ДОПОВІДІ
НАЦІОНАЛЬНОЇ
АКАДЕМІЇ НАУК
УКРАЇНИ

ІНФОРМАТИКА ТА КІБЕРНЕТИКА
INFORMATICS AND CYBERNETICS

https://doi.org/10.15407/dopovidi2021.01.009
УДК 004.047

**А.Б. Качинський[1], М.С. Стремецька[2]**

[1] Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ
[2] Фізико-технічний інститут НТУ України "Київський політехнічний інститут ім. Ігоря Сікорського"
E-mail: akachynsky@gmail.com, mira.stremetska@gmail.com

# Операційна аналітика як інструмент моніторингу даних та управління подіями систем забезпечення кібербезпеки

*Представлено членом-кореспондентом НАН України О. М. Трофимчуком*

*В умовах зростання попиту на цифровізацію процесів збору, передачі, обробки і зберігання даних у всіх сферах життєдіяльності особи, суспільства та держави виникла гостра необхідність в побудові інфраструктури мереж передачі інформації, що можуть забезпечити захищене з'єднання між центрами обробки даних та кінцевими користувачами. Ці мережі повинні мати високу відмовостійкість, забезпечувати швидку та ефективну обробку інформаційних запитів, особливо якщо це стосується критичної інфраструктури. В статті запропонована оригінальна структурно-функціональна схема управління даними для SIEM-систем, що враховує прямі та зворотні зв'язки фізичного, математичного й аналітичного рівнів, заснована на теорії страт М. Месаровича. Побудовано модель багаторівневої системи моніторингу даних та управління подіями безпеки, що дозволяє реалізувати системний підхід до підтримки стану захищеності складних систем, а також забезпечення механізмів для оперативного реагування на потенційні інциденти кібербезпеки в режимі реального часу.*

***Ключові слова:*** *Security Information and Event Management (SIEM), Threat Intelligence Platform (TIP), управління подіями кібербезпеки, моніторинг даних, теорія страт.*

Безпека складних багаторівневих ієрархічних систем, у тому числі таких як національна система оплати комунальних послуг [1], визначається як здатність систем мінімізувати ризики для основних об'єктів захисту не тільки в умовах виникнення та існування загроз, але й у разі їх реалізації, зберігаючи при цьому свою структуру незмінною [2, ст. 489]. Ключовою умовою функціонування таких мереж є постійна та всеохоплююча підтримка стану захищеності, а також забезпечення механізмів для оперативного реагування на потенційні інциденти кібербезпеки в режимі реального часу.

Цитування: Качинський А.Б., Стремецька М.С. Операційна аналітика як інструмент моніторингу даних та управління подіями систем забезпечення кібербезпеки. *Допов. Нац. акад. наук Укр.* 2021. № 1. С. 9—16. https://doi.org/10.15407/dopovidi2021.01.009

Першочерговою виступає задача моніторингу інформації, що стосується подій безпеки комп'ютерних мереж та систем, зібраної за допомогою різноманітних сенсорів. Всі вони генерують різні види даних, а також створені фахівцями для різних цілей. Тому кожен лише частково відображає стан захищеності об'єктів мережі.

Відомості щодо подій безпеки розподіляються по величезній кількості мережевого трафіку та системних журналів, тому ефективний аналіз стану захищеності мережі потребує швидкої обробки великих об'ємів даних.

Переважна кількість мережевого трафіку не містить ніякої потенційної інформації про аномальну поведінку, що регулярно повторюється, на тлі чого реально корисну інформацію має вкрай незначна її кількість. Таким чином, збір необхідного обсягу інформації для відтворення рідкісних подій безпеки є критичним технічним зусиллям.

Розробка адекватної структури обробки даних мережевого трафіку вимагає розуміння того, яким чином різні сенсори здійснюють збір даних, як вони доповнюють, дублюють і взаємодіють один з одним, а також розуміння принципів ефективного зберігання даних, що забезпечують їх ефективний аналіз.

