УДК 004.8:[37.013.2] https://doi.org/10.32515/2664-262X.2025.11(42).2.11-22

М. В. Лахно
Національний університет біоресурсів і природокористування України, Київ, Україна
e-mail: valss725@gmail.com

Контекстуальні характеристики цифрових слідів та їхній вплив на інформаційну безпеку університету

Для галузі управління інформаційною безпекою (ІБ) інформаційно-освітнього середовища університету (ІОСУ) показано, що ефективне використання даних, отриманих шляхом збирання та аналізу різних цифрових слідів (ЦС) користувачів, включно зі студентами, викладачами та співробітниками, можливе для підвищення рівня ІБ і ступеня його захищеності від зовнішніх і внутрішніх загроз. Розроблено концептуальну модель модуля системи підтримки прийняття рішень (СППР), що базується на аналізі ЦС користувачів ІОСУ. Модель враховує як контекстно-залежні, так і контекстно-незалежні характеристики ЦС (відповідно КЗХ, КНХ), що впливають на стан ІБ в ІОСУ. Аналіз ЦС у межах ІОСУ надає фахівцям з ІБ університету можливість не тільки визначати рівні компетентності та відстежувати індивідуальні освітні траєкторії студентів, а й розробляти відповідні керуючі впливи для забезпечення стійкості ІБ. Розглядається використання програмних продуктів, як-от Splunk і ELK Stack (що включає Elasticsearch, Logstash і Kibana), що дають змогу не тільки ефективно аналізувати, а й виявляти потенційні загрози ІБ ІОСУ. Ці інструменти забезпечують університетам високу ефективність в ідентифікації, відстеженні та аналізі проблем, пов'язаних із ЦС користувачів, тим самим сприяючи поліпшенню захисту від несанкціонованого доступу до ресурсів ІОСУ.
інформаційно-освітнє середовище, освітнє середовище університету, інформаційна безпека, цифрові сліди, прийняття рішень

Постановка проблеми. Активне впровадження інформаційних технологій (ІТ) і систем (ІС) у різні сфери життя, зокрема й освітній процес, призводить до часткового або повного перенесення діяльності людини у віртуальне середовище. В освіті ця тенденція особливо яскраво себе проявила під час глобальної пандемії коронавірусу Covid-19, а також в умовах збройної агресії з боку РФ, коли чимало освітніх процесів було перенесено у віртуальне середовище, і практично всі заняття проводили з використанням відповідного програмного забезпечення (ПЗ) – відеоконференції, системи дистанційного навчання, віртуальні лабораторії, класи, які створюють у хмарному середовищі тощо. Таким чином, у сучасному світі, де ІТ та системи стали відігравати домінуючу роль у багатьох процесах, зокрема й у сфері освіти, інформаційно-освітнє середовище університетів (ІОСУУ) стало необхідною умовою реалізації якісного навчального процесу. Під час навчання студенти залишають різноманітні т. зв. цифрові сліди (далі використовуємо абревіатуру – ЦС). Ось лише невеликий перелік таких ЦС, які залишають студенти під час навчання та взаємодій з ІОСУ [1, 2, 3]. Ці ЦС можуть бути корисними студентам, викладачам та університетським адміністраторам для відстеження навчального прогресу, оцінювання робіт, комунікації та аналізу даних. Однак не менш релевантним є і завдання щодо забезпечення конфіденційності та захисту цих ЦС і персональних даних учнів і викладачів, оскільки ЦС у багатьох випадках містять особисту інформацію. Отже, інформація, що міститься в ЦС, може бути об'єктом кібератак або зловживань. В умовах глобалізації університети мають максимально адаптувати свою політику інформаційної безпеки (ІБ) до захисту даних у ЦС, а також інформувати студентів про стратегію закладу освіти щодо захисту даних користувачів, політику його ІБ та конкретні практики захисту їхніх ЦС. Інформаційна безпека ІОСУ являє собою складну

систему, що передбачає захист наявного в закладі освіти інформаційного простору та унеможливлює пошкодження або викрадення персональних даних усіх учасників навчального процесу, а також інформації, яка має фінансову, освітню, інтелектуальну цінність тощо. Забезпечення ефективного функціонування системи ІБ ІОСУ передбачає витрати певних грошових ресурсів у рамках обраної навчальним закладом стратегії захисту даних. Під час розроблення такої стратегії доцільно врахувати фактори зовнішнього та внутрішнього середовища, оскільки досягнення оптимального результату можливе лише за умови знаходження рівноваги між наявними можливостями та бажаними результатами. А в такій ситуації для менеджменту навчального закладу і персоналу, що відповідає за політику ІБ, може стати затребуваним контекстно-керований підхід до інтелектуальної підтримки прийняття рішень щодо забезпечення ІБ ІОСУ на основі аналізу ЦС користувачів. Усе вищесказане й зумовило інтерес до досліджень у цій галузі.

Аналіз останніх досліджень і публікацій.. У [4, 5, 6] розглядалися криміналістичні аспекти поняття "доріжка цифрових слідів". Авторами зазначається, що місцем виявлення ЦС можуть бути не тільки матеріальні, а й нематеріальні об'єкти, як-от ресурси мережі Інтернет, профіль особи в соціальній мережі тощо. У [5] показано, що ЦС є значущою комп'ютерною інформацією про події або дії, відображені в матеріальному середовищі, у процесі її виникнення, оброблення, зберігання та передавання. У [6, 7, 8] як приклад ЦС автори розглядали доволі широкий перелік об'єктів - різного роду інформацію, зафіксовану на носіях: дампи оперативної пам'яті; дамп трафіку; ідентифікаційні ознаки, які дають змогу унікально ідентифікувати користувача, телекомунікаційну мережу, кінцеве обладнання (IP-адресу комп'ютера в мережі, MAC-адресу мережевого обладнання, адресу електронної пошти, ідентифікатор соціальної мережі, номер банківської картки, здійснені з неї транзакції, номер телефону, дані систем геолокації тощо). Окремий напрям досліджень пов'язаний із пошуком та аналізом ЦС в інформаційних системах (ІС) університетів, зокрема тих, що відповідають за дистанційну взаємодію студентів і викладачів. У роботах [9, 10] автори зосередили увагу на описовій аналітиці освітнього середовища, використовуючи дані, отримані, зокрема, за допомогою ЦС учнів. У роботах [11, 12] автори розглядають функції ЦС не тільки для контролю індивідуальної освітньої траєкторії студентів, а й як частину системного підходу в процесі оцінювання якості освіти з використанням технології аналізу ЦС. Такий підхід дає змогу аналізувати особисті нахили студентів щодо їхнього професійного зростання [12]. На думку авторів праць [11, 13], у межах освітньої аналітики актуальними є питання з'ясування змісту таких структурних елементів ІОСЗУ, як "цифровий слід", "цифровий портрет студента", "цифровий профіль/паспорт", "маршрут/траєкторія", а також як їхні специфічні характеристики

Однак багато авторів, як наприклад, у роботах [10, 11, 12, 13] зосередили більше уваги на емпіричних дослідженнях, не приділяючи уваги аспектам моделювання процесів, пов'язаних із вивченням впливу ЦС на інформаційну безпеку учнів. Таким чином, актуальним є завдання щодо розвитку моделей інтелектуальної підтримки ухвалення рішень із забезпечення ІБ користувачів ІОСУ на основі аналізу ЦС.

Постановка завдання. Таким чином, метою дослідження є розроблення концептуальної моделі аналітичного модуля для системи підтримки ухвалення рішень з інформаційної безпеки ІОСУ на основі аналізу цифрових слідів користувачів.

Виклад основного матеріалу. Опишемо концептуально інформаційну модель аналітичного модуля ядра системи підтримки прийняття рішень щодо ІБ ІОСУ на основі аналізу ЦС користувачів. Схематично така модель показана на рисунку 1. У

такому поданні можна наочно виокремити зовнішні та внутрішні ресурси ІОСУ. Зовнішніми ресурсами вважатимемо ресурси, які існують поза концептуальною моделлю, що показана на рисунку 1. До таких ресурсів можна віднести, наприклад: профілі користувачів; моделі життєвого циклу користувачів в ІОСУ; методи профілювання користувачів в ІОСУ. Як було зазначено в огляді літератури, більшість робіт у сфері аналізу ЦС студентів в ІОСУ присвячено проблематиці аналізу поведінкової моделі учнів, що, відповідно, передбачає необхідність групування користувачів із подібною поведінкою [3, 4, 5]. При виконанні такого роду аналізу для ухвалення рішень, що стосуються як ІБ, так і формування індивідуальної освітньої траєкторії учнів, розглядаються сегменти користувачів, отримані в результаті поведінкової сегментації користувачів на основі моделей життєвого циклу користувача в ІОСУ.

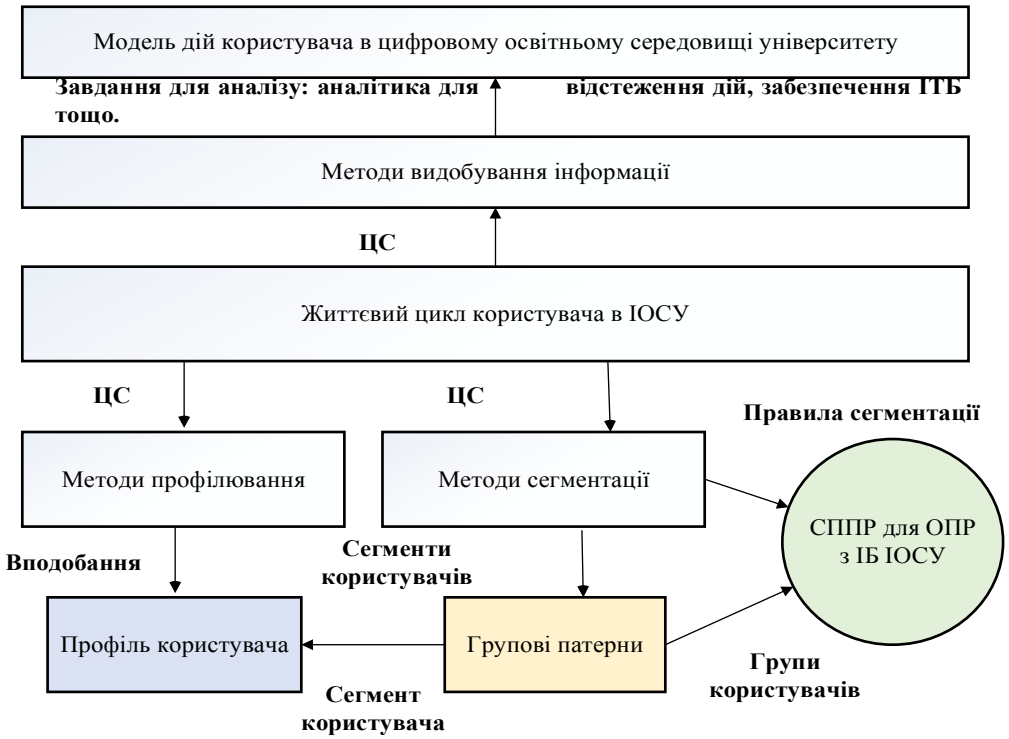

Рисунок 1 – Концептуальна інформаційна модель аналітичного модуля ядра СППР з ІБ ІОСУ на основі аналізу ЦС користувачів

Джерело: розроблено автором

Зауважимо, що ЦС студентів в ІОСУ починають формуватися з моменту їхнього вступу до навчальних закладів. І вже фактично з моменту зарахування міститимуть великий масив даних, починаючи від персональної інформації про студента і закінчуючи даними про проходження студентами всіляких онлайн-курсів. Однак у розрізі нашого дослідження основним завданням було розроблення моделі, що дає змогу вивчати вплив ЦС на ІБ учнів. Якщо розглянути завдання аналізу ЦС в ІОСУ, то вони безпосередньо пов'язані з процесом управління ІБ ІОСУ. Отже, під час вибудовування політики ІБ ІОСУ слід враховувати вимоги до безпеки зберігання, оброблення, синтезу й аналізу даних освітнього контенту та ЦС. Один зі шляхів розв'язання такого завдання полягає в організації безперервного моніторингу ІОСУ, в якій переважно і знаходиться освітній контент. Паралельно вирішується і завдання з управління комунікаціями для формування протоколів обміну даними ЦС, включно з аспектами їхньої ІБ. Наприклад, у системі дистанційного навчання (СДН) у будь-якого

студента є відповідний профіль. Тоді кортеж, що описує такий профіль у найпростішій його інтерпретації, можна представити так:

$$UP = \langle ID\_user, UP\_out, UP\_in \rangle, \quad (1)$$

де UP – профіль користувача; ID_user- унікальне ім'я користувача в ІОСУ, у тому числі в СДН, UP_out, UP_in- відповідно, множини КЗХ і КНХ користувача.

Очевидно, що користувач ІОСУ має як КЗХ, так і КНХ, що впливають на забезпечення його ІБ. До КЗХ можна, наприклад, віднести: рівень технічних знань користувача з питань ІБ. І справді користувачі, зокрема студенти, з більш глибокими пізнаннями в питаннях ІБ, як правило, краще обізнані про ризики та заходи щодо безпечної роботи в мережі. До цієї множини можна віднести і ступінь інформованості користувача про ризики ІБ. Адже глибший рівень інформованості впливає на поведінку користувача і, найчастіше, зумовлює вибір заходів ІБ. Також у цій множині буде присутній і стиль поведінки користувача в мережі університету. Відповідно, більш обережний користувач не стане відкривати небезпечні посилання на зовнішні ресурси або завантажувати ненадійне програмне забезпечення. До КНХ можна зарахувати, наприклад, ідентифікаційні дані користувача; заходи щодо забезпечення фізичної безпеки пристроїв (наприклад, використання сканера відбитків пальців на ноутбуці або смартфоні); своєчасне оновлення ПЗ; застосування стійких паролів; усвідомлення ризиків фішингу та здатність класифікувати підозрілі повідомлення, що можуть бути пов'язані з розголошенням особистої інформації; та ін. Зазначені КЗХ і КНХ, зрозуміло, можуть варіюватися в кожній ІОСУ з урахуванням її особливостей, наприклад, пов'язаних з архітектурою обчислювальної мережі, обраною стратегією забезпечення ІБ та ін. факторів. Для аналізу ЦС фахівці з ІБ використовують різні програмні засоби [14-22]. Один із результатів, одержуваний під час аналізу ЦС в ІОСУ, - це профіль завдань, які вирішував учень. Під профілем завдань в ІОСУ розуміється формалізований опис процесу взаємодії учнів з елементами ІОСУ, наприклад, сайтом системи дистанційного навчання або з інформаційною системою університету. Така взаємодія відбувається під час розв'язання певного завдання. Отже, на підставі профілів завдань у розрізі забезпечення ІБ ІОСУ фахівці із захисту інформації можуть визначати види завдань, у різних ракурсах ІБ, і, наприклад, за допомогою, відстежувати прийняті користувачами рішення. Тоді справедливим є наступний вираз, що описує модель життєвого циклу користувача в ІОСУ (MC), під час розв'язання конкретного завдання:

$$MC = (TP, Type\_P(t_o, t_d), PA, D(t_d), R_1, R_2), \quad (2)$$

де $TP-$ ідентифікатор, який привласнюють завданню певного профілю; $Type\_P(t_o, t_d)-$ тип (вид), розв'язуваної учнем задачі, починаючи з моменту часу (t_o) і до моменту ухвалення рішення (t_d) (наприклад, завантаження виконаної роботи в систему дистанційного навчання); PA – проблемна область для розв'язуваної задачі; $D(t_d)-$ ухвалення рішення в момент часу (t_d); $R_1 \in Type\_P \times PA,\ R_2 \in Type\_P \times D$.

Якщо ставиться завдання виконати сегментацію користувачів, наприклад, для опису поведінки користувачів під час ухвалення рішень (G), можна розбити учнів, зареєстрованих в ІОСУ, на види сегментів за рівнем технічних знань користувача з питань ІБ. Тоді студенти з глибшими пізнаннями в питаннях ІБ, відповідальніше поводяться в мережі університету, усвідомлюючи ризики ІБ і вживаючи заходів щодо безпечної роботи в мережі навчального закладу. Відповідно:

$$G = \langle NG, Type\_P, PA, TB, PR_g \rangle, \quad (3)$$

де G – поведінка користувачів під час ухвалення рішень; NG – назва (ім'я) групи; TB – вид поведінки; PR_g – близькість вподобань, що стосуються дотримання правил безпечної поведінки в мережі університету (для групи).

Інформацію про тип користувачів як осіб, які приймають рішення (ОПР), можна описати такою залежністю (4) – (CON)

$$CON = (ID\_user, Type\_user(t), Type\_P, PA, PR_u, PR_u(t), R_1, R_3), \quad (4)$$

де $Type\_user(t)$ – тип користувача за критеріями безпечної поведінки в ІОСУ; $PR_u, PR_u(t)$ – відповідно, множина вподобань користувача за критеріями ІБ, і множина вподобань на основі (4) (CON) у моменти часу (t); $R_3 \in PR_u \times PA$.

Так, наприклад, за критеріями безпечної поведінки в ІОСУ можна виділити такі типи користувачів:

Обізнані користувачі. До цієї групи можна віднести користувачів, які добре обізнані про ризики в мережі університету (або ІОСУ загалом) і вживають активних заходів для забезпечення безпеки своїх даних та акаунтів. Такого роду користувачі завжди дотримуються рекомендацій зі створення складних паролів, регулярно оновлюють програмне забезпечення, не відкривають підозрілих посилань або вкладень в електронних листах і використовують надійні антивірусні програми.

Недбалі користувачі. До цієї групи можна віднести користувачів, які не звертають належної уваги на заходи ІБ і можуть бути більш уразливими для атак. Зазвичай користувачів у даній групі характеризує використання слабких паролів, повтор паролів для різних акаунтів, несвоєчасне оновлення ПЗ, ігнорування підозрілої активності та заходів для захисту своїх даних в ІОСУ.

Користувачі, які не знають. До цієї групи можна віднести користувачів, які не володіють достатнім рівнем знань про заходи ІБ під час роботи в мережі. Вони можуть і не знати про ризики, пов'язані з відкриттям підозрілих посилань, не знайомі з правилами використання загальнодоступних Wi-Fi мереж. Такі користувачі можуть встановлювати ненадійне ПЗ, а також часто передають конфіденційні дані через незахищені канали зв'язку.

