УДК 004.056.53 DOI 10.31548/energiya2022.01.072

ВИЯВЛЕННЯ АНОМАЛІЙ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ НА ОСНОВІ АНАЛІЗУ ЕНТРОПІЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

М. В. Панченко, студент

А. М. Бігдан, асистент

Т. В. Бабенко, доктор технічних наук, професор

Київський національний університет імені Тараса Шевченка

E-mail: maxim.panchencko012@gmail.com

Д. С. Тимофєєв, старший викладач

Національний технічний університет «Дніпровська Політехніка»

E-mail: tymofieiev.d.s@nmu.one

Анотація. *Заходи захисту від кібератак не можуть надавати стовідсоткову гарантію неможливості проникнення зловмисника в інформаційну систему. Якщо зловмисник отримав доступ до системи, то потрібно якомога швидше виявити такі дії та перервати доступ, а також провести розслідування для того, щоб виправити прогалини в безпеці. Для виявлення атак використовуються методи, які поділяються на ті, що виявляють зловживання та ті, що виявляють аномалії. У цій роботі досліджується можливість використання частотного методу, що виявляє аномалії в роботі системи шляхом аналізу ентропії журналу подій. Зазвичай цей метод використовується для виявлення аномалій у мережевому трафіку, але на наявність несанкціонованих дій також можуть вказувати аномалії в журналі подій на хостах. Виконані дослідження на прикладі журналу подій операційної системи Windows показали, що шляхом аналізу ентропії можна виявити перевищення порогів безпеки щодо кількості різних повідомлень в журналі подій. Це може вказувати на аномалії в роботі інформаційної системи. Запропонований у роботі метод може бути інтегрований у системи виявлення вторгнень, що сповіщатимуть адміністратора безпеки про можливі зловживання.*

Ключові слова: ***ентропія, аномалії, журнал подій, інформаційна безпека***

Актуальність. Враховуюче те, що збитки від кібератак постійно зростають, зокрема, збитки від кіберзлочинності у світі за 2020 рік склали 945 мільярдів доларів [1], засоби захисту від кібератак не можуть протидіяти існуючим загрозам. Тому актуальними є дослідження з синтезу моделей, що дозволяють вчасно виявити

аномалії в роботі процесів, які в майбутньому можуть призвести до реалізації певного вектору атаки і компрометації інформаційних систем.

Аналіз останніх досліджень та публікацій. Нині існує багато досліджень щодо методів виявлення аномалій в інформаційних системах [2-5]. У розглянутих дослідженнях [2,3] зазначено, що достатній рівень захищеності може бути досягнутий за рахунок використання комбінованого підходу, тобто застосування одночасно декількох методів виявлення вторгнень разом із засобами захисту. Розглянутий у цій роботі метод зазвичай використовується для виявлення аномалій у мережевому трафіку [6].

Мета дослідження – проаналізувати можливість виявлення аномалій та інцидентів інформаційної безпеки (ІБ) в ІС на основі аналізу ентропії журналу подій процесів/системи.

Матеріали та методи дослідження. Для виявлення атак використовуються методи, що поділяються на такі групи [2]: ті, що базуються на зберіганні прикладів поведінки; частотні; нейромережеві; ті, що будують скінченні автомати; інші спеціальні. Ці методи також поділяють на ті, що виявляють зловживання та ті, що виявляють аномалії. Зловживання засновані на використанні існуючих недоліків інформаційної системи, а аномалія – це відхилення від нормального стану системи, незвичайна активність в ній, що може свідчити про певні атакуючі дії [3]. Перевага методів, що виявляють аномалії – можливість виявлення нових атак без модифікації або поновлення параметрів моделі в тому випадку, коли поведінка системи в процесі атаки є статистично відмінною від нормальної поведінки системи [4]. Використовуючи одночасно декілька методів, тобто комплексний підхід щодо виявлення інцидентів інформаційної безпеки, можна підвищити ефективність протидії вторгненням [2].

