© Koval M. Structured Modeling of Cyber Threat Intelligence Data: Methodological Foundations and Practical Applicability. // *Ukrainian Scientific Journal of Information Security, 2026, vol. 32, issue 2, pp. 51-59. © This is an open access article under the CC BY 4.0 International license.*

DOI: 10.18372/2225-5036.32.21658

СТРУКТУРОВАНЕ МОДЕЛЮВАННЯ ДАНИХ ПРО КІБЕРЗАГРОЗИ: МЕТОДИЧНІ ЗАСАДИ ТА ПРАКТИЧНА ЗАСТОСОВНІСТЬ

Микола Коваль

Національний університет «Київський авіаційний інститут»,
Державна служба спеціального зв'язку та захисту інформації України



КОВАЛЬ Микола Васильович
Рік та місце народження: 1988, м. Київ
Освіта: ІСЗЗІ НТУУ "КПІ", 2010
Посада: *фахівець Державного центру кіберзахисту Державної служби спеціального зв'язку та захисту інформації України*
Наукові інтереси: кібербезпека, інформаційна безпека, технічний захист інформації, реагування на кіберінциденти, комп'ютерно-технічні дослідження, кіберрозвідка
E-mail: uac.3333@gmail.com
Orcid ID: 0009-0004-9699-6849

Анотація. *У статті розглянуто структуроване моделювання даних про кіберзагрози як передумову їх накопичення, зіставлення, аналізу та повторного використання. Обґрунтовано, що зберігання окремих індикаторів у таблицях, системах обліку звернень або текстових звітах без формалізованих зв'язків з подіями, інфраструктурою, програмними засобами реалізації кіберзагроз, кампаніями й аналітичними висновками обмежує практичну цінність таких даних. Розмежовано платформу як прикладний засіб роботи з інформацією та модель подання даних як формалізований спосіб опису сутностей, їхніх властивостей і зв'язків. На основі структурно-функціонального аналізу офіційної документації простежено розвиток формалізованого об'єктного подання від CRITs/CybOX до STIX 2.1 та окремо розглянуто гіперграфовий підхід Synapse. Визначено спільні предметні категорії, за якими оцінено здатність розглянутих підходів подавати спостережувані технічні об'єкти, первинні повідомлення, атаки та факти компрометації, кампанії, інфраструктуру, програмні засоби, тактики, техніки, процедури, об'єкти впливу й аналітичні твердження. Показано, що розглянуті підходи не є взаємозамінними: CRITs/CybOX орієнтований на практичне накопичення об'єктів і похідних ознак, STIX 2.1 – на стандартизоване подання та обмін, Synapse – на графове поєднання фактичного, подієвого, контекстного й аналітичного шарів. Практичне значення результатів полягає у формуванні методичних орієнтирів для переходу від фрагментарного зберігання індикаторів до системної роботи з пов'язаними даними про кіберзагрози.*

Ключові слова: *моделювання даних про кіберзагрози, threat intelligence, індикатори кіберзагроз, структуровані дані, граф знань, STIX 2.1, CybOX, Synapse, CRITs.*

Вступ

Відповідно до пункту 1 частини третьої статті 9 Закону України «Про основні засади забезпечення кібербезпеки України» [1] до завдань національної команди реагування на кіберінциденти, кібератаки, кіберзагрози CERT-UA належать моніторинг, накопичення та проведення аналізу даних про кіберінциденти, кібератаки, кіберзагрози на національному, галузевому та регіональному рівнях, а також проведення динамічного аналізу ризиків і забезпечення ситуаційної обізнаності. У 2026 році Порядок обміну інформацією про кіберінциденти, кібератаки та кіберзагрози [2] конкретизував предмет такого обміну, до якого належать повідомлення про кіберінциденти й кібератаки, інформація про події кібербезпеки, тактики, техніки та процедури, індикатори кіберзагроз, відомості про вразливості й ризики, звіти та рекомендації. За наявності технічної сумісності обмін може здійснюватися в автоматизованому й машинозчитуваному форматах. Отже, робота з даними про кіберзагрози є не допоміжним, а базовим елементом функціонування національної системи реагування.

Разом з тим саме словосполучення «дані про кіберінциденти, кібератаки, кіберзагрози» у правозастосовній та аналітичній практиці є значно ширшим за окремі індикатори кіберзагроз. Закон визначає індикатори кіберзагроз як показники (технічні дані), що використовуються для виявлення та реагування на кіберзагрози. Проте для динамічного аналізу ризиків та забезпечення ситуаційної обізнаності недостатньо лише переліку технічних показників без зазначення їхнього походження, зв'язків, часових характеристик і належності до події, кампанії або кластера кіберзагрози. У такому вигляді дані придатні до локального виявлення, але обмежено придатні до зіставлення, узагальнення та пояснення.

У практичній діяльності відомості про кіберзагрози нерідко накопичуються у вигляді окремих таблиць, записів у системах обліку звернень, переліків індикаторів або текстових описів. Такі засоби можуть бути зручними для локального обліку, однак вони не забезпечують сталого відтворення семантики об'єктів і зв'язків між ними. Унаслідок цього одна й та сама IP-адреса, домен, файл або адреса електронної пошти можуть багаторазово фіксуватися без збереження ролі в конкретній події,

джерела отримання, часової чинності, належності до програмного засобу, кампанії чи аналітичного висновку. Проблема полягає не у виборі формату файла або програмного продукту, а у відсутності послідовної моделі роботи з пов'язаними даними.

Проблематика структурованого подання даних про кіберзагрози досліджувалася у працях, присвячених таксономіям, стандартам обміну, онтологіям і моделям даних [3], [4]. Окремо розглядалися практичні наслідки обміну такими даними та питання їх операціоналізації [5]. Водночас для вітчизняної практики актуальним залишається не лише вибір стандарту обміну, а й методичний перехід від фрагментарного обліку індикаторів до накопичення взаємопов'язаних фактів, контексту й аналітичних тверджень, придатних до повторного використання різними командами та інформаційними системами.