Байдужі користувачі. До цієї групи можна віднести користувачів, які не виявляють інтересу до питань ІБ у мережі, і, відповідно, не дотримуються жодних заходів ІБ. Ця категорія користувачів не перевіряє свої акаунти на наявність несанкціонованої активності та не звертає увагу на попередження про можливі загрози.

Безвідповідальні користувачі. Найнебезпечніша група користувачів в ІОСУ, яка порушує правила та політики ІБ у мережі університету. Вони можуть спробувати отримати несанкціонований доступ до систем ІОСУ, поширювати шкідливі програми, порушувати конфіденційність даних або вести недобросовісну активність в ІОСУ.

Зауважимо, що наведена вище категоризація типів користувачів доволі умовна, оскільки найчастіше немає чітких меж між згаданими вище категоріями користувачів, а також можуть існувати відтінки між ними. Крім того, у міру набуття знань, наприклад, завдяки відповідним курсам у навчальних програмах університетів, користувачі можуть переходити від одного типу до іншого, усвідомлюючи важливість ІБ у мережі та вживаючи відповідних заходів для захисту своїх даних та акаунтів.

Залежності (1) – (4) за їхньої відповідної алгоритмізації та програмної реалізації можна використовувати як базис під час генерації рекомендацій користувачеві ІОСУ для, наприклад, підвищення ступеня його ІБ під час роботи в мережі університету.

Під час аналізу ЦС в ІОСУ фахівцям з ІБ найчастіше доводиться стикатися саме з аналізом. Причому, лог-файли і ЦС з'єднані для аналізу цифрових доказів дій користувачів. Таким чином, лог-файли і ЦС взаємопов'язані, оскільки лог-файли містять інформацію про події та дії, які можуть бути проаналізовані для виявлення, ідентифікації та інтерпретації ЦС у світлі аналізу цифрових доказів активностей користувачів ІОСУ. По суті лог-файли – це незалежні характеристики роботи користувача в мережі університету. Вони містять інформацію про входи в систему, використання ресурсів, помилки, мережеву активність та інші події в ІОСУ.

Однак, щоб повністю зрозуміти нюанси роботи користувача, лог-файли зазвичай потребують аналізу та інтерпретації іншими інструментами або фахівцями. КЗХ роботи користувача в ІОСУ можуть містити інформацію про час, місцезнаходження, використані додатки та інші фактори, які можуть бути пов'язані з конкретною ситуацією або завданням користувача в ІОСУ. Ці характеристики можуть бути витягнуті з, але зазвичай потребують додаткового аналізу та контексту, що можна у світлі завдань забезпечення ІБ ІОСУ, наприклад, реалізувати в таких програмних продуктах як: Splunk, ELK Stack (стек програмних продуктів, що містить Elasticsearch, Logstash і Kibana) тощо. Якщо розглядати подію, пов'язану з ІБ ІОСУ, і зафіксовану, наприклад, у системі Splunk, то подію можна трактувати як факт, що був зафіксований у певний момент часу під час взаємодії певного користувача на певному пристрої з ІОСУ. Події матимуть непорожні унікальні набори атрибутів. До таких атрибутів можна віднести: ім'я користувача, пристрої взаємодії з ІОСУ, час взаємодії, типи події та ін. Також можливі й спеціальні атрибути, які залежать від типу події.

Для забезпечення ІБ ІОСУ до спеціальних атрибутів можна, наприклад, рівень доступу – визначає рівень привілеїв і доступу до інформації або ресурсів в ІОСУ; аудит і моніторинг – визначає, які дії та події записуватимуться і моніторитимуться системою, наприклад, Splunk, Elasticsearch, Logstash та Kibana; багатофакторна автентифікація - визначає, які методи автентифікації будуть використовуватись для перевірки особистості користувачів в ІОСУ; безпека мережі ІОСУ - визначає заходи безпеки, які застосовуються до мережі ІОСУ; оновлення і патчі - визначає, яким чином будуть оновлюватись і оновлюватися дані в мережі. Це лише невеликий перелік, що включає приклади спеціальних атрибутів, які можуть залежати від типу події ІБ в ІОСУ. Реальні атрибути можуть відрізнятися залежно від конкретних вимог і політик ІБ університету.

Введемо такі позначення: $E = \{e_1,...,e_n\}$ – множина подій, які зафіксовані в системах ІБ ІОСУ (антивірусне ПЗ, фаєрвол, IPS/IDS – системи виявлення та запобігання вторгненням, SIEM та ін.); $S = \{s_1,...,s_m\}$ – множина усіх зафіксованих сесій під час роботи студентів і викладачів з ІОСУ; $CL_e = \{cl_1,...,cl_k\}$ – множина класів подій ІБ, які мають певні властивості або ознаки. Тут $\{cl_i\}, i = \overline{1,k}$ – окремий клас ІБ. Наприклад, можна виконати таке розбиття на класи: "Адміністратори"; "Викладачі"; "Студенти"; "Гості". Відповідно, кожен клас для забезпечення ІБ матиме свої права доступу та обмеження, що дасть змогу керувати доступом до різних рівнів інформації в ІОСУ та захищати дані від несанкціонованого використання. Отже, патерн являтиме собою розміщення елементів множини $CL_e = \{cl_1,...,cl_k\}$, у якій є повторення: $rep_i = \langle cl_{i1},...,cl_{ij} \rangle$, де $\{cl_{ij}\}, i = \overline{1,r}, j = \overline{1,q_i}, cl_{ij} \in CL_e$ - одинична подія для i – го патерну;