До методів, що виявляють аномалії, зокрема, належать частотні (статистичні) методи. Основними статистичними методами, що використовуються в інформаційній безпеці є [5]: аналіз виживання; аналіз часового ряду; дерева рішень; кластерний аналіз; аналіз ентропії. Серед цих методів одним із найбільш ефективних є аналіз ентропії [5], що зазвичай використовується для виявлення аномалій в

мережевому трафіку [6]. Окрім аномалій у мережі на наявності несанкціонованих дій також можуть вказувати аномалії в журналі подій на хостах. На аномалії в роботі системи може вказувати кількість різних записів, яку можна аналізувати за допомогою обчислення ентропії.

Результати досліджень та їх обговорення. Ентропію джерела повідомлень розраховували за формулою Шеннона [7, с. 32]:

$$H(A) = -\sum_{i=1}^{k} p_i \log_2 p_i, \tag{1}$$

де A – множина повідомлень, k – кількість повідомлень, p_i – імовірність появи кожного повідомлення.

Для розрахунку ймовірності появи повідомлення використовували формулу:

$$p_i = \frac{n_i}{N}, \tag{2}$$

де n_i – кількість повідомлень певного типу, N – загальна кількість повідомлень.

Для розрахунку ентропії журналу подій використовували метод рухомого вікна [8]. Ентропія розраховувалась для повідомлень у межах обраного вікна розміром W. Для перших W повідомлень ентропія розраховується за формулою (1), далі із появою нових повідомлень вікно зсувається на величину dW.

На рис. 1 показано схему методу рухомого вікна зі значеннями $W = 5$ та $dW = 1$, різні повідомлення позначені цифрами.

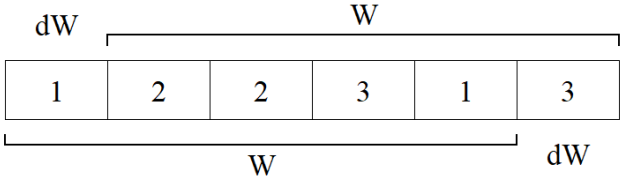

Рис. 1. Схема методу рухомого вікна

Після зсуву вікна розраховували поточне значення ентропії:

$$H_i = H_{i-1} + \Delta H.$$

Зміна ентропії розраховувалася за формулою:

$$\Delta H = H_i - H_{i-1}. \tag{3}$$

Підставивши формулу (1) у формулу (2), отримали вираз:

$$\Delta H = -\sum_{i=1}^{k} p_i \log_2 p_i + \sum_{j=1}^{l} p'_j \log_2 p'_j, \quad (4)$$

де літери без штриха – ймовірності значення повідомлень для поточного вікна, зі штрихом – для попереднього вікна.

Ті елементи, імовірність яких у поточному і попередньому вікні не змінилась, не будуть впливати на зміну ентропії, отже вона буде залежати лише від тих елементів, що надійшли та вибули з вікна. На рис. 1 після зсуву вікна зміниться ймовірність елементів 1 і 3, а для елементу 2 залишиться незмінною. Якщо позначити попередні значення ймовірності елементів 1 і 3 як p_{11} та p_{31}, а поточні – p_{12} та p_{32}, тоді зміна ентропії буде розраховуватись за формулою:

$$\Delta H = -p_{12}\log_2 p_{12} - p_{32}\log_2 p_{32} + p_{11}\log_2 p_{11} + p_{31}\log_2 p_{31}. \quad (5)$$

Тобто, для елементів, значення ймовірності яких змінилось, потрібно відняти ентропію до зсуву вікна та додати ентропію після зсуву.

Аналіз можливості використання ентропії для виявлення аномалій у роботі процесу/системи виконували на основі порівняння еталонної ентропії процесу/системи з ентропією процесу/системи після виконання несанкціонованих дій. Вибірки для подальших досліджень формували на основі записів журналів подій при нормальній та аномальній роботі системи.

Аналіз вмісту журналу подій показав, що за замовчуванням операційна система Windows не записує в журнал повідомлення про створення, зміну або видалення файлів, створення процесів, модифікацію реєстру. Але реєстрацію цих подій можна активувати, встановивши службу Sysmon.

У рамках цього дослідження для розрахунку ентропії було використано журнал протоколювання роботи прикладних процесів, журнал служби Sysmon, а також журнал PowerShell Operational, в якому зберігається інформація про події PowerShell, зокрема виконані в консолі команди.