Наведені вище факти вказують на те, що система моніторингу даних та управління подіями безпеки (СМДУПБ) складається з сукупності взаємодіючих ієрархічно впорядкованих елементів з наданим правом ухвалення рішень, що описується з позиції теорії систем [3, 4].

Метою цієї роботи є визначення структурно-функціональних особливостей формалізації моделі багаторівневого управління кібербезпекою складних систем з ієрархічною структурою на основі теоретико-множинного підходу.

Для моделі обробки даних мережевого трафіку такою багаторівневою структурою може бути страта. Стратифікована система або стратифікований опис — це сімейство моделей, кожна з яких описує поведінку системи з погляду різних рівнів абстрагування. На кожній страті в ієрархії структур існує свій власний набір змінних, що дозволяють значною мірою обмежити дослідження лише однією стратою [5, с. 56, 103—106].

Узагальнена структурна схема системи моніторингу даних та управління подіями безпеки є ієрархічною структурою, що пропонується для подальшого розгляду, зображена на рис. 1, тут кожна зі страт має свої характерні особливості, свій набір змінних, законів і принципів. Тому головним завданням є розробка низки математичних моделей, кожна з яких описувала би поведінку СМДУПБ з погляду різних рівнів абстрагування: фізичного, математичного й аналітичного.

Відправним пунктом для стратифікованого опису технологічного стеку системи моніторингу даних та управління подіями безпеки $S : X \to Y$ є припущення про те, що множина зовнішніх стимулів $X$ і множина відгуків $Y$ представляються у вигляді декартових добутків, а саме вважаються заданими два сімейства множин: $X^i$, $1 \leqslant i \leqslant 3$, $Y^i$, $1 \leqslant i \leqslant 3$, $i \in \mathbb{N}$, таких що

$$X = X^1 \times ... \times X^3 \text{ і } Y = Y^1 \times ... \times Y^3 \qquad (1)$$

Це припущення означає можливість розбиття відгуків і вхідних стимулів на компоненти. Якщо множини $X$ і $Y$ можуть бути представлені у вигляді (1), то кожна пара $(X^i, Y^i)$, $1 \leqslant i \leqslant 3$, $i \in \mathbb{N}$, приписується певній страті.

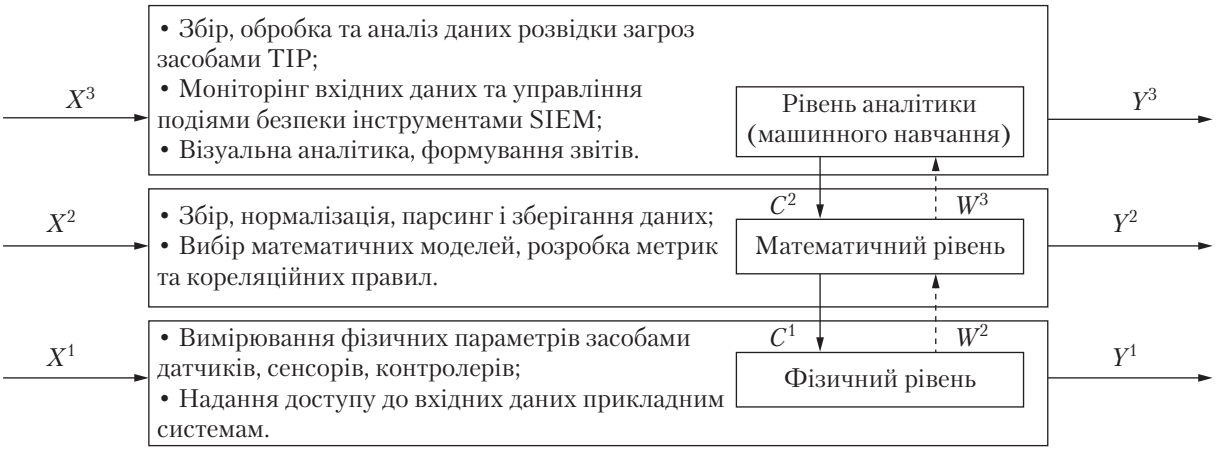


***Рис. 1.*** Ієрархія страт системи моніторингу даних та управління подіями безпеки

Формалізуючи задану модель, визначаємо:

$X^i = [X_0^i, X_1^i, ..., X_{n_i}^i]$ — вектор вхідних стимулів підсистеми $i$-го ієрархічного рівня;

$Y^i = [Y_0^i, Y_1^i, ..., Y_{n_i}^i]$ — вектор вихідних відгуків підсистеми $i$-го ієрархічного рівня.

Побудуємо за правилами теоретико-множинного підходу розбиття множини $\hat{S}$ структурно-функціональних елементів, блоків, зв'язків та змінних системи $S$ на упорядковану сукупність $\overline{S} = \langle S^1, S^2, S^3 \rangle$.

Процедура формалізованого розбиття множини $\hat{S}$ повинна задовольняти наступним трьом умовам:

$$\forall S^i \in \overline{S} : S^i \subseteq \hat{S}; \tag{2}$$

$$\forall S^i, S^j \in \overline{S} : S^i \neq S^j \to S^i \cap S^j = \varnothing; \tag{3}$$

$$\bigcup_{S^i \in \overline{S}} S^i = \hat{S}. \tag{4}$$

Введемо додаткову умову (5), що задає порядок взаємодії стимулів і відгуків стратифікованої за умовами (2)—(4) багаторівневої системи $S$ типу "вхід—вихід":

$$\forall (X^i, Y^i) \in S^i. \tag{5}$$

Таким чином $i$-а страта системи $S$ — це підсистема, що представлена як набір відображень $S^i$, $1 \leqslant i \leqslant 3$, $i \in \mathbb{N}$:

1) $S^3 : X^3 \times W^3 \to Y^3$;
2) $S^2 : X^2 \times C^2 \times W^2 \to Y^2$;
3) $S^1 : X^1 \times C^1 \to Y^1$.

Сімейство визначених таким чином підсистем $S^i$ називається стратифікацією системи $S$, якщо існує два сімейства відображень

$$h^i : Y^i \to W^{(i+1)},\ 1 \leqslant i < 3 \quad \text{та } c^i : Y^i \to C^{i-1},\ 1 < i \leqslant 3,\ i \in \mathbb{N},$$

такі, що для кожного елемента $x$ з $X$ і $y = S(x)$:

1) $y^3 = S^3(x^3, h^2(y^2))$;

2) $y^2 = S^2(x^2, c^3(y^3), h^1(y^1))$;

3) $y^1 = S^1(x^1, c^2(y^2))$.

Множина $Y^i$ складається з відгуків $i$-ої страти: $C^i$ і $W^i$ представляють собою множину стимулів, що надходять зі страт, дотичних до $i$-ої страти згори і знизу відповідно. Відображення $h^i$ і $c^i$ називаються інформаційною функцією і розподільною функцією $i$-ої страти відповідно. Вони зв'язують страти разом, утворюючи систему $S$.

**1. Страти. Рівні опису та абстрагування.** Стратифікуємо розглянуту вище багаторівневу ієрархічну систему з точки зору організації системи моніторингу даних та аналізу подій безпеки (див. рис. 1).

Перші два рівні відповідають за збір, нормалізацію, парсинг та зберігання даних з мереж, систем безпеки, хостів та надання доступу до цих даних визначеним інструментам моніторингу та звітності.

Фізичний рівень включає моніторинг мережевої інфраструктури, а саме моніторинг:

— зовнішніх факторів (температура, електроживлення і т. д.);

— портів (поточний стан, доступність);

— стану процесорів;

— пам'яті;

— інших параметрів в залежності від особливостей мережевого обладнання.

Математичний рівень включає моніторинг прикладних систем, телеметрії комп'ютерних мереж та систем безпеки, серед яких

— міжмережеві екрани;

— комутатори, маршрутизатори, безпровідні точки доступу;

— VPN шлюзи;

— системи IPS/IDS;

— системи управління IAM/IAD/PAM;

— системи захисту від DDOS;

— системи аналізу мережевого трафіку;

— системи захисту від шкідливого ПЗ;

— веб-проксі/шлюзи;

— поштові проксі/шлюзи;

— системні журнали (логи) серверів, операційних систем та прикладних сервісів;

— Active Directory;

— пісочниці (Sandboxes) та ін.

Третій рівень відповідає за аналітику даних, сформованих на основі двох попередніх рівнів та включає наступні компоненти вхідних даних.

Дані аналітики безпеки:

— дані розвідки загроз (Threat Intelligence, TI);

— дані розвідки вразливостей (Vulnerability Intelligence, VI).

Індикатори компрометації (Indicators of Compromise, IOCs):

сигнатури вірусів;

IP-адреси;

MD-5 хеші файлів шкідливого ПЗ;

URL-адреси або доменні імена серверів контролю та управління ботнет—мережами.

Стандарти і політики безпеки.

**2. Стратифікований опис системи.** Базовим терміном для пояснення ролі фізичної страти виступає датчик як елемент контролюючого пристрою системи, що виконує безпосереднє вимірювання значень параметрів системи та інтерпретує отриману інформацію в зручний для передачі сигнал.

Відомо, що датчик (сенсор, від англ. sensor) — це поняття в системах управління, первинний перетворювач, елемент вимірювального, сигнального, регулюючого або управляючого пристрою системи, що перетворює контрольовану величину в зручний для використання сигнал. Датчики перетворюють контрольовану величину (температуру, напругу, частоту процесора тощо в сигнал (електричний, оптичний тощо), зручний для вимірювання, передачі, перетворення, зберігання і реєстрації інформації про стан об'єкта вимірювань [6].

Услід за Коллінсом [7], під категорією "сенсор" ми розуміємо будь-що: тобто від дзеркалювання трафіку до системного журналу міжмережевого екрану, — того, що фіксує дані про фільтрацію мережевого трафіку та може бути використаним для оцінки як інформаційної, так і кібернетичної безпеки.

Мережа таких датчиків: як апаратних (на фізичному рівні), так і програмних (на математичному рівні), дозволяє сформувати вектори вхідних даних $x^1$, $x^2$ множини зовнішніх стимулів $X$ для фізичної $S^1$ та математичної $S^2$ страти системи $S$ відповідно. Нехай $x^1 = [x_0^1,\ x_1^1,\ x_2^1, \ldots,\ x_{n_1}^1]$ — вектор вхідних параметрів підсистеми $S^1$, $x^2 = [x_0^2,\ x_1^2,\ x_2^2, \ldots,\ x_{n_2}^2]$ — вектор вхідних параметрів підсистеми $S^2$, тоді $y^1 = [y_0^1,\ y_1^1,\ y_2^1, \ldots,\ y_{n_1}^1]$ — вектор вихідних параметрів підсистеми $S^1$, а $y^2 = [y_0^2,\ y_1^2,\ y_2^2, \ldots,\ y_{n_2}^2]$ — вектор вихідних параметрів підсистеми $S^2$.

**Фізичний рівень.** Розглянемо архітектуру фізичного рівня ієрархії (рис. 2).

На фізичному рівні ієрархії мережа апаратних датчиків виконує вимірювання фізичних параметрів системи $x^1 = [x_0^1,\ x_1^1,\ x_2^1, \ldots,\ x_{n_1}^1]$. Розглянемо детальніше вектор вхідних параметрів першої страти:

$x_0^1$ — стан процесора (завантаженість, температура, споживання електроенергії тощо);

$x_1^1$ — стан пам'яті (заповненість, швидкість зчитування/запису тощо);

$x_2^1$ — поточній стан портів, їх доступність (0 або 1);

$x_3^1$ — стан іншого мережевого обладнання.

