

КОВАЛЬЧУК ОЛЬГА

Західноукраїнський національний університет

<https://orcid.org/0000-0001-6490-9633>e-mail: olhakov@gmail.com**БАБАЛА ЛЮДМИЛА**

Західноукраїнський національний університет

<https://orcid.org/0000-0002-2388-270X>e-mail: ludaduma7@gmail.com**ІВАНИЦЬКИЙ РОМАН**

Тернопільський національний педагогічний

університет імені Володимира Гнатюка

<https://orcid.org/0009-0004-1687-9539>e-mail: romikiv@ukr.net

ІНТЕЛЕКТУАЛЬНА МОДЕЛЬ ВИЯВЛЕННЯ АСОЦІАТИВНИХ ПРАВИЛ У БАЗАХ ДАНИХ КРИМІНАЛЬНИХ ПРАВОПОРУШЕНЬ

Стаття присвячена виявленню закономірностей та тенденції скоєння кримінальних правопорушень за допомогою аналізу асоціативних правил. Емпіричною базою дослідження є інформація про 1975 кримінальних правопорушень, скоєних у Тернопільській області у період з 2012 по 2023 роки. Аналіз проведено на основі даних про крадіжки, пограбування та незаконне заволодіння транспортними засобами. Враховано наступні фактори: локація та час скоєння кримінальних правопорушень, освітлення, участь кількох осіб. У середовищі для статистичної обробки даних методами інтелектуального аналізу RapidMiner Studio застосовано алгоритм FP-Growth. Згенеровано асоціативні правила, які виявляють неочевидні взаємозв'язки та закономірності у даних про скоєні кримінальні правопорушення.

Ключові слова: інтелектуальний аналіз даних, асоціативні правила, аналіз даних про кримінальні правопорушення, машинне навчання, прогнозування правопорушень, прийняття рішень.

KOVALCHUK OLHA

West Ukrainian National University

BABALA LUDMILA

West Ukrainian National University

IVANYTSKYI ROMAN

Ternopil Volodymyr Hnatiuk National Pedagogical University

AN INTELLIGENT MODEL FOR DETECTING ASSOCIATION RULES IN CRIMINAL DATABASES

This study explores patterns in criminal activities through association rule mining analysis of criminal offenses in Ukraine's Ternopil region from 2012 to 2023. Using the FP-Growth algorithm and RapidMiner Studio, the research reveals significant patterns in criminal behavior by examining various factors including crime type, lighting conditions, and group dynamics. The analysis uncovered several key patterns: groups frequently commit theft regardless of lighting conditions, with Ternopil as a primary location for these crimes. Group criminal activities show peak occurrence in April and during the latter part of the week, particularly on Thursdays and Fridays. A notable finding is the close association between theft and illegal vehicle appropriation, with vehicle theft commonly executed by groups under dark conditions.

This study's comprehensive approach incorporates multiple factors to provide a more complete understanding of criminal activity patterns. The discovered association rules offer valuable insights for law enforcement agencies, enabling them to develop targeted prevention strategies and optimize resource allocation. The findings particularly highlight the connection between theft and illegal vehicle appropriation, with the latter typically occurring during nighttime hours, suggesting the need for increased patrol presence in streets and parking areas during these times. The study acknowledges that crime patterns are often specific to particular localities, indicating the need for additional research across different regions to develop a more comprehensive understanding of criminal behavior and its influencing factors. The research demonstrates that data mining techniques, specifically association rule mining, can effectively uncover non-obvious yet meaningful patterns in crime data. These insights can inform proactive policing efforts, ultimately contributing to enhanced crime prevention and safer communities. As data-driven approaches continue to evolve, their integration into law enforcement practices will become increasingly crucial in the effective combat against criminal activities.

Keywords: data mining, association rules, crime analysis, machine learning, prediction of offenses, decision-making.

Постановка проблеми

У сучасному суспільстві, що характеризується стрімким розвитком технологій, аналіз злочинності набуває нового виміру завдяки використанню передових методів обробки даних. Всебічний аналіз інформації про скоєні злочини, зокрема просторового розподілу, часових закономірностей та пов'язаних з фактом скоєння злочину факторів, створює основу для розробки прогностичних моделей оцінювання ризику скоєння правопорушень у майбутньому [1]. