КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.32.1197
УДК [004.738.5:351]621.396.6

Шелест Михайло Євгенович
доктор технічних наук, професор, професор кафедри кібербезпеки та математичного моделювання
місце роботи: Національний університет «Чернігівська політехніка», Чернігів, Україна
ORCID: 0000-0001-7110-4876
mishel3141@gmail.com

Ткач Юлія Миколаївна
доктор педагогічних наук, кандилат технічних наук, професор, завідувач кафедри кібербезпеки та математичного моделювання
місце роботи: Національний університет «Чернігівська політехніка», Чернігів, Україна
ORCID ID: 0000-0002-8565-0525
tkachym79@gmail.com

КЛЕПТОРИЗИК ЯК ОКРЕМИЙ КЛАС РИЗИКУ ЦИФРОВОЇ ДОВІРИ

Анотація. У цій статті представлено клепторизик як окремий клас ризиків цифрової довіри, що виникають через навмисно вбудовані, контрольовані слабкі місця в архітектурі криптографічних та інформаційних систем. На відміну від традиційних ризиків кібербезпеки, які виникають через недоліки впровадження або операційні вразливості, клепторизик виникає на етапі проектування та існує незалежно від його активації. У статті формалізовано концепцію клепторизику, визначено його ключові властивості та відрізнено його від традиційних категорій ризиків. Запропоновано компактну модель життєвого циклу, яка описує формування, легітимізацію, приховане існування та потенційну активацію клепторизику. Історичні тематичні дослідження, включаючи криптографічні пристрої Crypto AG та генератор випадкових чисел Dual_EC_DRBG, демонструють, що такі ризики можуть існувати в рамках формально сумісних та широко розгорнутих систем. Результати дослідження показують, що клепторизик є архітектурною характеристикою, а не операційною подією. У цій роботі обґрунтовується перехід від інцидентно-орієнтованої кібербезпеки до архітектурно-орієнтованої парадигми аналізу довіри та окреслено наслідки для управління довірою, криптографічної гарантії та розробки систем безпеки, що враховують клепторизики.

Ключові слова: клепторизик; клептографія; цифрова довіра; криптографічні бекдори; архітектура довіри; ризики кібербезпеки, інформаційна безпека.

ВСТУП

Сучасна цифрова екосистема функціонує на основі фундаментального припущення про коректність, доброчесність та відсутність прихованих механізмів контролю в криптографічних системах. Довіра до алгоритмів, протоколів і стандартів є базовою передумовою функціонування електронного урядування, фінансових систем, хмарних платформ, оборонної інфраструктури та міжнародних цифрових комунікацій. Саме криптографічні механізми формують технологічну основу цифрової довіри, яка визначає можливість безпечної взаємодії суб’єктів у цифровому середовищі.

Класичні моделі управління ризиками, зокрема ISO 31000, NIST Risk Management Framework та COSO ERM, розглядають ризик як імовірнісну подію, що виникає внаслідок помилки, збою або реалізації зовнішньої загрози. У сфері кібербезпеки ця логіка трансформується у модель «вразливість – загроза – інцидент», у межах якої ризик виникає як наслідок експлуатації технічного дефекту або недосконалості системи.

Водночас сучасна практика інформаційної безпеки демонструє випадки навмисного проєктування криптографічних або протокольних слабкостей, які не є результатом випадкових помилок, а інтегруються на етапі архітектурного дизайну системи [13]. Такі слабкості можуть залишатися латентними протягом тривалого часу, формально відповідати стандартам сертифікації, активуватися лише за наявності специфічного знання та створювати структурну асиметрію доступу до інформації між суб'єктами.

У цьому контексті класична ризикова парадигма виявляється недостатньою, оскільки вона не враховує можливості навмисного формування архітектурних механізмів контролю. Це обумовлює необхідність концептуального виокремлення окремої категорії ризиків (ми назвали їх клепторизиками), пов'язаних не з помилками або інцидентами експлуатації, а з навмисно сформованими властивостями архітектури системи.

У цій роботі запропоновано концептуальну модель клепторизику як окремої категорії ризиків цифрових систем.

Наукова новизна роботи полягає у формалізації поняття клепторизику як окремої категорії ризиків цифрових систем, що має архітектурну природу і не зводиться до експлуатаційних вразливостей або програмних дефектів. У роботі вперше:

- обґрунтовано трактування клепторизику як структурної характеристики системи, що формується на етапі проектування і може існувати незалежно від факту експлуатації;
- запропоновано концептуальну модель життєвого циклу клепторизику, яка включає етапи формування, легітимізації, латентного існування та активації;
- показано принципову відмінність клепторизиків від класичних моделей кіберризиків, що базуються на подієвій та ймовірнісній природі загроз;
- закладено теоретичну основу для подальшої розробки методів клептоаудиту та архітектурного аналізу довіри.

Постановка проблеми. У сучасних цифрових системах ризики традиційно трактуються як наслідок помилок, збоїв або експлуатаційних вразливостей, що виникають у процесі функціонування інформаційно-комунікаційної інфраструктури. Класичні підходи до управління ризиками ґрунтуються на імовірнісній моделі, у межах якої ризик розглядається як результат невизначеності або реалізації загрози щодо певної вразливості. Відповідно, управління ризиками орієнтується на виявлення дефектів, їх оцінювання та реагування після впровадження системи. Однак такі підходи виходять із припущення про ненавмисну природу технічних недоліків і не враховують можливості закладення керованих слабкостей на етапі архітектурного проєктування.

Проблема полягає в тому, що навмисно спроєктовані криптографічні або протокольні слабкості не є класичними вразливостями й не можуть бути коректно інтерпретовані як інциденти інформаційної безпеки. Вони характеризуються стратегічним наміром, асиметрією знання між суб'єктами, прихованістю на етапі сертифікації та потенційною неможливістю виявлення стандартними методами аудиту. Така природа ризику виходить за межі традиційної парадигми risk management і вимагає концептуального виокремлення окремого класу ризиків цифрової довіри – клепторизиків, що мають архітектурний, системний та інституційний характер.

Аналіз останніх досліджень і публікацій. Проблематика навмисно закладених криптографічних та програмних слабкостей активно обговорюється в сучасних дослідженнях інформаційної безпеки. У науковій літературі бекдори розглядаються як механізми прихованого доступу до інформаційних систем, які можуть бути інтегровані на рівні програмного коду, апаратного забезпечення або криптографічних параметрів[1].

У цих дослідженнях наголошується, що навіть легітимізоване впровадження механізмів спеціального доступу створює системну вразливість, оскільки ослаблення криптографічного захисту не може бути обмежене лише контрольованим використанням.

Питання довіри до цифрових технологій також розглядається в роботах, присвячених цифровому урядуванню та інформаційній стійкості держав. У міжнародних аналітичних документах цифрова довіра визначається як фундаментальна умова функціонування електронних сервісів, хмарної інфраструктури та транскордонного обміну даними [8]. Підкреслюється, що довіра формується не лише через технічні характеристики системи, а й через прозорість архітектури, незалежність аудиту та відсутність прихованих механізмів контролю.

Водночас у межах класичних моделей управління ризиками ризик трактується як результат невизначеності або як функція ймовірності реалізації загрози та масштабу її впливу [3]. Подібна логіка відображена і в підходах NIST Risk Management Framework, де оцінювання ризику ґрунтується на аналізі загроз, вразливостей і потенційного впливу на систему [5].

Таким чином, у сучасних дослідженнях можна виокремити три домінуючі напрями: аналіз технічних механізмів бекдорів, вивчення феномену цифрової довіри та розроблення методологій управління ризиками. Проте ці напрями існують переважно ізольовано. Технічні роботи розглядають бекдори як окремі вразливості або інструменти спеціального доступу; дослідження цифрової довіри зосереджуються на інституційних і соціальних аспектах; класичні моделі risk management базуються на імовірнісній природі ризику та не враховують навмисний характер архітектурних слабкостей.

Відсутність інтегрованого підходу до оцінювання навмисно спроєктованих слабкостей як системного чинника підриву цифрової довіри створює концептуальну прогалину в науковому дискурсі. Саме ця концептуальна прогалина обумовлює необхідність формування окремої категорії ризиків, що мають архітектурний, навмисний та інституційний характер. У межах цього дослідження запропоновано концептуалізацію такого класу ризиків – клепторизиків – як системного феномену цифрової довіри, що не зводиться до традиційних кіберінцидентів або експлуатаційних вразливостей.

Мета статті. Теоретичне обґрунтування та концептуалізація клепторизику як окремого класу ризику цифрової довіри шляхом аналізу обмеженості класичних моделей управління ризиками щодо навмисно спроєктованих архітектурних слабкостей, визначення його сутнісних ознак та місця у сучасній парадигмі забезпечення цифрової безпеки.

Наукова новизна дослідження. Наукова новизна роботи полягає в такому:

– вперше обґрунтовано доцільність виокремлення клепторизику як окремого класу ризиків цифрової довіри;
– запропоновано трирівневу модель клепторизику (архітектурний, функціональний, інституційний рівні);
– розроблено модель життєвого циклу клепторизику;
– здійснено емпіричну верифікацію концепції на прикладах криптографічних систем;
– обґрунтовано необхідність переходу до парадигми Kleptorisk Management.

МЕТОДИКА ДОСЛІДЖЕННЯ

Методологічною основою дослідження є концептуальний та системний аналіз архітектурних характеристик криптографічних і інформаційно-комунікаційних систем з позицій цифрової довіри. На відміну від класичних підходів, орієнтованих на виявлення експлуатаційних вразливостей, у роботі акцент зроблено на дослідженні властивостей систем, сформованих на етапі їх архітектурного проєктування.

У процесі дослідження застосовано метод системного аналізу для виявлення структурних механізмів формування керованих слабкостей, порівняльний аналіз для розмежування клепторизику та традиційних категорій ризику, а також кейс-аналіз історичних прикладів криптографічних систем, що містили ознаки навмисно сформованої асиметрії контролю. Використання кейс-аналізу дозволило дослідити клепторизик як реальний феномен, а не лише теоретичну конструкцію.

На основі узагальнення результатів застосованих методів запропоновано концептуальну модель життєвого циклу клепторизику, яка описує його формування, легітимацію, латентне існування та потенційну активацію. Такий методологічний підхід забезпечує можливість розгляду клепторизику як окремого класу архітектурного ризику цифрової довіри, що має архітектурний і системний характер.

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Теоретичні основи дослідження базуються на розгляді цифрової довіри як архітектурної властивості криптографічних та інформаційно-комунікаційних систем. У цьому контексті довіра визначається не лише як впевненість у математичній стійкості алгоритмів або відсутності експлуатаційних вразливостей, а як системна характеристика, що включає відсутність прихованих механізмів контролю, асиметрії доступу або керованих слабкостей, сформованих на етапі проєктування.

На відміну від класичних підходів, у межах яких ризик розглядається як наслідок помилки або випадкової вразливості, у цьому дослідженні розглядається можливість існування ризиків, що мають навмисний і архітектурний характер. Такі ризики не є результатом дефекту реалізації, а формуються як властивість архітектури системи та можуть залишатися латентними незалежно від факту їх експлуатації.

У межах запропонованого підходу вводиться поняття клепторизику.

Клепторизик визначається як системний архітектурний ризик цифрової довіри, що виникає внаслідок навмисного закладення керованої слабкості або асиметрії контролю в архітектуру криптографічної або інформаційної системи та створює потенційну можливість прихованого доступу до захищеної інформації.

Принциповою відмінністю клепторизику від традиційного ризику є його доінцидентна природа. Якщо класичний кіберризик виникає як результат експлуатації вразливості, то клепторизик існує незалежно від факту його використання, оскільки є властивістю самої архітектури системи.

Клепторизик характеризується такими ключовими властивостями:

- навмисний характер формування;
- латентність та відсутність явних проявів у процесі експлуатації;
- асиметрія контролю між суб’єктами;
- можливість формальної відповідності стандартам безпеки;
- інституційний характер потенційних наслідків.

Таким чином, клепторизик є не подією або інцидентом, а структурною характеристикою системи, яка визначає архітектурний рівень цифрової довіри.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Сучасна парадигма управління ризиками в інформаційній безпеці базується на імовірнісній природі загроз. У межах стандарту ISO 31000 ризик визначається як ефект невизначеності щодо досягнення цілей, а в рамках NIST Risk Management Framework – як функція загрози, вразливості та потенційного впливу. Таким чином, класичні підходи виходять із припущення, що ризик виникає внаслідок зовнішнього впливу або експлуатації ненавмисної вразливості.

Результати проведеного дослідження показали, що в цифрових системах існує інший тип ризику, який не може бути коректно описаний у межах цієї моделі. Йдеться про навмисно спроєктовані слабкості, закладені на етапі архітектурного проєктування системи.

У дослідженнях криптографічних бекдорів підкреслюється, що впровадження механізмів спеціального доступу створює системну вразливість, навіть якщо такі механізми формально обґрунтовуються потребами безпеки. Проте в існуючих підходах такі явища розглядаються як окремі технічні або політичні випадки, а не як системний клас ризику.

У межах проведеного дослідження обґрунтовано доцільність трактування клепторизику як окремого класу ризиків цифрової довіри, що виникає внаслідок навмисного закладення керованої слабкості в архітектуру криптографічної, протокольної або інфраструктурної системи.

На відміну від традиційної вразливості, клепторизик не є наслідком помилки чи недосконалості реалізації. Він формується в момент прийняття архітектурного рішення та має навмисний характер.

Теоретична конструкція клепторизику, сформована в межах проведеного дослідження, базується на трирівневому підході, який дозволяє розкрити його системний характер.

Перший – архітектурний рівень. Результати аналізу показали, що саме на цьому рівні формується потенціал асиметричного контролю через закладення механізму доступу або ослаблення криптографічної стійкості. Таким чином, клепторизик виникає не як наслідок експлуатації системи, а як результат прийняття архітектурного рішення.

Другий – функціональний рівень. На цьому етапі система може проходити формальні процедури сертифікації та відповідати технічним стандартам, при цьому зберігаючи латентну керовану слабкість. Це підтверджує, що формальна відповідність стандартам безпеки не гарантує відсутності архітектурного ризику.

Третій – інституційний рівень. На цьому рівні проявляється вплив клепторизику на політику довіри, цифровий суверенітет та залежність користувачів від суб’єкта контролю. Результати дослідження показали, що саме інституційний рівень визначає стратегічний характер наслідків клепторизику.

Таким чином, отримані результати дозволили встановити, що клепторизик є багаторівневим феноменом, який формується на архітектурному рівні, легітимується на функціональному рівні та реалізує свої стратегічні наслідки на інституційному рівні.

У межах проведеного дослідження сформовано понятійний апарат, необхідний для опису клепторизику як окремого класу ризиків цифрової довіри.

Ключовим результатом стало уточнення поняття керованої слабкості. На відміну від класичної вразливості, яка виникає ненавмисно як наслідок помилки або обмеження реалізації, керована слабкість інтегрується свідомо як елемент архітектурного рішення та передбачає потенційну можливість доступу до інформації або контролю над функціонуванням системи для суб'єкта, який володіє відповідним знанням або ресурсом.

Результати аналізу показали, що наявність керованої слабкості створює архітектурну асиметрію довіри, за якої різні суб'єкти перебувають у нерівних умовах щодо можливості контролю або інтерпретації криптографічних процесів. При цьому один із суб'єктів має приховану перевагу, тоді як інші продовжують функціонувати в межах припущення про повну безпечність системи [13].

У цьому контексті клепторизик визначено як системний ризик цифрової довіри, що виникає внаслідок наявності керованої слабкості та пов'язаної з нею архітектурної асиметрії контролю. Такий ризик існує незалежно від факту його реалізації та є властивістю самої архітектури системи.

Важливим результатом дослідження стало встановлення того, що клепторизик має доексплуатаційний характер. Він формується на етапі проєктування системи і може залишатися латентним протягом тривалого часу, не проявляючись у вигляді інцидентів або порушень функціонування.

Це дозволило сформулювати узагальнююче положення, відповідно до якого клепторизик є не подією, а структурною характеристикою системи, що відображає наявність асиметрії контролю в її архітектурі.

На основі отриманих результатів також обґрунтовано доцільність використання поняття управління клепторизиком (Kleptorisk Management) як окремого напряму управління ризиками цифрової довіри. На відміну від традиційного підходу, який орієнтований на виявлення та усунення вразливостей, управління клепторизиком спрямоване на аналіз архітектурних рішень, оцінювання асиметрій контролю та забезпечення прозорості механізмів функціонування системи.

Практичну верифікацію запропонованої концепції клепторизику здійснено на основі аналізу відомого випадку криптографічних систем компанії Crypto AG.

Результати проведеного дослідження показали, що криптографічні пристрої цієї компанії, які використовувалися державними структурами багатьох країн світу, містили механізми, що створювали потенційну можливість асиметричного доступу до захищеної інформації. При цьому відповідні системи функціонували коректно, забезпечували криптографічний захист відповідно до заявлених характеристик і використовувалися як довірені засоби захисту інформації.

Принципово важливим результатом аналізу є встановлення того, що потенційна можливість контролю була інтегрована на рівні архітектурного проєктування системи. Таким чином, ризик існував як властивість архітектури незалежно від факту його реалізації.

Це дозволило підтвердити ключове положення дослідження, відповідно до якого клепторизик формується на етапі створення системи і не є наслідком технічної помилки або випадкової вразливості.

Отримані результати також показали, що користувачі систем Crypto AG перебували в умовах асиметрії довіри. Вони використовували криптографічні засоби, які вважалися безпечними, не маючи можливості встановити наявність керованої слабкості або оцінити пов'язаний з нею ризик.

Таким чином, аналіз даного випадку підтверджує, що клепторизик може існувати в системах, які формально відповідають стандартам безпеки і функціонують у штатному

режимі. Це узгоджується із запропонованою в межах дослідження моделлю клепторизику як архітектурної характеристики цифрової системи.

Подальшу емпіричну верифікацію запропонованої концепції клепторизику здійснено на основі аналізу генератора псевдовипадкових чисел Dual_EC_DRBG.

Результати проведеного дослідження показали, що архітектурні особливості цього генератора передбачали потенційну можливість прогнозування його вихідних значень за умови володіння спеціальними параметрами. При цьому генератор був стандартизований і розглядався як легітимний криптографічний механізм.

Принципово важливим результатом аналізу є встановлення того, що потенційна можливість асиметричного контролю була обумовлена не помилкою реалізації, а особливостями архітектурної конструкції алгоритму. Це означає, що відповідний ризик існував на рівні математичної моделі генератора.

Отримані результати підтвердили, що система могла відповідати формальним вимогам стандарту і при цьому містити латентний механізм, який створював асиметрію контролю між суб'єктами.

Цей випадок підтверджує ключове положення дослідження, відповідно до якого клепторизик може бути інтегрований безпосередньо в алгоритмічну конструкцію криптографічного механізму і не залежить від помилок програмної реалізації або конфігурації.

Таким чином, аналіз генератора Dual_EC_DRBG підтверджує, що клепторизик може виникати на рівні математичної архітектури криптографічного алгоритму, що розширює запропоновану модель клепторизику і підтверджує її універсальний характер.

У межах проведеного дослідження розроблено модель життєвого циклу клепторизику, яка дозволяє описати процес його формування, існування та потенційної реалізації (рис. 1).

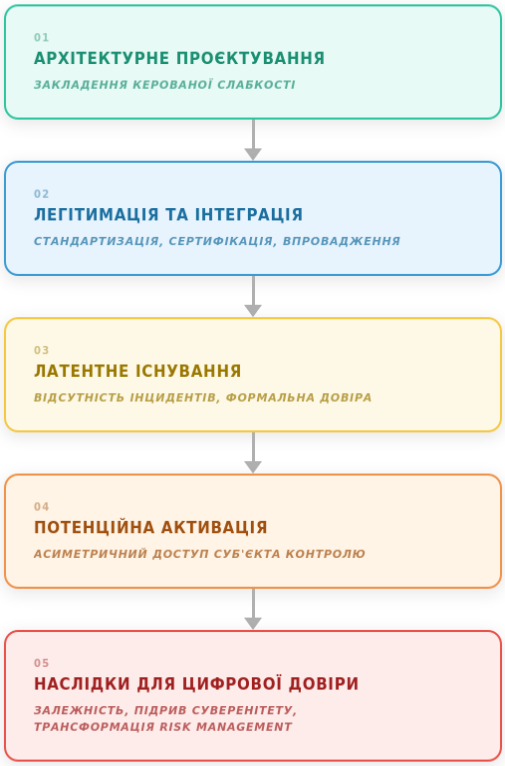

Рис. 1. Модель життєвого циклу клепторизику

Результати аналізу показали, що на відміну від традиційних кіберризиків, які виникають у процесі експлуатації системи, клепторизик формується на етапі її архітектурного проєктування. Саме на цьому етапі приймається рішення про інтеграцію механізму, що створює потенційну можливість асиметричного контролю.

Наступним етапом є легітимація, під час якої система впроваджується в експлуатацію, проходить процедури стандартизації, сертифікації або прийняття користувачами. Отримані результати показали, що саме на цьому етапі відбувається інституційне закріплення довіри до системи, незалежно від наявності в її архітектурі керованої слабкості.

Після цього клепторизик переходить у фазу латентного існування, яка може тривати протягом значного часу. У цей період система функціонує у штатному режимі, а закладений механізм контролю не проявляється у вигляді інцидентів або збоїв. Результати дослідження показали, що латентність є однією з ключових властивостей клепторизику, яка ускладнює його виявлення традиційними методами аналізу безпеки.

Завершальним етапом є потенційна активація, яка може відбутися за наявності відповідного суб'єкта контролю та необхідних умов. При цьому важливою характеристикою клепторизику є те, що його існування не залежить від факту активації. Ризик існує як властивість архітектури системи незалежно від того, чи був відповідний механізм використаний.

Таким чином, отримані результати дозволили встановити, що клепторизик має повний життєвий цикл, який починається на етапі проєктування системи і може завершуватися його потенційною реалізацією. Це підтверджує, що клепторизик є не подією, а процесом, інтегрованим у життєвий цикл цифрової системи.

Запропонована модель життєвого циклу узгоджується з результатами аналізу розглянутих криптографічних систем і підтверджує її придатність для опису механізмів формування клепторизику.

Отримані результати дозволили встановити, що клепторизик має принципово іншу природу порівняно з традиційним кіберризиком. Для систематизації виявлених відмінностей у межах дослідження проведено порівняльний аналіз цих двох типів ризиків, результати якого наведено в табл. 1.

Проведений аналіз показав, що традиційний кіберризик виникає як наслідок експлуатації вразливості та має подієвий характер. Його існування пов'язане з імовірністю реалізації загрози та проявляється у вигляді інциденту.

На відміну від цього, клепторизик формується на етапі архітектурного проєктування системи і є її структурною характеристикою. Він існує незалежно від факту реалізації та не потребує експлуатації вразливості для свого існування.

Результати дослідження також показали, що джерело традиційного кіберризику має, як правило, зовнішній характер, тоді як джерело клепторизику інтегроване в саму архітектуру системи.

Порівняльний аналіз підтвердив, що клепторизик не може бути коректно описаний у межах класичної моделі «загроза – вразливість – вплив», оскільки він виникає не внаслідок порушення функціонування системи, а як результат її архітектурної організації.

Таким чином, результати, наведені в табл. 1, підтверджують доцільність розгляду клепторизику як окремого класу ризиків цифрової довіри.

Таблиця 1

Порівняльна характеристика традиційного кіберризику та клепторизику

Критерій	Традиційний кіберризик	Клепторизик
Об'єкт аналізу	Подія або інцидент	Архітектурна конфігурація системи
Джерело ризику	Зовнішня загроза або експлуатація вразливості	Навмисно закладена архітектурна слабкість
Природа виникнення	Ненавмисна помилка або недосконаліть реалізації	Свідоме проєктне рішення
Час виникнення	Під час експлуатації системи	На етапі проєктування
Виявлення	Через аудит, тестування, моніторинг інцидентів	Ускладнене або неможливе стандартними методами
Імовірність	Оцінюється статистично	Не може бути коректно описаний у межах класичних імовірнісних моделей
Інцидентність	Реалізується у вигляді події	Може існувати без реалізації
Суб'єкт контролю	Зовнішній атакуючий	Суб'єкт проєктування або контролю
Наслідки	Порушення конфіденційності, цілісності, доступності	Асиметрія довіри, стратегічна залежність
Управління	Реактивне (Incident Response)	Превентивне (архітектурний аналіз і контроль проєктних рішень)
Рівень впливу	Технічний або операційний	Технічний, інституційний, стратегічний

Порівняльний аналіз, наведений у табл. 1, демонструє, що клепторизик не може бути зведений до категорії кіберінциденту. Результати дослідження показали, що він не є подією, а є структурною характеристикою системи. Відповідно, застосування класичних методів оцінювання ризику, зокрема моделі «ймовірність × вплив», є методологічно некоректним для його опису.