У попередній роботі автора активний кіберзахист було розглянуто через сукупність технологічних спроможностей, серед яких виявлення, ідентифікація та атрибуція кіберзагроз передують вибору й застосуванню заходу активної протидії [6]. Структуроване моделювання даних є необхідною інформаційною основою для реалізації цих спроможностей у масштабованому та відтворюваному вигляді.

Метою цієї роботи є обґрунтування структурованого моделювання даних про кіберзагрози як самостійної методичної основи їх накопичення, зіставлення, аналізу й повторного використання, а також визначення практичної застосовності основних підходів до подання таких даних. Для досягнення мети розмежовано поняття платформи та моделі подання даних; визначено методику аналізу сутностей і зв'язків; простежено еволюцію підходів CRITs/CybOX, STIX 2.1 і Synapse; узагальнено їхні можливості за спільними предметними категоріями.

Наукова новизна одержаних результатів полягає у розвитку методичних засад організації роботи з даними про кіберзагрози шляхом розмежування технологічної платформи та її понятійно-структурної основи, а також оцінювання практичної придатності моделей не за кількістю підтримуваних індикаторів, а за здатністю зберігати зміст сутностей і зв'язків між фактичним, подієвим, контекстним та аналітичним шарами даних.

Основна частина дослідження

Необхідність структурованого моделювання даних про кіберзагрози

Сам по собі перелік індикаторів кіберзагроз не дає змоги повноцінно здійснювати зіставлення подій, виявляти повторювані закономірності, встановлювати зв'язки між кампаніями, кіберінцидентами та формувати обґрунтовані висновки щодо належності активності до певного кластера кіберзагроз. Індикатор, розглянутий поза контекстом, придатний насамперед для локального виявлення або блокування. Натомість для аналітичної роботи значення мають також джерело походження відомостей, часові характеристики, роль конкретного об'єкта в події, його зв'язок з іншими об'єктами, належність до кампанії, повторюваність у різних епізодах та співвіднесення зі вже відомими кластерами активності. У цьому сенсі структуроване моделювання є способом упорядкування даних, за якого технічні ознаки, пов'язані об'єкти та аналітичні висновки розглядаються не ізольовано, а як елементи єдиної системи [4].

Під даними про кіберзагрози доцільно розуміти сукупність спостережуваних технічних ознак, відомостей про події, індикаторів, інфраструктуру, програмні засоби реалізації кіберзагроз, тактики, техніки, процедури, об'єкти впливу, часові межі та аналітичні висновки, що використовуються для виявлення, класифікації, зіставлення й пояснення діяльності у кіберпросторі. Саме така сукупність дає змогу не лише виявити окрему подію, а й пов'язати її з іншими подіями, кампаніями та відомими кластерами активності.

Потреба в такому підході підтверджується сучасною європейською практикою. CERT-EU визначає власний Cyber Threat Intelligence Framework [7] як спільну еталонну модель для послідовної класифікації, оцінювання та пріоритизації кіберзагроз і наголошує, що перетворення спостережень та результатів аналізу на структуровані дані посилює ситуаційну обізнаність, узгодженість дій і швидкість реагування. Структурованість у цьому разі є не технічною зручністю, а умовою аналітичної послідовності.

Подібну логіку закладено в міжнародних моделях подання інформації про кіберзагрози. STIX 2.1 [8] передбачає подання різних аспектів таких даних за допомогою об'єктів і описових зв'язків між ними — від спостережуваних технічних ознак та індикаторів до кампаній, кластерів кіберзагроз, програмних засобів, шаблонів атак і аналітичних тверджень. Вихідною одиницею роботи стає не ізольований показник, а формалізований набір сутностей і відношень, придатний для сприйняття аналітиком і машинної обробки.

У Synapse модель даних (Data Model) [9] і пов'язана з нею аналітична модель (Analytical Model) [10] призначені для узгодженого подання як самих даних, так і аналітичних тверджень. Предметний каркас Vertex Threat Intel Model [11] дає змогу ставити запитання безпосередньо до пов'язаних даних, не зводячи аналітичну роботу до читання та повторного переказу масивів текстових звітів.

Практика використання спеціалізованих платформ також свідчить, що робота з даними про кіберзагрози вийшла за межі простих таблиць або неформалізованих записів. CRITs [12] спирається на об'єкти верхнього рівня, похідні об'єкти-метадані та зв'язки між ними, а CybOX [13] задає значну частину словника спостережуваних технічних об'єктів. Така

організація забезпечує повторний пошук і аналіз пов'язаних ознак.

Слід водночас розрізняти моделювання та обмін. Платформа обміну може забезпечувати формування, маркування й поширення підготовленого набору даних про подію або кіберзагрозу, однак сама наявність механізму обміну не гарантує, що внутрішні дані організації подані як зв'язана й семантично узгоджена модель. Зокрема, MISP [15] має важливе операційне призначення для обміну подіями та індикаторами, однак у межах цього дослідження не розглядається як рівнозначний предмет аналізу поряд із моделями, орієнтованими на детальне відтворення сутностей і зв'язків.

Отже, структуроване моделювання даних про кіберзагрози слід розглядати як необхідну умову не лише для їх накопичення і зберігання, а насамперед для зіставлення, повторного використання, перевірки, обміну та формування аналітичних висновків. Подальший виклад зосереджено на підходах, що забезпечують формалізований опис сутностей, їхніх властивостей, зв'язків та аналітичного контексту.

Методика дослідження

Об'єктом дослідження є способи структурованого подання даних про кіберзагрози, а предметом — семантичні й структурні можливості об'єктів, властивостей та зв'язків, які використовуються для відтворення спостережуваних фактів, подій, контексту й аналітичних висновків. Методологічну основу становили структурно-функціональний, порівняльний та історико-логічний методи, а також моделювання предметної області на основі офіційної документації та практичного досвіду автора.

Для аналізу обрано CRITs у поєднанні з CybOX 2.1.1, STIX 2.1 у статусі OASIS Standard від 10 червня 2021 року, а також актуальні станом на 5 вересня 2026 року документи Synapse 2.251.0 — Data Model і Analytical Model — разом із керівництвом Vertex Threat Intel Model. Метою дослідження не є вичерпний огляд усіх наявних платформ і моделей подання даних про кіберзагрози. Обрані підходи дають змогу, з одного боку, простежити розвиток формалізованого об'єктного подання від CybOX та його практичного використання у CRITs до STIX 2.1, а з іншого — розглянути відмінний гіперграфовий підхід Synapse, у якому модель даних доповнено аналітичною моделлю.

Одиницею аналізу визначено предметну сутність або смислове відношення, необхідне для відтворення конкретного фрагмента даних про кіберзагрозу. Під час зіставлення розрізнялися: 1) пряме штатне подання окремим об'єктом або зв'язком; 2) подання властивістю іншого об'єкта; 3) похідне значення, яке обчислюється або виводиться з наявних даних; 4) платформно-специфічне чи розширюване подання; 5) відсутність штатного відповідника. Такий поділ потрібний, оскільки однакова за змістом інформація в різних моделях може існувати на різних структурних рівнях.

Найближчим відповідником вважався такий об'єкт, набір властивостей або поєднання зв'язків, що дає змогу відтворити зміст вихідної сутності без зміни її ролі в контексті модельованої події. Коректність зіставлення оцінювалася шляхом повторного звернення до канонічних специфікацій і документації, контролю внутрішньої узгодженості та послідовної перевірки можливості подання наскрізного фрагмента моделі — від первинної події й спостережуваних технічних об'єктів через події виконання, мережеві та інфраструктурні відомості до програмного засобу, кампанії й кластера кіберзагрози. Остаточна верифікація мала експертний характер і здійснювалася автором з урахуванням практичного досвіду накопичення та аналізу даних про кіберзагрози; тому результати слід розглядати як методично обґрунтоване змістове зіставлення, а не як автоматичну схему взаємної конвертації.

Еволюція платформ і моделей подання даних про кіберзагрози

Для коректного розгляду засобів роботи з даними про кіберзагрози необхідно розрізняти платформу як прикладний інструмент накопичення, опрацювання та аналізу відомостей і модель подання даних як формалізований спосіб опису сутностей, їхніх властивостей та зв'язків. У межах цієї роботи CRITs і Synapse розглядаються як платформи, тоді як CybOX, STIX 2.1 та модель даних Synapse — як формалізовані мови або моделі подання даних.

На відміну від платформи, яка може поєднувати сховище, інтерфейс, механізми пошуку, автоматизації та обміну, модель подання даних визначає, які саме об'єкти можуть бути подані, які атрибути вони містять, якими способами вони пов'язуються між собою та як у формалізованому вигляді виражаються аналітичні твердження. Саме тому для задач технічної атрибуції вирішальне значення має не лише наявність платформи, а й те, наскільки її модель даних придатна до подання інфраструктури, програмних засобів реалізації кіберзагроз, тактик, технік, процедур, кампаній і кластерів кіберзагроз з можливістю осмисленого відображення відповідних зв'язків між об'єктами.

3.1. CRITs і CybOX

CRITs (Collaborative Research Into Threats) [12] є відкритою платформою для накопичення та спільного аналізу відомостей про загрози і програмні засоби реалізації кіберзагроз. Її архітектура побудована навколо об'єктів верхнього рівня (top-level objects, TLO), до яких належать, зокрема, Actors, Campaigns, Domains, Emails, Events, Indicators, IPs, Samples і Targets. Документно орієнтоване зберігання реалізоване з використанням MongoDB та MongoEngine.

У CRITs важливим є розрізнення між об'єктом верхнього рівня (top-level object, TLO) та похідним

об'єктом-метаданими (object). Перший виступає самостійною аналітичною сутністю верхнього рівня, тоді як другий описується як фрагмент метаданих, отриманий під час аналізу відповідного об'єкта верхнього рівня. Офіційна документація наводить типові приклади таких похідних об'єктів: IP-адреса, знайдена в результаті аналізу бінарного файлу; адреси електронної пошти, виявлені в листі; шлях до файлу, створеного на хості; URL, до якого здійснювалося звернення. Такий підхід добре відповідає практиці технічного аналізу, коли первинна сутність доповнюється похідними технічними ознаками.

Формально модель CRITs спирається на CybOX. Документація платформи передбачає додавання до об'єкта верхнього рівня об'єктів CybOX і використання типів зв'язків CybOX, але CRITs одночасно має власні словники подій, індикаторів, об'єктів і зв'язків. Через це канонічна назва елемента CybOX та його фактичне відображення у вебінтерфейсі CRITs можуть не збігатися буквально.

CybOX 2.1.1 [13] є стандартизованою мовою для кодування й передавання деталізованих відомостей про спостережувані технічні об'єкти та події у кіберпросторі. Об'єкти описуються як набори схематично визначених властивостей, а взаємозв'язки можуть виражатися через Related Objects; окремо підтримуються Actions і Events як конструкції для подання динамічних змін та операцій над об'єктами.

Таким чином, CRITs і CybOX утворюють зв'язану пару: CRITs забезпечує прикладну платформу для роботи аналітика, тоді як CybOX задає значну частину базового словника об'єктів і зв'язків. Для задач технічної атрибуції цей підхід є цінним насамперед тим, що дозволяє прямо пов'язувати похідні технічні ознаки з подіями, кампаніями, зразками програмних засобів реалізації кіберзагроз та кластерами активності. Водночас обмеженням є історичний характер цієї моделі та часткова залежність від власних словників платформи, через що відповідність між канонічним стандартом і фактичним відображенням у CRITs не завжди є буквальною.

3.2. STIX 2.1

STIX 2.1 є стандартизованою мовою та форматом серіалізації даних для обміну інформацією про кіберзагрози. У специфікації OASIS [8] його визначено як графову мову, що забезпечує гнучке, модульне, структуроване та узгоджене подання. STIX не прив'язаний до однієї платформи, а задає спільну мову, якою можуть обмінюватися різні системи й організації.

Для цієї роботи принципово важливо, що STIX 2.1 уже містить розвинену типологію об'єктів різного рівня узагальнення. У ньому розрізняються основні об'єкти STIX (STIX Core Objects), до яких належать об'єкти предметної області STIX (STIX Domain Objects, SDO), об'єкти спостережуваних технічних ознак STIX (STIX Cyber-observable Objects, SCO) та об'єкти зв'язків STIX (STIX Relationship Objects, SRO). Крім того, STIX 2.1 передбачає метаоб'єкти (STIX Meta Objects), зокрема об'єкти визначення розширень, мовного вмісту та маркування даних. Саме через таку побудову STIX 2.1 дозволяє в одному просторі даних подавати як технічні спостереження, так і кампанії, кластери кіберзагроз, суб'єктів кіберзагроз, програмні засоби реалізації кіберзагроз, шаблони атак та інші сутності, необхідні для аналізу і технічної атрибуції.

Узагальнену структуру основних типів об'єктів STIX 2.1 наведено на рис. 1. Об'єкти предметної області (STIX Domain Objects, SDO) описують аналітичні сутності, зокрема кампанії, кластери кіберзагроз, суб'єктів кіберзагроз, програмні засоби, індикатори, звіти та шаблони атак. Об'єкти спостережуваних технічних ознак (STIX Cyber-observable Objects, SCO) подають файли, процеси, доменні імена, IP-адреси, URL, електронні листи, мережевий трафік або ключі реєстру Windows. Об'єкти зв'язків (STIX Relationship Objects, SRO) формалізують відношення між SDO, між SCO, а також між SDO і SCO; Sighting фіксує факт спостереження певного STIX-об'єкта.

Об'єкт Attack Pattern може пов'язуватися із зовнішніми таксономіями шаблонів атак, зокрема CAPEC [14], через зовнішні посилання. Це дає змогу поєднувати формальну модель STIX із детальнішими каталогами способів здійснення атак.

Окрему увагу слід звернути на історичний зв'язок STIX і CybOX. Офіційна документація OASIS прямо зазначає, що у 2016 році було ухвалено рішення об'єднати STIX і CybOX в один стандарт, унаслідок чого об'єкти CybOX у межах STIX 2.0 були відображені як спостережувані технічні об'єкти STIX. З цієї причини порівняння CRITs/CybOX зі STIX 2.1 не є зіставленням двох цілком незалежних підходів; навпаки, STIX 2.1 є подальшим розвитком і узагальненням того напряму формалізації, який раніше реалізовувався окремо через CybOX.

Для аналітичної роботи важливо, що STIX 2.1 вже на рівні моделі розводить поняття threat-actor, intrusion-set, campaign, malware, attack-pattern, indicator та observed-data. Це дає змогу не зводити кластер кіберзагрози, суб'єкта, кампанію, програмний засіб і технічне спостереження до однієї універсальної категорії. Водночас стандарт орієнтований насамперед на переносимість і обмін, тому частина подієвих або платформно-специфічних деталей потребує поєднання кількох об'єктів чи розширень.

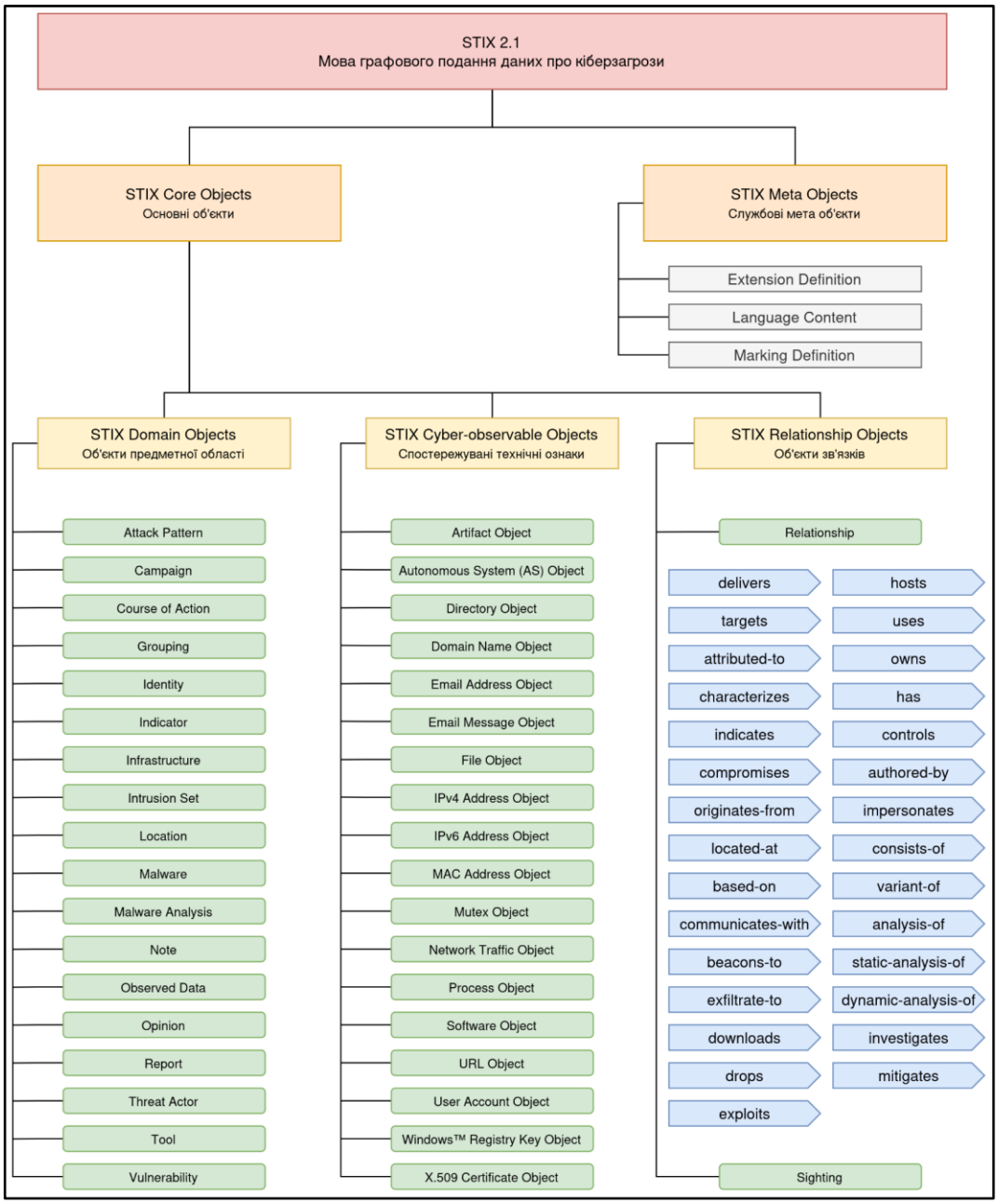


Рис. 1. Узагальнена структура основних типів об'єктів STIX 2.1

3.3. Synapse

Synapse є платформою та аналітичним середовищем, побудованим як граф знань. Її граф спирається на розширювану модель даних (Data Model) [9] і пов'язану з нею аналітичну модель (Analytical Model) [10], яка призначена для структурованого подання не лише даних, а й аналітичних тверджень та оцінок. Сховище Cortex реалізує гіперграфову модель [18].

На відміну від STIX 2.1, Synapse не організовує предметну область у вигляді закритого переліку класів об'єктів. Базовими конструкціями його моделі даних є тип (type), форма (form) і вузол (node). Тип визначає допустимий вигляд і правила нормалізації значення; форма є шаблоном об'єкта з унікальною первинною властивістю; вузол — конкретним екземпляром форми. Форми можуть містити вторинні, універсальні та розширені властивості, а інтерфейси задають спільні набори вторинних властивостей для груп форм. Вузли можуть представляти об'єкти або сутності, зв'язки та події. Узагальнену структуру моделі даних і пов'язаної з нею аналітичної моделі Synapse наведено на рис. 2.

Гіперграфова організація Synapse полягає в тому, що всі об'єкти й більшість зв'язків подаються як вузли, а переходи між ними переважно здійснюються за спільними типізованими значеннями властивостей. Окремі зв'язки та події можуть самі бути вузлами. Для простих спрямованих відношень використовуються "легкі" ребра виду n1 -(verb)> n2, які не мають власних властивостей і тегів. Поєднання цих механізмів дає змогу переходити між технічними, подієвими та аналітичними сутностями без обмеження аналізу лише наперед визначеними парними зв'язками.

© Koval M. Structured Modeling of Cyber Threat Intelligence Data: Methodological Foundations and Practical Applicability. // Ukrainian Scientific Journal of Information Security, 2026, vol. 32, issue 2, pp. 51-59. © This is an open access article under the CC BY 4.0 International license.

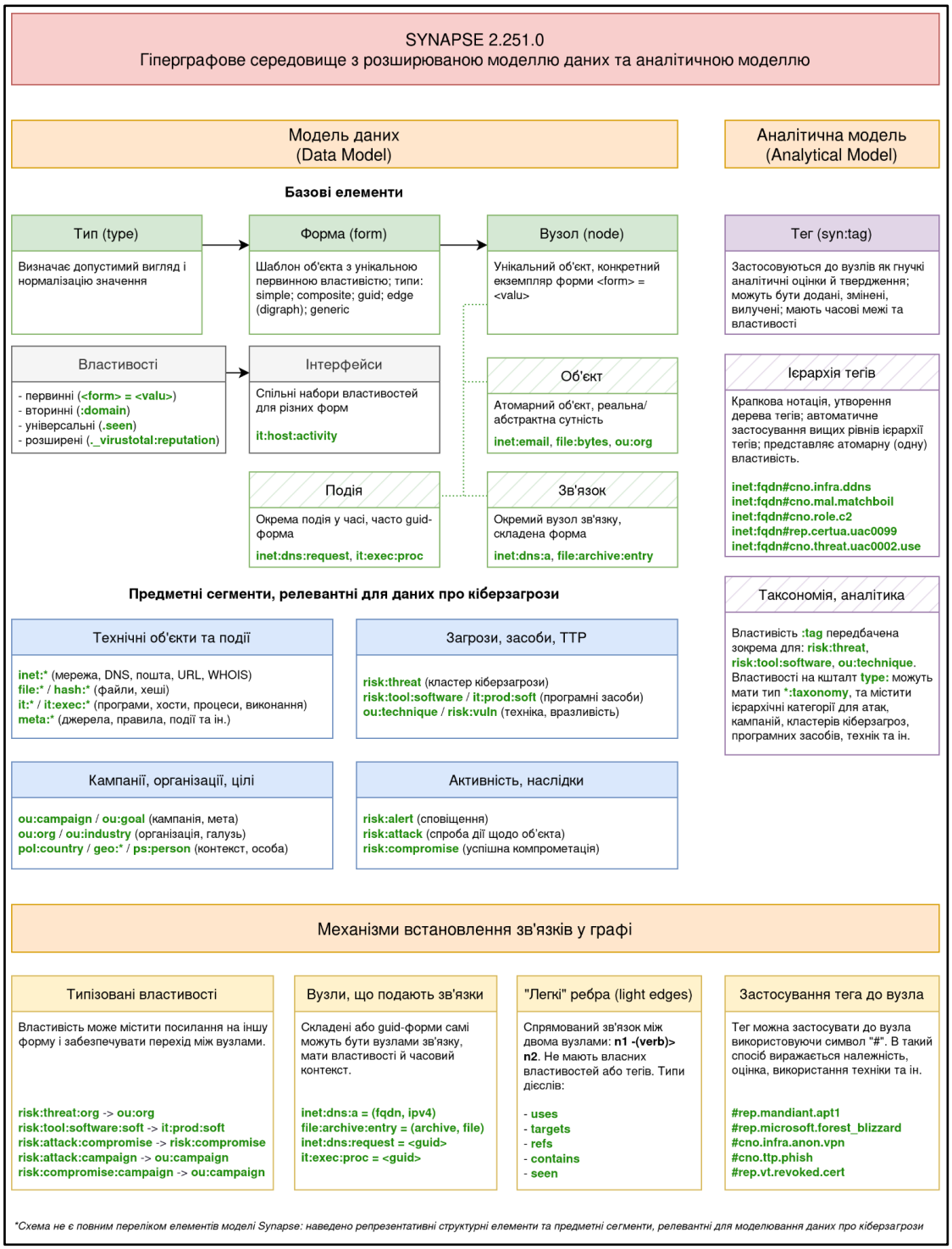


Рис. 2. Узагальнена структура моделі даних та аналітичної моделі Synapse (сформовано автором за [9]–[11], [18])

Аналітична модель доповнює модель даних тегами й таксономіями. Тег одночасно існує як вузол syn:tag і як позначка, застосована до іншого вузла, та може відображати змінювану аналітичну оцінку. Властивість :tag, передбачена, зокрема, для risk:threat, risk:tool:software та ou:technique, пов'язує відповідний тег із формалізованим описом кластера кіберзагрози, програмного засобу або техніки. Таксономії реалізуються як ієрархічні значення властивостей, зазвичай :type; наперед визначених тегів або таксономій Synapse не встановлює.

Для моделювання даних про кіберзагрози використовуються базові технічні форми сегментів inet:*, file:*, hash:* та it:*, а також предметні сутності Vertex Threat Intel Model [11], зосереджені передусім у сегментах risk:*, ou:* та it:*. Через них подаються сповіщення, атаки, факти компрометації, кампанії, кластери кіберзагроз, організації, програмні засоби, техніки, вразливості, цілі та об'єкти впливу. Особливістю цього підходу є можливість відокремити спостережувані факти від змінюваних аналітичних оцінок, водночас зберігаючи безпосередні зв'язки між ними.

Практична застосовність підходів до моделювання даних

Отже, CRITs, STIX 2.1 і Synapse не є взаємозамінними елементами одного рівня. CRITs і Synapse слід розглядати насамперед як платформи, хоча Synapse одночасно має і власну розвинену модель даних. CybOX і STIX 2.1 є, в першу чергу, моделями або мовами подання даних, при цьому STIX 2.1 становить наступний етап розвитку після CybOX. Для CRITs типовим є підхід, зорієнтований на об'єкти верхнього рівня (top-level objects, TLO), похідні об'єкти та словник типів зв'язків (relationship vocabulary); для STIX 2.1 — графове подання через стандартизовані типи об'єктів і об'єкти зв'язків (relationship objects); для Synapse — підхід графа знань і гіперграфа (knowledge graph / hypergraph) з поєднанням моделі даних (Data Model) та аналітичної моделі (Analytical Model).

Для переходу від описового огляду до узагальнення недостатньо зіставити назви окремих об'єктів, оскільки розглянуті платформи та моделі мають різне призначення, рівень деталізації й логіку організації даних. Напрями аналізу сформовано з урахуванням Діамантової моделі аналізу вторгнень [16], FIRST CSIRT Services Framework [17], предметного каркаса Synapse та виконаного автором детального зіставлення сутностей і зв'язків.

Виділені категорії не є послідовними етапами єдиного процесу й не передбачають однозначної відповідності між елементами різних моделей. Вони слугують спільними напрямами оцінювання того, як кожен підхід подає спостережувані технічні об'єкти, первинні повідомлення, прояви та наслідки діяльності, засоби її реалізації, об'єкти впливу, аналітичні узагальнення й атрибуційні твердження.