$r-$ потужність множини патернів (P); q_i – сукупність подій для патерну. Кожну подію ІБ опишемо за допомогою непорожньої множини ознак (набору атрибутів) – $M=\{m_1,...,m_w\}$. Відповідно, ознаки, властиві об'єктам множини $E=\{e_1,...,e_n\}$. Вважаємо, що об'єкти в $E=\{e_1,...,e_n\}$ і ознаки в $M=\{m_1,...,m_w\}$ розрізняються. Відповідно до [22] має місце таке співвідношення $I|E\times M$, яке відповідає інцидентності між множинами $E=\{e_1,...,e_n\}$ і $M=\{m_1,...,m_w\}$. Наявність у I пари $(e,m), e|M$ і $m|M$, відповідає тому, що об'єкт (e) має ознаку (m). І навпаки, ознака (m) буде властива об'єкту (e). Відповідно до [22], є множина ознак контексту $K=(E,M,I)$. Тоді, на множині ознак контексту $K=(E,M,I)$ буде справедливим наступне асоціативне правило $AR=(X,Y)$, яке має вигляд $X\Rightarrow Y$. Це правило буде кількісно за допомогою двох числових функцій характеризувати підтримку $\sup(X\Rightarrow Y)$ і достовірність $conf(X\Rightarrow Y)$ асоціативного правила. У розрізі ІБ ІОСУ розв'язання задачі з пошуку послідовних патернів можна звести до виявлення максимальних послідовностей, що мають підтримку, вищу за заданий поріг. Іншими словами, для заданого $K=(E,M,I)$ необхідно знайти множину $AR=(X,Y)$ усіх асоціативних залежностей множини патернів (P). Зазначимо, що набір $AR=(X,Y)$ слід описати відносно порогових значень $\sup(X\Rightarrow Y)$ і $conf(X\Rightarrow Y)$. Для того, щоб класифікувати події ІБ в ІОСУ можна обчислити цільову функцію на основі такого правила:

$$conf(X\Rightarrow Y)=\max f(E\to CL_e). \quad (5)$$

Як вище зазначалося, у процесі реалізації заходів, спрямованих на підвищення ІБ ІОСУ, і, відповідно, під час розв'язання локальних завдань з аналізу ЦС фахівцям з ІБ доводиться стикатися з аналізом. Практична реалізація задачі аналізу патернів (шаблонів), що містять необліковані персональні дані, була виконана за допомогою лог-файлів кількох курсів системи дистанційного навчання (СДН) Національного університету біоресурсів і природокористування України. Запити для візуалізації такої карти у Splunk, Elasticsearch, Logstash та Kibana, виглядають, наприклад, так:

sourcetype= access_combined_wcookie| iplocation clientip| stats count by Country| geom geo_countries featureIdField="Country".

Покладемо, наприклад, що умовний ієрархічний лог-файл СДН складається з таких частин: процес; випадок; подія. Можна використовувати й складнішу ієрархічну структуру. Подію можна описати у форматі кортежу $\langle doc, op, per\rangle$, де doc – відповідний документ в ІОСУ; op – вид операції в ІОСУ; per – особа, яка виконала операцію (ОВО). Наприклад, розглянемо три компоненти процесу обігу документів у СДН університету: $Op\times Op$, $Per\times Per$ та $Per\times Op$. Отже, складову $Op\times Op$ можна описати відповідним лог-файлом Q. Цей файл відповідатиме сукупності послідовних траєкторій у такому вигляді $Q=\{tr_i | tr_i\in Op^*\}$. Тут траєкторія tr – випадок, для якого характерна така послідовність операцій, що $tr\in Op^*$. Де Op^* – сукупність усіх сполучень з алфавіту типів операцій у СДН університету – Op.

Друга компонента (складова) дає змогу описати організаційну структуру обігу документів у СДН, тобто - $Per(Q)$. Другу компоненту зручно відображати у вигляді

графа спільної роботи ОВО. Такий граф Gr має N вершин, де кожна вершина відповідатиме конкретному ОВО. Третя складова відображатиме розподіл ОВО за операціями відповідних процесів.

Ми використовували метод кластеризації для аналізу даних у, що характеризують стани ІБ СДН університету. Приймаємо, що вихідними даними будуть: множина ОВО $\{per_1, per_2,..., per_N\}$, що беруть участь у процесі документообігу, наприклад, для конкретних курсів у СДН; матриця близькості між діями $\{per_1, per_2,..., per_N\}$ – W_{ij}; склад ОВО, що беруть участь у поодиноких випадках опрацювання документів у СДН – $per(doc)$. Нижче подано покроковий алгоритм, див. рис. 2, виявлення загрози порушення складу ОВО, які, наприклад, беруть участь у процесі реалізації спільного навчального проєкту, розміщеного в СДН.

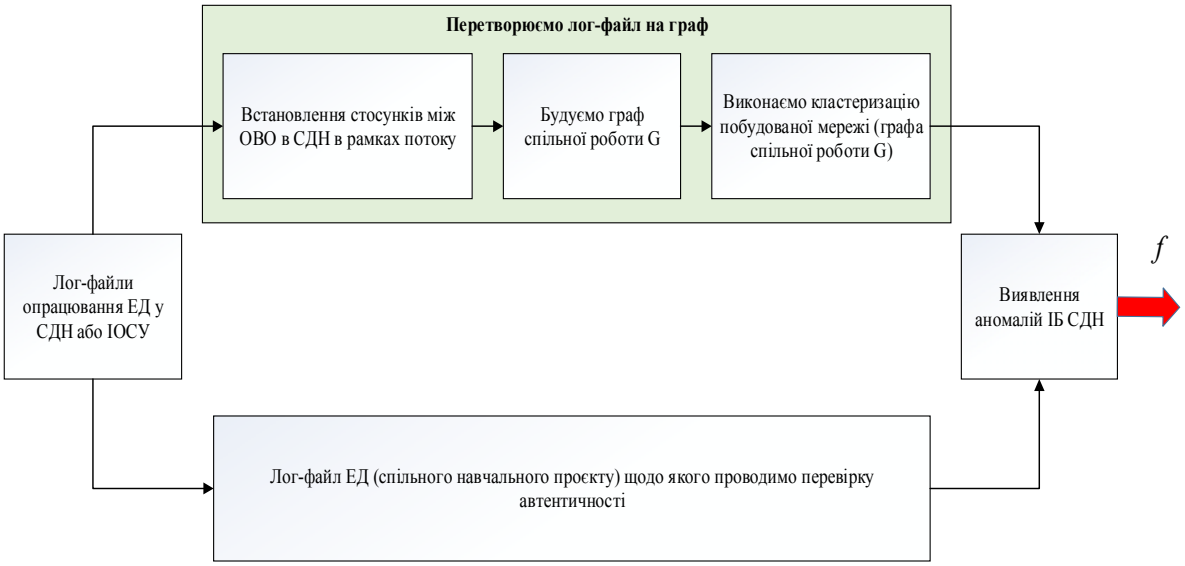

Рисунок 2 – Схема опрацювання лог-файлів під час аналізу порушення складу виконавців спільного навчального проєкту в системі дистанційного навчання університету

Джерело: розроблено автором

Крок 1. Створюємо N груп з однією ОВО в кожному $c_1 = \{per_1\}, c_2 = \{per_2\},..., c_n = \{per_n\}.$

Крок 2. Обчислимо метрики близькості:

$$D(c_i, c_j) = \frac{1}{|c_i||c_j|} \cdot \sum_{x \le |c_i|} \sum_{y \le |c_j|} r_{xy}, \quad (6)$$

де $|c_i|$ – число вершин ОВО, що входять до групи i; $|c_j|$ – число вершин ОВО, що входять до групи j; r_{xy} – потужність зв'язку між вершинами графа.