Отримані результати підтвердили, що клепторизик трансформує сам об'єкт ризикового аналізу: замість події аналізується архітектурна конфігурація системи; замість статистичної частоти – структура контролю; замість реактивного реагування – архітектурна профілактика на етапі проєктування.

У межах запропонованої концепції особливого значення набуває інституційний вимір клепторизику. Результати дослідження показали, що наявність архітектурної асиметрії контролю в критичній цифровій інфраструктурі формує довгострокову структурну залежність, навіть за відсутності її активного використання.

Це безпосередньо пов'язано з поняттям цифрового суверенітету, який передбачає можливість держави або організації самостійно контролювати ключові параметри цифрової інфраструктури, зокрема:

- криптографічні механізми;
- процедури сертифікації;
- алгоритмічні параметри;
- політику оновлень і доступу.

Результати проведеного аналізу показали, що у випадку наявності клепторизику цифровий суверенітет набуває умовного характеру, оскільки формальна автономія може поєднуватися зі структурною залежністю від суб'єкта, який володіє знанням або контролем над архітектурною слабкістю.

Інституційні наслідки клепторизику проявляються в кількох взаємопов'язаних площинах.

По-перше, це політика довіри. Довіра до цифрових систем базується на припущенні про відсутність прихованих механізмів контролю. Результати дослідження показали, що

наявність потенційної керованої слабкості трансформує довіру в асиметричну категорію, яка може бути об'єктом прихованого контролю.

По-друге, це технологічна залежність. Архітектурна асиметрія створює довгострокову залежність від постачальника технології або суб'єкта, який контролює архітектурні параметри системи. При цьому навіть відсутність активного використання відповідного механізму не усуває пов'язаного з ним ризику.

По-третє, це геополітичний вимір. Отримані результати показали, що контроль над криптографічними стандартами, алгоритмічними параметрами та інфраструктурними рішеннями може виступати інструментом стратегічного впливу. У цьому контексті клепторизик стає фактором технологічної політики.

По-четверте, це політика стандартизації та сертифікації. Проведений аналіз підтвердив, що формальна відповідність стандартам не гарантує відсутності клепторизику. Це означає, що процедури стандартизації повинні включати аналіз архітектурної прозорості та можливості незалежної криптографічної верифікації.

Таким чином, результати дослідження підтвердили, що клепторизик має не лише технічний, а й системно-інституційний характер. Його ігнорування призводить до підміни поняття безпеки формальною відповідністю технічним вимогам.

У межах запропонованої концепції цифровий суверенітет набуває нового змісту – як здатність забезпечити відсутність архітектурної асиметрії контролю. Це передбачає не лише технічну незалежність, а й прозорість архітектурних рішень, контроль над ключовими параметрами системи та можливість незалежної криптографічної верифікації.

Отже, результати проведеного дослідження підтвердили, що клепторизик є самостійною категорією ризиків цифрової довіри, яка має відмінну природу виникнення, багаторівневу структуру та власний життєвий цикл. Запропонований понятійний апарат, трирівнева модель, модель життєвого циклу та порівняльний аналіз із традиційним кіберризиком забезпечують теоретичну та методологічну основу для подальшого розвитку підходів до управління клепторизиками. Емпірична верифікація на прикладах Crypto AG та Dual_EC_DRBG демонструє, що клепторизик може існувати в системах, які формально відповідають стандартам і функціонують у штатному режимі, що обмежує застосовність класичних інструментів оцінювання ризику. Таким чином, отримані результати обґрунтовують необхідність переходу від інцидентно-орієнтованого аналізу безпеки до архітектурно-орієнтованої парадигми оцінювання довіри. Отримані результати формують основу для практичного застосування концепції клепторизику та визначають напрями подальших досліджень.

ПРАКТИЧНЕ ЗНАЧЕННЯ

Практичне значення отриманих результатів полягає у можливості використання запропонованої концепції клепторизику як аналітичного інструменту при проектуванні, аудиті та експертизі цифрових систем.

Запропонована модель дозволяє:

- ідентифікувати потенційні клептографічні механізми на ранніх стадіях життєвого циклу системи;
- аналізувати архітектурні рішення з позицій асиметрії контролю та прихованого впливу;
- формувати вимоги до прозорості, верифікованості та незалежного контролю;

- використовувати концепцію клепторизику у процедурах сертифікації та оцінки довіри.