На рис. 2 показано графік зміни ентропії під час звичайної роботи системи з увімкненою службою Sysmon. Значення ентропії обчислювали, базуючись на імені (ID) процесу, а для повідомлень Sysmon – на основі коду події. У цій роботі не розглядається аналіз мережевих підключень, тому ці записи були вимкнені. Розмір вікна дорівнює 100, що дозволяє при штатній роботі системи максимально повно

охопити події, що підлягають протоколюванню. Середнє значення ентропії на графіку – 1,53299, мінімальне – 0,58508.

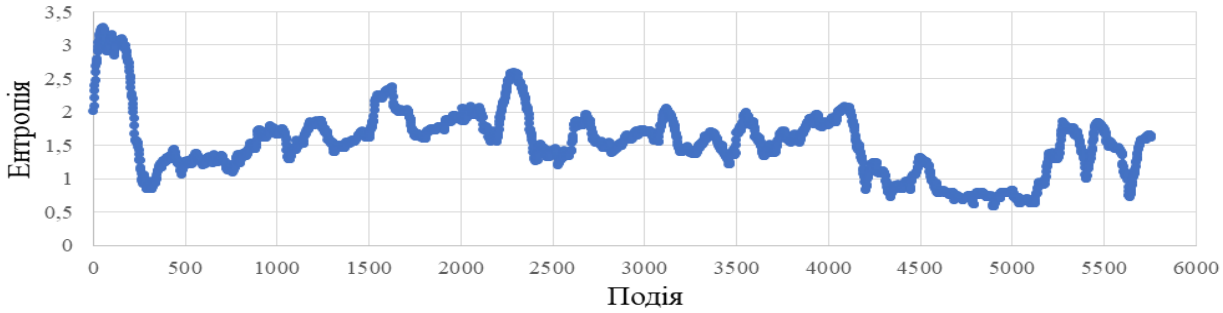

Рис. 2. Графік зміни ентропії під час нормальної роботи системи

На наступному етапі досліджень виконали моделювання несанкціонованих дії в системі з активною службою Sysmon. У рамках реалізації даного етапу було виконано запис великої кількості файлів та створено велику кількість процесів.

Створення великої кількості файлів або процесів може вказувати на локальні DoS атаки, що спрямовані на деградацію продуктивності програмно-апаратного комплексу, і, як наслідок, порушення доступності. Одночасна зміна великої кількості файлів також може свідчити про несанкціоновані дії, наприклад роботу віруса-шифрувальника (ransomware). Якщо журнал містить багато однакових записів, то ентропія буде близька до нуля, а якщо кількість записів буде більшою за розмір вікна, то ентропія буде нульовою.

На рис. 3 показано графік зміни ентропії під час запису великої кількості файлів. Такий же графік отримано у випадку створення великої кількості процесів. Під час запису файлів ентропія дорівнює нулю, бо у вікні є лише повідомлення про цю подію і її ймовірність дорівнює 1.

Рис. 3. Графік зміни ентропії під час запису великої кількості файлів

Таким чином, отримані результати дозволили зробити висновок, що аналіз значень ентропії системи, які розраховані на основі записів із журналу служби Sysmon, дозволяє одразу виявити аномальні події в системі, зокрема запис великої кількості файлів, створення великої кількості процесів в системі.

Крім того, велику кількість однакових записів в журналі можуть генерувати і інші події в системі, що не відносяться до несанкціонованих, наприклад, оновлення чи встановлення програм. Також в розрахунках ентропії не були враховані повідомлення аудиту безпеки, які теж можуть генерувати велику кількість однакових повідомлень.

Для того, щоб уникнути нульових значень ентропії під час нормальної роботи системи, можна збільшити розмір вікна, для якого буде розраховуватись ентропія. На попередніх графіках розмір вікна був 100. Якщо збільшити його до 300, то повідомлення безпеки більше не будуть спричиняти нульові значення ентропії, але вони все ще близькі до нуля.

Для встановлення ефективного розміру вікна потрібно визначити, яка кількість однакових записів в журналі є нормою, а яка свідчить про проблеми в роботі системи. З цією метою виконали перевірку, як впливають на ентропію атаки, що направлені на отримання несанкціонованого доступу до системи. Було виконано атаку за допомогою Metasploit, яка включає в себе віддалене підключення до цільової системи та виконання ескалації привілеїв «UAC bypass», що дозволить отримати права користувача «Система». Результат атаки включає в себе віддалене підключення до цільової системи та виконання ескалації привілеїв. На рис. 4 показано графік зміни ентропії із розміром вікна 300 під час нормальної роботи системи та під час виконання атаки. Середнє значення – 1,89971, мінімальне – 0,09664, максимальне – 3,14670.