Крок 3. Об'єднаємо пару груп, що характеризуються максимальною метрикою близькості. Кількість нових $N' = N - 1.$

Крок 4. Кроки 2,3,4 повторюються доти, $N' = 1.$

Крок 5. Виділяються групи для $G = \{g_1, g_2,..., g_{n-1}\}.$

Крок 6. Обчислюються коефіцієнти модулярності V [23] для кожної групи $g_i \in G$. Для чого було застосовано таку залежність [24]:

$$V = \left(1 \middle/ (2 \cdot m) \right) \cdot \sum_{xy} \left(W_{xy} - \frac{k_x \cdot k_y}{2m} \right) \cdot \beta(u_x, u_y), \tag{7}$$

де $m -$ v сума ваг дуг графа; W_{xy} – матриця близькості; k_x, k_y – показники ступеня груп x і y (кількість інцидентних для них дуг); $\beta(u_x, u_y) = 1$, якщо вершини x і y належать одній групі і дорівнює 0 у протилежному випадку.

Крок 7. Визначимо, в якому випадку коефіцієнт модулярності V [23] матиме максимальне значення. Як результат отримуємо найкращий поділ на кластери, $Per(Q) = \{Cl_1, Cl_2, ..., Cl_n\}$, де $Per(Q) -$ сукупність учнів, наприклад, на факультеті; Cl – кластери.

Крок 8. Знайдемо в базі довірених лог-файлів випадки, для яких ОВО, що беруть участь у спільному навчальному проєкті, перебуватимуть у різних кластерах. Для цього дивимося результати, отримані на кроці 7. Для $U_q = \{U_{q_1}, U_{q_2}, ..., U_{q_n}\}$ визначимо метрики близькості для кожного набору U_q, застосувавши формулу (6).

Крок 9. Визначимо найменшу метрику близькості для набору $U_q = \{U_{q_1}, U_{q_2}, ..., U_{q_n}\}$. Потім встановимо найменшу метрику близькості для набору як межу для виявлення випадків аномальностей для $D_{\min}$.

Крок 10. Вважаємо, що лог-файл, який перебуває на аналізі, характеризує набір ОВО, які беруть участь у спільному навчальному проєкті, тобто $per(doc)$. Визначимо найменше значення метрики близькості між $per(doc)$ і кластерами. Вираз має такий вигляд:

$$\min D(per(doc), K_i) = \min(D(per(doc), K_1), D(per(doc), K_2), ..., D((per(doc), K_i)). \tag{8}$$

Крок 11. Визначимо значення функції f (на рисунку 2 показана стрілкою червоного кольору). Ця функція повідомляє про можливу аномальність набору ОВО, що бере участь у документопотоці $per(doc)$. Вважаємо, що цей документопотік належить до відносно $Per(Q)$.

Отримані результати можна обробляти спільно з текстовими, наприклад, використовуючи ПЗ Elasticsearch.

Під час реалізації експериментальної частини роботи було розглянуто основні види лог-файлів в ІОСУ, зокрема в СДН, та особливості зберігання й аналізу таких файлів з погляду забезпечення ІБ ІОСУ. Було проведено порівняльний аналіз систем журналювання подій серверів Windows і Linux, що дало змогу точніше визначити їхні відмінності та спільні риси. Експериментально було досліджено основні компоненти системи Elasticsearch, Logstash і Kibana для серверів двох зазначених вище університетів. Особливу увагу приділено розробці фільтрів і патернів для Logstash і шаблонів для Elasticsearch, які є ключовим елементом ефективного аналізу лог-файлів в ІОСУ. Таким чином, використовуючи наведені вище викладки, що стосуються схем

обробки лог-файлів і формування патернів ЦС, було зроблено висновок, що стек ELK є потужним інструментом для аналізу лог-файлів, як частини цифрового сліду студентів, і виявлення потенційних загроз безпеці ІОСУ

Висновки. Проаналізовано можливості використання даних цифрових слідів (ЦС) користувачів інформаційно-освітньої системи університету (ІОСУ) для інтелектуалізації ухвалення рішень, спрямованих на поліпшення інформаційної безпеки (ІБ). Показано, що для управління ІБ ІОСУ можна використовувати дані, отримані під час аналізу ЦС учнів і співробітників. Запропоновано концептуальну модель модуля системи підтримки ухвалення рішень (СППР), що ґрунтується на аналізі ЦС, яка враховує як контекстно-залежні, так і контекстно-незалежні характеристики (КЗХ і КНХ), що впливають на стан ІБ. Аналіз ЦС дає змогу менеджменту навчальних закладів не тільки відстежувати освітні траєкторії учнів і рівні їхніх компетенцій, а й надавати рекомендації для поліпшення ІБ через СППР для підрозділів, які опікуються захистом ІОСУ. Додаткові дослідження в галузі деталізованих моделей характеристик ЦС можуть стати основою для обчислювального ядра СППР, що сприяє побудові багатошарової системи ІБ. Експерименти показали, що стек ELK є ефективним інструментом для аналізу лог-файлів і виявлення загроз безпеці ІОСУ, забезпечуючи високий рівень захисту від несанкціонованого доступу.

