGLE, font=f8, width=30, height=15)

boxNameFloat.place(x=665, y=30)

entNameFloat = Entry(root, font=f8, width=30)

entNameFloat.place (x=665, y=258)

FloatOk = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok')

FloatOk.place(x=665, y=288)

FloatPlot = Button(root, width=8, height=1, bg="white", fg="blue", text='Plot')

FloatPlot.place(x=789, y=288)

entNameFloat2 = Entry(root, font=f8, width=30)

entNameFloat2.place (x=665, y=323)

FloatOk2 = Button(root, width=8, height=1, bg="white", fg="blue", text='Ok-2')

FloatOk2.place(x=665, y=353)

FloatPlot2 = Button(root, width=8, height=1, bg="white", fg="blue", text='Plot-2')

FloatPlot2.place(x=789, y=353)

####### Кнопка выхода из программы

textAboutColumn = Text(root, width=30, height=11, wrap=WORD)

textAboutColumn.place(x=665, y=393)

ListOk = Button(root, width=8, height=1, bg="white", fg="blue", text='Info')

ListOk.place(x=665, y=560)

QuitButton = Button(root, width=8, height=1, bg="white", fg="red", text='Quit')

QuitButton.place(x=789, y=590)

###########################################################################

def makeNameTypeColumn():

global NameType

NameCol = pd.Series(df.columns)

listTypeOfObject = []

for tmp in NameCol:

strType = df[tmp].dtype

listTypeOfObject.append(strType)

TypeCol = pd.Series(listTypeOfObject)

NameType = pd.DataFrame([NameCol, TypeCol])

NameType = NameType.T

NameType.rename(columns={0: 'Name', 1: 'Type'},

inplace=True)

def fillBoxes():

for i in NameType['Name']:

boxNameColumns.insert(END, i)

for i in NameType[NameType.Type == 'object']['Name']:

boxNameObject.insert(END, i)

for i in NameType[NameType.Type == 'int64']['Name']:

boxNameInt.insert(END, i)

for i in NameType[NameType.Type == 'float64']['Name']:

boxNameFloat.insert(END, i)

for i in cb['attack']:

boxNameAttacks.insert(END, i)

############################################################################

def ToEntNameObject(event):

s = list(boxNameObject.curselection())

if len(s) != 1:

return

volNameObject = boxNameObject.get(s[0])

entNameObject.delete(0, END)

entNameObject.insert(0, volNameObject)

def plotterObject(event):

s = entNameObject.get()

plt.figure(figsize=(6, 6))

if s == 'service':

plt.figure(figsize=(6, 9))

m = df[s].value\_counts()

m.plot(kind='barh')

plt.title(s)

plt.title(s)

plt.xscale('log')

plt.grid(True)

plt.show()

############################################################################

def ToEntNameAttack(event):

s = list(boxNameAttacks.curselection())

if len(s) != 1:

return

volNameAttack = boxNameAttacks.get(s[0])

entAttack.delete(0, END)

entAttack.insert(0, volNameAttack)

def cleanerAttack(event):

entAttack.delete(0, END)

############################################################################

def ToEntNameInt(event):

s = list(boxNameInt.curselection())

if len(s) != 1:

return

volNameInt = boxNameInt.get(s[0])

entNameInt.delete(0, END)

entNameInt.insert(0, volNameInt)

def ToEntNameInt2(event):

s = list(boxNameInt.curselection())

if len(s) != 1:

return

volNameInt = boxNameInt.get(s[0])

entNameInt2.delete(0, END)

entNameInt2.insert(0, volNameInt)

def plotterInt(event):

s = entNameInt.get()

if s == '':

return

plt.figure('Графік зміни значень параметру', figsize=(6, 6), frameon=False)

plt.grid(True)

plt.xlabel('Ordinal number KDD-99')

plt.ylabel(s)

try:

a = entAttack.get()

except:

pass

if a == '':

df[s].plot()

plt.title(s)

else:

df[(df.attack==a)][s].plot()

plt.title(a +': '+ s)

plt.show()

def scatterIntInt(event):

i = entNameInt.get()

if i == '':

return

j = entNameInt2.get()

if j == '':

return

plt.figure('Результати кластерного аналізу', figsize=(3, 3), frameon=False)

plt.xlabel(i)

plt.ylabel(j)

plt.grid(True)

for s in cb['attack']:

x = np.mean(df[(df.attack==s)][i])

y = np.mean(df[(df.attack==s)][j])

c = cb[(cb.attack==s)]['color']

plt.scatter(x, y, color=c)

plt.show()