Під час виконання атаки значення ентропії було максимальним за весь період спостережень, що зумовлено протоколюванням подій безпеки, зокрема таких, як модифікація реєстру та виконання команд в консолі. Також на графіку під час штатної роботи системи можна спостерігати значення ентропії, що близьке до нуля, яке спричинено значною кількістю подій безпеки, що пов’язані з читанням

облікових даних. Порівняємо цей графік із графіком зміни ентропії під час нормальної роботи системи (рис. 5).

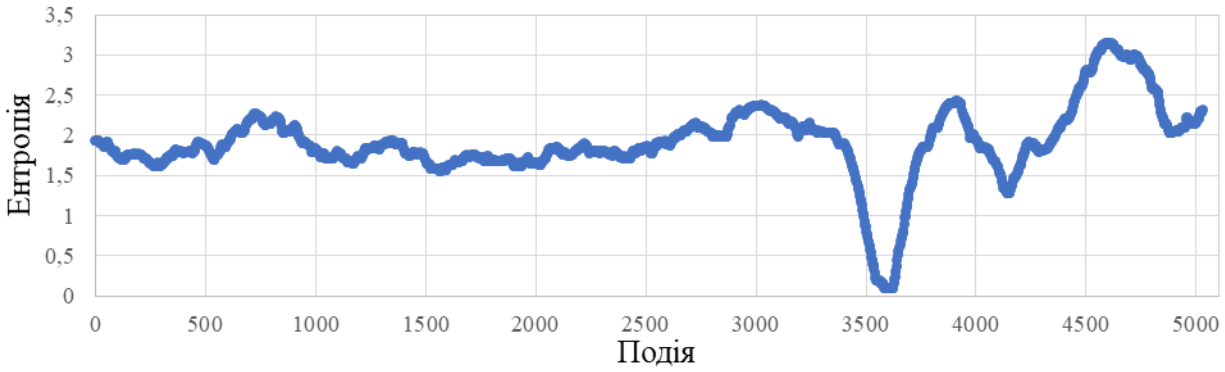

Рис. 4. Графік зміни ентропії під час виконання атаки

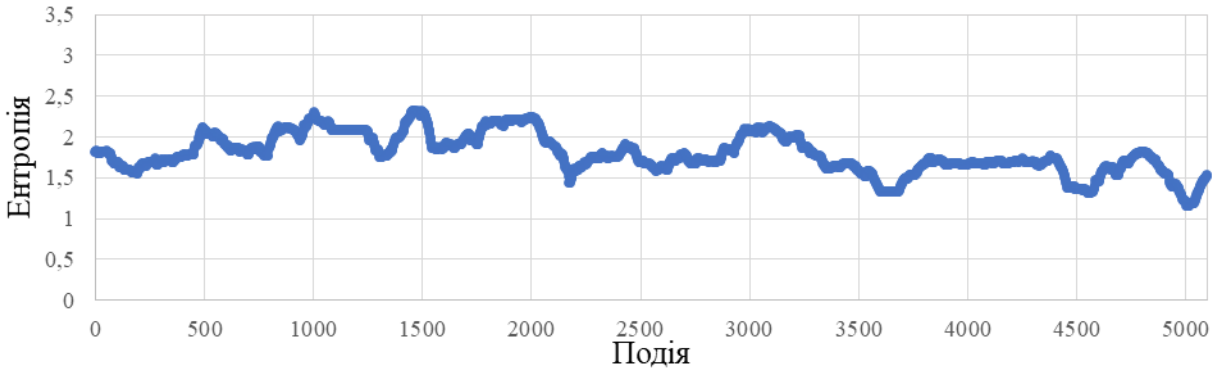

Рис. 5. Графік зміни ентропії під час нормальної роботи системи

Різниця цих графіків за модулем показана на рис. 6. Під час нормальної роботи системи відхилення ентропії не перевищує значення 1, а під час аномалій становить 1,38 та 1,78.