Список літератури

1. Лахно В., Волошин С., Мамченко С., Кулініч О., Касаткін Д. Кластерний аналіз для дослідження цифрових слідів студентів закладів освіти. *Кібербезпека: освіта, наука, техніка*. 2024. № 3(23). С. 31–41.
2. Пічкур М. О., Полудень Л. І., Демченко І. І., Сотська Г. І. Моніторинг цифрових слідів образотворчої підготовки здобувачів вищої мистецької освіти. *Інформаційні технології та засоби навчання*. 2023. № 2(94). С. 128–149.
3. Морзе Н., Кузьмінська О., Мазорчук М. Ставлення до цифрового навчального середовища в українських університетах. ICT in Education, Research, and Industrial Applications. *Proc. 15th Int. Conf. ICTERI 2019. Volume II: Workshops*. Запоріжжя : ЗНУ, 2019. Т. 2393. С. 53–67.
4. Andersen J. V., Lindberg A., Lindgren R., Selander L. Алгоритмічна агентура в інформаційних системах: дослідницькі можливості для аналітики даних цифрових слідів. *49th Hawaii International Conference on System Sciences (HICSS)*. 2016. С. 4597–4605.
5. Hedman J., Srinivisan N., Lindgren R. Digital Traces of Information Systems: Sociomateriality Made Researchable. *Proceedings of the 34th International Conference on Information Systems*. Milan, 2013. Vol. 38. P. 809–830.
6. Ollagnier-Beldame M. Використання цифрових слідів: перспективна основа для проектування адаптованих інформаційних систем? *International Journal on Computer Science and Information Systems*. 2011. Vol. 6(2). P. 24–45.
7. Sen I., Flöck F., Weller K., Weiß B., Wagner C. Структура тотальної помилки для цифрових слідів поведінки людини на онлайн-платформах. *Public Opinion Quarterly*. 2021. Vol. 85(S1). P. 399–422.
8. Mohssine B., Mohammed A., Abdelwahed N., Mohammed T. Адаптивна система допомоги на основі "Цифрових слідів" учнів та стилів навчання. *International Journal of Emerging Technologies in Learning (iJET)*. 2021. Vol. 16(10). P. 288–294.
9. Kendall J. S., Dandapani N., Cicchinelli L. F. Бенчмаркінг інформаційної системи управління освітою штату Понпей. *REL 2017-175*. Pacific Regional Educational Laboratory, 2016.
10. Билієва Д. С., Лобатюк В. В., Рубцова А. В. Інформаційно-комунікаційні технології як активний принцип соціальних змін. *IOP Conference Series: Earth and Environmental Science*. 2019. Vol. 337, № 1. P. 012054.
11. Билієва Д., Лобатюк В., Сафонова А., Рубцова А. Співвідношення між практичним аспектом курсу та прогресом електронного навчання. *Науки про освіту*. 2019. Т. 9, № 3. С. 167.
12. Смирнов А. В. Т., Левашова Т. В. Контекстно-свідомий підхід до інтелектуальної підтримки прийняття рішень на основі цифрових слідів користувачів. *Informatics and Automation*. 2020. Vol. 19(5). P. 915–941.
13. Almarghani E. M., Mijatovic I. Фактори, що впливають на залученість студентів у вищих навчальних закладах. *Teaching in Higher Education*. 2017. Vol. 22(8). P. 940–956.

14. Tara G., Mishra A. Порівняльне дослідження інструментів цифрової криміналістики для вилучення даних з електронних пристроїв. *2021*. С. 97–104.
15. Yulianto S., Soewito B. Дослідження впливу на відновлення даних у комп'ютерній криміналістиці. *IEEE Int. Conf. on Cryptography, Informatics, and Cybersecurity (ICoCICs)*. 2023. P. 109–114.
16. Javed A. R., Ahmed W., Alazab M., Jalil Z., Kifayat K., Gadekallu T. R. Всебічний огляд комп'ютерної криміналістики: сучасний стан, інструменти, методи, виклики та майбутні напрямки. *IEEE Access*. 2022. Vol. 10. P. 11065–11089.
17. Delija D., Mohenski I., Sirovatka G. Порівняльний аналіз інструментів мережевої криміналістики на різних операційних системах. *44th Int. Convention on Information, Communication and Electronic Technology (MIPRO)*. 2021. P. 1231–1235.
18. Mehrotra T., Mehtre B. M. Автоматизований інструмент криміналістичної експертизи метаданих зображень і Windows 7. *Int. Conf. on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*. 2014. P. 419–425.
19. Huntington P., Nicholas D., Jamali H. R., Watkinson A. Отримання суб'єктних даних із лог-файлів за допомогою глибокого лог-аналізу: case study OhioLINK. *Journal of Information Science*. 2006. Vol. 32(4). P. 299–308.
20. Zhu J., Wu H., Gao G. Ефективний метод видобутку веб-послідовних шаблонів на основі фільтра сеансів та ідентифікації транзакцій. *Journal of Networks*. 2010. Vol. 5(9). P. 1017.
21. Wang J. Z., Huang J. L., Chen Y. C. Про ефективний видобуток послідовних шаблонів високої корисності. *Knowledge and Information Systems*. 2016. Vol. 49. P. 597–627.
22. Wedyan S. Огляд і порівняння підходів до видобутку даних асоціативної класифікації. *International Journal of Industrial and Manufacturing Engineering*. 2014. Vol. 8(1). P. 34–45.
23. Goswami A., Polly P. D. Методи вивчення морфологічної інтеграції та модульности. *Paleontological Society Papers*. 2010. Vol. 16. P. 213–243.
24. Tsibulya A. N., NGIA H. M. Algoritm analiza zhurnala registratsii sistemnyih protsessov sistemyi elektronnogo dokumentooborota s ispolzovaniem metoda klasterizatsii sotsialьnuyih setey. *Informatsionnyie sistemyi i tehnologii*. 2016. № 94(2). C. 148.