##############################################################################

def ToEntNameFloat(event):

s = list(boxNameFloat.curselection())

if len(s) != 1:

return

volNameFloat = boxNameFloat.get(s[0])

entNameFloat.delete(0, END)

entNameFloat.insert(0, volNameFloat)

def ToEntNameFloat2(event):

s = list(boxNameFloat.curselection())

if len(s) != 1:

return

volNameFloat = boxNameFloat.get(s[0])

entNameFloat2.delete(0, END)

entNameFloat2.insert(0, volNameFloat)

def plotterFloat(event):

s = entNameFloat.get()

if s == '':

return

plt.figure('Графік зміни значень параметру', figsize=(6, 6), frameon=False)

plt.grid(True)

plt.xlabel('Ordinal number KDD-99')

plt.ylabel(s)

try:

a = entAttack.get()

except:

pass

if a == '':

df[s].plot()

plt.title(s)

else:

df[(df.attack==a)][s].plot()

plt.title(a +': '+ s)

plt.show()

def scatterFloatFloat(event):

i = entNameFloat.get()

if i == '':

return

j = entNameFloat2.get()

if j == '':

return

plt.figure('Результати кластерного аналізу', figsize=(3, 3), frameon=False)

plt.xlabel(i)

plt.ylabel(j)

plt.grid(True)

for s in cb['attack']:

x = np.mean(df[(df.attack==s)][i])

y = np.mean(df[(df.attack==s)][j])

c = cb[(cb.attack==s)]['color']

plt.scatter(x, y, color=c)

plt.show()

#############################################################################

def ToTextAboutColumn(event):

s = list(boxNameColumns.curselection())

if len(s) != 1:

return

NameColumn = boxNameColumns.get(s[0])

seriesColumn = df[NameColumn]

textAboutColumn.delete(1.0, END)

textAboutColumn.insert(1.0, seriesColumn.describe())

def quiter(event):

root.destroy()

#############################################################################

from tkinter import \*

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import csv

pd.set\_option('display.max\_columns', 8)

pd.set\_option('display.max\_rows', 10)

pd.set\_option('display.width', 80)

df = pd.read\_csv('./kdd99-10.csv', index\_col=False)

cb = pd.read\_csv('./attackcolor.csv', index\_col=False)

root = Tk()

root.title('Інтерфейс датафрейма KDD-99. 494 021 записів')

sizeMainWindow()

MainWindow()

makeNameTypeColumn()

fillBoxes()

plt.rcParams['toolbar'] = 'None'

AttackOk.bind('<Button-1>', ToEntNameAttack)

AttackClear.bind('<Button-1>', cleanerAttack)

ObjectOk.bind('<Button-1>', ToEntNameObject)

ObjectPlot.bind('<Button-1>', plotterObject)

IntOk.bind('<Button-1>', ToEntNameInt)

IntPlot.bind('<Button-1>', plotterInt)

IntOk2.bind('<Button-1>', ToEntNameInt2)

IntPlot2.bind('<Button-1>', scatterIntInt)

FloatOk.bind('<Button-1>', ToEntNameFloat)