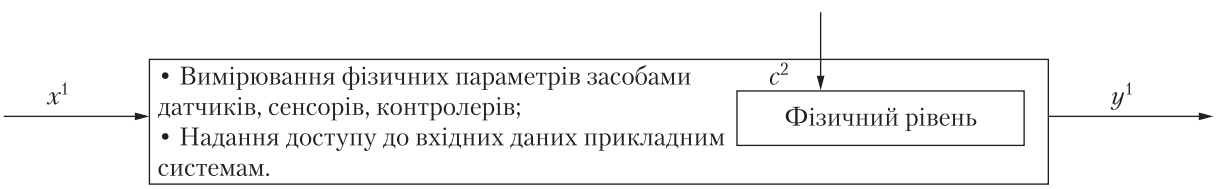


***Рис. 2.*** Стратифікований опис фізичного рівня ієрархії

На основі цих даних з урахуванням керуючого впливу розподільної функції другого рівня ієрархії $c^2(y^2)$ отримуємо вихідні значення $y^1$:

$$y^1 = S^1(x^1,\ c^2(y^2)).$$

Інформаційна функція $h^1(y^1)$ передає значення вихідних параметрів підсистеми $S^1$, впливаючи на значення вихідних параметрів підсистеми $S^2$.

**Математичний рівень.** Розглянемо архітектуру математичного рівня ієрархії (рис. 3).

На математичному рівні ієрархії мережа прикладних сенсорів виконує вимірювання параметрів прикладних системи $x^2 = [x_0^2,\ x_1^2,\ x_2^2, \ldots,\ x_{n_2}^2]$.

Загалом може існувати багато різних джерел даних для цього рівня, кожен із яких параметризується заздалегідь визначеним чином (кожному потоку відповідає хост, джерело та тип джерела як, наприклад, для Splunk ES), а також індексується місце, куди надсилаються дані з кожного такого джерела (наприклад, індексер для Splunk ES). Всі ці значення допомагають визначити походження тих чи інших даних при потраплянні до SIEM-системи, проте необхідно виділити серед них такі:

$x_0^2$ — дані щодо подій безпеки та системні журнали (для прикладу, можуть передаватись у вигляді Syslog трафіку);

$x_1^2$ — телеметрія мережевих пристроїв (типово передається за стандартом NetFlow);

$x_2^2$ — захоплення мережевих пакетів (PCAP файли).

На основі цих даних з урахуванням керуючого впливу розподільної функції третього рівня ієрархії $c^3(y^3)$ та інформаційної функції першого рівня ієрархії $h^1(y^1)$ отримуємо вихідні значення $y^2$:

$$y^2 = S^2(x^2,\ c^3(y^3),\ h^1(y^1)).$$

Інформаційна функція $h^2(y^2)$ передає значення вихідних параметрів підсистеми $S^2$, впливаючи на значення вихідних параметрів підсистеми $S^3$.

**Рівень аналітики.** Розглянемо архітектуру рівня аналітики (рис. 4).

На рівні аналітики виконується збір, обробка та аналіз даних розвідки загроз засобами платформи розвідки загроз (TIP), а також проводиться моніторинг та аналітична обробка вхідних даних фізичного та математичного рівня з залученням даних розвідки вразливостей (SIEM), індикаторів компрометації та політик безпеки $x^3 = [x_0^3,\ x_1^3,\ x_2^3, \ldots,\ x_{n_3}^3]$, що постійно оновлюються.