Рис. 6. Відхилення ентропії від еталонних значень

На основі вище означеного було зроблено висновок, що такі події суттєво вливають на значення ентропії системи, відповідно велике відхилення від еталонних значень ентропії може вказувати на виконання несанкціонованих дій у системі.

Висновки і перспективи. Таким чином, запропонований підхід дозволяє виявити аномальні стани системи, такі як запис великої кількості файлів, створення великої кількості процесів, отримання несанкціонованого віддаленого доступу до системи та ескалація привілеїв. Під час реалізації цих атак значення ентропії суттєво відрізняються від середніх. Тому можна зробити висновок, що цей метод аналізу протоколювання подій безпеки дозволяє виявляти перевищення визначених порогів безпеки, які можуть вказувати на аномалії в роботі системи. Запропонований у роботі метод може бути інтегрований у системи виявлення вторгнень, що сповіщатиме адміністратора безпеки про можливі зловживання. Також він може бути використаним при реалізації послуг безпеки, зокрема, забезпечення критеріїв рівня гарантій «НР-3. Сигналізація про небезпеку» сервісу «Реєстрація» за НД ТЗІ 2.5-004-99. Також треба зазначити, що запропонований підхід потребує подальшого розвитку, оскільки існують стани системи, за яких значення ентропії є близькими до тих, які спостерігаються під час реалізації вектору атаки.

Список використаних джерел

1. Smith Z. M., Lostri E., Lewis J. A. The hidden costs of cybercrime. McAfee. 2020. URL: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-hidden-costs-of-cybercrime.pdf

2. Колодчак О. М. Сучасні методи виявлення аномалій в системах виявлення вторгнень. Вісник Національного ун-ту «Львівська політехніка». Комп'ютерні системи та мережі. 2012. № 745. С. 98-104. URL: https://science.lpnu.ua/sites/default/files/journal-paper/2017/nov/6726/16-98-104.pdf

3. Казмірчук С., Корченко А., Паращук Т. Аналіз систем виявлення вторгнень. Захист інформації: науковий журнал. 2018. Т. 20, № 4. С. 259-276. URL: https://doi.org/10.18372/2410-7840.20.13425

4. Рубан І. В., Мартовицький В. О., Партика С. О. Класифікація методів виявлення аномалій в інформаційних системах. Системи озброєння і військова техніка. 2016. № 3. С. 100-105. URL: http://nbuv.gov.ua/UJRN/soivt_2016_3_24

5. Радівілова Т. Виявлення аномалій в телекомунікаційному трафіку статистичними методами. Кібербезпека: освіта, наука, техніка. 2021. – Т. 11, № 3. С. 183–194. URL: https://doi.org/10.28925/2663-4023.2021.11.183194

6. Yu Gu, Andrew McCallum, Don Towsley. Detecting anomalies in network traffic using maximum entropy estimation. The 5th ACM SIGCOMM conference, Berkeley, California, 19–21 October 2005 – New York, New York, USA, 2005. URL: https://doi.org/10.1145/1330107.1330148

7. Жураковський Ю. П., Полторак В. П.. Теорія інформації та кодування. К.: «Вища шк.», 2001. 255 с.

8. Gudkov O. Calculation algorithm for network flow parameters entropy in anomaly detection. IT security for the next generation. International Round, Delft University of Technology, 11–13 May 2012. URL: https://silo.tips/download/calculation-algorithm-for-network-flow-parameters-entropy-in-anomaly-detection

9.

References

1. Smith, Z. M., Lostri, E., Lewis, J. A. (2020). The Hidden Costs of Cybercrime. McAfee. Available at: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-hidden-costs-of-cybercrime.pdf

2. Kolodchak, O. (2012). Suchasni metody vyiavlennia anomalii v systemakh vyiavlennia vtorhnen. [Modern methods of detecting anomalies in intrusion detection systems]. Visnyk Natsionalnoho un-tu «Lvivska politekhnika». Komp'iuterni systemy ta merezhi, 745, 98–104. Available at: https://science.lpnu.ua/sites/default/files/journal-paper/2017/nov/6726/16-98-104.pdf