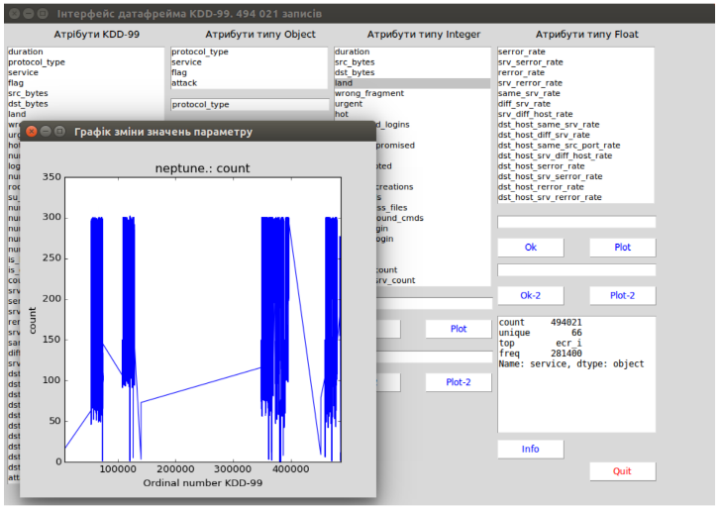
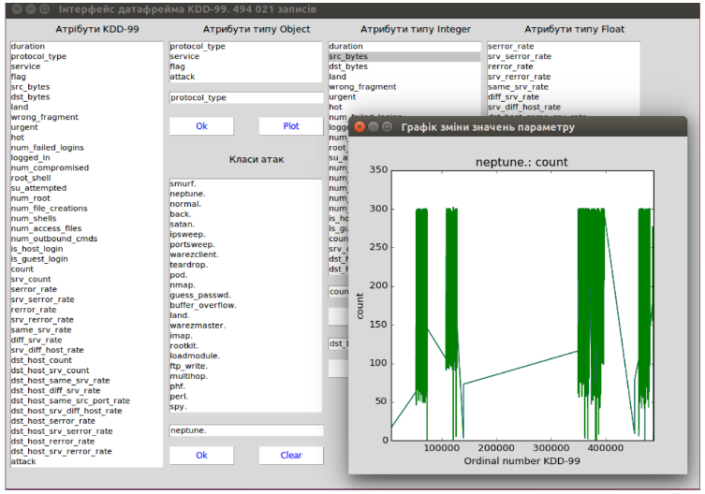
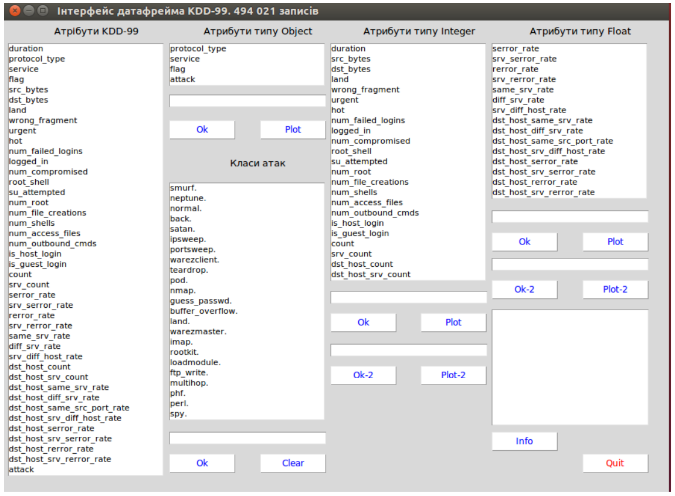
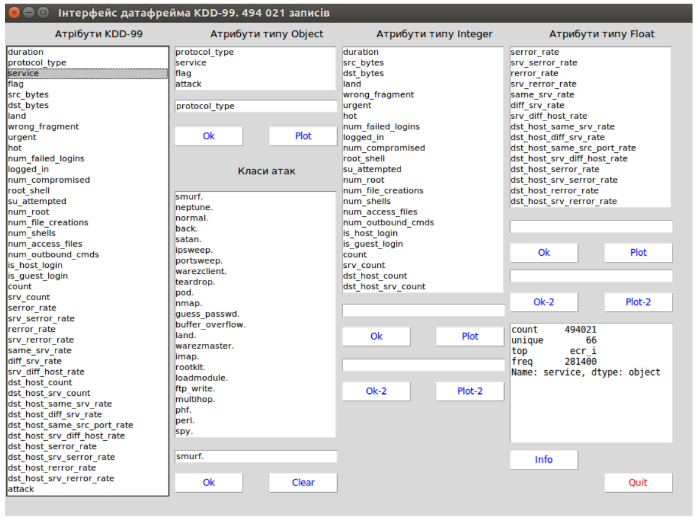
FloatPlot.bind('<Button-1>', plotterFloat)

FloatOk2.bind('<Button-1>', ToEntNameFloat2)

FloatPlot2.bind('<Button-1>', scatterFloatFloat)

ListOk.bind('<Button-1>', ToTextAboutColumn)

QuitButton.bind("<Button-1>", quiter)

root.mainloop()