References

1. Lakhno, V., Voloshyn, S., Mamchenko, S., Kulinich, O., & Kasatkin, D. (2024). Cluster analysis for the study of digital footprints of students of educational institutions. Electronic professional scientific publication “Cybersecurity: Education, Science, Technology”, 3(23), 31-41 [in Ukrainian].
2. Pichkur, M. O., Poluden, L. I., Demchenko, I. I., & Sotska, G. I. (2023). Monitoring of digital traces of fine arts training of applicants for higher art education. Information Technologies and Learning Tools, 2(94), 128-149 [in Ukrainian].
3. Morse, N., Kuzminska, O., & Mazorchuk, M. (2019). Attitudes towards the digital learning environment in Ukrainian universities. In ICT in Education, Research, and Industrial Applications. Proc. 15th Int. Conf. ICTERI 2019. Volume II: Workshops. (Vol. 2393, pp. 53-67). Zaporizhzhia National University, Department of Computer Science [in Ukrainian].
4. Andersen, J. V., Lindberg, A., Lindgren, R., & Selander, L. (2016, January). Algorithmic agency in information systems: research opportunities for analyzing digital footprint data. In 2016 49th Hawaii International Conference on System Sciences (HICSS) (pp. 4597-4605). IEEE.
5. J. Hedman, N. Srinivisan and R. Lindgren, “Digital Traces of Information Systems: Sociomateriality Made Researchable,” Thirty-fourth International Conference on Information Systems Milan 2013, vol. 38, pp. 809-830, 2013.
6. Ollagnier-Beldame, M. (2011). Utilizing digital footprints: a promising framework for designing adapted information systems? International Journal on Computer Science and Information Systems. Special issue “Users and Information Systems”, 6(2), 24-45.
7. Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. Public Opinion Quarterly, 85(S1), 399-422.
8. Mohssine, B., Mohammed, A., Abdelwahed, N., & Mohammed, T. (2021). An adaptive tutoring system based on students' digital footprints and learning styles. International Journal of Emerging Technologies in Education (iJET), 16(10), 288-294.
9. Kendall, J. S., Dandapani, N., & Cicchinelli, L. F. (2016). Benchmarking the Pohnpei State Education Management Information System. REL 2017-175. Regional Education Laboratory for the Pacific.
10. Bilyeva DS, Lobatyuk VV, Rubtsova AV (2019). Information and communication technologies as an active principle of social change. In IOP Conference Series: Earth and Environmental Science (Vol. 337, No. 1, p. 012054). IOP Publishing.

11. Bilyeva DS, Lobatyuk VV, Rubtsova AV (2019). Information and communication technologies as an active principle of social change. In IOP Conference Series: Earth and Environmental Science (Vol. 337, No. 1, p. 012054). IOP Publishing.
12. Bilyeva D., Lobatyuk V., Safonova A., Rubtsova A. (2019). The relationship between the practical aspect of the course and the progress of e-learning. Education Sciences, 9(3), 167.
13. Smirnov, A. V. T., & Levashova, T. V. (2020). Context-aware approach to intelligent decision support based on digital user footprints. Informatics and Automation, 19(5), 915-941.
14. Almarghani, E. M., & Mijatovic, I. (2017). Factors influencing student engagement in higher education - it's all about good teaching. Teaching in Higher Education, 22(8), 940-956.
15. Tara, Harshita and Amarnath Mishra. "A comparative study of digital forensics tools for data extraction from electronic devices." (2021): 97-104.
16. Yulianto, S., & Soewito, B. (2023, August). Research on the impact on data recovery in computer forensics. In 2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs) (pp. 109-114). IEEE.
17. Javed, A. R., Ahmed, W., Alazab, M., Jalil, Z., Kifayat, K., & Gadekallu, T. R. (2022). A comprehensive review of computer forensics: state of the art, tools, techniques, challenges, and future directions. IEEE Access, 10, 11065-11089.
18. Delija, D., Mohenski, I., & Sirovatka, G. (2021, September). Comparative analysis of network forensics tools on different operating systems. In 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 1231-1235). IEEE.
19. Mehrotra, T., & Mehre, B. M. (2014, July). An automated forensic tool for image metadata and Windows 7. In 2014 International Conference on Control, Instrumentation, Communication and Computing Technologies (ICCICCT) (pp. 419-425). IEEE.
20. Huntington, P., Nicholas, D., Jamali, H. R., & Watkinson, A. (2006). Extracting subject data from log files using deep log analysis: The OhioLINK case study. Journal of Information Science, 32(4), 299-308.
21. Zhu, J., Wu, H., & Gao, G. (2010). An efficient method for web sequential pattern mining based on session filter and transaction identification. Journal of Networks, 5(9), 1017.
22. Wang, J. Z., Huang, J. L., & Chen, Y. C. (2016). On efficient mining of sequential high utility patterns. Knowledge and Information Systems, 49, 597-627.
23. Wedyan, S. (2014). A survey and comparison of approaches to mining associative classification data. International Journal of Industrial and Manufacturing Engineering, 8(1), 34-45.
24. Goswami, A., & Polly, P. D. (2010). Methods for studying morphological integration and modularity. Transactions of the Palaeontological Society, 16, 213-243.
25. Tsibulya, A. N., & NGIA, H. M. (2016). Algoritm analiza zhurnala registratsii sistemnyih protsessov sistemy elektronnogo dokumentooborota s ispolzovaniem metoda klasterizatsii sotsial'nyih setey. Informatsionnyie sistemii i tehnologii, 94(2), 148.

Miroslav Lakhno
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine

Contextual Characteristics of Digital Footprints and Their Impact on University Information Security

For the studied area of information security management (ISM) of the university information and educational environment (IEE), it is shown that the effective use of data obtained by collecting and analyzing various digital footprints (DF) of users, including students, teachers and staff, is possible to increase the level of ISM and the degree of its protection against external and internal threats. A conceptual model of a decision support system (DSS) module based on the analysis of IOSS users' CS is developed. The model takes into account both context-dependent and context-independent characteristics of the CA (respectively, CCA, CCA) that affect the state of the IS in the IIS. The analysis of CS within the IES provides university IS specialists with the opportunity not only to determine the levels of competence and track individual educational trajectories of students, but also to develop appropriate controlling influences to ensure the sustainability of IS. We consider the use of software products such as Splunk and ELK Stack (which includes Elasticsearch, Logstash, and Kibana), which allow not only to effectively analyze but also to identify potential threats to IES. These tools provide universities with high efficiency in identifying, tracking and analyzing problems related to user's CA, thereby contributing to improved protection against unauthorized access to IPS resources.

information and educational environment, university educational environment, information security, digital footprints, decision-making

Одержано (Received) 09.04.2025 *Прорецензовано (Reviewed) 30.04.2025*
Прийнято до друку (Approved) 02.05.2025