3. Kazmirchuk, S. V., Korchenko, A. O., Parashchuk, T. I. (2018). Analiz system vyiavlennia vtorhnen [Analysis of intrusion detection systems]. Ukrainian Information Security Research Journal, 20(4). Available at: https://doi.org/10.18372/2410-7840.20.13425

4. Ruban, I. V., Martovytskyi, V. O., Partyka, S. O. (2016). Klasyfikatsiia metodiv vyiavlennia anomalii v informatsiinykh systemakh [Classification of methods for detecting anomalies in information systems]. Systemy ozbroiennia i viiskova tekhnika, (3), 100-105. Available at: http://nbuv.gov.ua/UJRN/soivt_2016_3_24

5. Radivilova, T., Kirichenko, L., Tawalbeh, M., Ilkov, A. (2021). Vyiavlennia anomalii v telekomunikatsiinomu trafiku statystychnymy metodamy [Detection of anomalies in the telecommunications traffic by statistical methods]. Cybersecurity: Education, Science, Technique, 11(3), 183–194. Available at: https://doi.org/10.28925/2663-4023.2021.11.183194

6. Gu, Y., McCallum, A., Towsley, D. (2005). Detecting anomalies in network traffic using maximum entropy estimation. In the 5th ACM SIGCOMM conference. ACM Press. Available at: https://doi.org/10.1145/1330107.1330148

7. Zhurakovskyi, Y. P., Poltorak, V. P. (2001). Teoriia informatsii ta koduvannia [Information theory and coding]. Kyiv: Vyshcha shkola, 255.

8. Gudkov, O. (2012), Calculation Algorithm for Network Flow Parameters Entropy in Anomaly Detection. IT Security for the Next Generation. International Round, Delft University of Technology, May 11–13, 2012.

ВЫЯВЛЕНИЕ АНОМАЛИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ АНАЛИЗА ЭНТРОПИИ ИНФОРМАЦИОННОЙ СИСТЕМЫ

М. В. Панченко, А. Н. Бигдан, Т. В. Бабенко, Д. С. Тимофеев

Аннотация. *Меры защиты от кибератак не могут предоставлять стопроцентную гарантию невозможности проникновения злоумышленника в информационную систему. Если злоумышленник получил доступ к системе, то нужно как можно быстрее обнаружить такие действия и прервать доступ, а также провести расследование, чтобы исправить пробелы в безопасности. Для выявления атак используются методы, делящиеся на те, что выявляют злоупотребления и те, что выявляют аномалии. В данной работе исследуется возможность использования частотного метода, который выявляет аномалии в работе системы путем анализа энтропии журнала событий. Обычно этот метод используется для обнаружения аномалий в сетевом трафике, но на наличие несанкционированных действий также могут указывать аномалии в журнале событий на хостах. Проведённые исследования на примере журнала событий операционной системы Windows показали, что путем анализа энтропии можно выявить превышение порогов безопасности по количеству различных сообщений в журнале событий. Это может указывать на аномалии в работе информационной системы. Предложенный в работе метод может быть интегрирован в системы обнаружения вторжений, уведомляющих администратора безопасности о возможных злоупотреблениях.*

Ключевые слова: энтропия, аномалии, журнал событий, информационная безопасность

DETECTING THE INFORMATION SECURITY ANOMALIES BASED ON AN ENTROPY ANALYSIS OF THE INFORMATION SYSTEM

M. Panchenko, A. Bigdan, T. Babenko, D. Tymofieiev

Abstract. *Measures to protect against cyberattacks are unable to provide a 100% guarantee that an attacker cannot penetrate an information system. If an attacker has gained access to the system, then such actions should be detected as soon as possible and interrupt access, as well as an investigation to fix security gaps. Methods used to detect attacks are divided into detecting misuse and detecting anomalies. This paper investigates the applicability of a frequency method that detects anomalies in the system by analyzing the entropy of the event log. This method is typically used to detect anomalies in network traffic, and unauthorized activities can also be indicated by anomalies in the hosts' event log. Studies on the Windows event log have shown that by analyzing the entropy, it is possible to detect exceeding the security thresholds by the number of different messages in the event log. This may indicate anomalies in the operation of the information system. The proposed method can be integrated into intrusion detection systems that notify the security administrator of possible violations.*

Key words: ***entropy, anomalies, event log, information security***