Таблиця 1

Узагальнена практична придатність розглянутих платформ і моделей

Предметна категорія	CRITs/CybOX	STIX 2.1	Synapse
Спостережувані технічні об'єкти та індикатори	TLO Indicator, IP, Domain, Sample; об'єкти CybOX і похідні метадані, пов'язані з TLO.	Indicator, Observed Data та SCO: File, IPv4 Address, Domain Name, URL, Email Message, Process, Registry Key тощо.	Вузли форм inet:*, file:*, hash:*, it:*; технічний об'єкт може безпосередньо пов'язуватися з тегами й іншими вузлами.
Первинні повідомлення, сповіщення та події	Event; окремої штатної сутності для сповіщення засобу захисту або первинного повідомлення немає.	Observed Data, Grouping або мінімальний Incident залежно від змісту; спеціалізованого alert-об'єкта немає.	risk:alert для первинного сповіщення; meta:event для подієвого контексту.
Атака, інцидент і факт компрометації	Переважно Event або Campaign; спроба атаки, інцидент і підтверджена компрометація чітко не розмежовані.	Incident є мінімальним об'єктом; перебіг і наслідки відтворюються поєднанням SDO, SCO, SRO або розширеннями.	Окремо розрізняються risk:attack і risk:compromise; підтримуються зв'язки зі сповіщеннями й кампаніями.
Кампанія	Campaign може бути аналітичною сутністю, контейнером кейса та одиницею експорту.	Окремий SDO Campaign із часовими межами, метою та визначеними зв'язками.	ou:campaign пов'язується з атаками, компрометаціями, техніками, цілями, кластерами й суб'єктами.
Інфраструктура, програмні засоби та вразливості	IP, Domain, Indicator, Backdoor, Exploit; значна частина контексту формується збагаченням і власними словниками.	Окремі SDO Infrastructure, Malware, Tool, Vulnerability та пов'язані SCO.	risk:tool:software, it:prod:soft, risk:vuln; розвинений WHOIS, ASN і мережевий контекст.
Тактики, техніки та процедури	Властивості, зв'язки, підписи й словники; окремий предметний шар TTP виражений обмежено.	Attack Pattern, kill-chain phases, зовнішні посилання на ATT&CK, CAPEC та інші таксономії.	ou:technique, ATT&CK-орієнтовані форми й теги; техніка може бути безпосередньо пов'язана з подією або об'єктом.
Об'єкти впливу та цілі діяльності	TLO Target; сектор і спрямованість можуть зберігатися у властивостях Campaign/Actor.	Identity, Location та relationship targets; залежно від типу цілі використовується відповідний SDO.	ou:org, ou:industry, pol:country, ps:person, ou:goal і відповідні властивості або зв'язки targets.
Кластер кіберзагрози та суб'єкт кіберзагрози	Actor може позначати як групу, так і умовно виділену активність; технічний кластер і реальний суб'єкт чітко не розведені.	Intrusion Set відокремлений від Threat Actor; їх асоціація подається окремим relationship attributed-to.	risk:threat відокремлений від ou:org; підтримуються власні таксономії, аліаси, зіставлення й перегляд оцінки.
Аналітичні твердження, походження та впевненість	Зв'язки, джерела, коментарі, статуси й словники; частина оцінок залишається в описових полях.	Note, Opinion, Report, confidence, created_by_ref, external references, markings, Relationship і Sighting.	Теги, властивості тегів, meta:source, часові межі, графові зв'язки; твердження може застосовуватися до конкретного вузла.

Проведений аналіз показує, що CRITs/CybOX, STIX 2.1 і Synapse вирішують різні, хоча й взаємопов'язані завдання. CRITs/CybOX насамперед забезпечує практичне накопичення об'єктів аналізу, індикаторів і похідних технічних ознак навколо об'єктів верхнього рівня; STIX 2.1 надає стандартизовану мову для подання та обміну відомостями між системами й організаціями; Synapse найпослідовніше поєднує у графовому середовищі фактичний, подієвий, контекстний та аналітичний шари, не ототожнюючи їх між собою.

Практичний вибір не повинен зводитися до пошуку одного «універсального» продукту. Для суб'єкта, який накопичує й аналізує дані про кіберзагрози, першочерговим є визначення власної понятійної моделі: які сутності вважаються самостійними, які зв'язки мають зберігатися, як фіксуються походження, часова чинність і впевненість, а також яким чином дані імпортуються,

експортуються та зіставляються з зовнішніми таксономіями. Платформа має обиратися або адаптуватися вже відносно цих вимог.

У межах підготовчого етапу дослідження виконано деталізоване зіставлення близько п'ятдесяти сутностей і тридцяти п'яти смислових зв'язків. Через обмеження обсягу в статті наведено узагальнені результати, тоді як саме зіставлення використано для виявлення типових розривів: відсутності штатного об'єкта, перенесення змісту у властивість, необхідності виведення значення або побудови ланцюга з кількох сутностей. Отже, практична сумісність моделей не є буквальною відповідністю назв.

Обмеження дослідження

Результати залежать від зафіксованих версій специфікацій і документації та не враховують усіх можливих платформних розширень. Частина відповідностей має експертний характер, а найближчий відповідник не завжди означає повну семантичну еквівалентність або автоматичну взаємну конвертацію. Дослідження також не оцінює продуктивність, вартість упровадження чи зручність інтерфейсів, оскільки його предметом є саме структура й практична виразність моделі даних.

Висновки

У роботі обгрунтовано, що структуроване моделювання даних про кіберзагрози є не окремою функцією сховища і не форматом переліку індикаторів, а способом збереження змісту спостережень, подій, контексту та аналітичних висновків разом зі зв'язками між ними. Без такої основи накопичені відомості втрачають повторну придатність, а їх зіставлення залежить від ручного відновлення контексту.

Простежено розвиток формалізованого об'єктного подання від CRITs/CybOX, де практична робота організована навколо об'єктів верхнього рівня й похідних ознак, до STIX 2.1 як стандартизованої графової мови подання та обміну. Synapse розглянуто як окремий розширюваний граф знань, у якому поєднуються фактичний, подієвий, контекстний та аналітичний шари.

Запропонована методика дає змогу оцінювати практичну придатність моделі за тим, чи може вона відтворити необхідну сутність або смисловий зв'язок безпосередньо, через властивість, похідне значення, платформно-специфічне чи розширюване подання, або ж не має штатного відповідника; за потреби зміст сутності чи зв'язку може відтворюватися поєднанням кількох елементів моделі. Такий спосіб аналізу зміщує увагу з формальної наявності назви об'єкта на збереження його ролі у конкретній події та в загальній аналітичній картині.

Практичне значення одержаних результатів полягає у формуванні орієнтирів для суб'єктів забезпечення кібербезпеки під час проєктування власних сховищ і процесів роботи з даними, вибору або адаптації платформи, побудови інтеграцій та машинозчитуваного обміну. Основним критерієм має бути перехід від фрагментарного обліку технічних показників до системного накопичення пов'язаних і перевірюваних даних.

Література

[1] Закон України «Про основні засади забезпечення кібербезпеки України» від 05.10.2017 № 2163-VIII [Електронний ресурс]. Режим доступу: https://zakon.rada.gov.ua/laws/show/2163-19.