Отримані результати можуть бути застосовані у сфері криптографічного захисту інформації, безпеки критичної інфраструктури, аналізу довіри до цифрових платформ та дослідження безпеки систем штучного інтелекту.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Проведене дослідження засвідчило методологічну недостатність класичних моделей управління ризиками для опису явищ, пов'язаних із навмисно закладеними архітектурними слабкостями. Традиційна парадигма risk management, що базується на імовірнісній природі загроз та інцидент-орієнтованій логіці, не охоплює ситуацій, у яких джерело ризику інтегроване в саму структуру системи та має навмисний характер.

У межах статті обґрунтовано, що клепторизик принципово відрізняється від традиційного кіберризику за такими ознаками: момент виникнення (етап проєктування), природа формування (архітектурне рішення), механізм прояву (латентність), суб'єкт контролю (інституційний або проєктний), а також стратегічний масштаб наслідків. На відміну від вразливості, що є результатом помилки або недосконалості реалізації, керована слабкість має свідомо інтегрований характер і створює асиметрію довіри між суб'єктами.

Доведено, що клепторизик не є подією або інцидентом, а становить структурну характеристику цифрової системи. Він може існувати без фактичної реалізації, що унеможливлює його адекватну оцінку через традиційні показники частоти чи ймовірності. Таким чином, застосування класичної моделі «загроза – вразливість – вплив» до аналізу клепторизику є концептуально обмеженим.

Сформульована трирівнева конструкція клепторизику (архітектурний, функціональний та інституційний рівні) дозволила інтегрувати технічний, процедурний та стратегічний виміри аналізу. На архітектурному рівні формується потенціал асиметричного контролю; на функціональному — забезпечується його латентність за умови формальної відповідності стандартам; на інституційному — виникає залежність, що впливає на політику довіри та цифровий суверенітет.

Порівняльний аналіз традиційного ризику та клепторизику продемонстрував, що останній має системний характер і не зводиться до технічної вразливості. Це обґрунтовує необхідність трансформації підходів до управління ризиками в напрямі Kleptorisk Management, що орієнтується на превентивний аналіз архітектурних рішень і структури контролю ще на етапі проєктування системи.

Отже, результати дослідження підтверджують доцільність концептуального виокремлення клепторизику як окремого класу ризиків цифрової довіри та формують теоретичне підґрунтя для розвитку методології його оцінювання й інтеграції у практику забезпечення цифрової безпеки.

Подальші дослідження можуть бути спрямовані на розроблення методів кількісного оцінювання клепторизиків, зокрема формування критеріїв вимірювання архітектурної асиметрії та рівня прозорості криптографічних механізмів. Перспективним є також інтегрування підходів до аналізу клепторизиків у процедури сертифікації та стандартизації цифрових технологій.

Окремого розвитку потребує дослідження механізмів мінімізації клепторизиків у критичній інфраструктурі, а також формування політики цифрової довіри, орієнтованої на забезпечення технологічної автономії та криптографічного суверенітету.

Таким чином, концепція клепторизику формує новий напрям досліджень у межах теорії цифрової безпеки, орієнтований на аналіз архітектурних основ довіри до криптографічних, програмних та інфраструктурних компонентів сучасних цифрових систем. Це обґрунтовує необхідність переходу від інцидент-орієнтованої до архітектурно-орієнтованої моделі аналізу цифрової довіри.

У цьому контексті клепторизик виступає не лише новою категорією ризику, але й новим об’єктом управління, що вимагає формування спеціалізованих підходів до його ідентифікації, оцінювання та контролю. Це відкриває перспективу становлення окремого напряму — Kleptorisk Management, орієнтованого на забезпечення архітектурної прозорості, криптографічного суверенітету та доказової довіри до цифрових систем.

Отримані результати можуть бути використані при проведенні клептоаудиту цифрових систем. Це відкриває перспективу формування стандартизованих підходів до оцінювання клепторизиків у цифрових системах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Abelson, H., Anderson, R., Bellovin, S., Benaloh, J., Blaze, M., Diffie, W., Gilmore, J., Green, M., Landau, S., Neumann, P., Rivest, R., & Schiller, J. (2015). Keys under doormats: Mandating insecurity by requiring government access to all data and communications. *Journal of Cybersecurity, 1*(1), 69–79. https://doi.org/10.1093/cybsec/tyv009
2. Anderson, R. (2020). *Security engineering: A guide to building dependable distributed systems* (3rd ed.). Wiley.
3. International Organization for Standardization. (2018). *ISO 31000: Risk management—Guidelines.* ISO.
4. Miller, G. (2020, February 11). The intelligence coup of the century: For decades, the CIA read the encrypted communications of allies and adversaries. *The Washington Post.*
5. National Institute of Standards and Technology. (2018). *Risk management framework for information systems and organizations: A system life cycle approach for security and privacy* (NIST SP 800-37 Rev. 2). https://doi.org/10.6028/NIST.SP.800-37r2
6. National Institute of Standards and Technology. (2012). *Recommendation for random number generation using deterministic random bit generators* (NIST SP 800-90A).
7. National Institute of Standards and Technology. (2014). *Dual EC in X9.82 and SP 800-90.*
8. Organisation for Economic Co-operation and Development. (2014). *Building digital government strategies: Principles and practices.* OECD Publishing. https://doi.org/10.1787/9789264223639-en
9. Schneier, B. (2015). *Applied cryptography: Protocols, algorithms, and source code in C* (20th anniversary ed.). Wiley.
10. Shumow, D., & Ferguson, N. (2007). On the possibility of a back door in the NIST SP800-90 Dual EC PRNG. In *Advances in cryptology—CRYPTO 2007* (Rump session).
11. Swissinfo.ch. (2020, November 10). *Swiss intelligence benefited from CIA-Crypto spying affair.*
12. Young, A., & Yung, M. (1997). Kleptography: Using cryptography against cryptography. In *Advances in cryptology—EUROCRYPT ’97* (Lecture Notes in Computer Science, Vol. 1233, pp. 62–74). Springer.
13. Tkach, Y. M., & Shelest, M. Y. (2025). *Kleptography: From backdoor to trust policy in the digital age.* Chernihiv Polytechnic National University.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (32), 2026
ISSN 2663 – 4023