Важливо, щоб усі джерела TI, VI, IOCs були достовірними, актуальними і належним чином підтримували процес автоматичного прийняття рішень щодо стану безпеки засобами

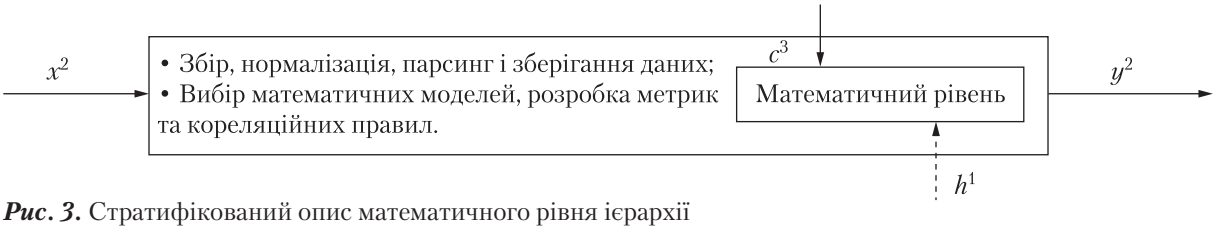


***Рис. 3.*** Стратифікований опис математичного рівня ієрархії

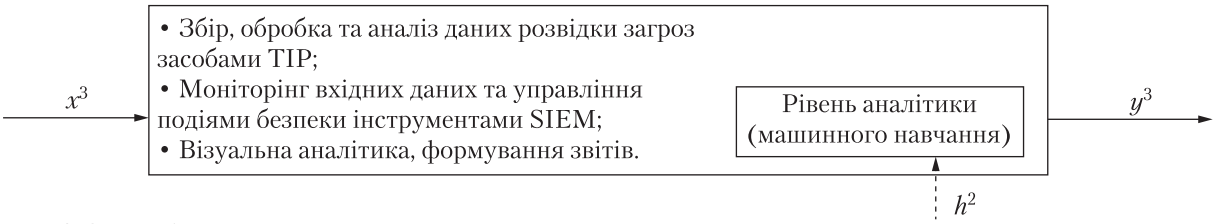


***Рис.4.*** Стратифікований опис рівня аналітики

SIEM на основі заздалегідь визначених метрик та кореляційних правил. Саме тому регулярне надходження та оновлення наступних даних є критично важливими для ефективної роботи системи моніторингу даних та управління подіями безпеки:

$x_0^2$ — дані аналітики безпеки (TI, VI);

$x_1^2$ — індикатори компрометації (IOCs);

$x_2^2$ — стандарти та політики безпеки.

На основі цих даних з урахуванням керуючого впливу інформаційної функції другого рівня ієрархії $h^2(y^2)$ отримуємо вихідні значення $y^3$:

$$y^3 = S^3(x^3,\ h^2(y^2)).$$

Розподільна функція $c^3(y^3)$ передає значення вихідних параметрів підсистеми $S^3$, впливаючи на значення вихідних параметрів підсистеми $S^2$.

**Висновки.** Теоретико-множинний підхід для визначення структурно-функціональних особливостей побудови моделі багаторівневої системи моніторингу даних та управління подіями безпеки є продуктивним як з погляду вибору оптимальних стратегій збору даних, так і узагальненої стратифікації організаційних і технологічних рівнів управління системами забезпечення безпеки.

Модель обробки мережевого трафіку удосконалюється за допомогою розбиття множини структурно-функціональних елементів, блоків, підсистем і зв'язків між ними на упорядковану сукупність підмножин, стратифікованих за фізичними, математичними та аналітичними ознаками, що дозволяє встановити порядок взаємодії впливів і відгуків у багаторівневих системах моніторингу даних та управління подіями безпеки.

Структурно-функціональна схема управління даними для SIEM-систем, що враховує прямі і зворотні зв'язки фізичного, математичного й аналітичного рівнів наразі є найбільш формалізованим і універсальним підходом математичного моделювання систем управління подібного типу.

Перспективи подальшого розвитку і застосування запропонованої багаторівневої моделі системи моніторингу даних та управління подіями безпеки полягають у можливостях прозорої формалізації цільових функцій та здійснення структурно-параметричної оптимізації систем забезпечення кібербезпеки.

ЦИТОВАНА ЛІТЕРАТУРА

1. Стремецька М. Оцінка пріоритетів механізмів кіберзахисту національної системи оплати комунальних послуг за допомогою методу аналізу ієрархій. *Захист інформації*. 2019. **21**, № 2. С. 129—137. http://doi.org/10.18372/2410-7840.21.13704
2. Качинський А. Б. Безпека складних систем. Київ: Юстон, 2017. 498 с.
3. Волкова В. Н., Денисов А. А. Теория систем. Учеб. пособие. Москва: Высшая школа, 2006. 511 с.
4. Капур К., Ламберсон Л. Надежность и проектирование систем. Москва: Мир, 1980. 608 с.
5. Месарович М., Мако Д., Такахара И. Теория иерархических многоуровневых систем. Москва: Мир, 1973. 344 с.
6. Датчики: справочное пособие. Шарапов В.М., Полищук Е.С. (ред.). Москва: Техносфера, 2012. 624 с.
7. Collins M. Network security through data analysis. O'Reilly Media, Inc., 2014. 327 p.

Надійшло до редакції 14.12.2020

REFERENCES

1. Stremetska, M. (2019). Priorities Evaluation of Cyber Defense Mechanisms of National Utilities Payment System Through the Use of the Analytic Hierarchy Process. Ukr. Inform. Security Research Journal, 21, No. 2, pp. 129-137 (in Ukrainian). http://doi.org/10.18372/2410-7840.21.13704
2. Kachynskyi, A. B. (2017). Bezpeka skladnykh system. Kyiv: Yuston. [in Ukrainian]
3. Volkova, V. N., & Denisov, A. A. (2006). Teoriya sistem. Moscow: Vysshaya shkola. (in Russian).
4. Kapur, K., & Lamberson, L. (1980). Nadezhnost' i proektirovanie sistem. Moscow: Mir. (in Russian).
5. Messarovich, M., Mako, D., & Takahara, I. (1973). Theory of hierarchical multilevel systems. Moscow: Mir. (in Russian).
6. Sharapov, V. M., Polishchuk, E. S., Koshevoy, N. D., Ishanin, G. G., Minaev, I. G., & Sovlukov, A. S. (2012). Datchiki: spravochnoe posobie. Moscow: Tekhnosfera. (in Russian).
7. Collins, M. (2014). Network security through data analysis. CA, USA: O'Reilly Media, Inc.

Received 14.12.2020

*A.B. Kachynskyi*[2], *M.S. Stremetska*[1]

[1] Institute of Telecommunications and Global Information Space of the NAS of Ukraine, Kyiv
[2] Institute of Physics and Technology NTU of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"
E-mail: akachynsky@gmail.com, mira.stremetska@gmail.com

OPERATIONAL ANALYTICS AS A DATA MONITORING AND EVENT MANAGEMENT TOOL OF THE CYBER SECURITY MANAGEMENT SYSTEMS

With growing demand for the digitalization of data collection, transmission, processing and storage processes in all life spheres of individual, society, and state, there is an urgent need to construct an infrastructure of information transmission networks which can provide a secure connection between endpoints and data centers. These networks must have high availability and provide the fast and efficient processing of information requests, especially in case of critical infrastructure networks. A structural functional scheme of data management for SIEM systems which includes straight and reverse relations between physical, mathematical and analytical levels is proposed, based on the stratum theory by M. Messarovich. A model of multilevel system for the data monitoring and cyber security event management is built in order to provide a systematic approach to maintain the safety state of complex systems and to ensure mechanisms for the operative real-time cyber security incident response.

***Keywords:*** *Security Information and Event Management (SIEM), Threat Intelligence Platform (TIP), cyber security event management, data monitoring, stratum theory.*