[2] Про затвердження Порядку обміну інформацією про кіберінциденти, кібератаки, кіберзагрози та Критеріїв визначення значного кіберінциденту: наказ Адміністрації Держспецзв'язку від 14.05.2026 № 351, зареєстрований у Міністерстві юстиції України 22.06.2026 за № 899/46293 [Електронний ресурс]. Режим доступу: https://zakon.rada.gov.ua/laws/show/z0899-26.

[3] Mavroeidis V., Bromander S. Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence. 2017 European Intelligence and Security Informatics Conference (EISIC). 2017. P. 91–98. doi: 10.1109/EISIC.2017.20.

[4] Bromander S., Swimmer M., Eian M., Skjøtskift G., Borg F. Modeling Cyber Threat Intelligence. Proceedings of the 6th International Conference on Information Systems Security and Privacy (ICISSP 2020). 2020. P. 273–280. doi: 10.5220/0008875302730280.

[5] Jin B., Hicks C., Ruohonen J., Naderi E., Nita-Rotaru C. Sharing Cyber Threat Intelligence: Does It Really Help? Proceedings of the Network and Distributed System Security Symposium (NDSS). 2024 [Електронний ресурс]. Режим доступу: https://www.ndss-symposium.org/wp-content/uploads/2024-228-paper.pdf.

[6] Коваль М. В. Концепція активного кіберзахисту в контексті технологічних спроможностей протидії кіберзагрозам. Український науковий журнал інформаційної безпеки. 2025. Т. 31, № 1. С. 37–42. doi: 10.18372/2225-5036.31.20636.

[7] CERT-EU. Cyber Threat Intelligence Framework. 2026 [Електронний ресурс]. Режим доступу: https://www.cert.europa.eu/publications/threat-intelligence/cyber-threat-intelligence-framework/.

[8] OASIS. STIX Version 2.1. OASIS Standard. 10 June 2021 [Електронний ресурс]. Режим доступу: https://docs.oasis-open.org/cti/stix/v2.1/os/stix-v2.1-os.html.

[9] The Vertex Project. Synapse Data Model. Synapse 2.251.0 documentation [Електронний ресурс]. Режим доступу: https://synapse.docs.vertex.link/en/latest/synapse/userguides/data_model.html.

[10]The Vertex Project. Synapse Analytical Model. Synapse 2.251.0 documentation [Електронний ресурс]. Режим доступу: https://synapse.docs.vertex.link/en/stable/synapse/userguides/analytical_model.html.

[11]The Vertex Project. Vertex Threat Intel Model [Електронний ресурс]. Режим доступу: https://synapse.docs.vertex.link/projects/rapid-powerups/en/latest/storm-packages/vertex-threat-intel/ugmodel.html.

[12]CRITs. Collaborative Research Into Threats [Електронний ресурс]. Режим доступу: https://github.com/crits/crits/.

[13]OASIS. CybOX Version 2.1.1. Part 01: Overview. 20 June 2016 [Електронний ресурс]. Режим доступу: https://docs.oasis-open.org/cti/cybox/v2.1.1/cybox-v2.1.1-part01-overview.html.

[14]MITRE. CAPEC List Version 3.9 [Електронний ресурс]. Режим доступу: https://capec.mitre.org/data/index.html.

[15]MISP Project. MISP core format [Електронний ресурс]. Режим доступу: https://www.misp-standard.org/rfc/misp-standard-core.html.

[16]Caltagirone S., Pendergast A., Betz C. The Diamond Model of Intrusion Analysis. 2013 [Електронний ресурс]. Режим доступу: https://www.activeresponse.org/wp-content/uploads/2013/07/diamond.pdf. doi: 10.13140/RG.2.2.31143.56481.

[17]FIRST. Computer Security Incident Response Team (CSIRT) Services Framework. Version 2.1 [Електронний ресурс]. Режим доступу: https://www.first.org/standards/frameworks/csirts/csirt_services_framework_v2.1.

[18]The Vertex Project. Background: Graphs and Hypergraphs [Електронний ресурс]. Режим доступу: https://synapse.docs.vertex.link/en/latest/synapse/userguides/background.html#bkd-graphs-hypergraphs

УДК 004.056.5:004.6

Koval M. Structured Modeling of Cyber Threat Intelligence Data: Methodological Foundations and Practical Applicability

Abstract. *This article examines structured modeling of cyber threat intelligence data as a prerequisite for its accumulation, correlation, analysis, and further use. It argues that storing individual indicators in spreadsheets, ticketing systems, or standalone reports without formalized relationships to events, infrastructure, software tools used in malicious cyber activity, campaigns, and analytical conclusions limits the practical value of the data. A distinction is made between a platform as an operational tool for working with information and a data representation model as a formal description of entities, their properties, and relationships. Based on a structural and functional analysis of official documentation, the study traces the development of formalized object representation from CRITs/CybOX to STIX 2.1 and separately examines the hypergraph approach of Synapse. Common subject-matter categories are identified to assess how these approaches represent observable technical objects, initial reports, attacks and confirmed compromises, campaigns, infrastructure, software tools, tactics, techniques, procedures, targets, and analytical assertions. The approaches are shown not to be interchangeable: CRITs/CybOX focuses on the practical accumulation of objects and derived attributes, STIX 2.1 on standardized representation and exchange, and Synapse on graph-based integration of factual, event-related, contextual, and analytical layers. The practical contribution consists in providing methodological guidance for moving from fragmented storage of indicators to systematic work with interconnected cyber threat data.*

Keywords: *cyber threat data modeling; cyber threat data analysis; threat intelligence; cyber threat indicators; structured data; knowledge graph; STIX 2.1; CybOX; Synapse; CRITs.*

КОВАЛЬ Микола Васильович, фахівець Державного центру кіберзахисту Державної служби спеціального зв'язку та захисту інформації України.

Koval Mykola, specialist at State Cyber Protection Center, State Service of Special Communication and Information Protection of Ukraine.

Дата надходження:	Дата прийняття:	Дата публікації:
02.07.2026	31.07.2026	21.08.2026