Mykhailo Shelest
Doctor of Technical Sciences, Professor
Chernihiv Polytechnic National University, Chernihiv, Ukraine
ORCID: 0000-0001-7110-4876
mishel3141@gmail.com

Yuliia Tkach
Doctor of Pedagogical Sciences, PhD Technical Sciences, Professor
Chernihiv Polytechnic National University, Chernihiv, Ukraine
ORCID: 0000-0002-8565-0525
tkachym79@gmail.com

KLEPTORISK AS A DISTINCT CLASS OF DIGITAL TRUST RISK

Abstract. This paper introduces kleptorisk as a distinct class of digital trust risk arising from intentionally embedded, controllable weaknesses within the architecture of cryptographic and information systems. Unlike conventional cybersecurity risks, which emerge from implementation flaws or operational vulnerabilities, kleptorisk originates at the design stage and persists independently of its activation. The paper formalizes the concept of kleptorisk, defines its key properties, and distinguishes it from traditional risk categories. A compact lifecycle model is proposed, describing kleptorisk formation, legitimization, latent existence, and potential activation. Historical case studies, including Crypto AG cryptographic devices and the Dual_EC_DRBG random number generator, demonstrate that such risks can exist within formally compliant and widely deployed systems. The findings indicate that kleptorisk represents an architectural characteristic rather than an operational event. This work argues for a shift from incident-centric cybersecurity toward an architecture-centric trust analysis paradigm and outlines implications for trust management, cryptographic assurance, and the development of kleptorisk-aware security frameworks.

Keywords: kleptorisk; kleptography; digital trust; cryptographic backdoors; trust architecture; cybersecurity risk.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Abelson, H., Anderson, R., Bellovin, S., Benaloh, J., Blaze, M., Diffie, W., Gilmore, J., Green, M., Landau, S., Neumann, P., Rivest, R., & Schiller, J. (2015). Keys under doormats: Mandating insecurity by requiring government access to all data and communications. *Journal of Cybersecurity*, *1*(1), 69–79. https://doi.org/10.1093/cybsec/tyv009
2. Anderson, R. (2020). *Security engineering: A guide to building dependable distributed systems* (3rd ed.). Wiley.
3. International Organization for Standardization. (2018). *ISO 31000: Risk management—Guidelines.* ISO.
4. Miller, G. (2020, February 11). The intelligence coup of the century: For decades, the CIA read the encrypted communications of allies and adversaries. *The Washington Post.*
5. National Institute of Standards and Technology. (2018). *Risk management framework for information systems and organizations: A system life cycle approach for security and privacy* (NIST SP 800-37 Rev. 2). https://doi.org/10.6028/NIST.SP.800-37r2
6. National Institute of Standards and Technology. (2012). *Recommendation for random number generation using deterministic random bit generators* (NIST SP 800-90A).
7. National Institute of Standards and Technology. (2014). *Dual EC in X9.82 and SP 800-90.*
8. Organisation for Economic Co-operation and Development. (2014). *Building digital government strategies: Principles and practices.* OECD Publishing. https://doi.org/10.1787/9789264223639-en
9. Schneier, B. (2015). *Applied cryptography: Protocols, algorithms, and source code in C* (20th anniversary ed.). Wiley.
10. Shumow, D., & Ferguson, N. (2007). On the possibility of a back door in the NIST SP800-90 Dual EC PRNG. In *Advances in cryptology—CRYPTO 2007* (Rump session).
11. Swissinfo.ch. (2020, November 10). *Swiss intelligence benefited from CIA-Crypto spying affair.*

12. Young, A., & Yung, M. (1997). Kleptography: Using cryptography against cryptography. In *Advances in cryptology—EUROCRYPT '97* (Lecture Notes in Computer Science, Vol. 1233, pp. 62–74). Springer.
13. Tkach, Y. M., & Shelest, M. Y. (2025). *Kleptography: From backdoor to trust policy in the digital age.* Chernihiv Polytechnic National University.

Отримано редакцією журналу / Received: 22.01.26
Прорецензовано / Revised: 15.02.26
Схвалено до друку / Accepted: 26.03.26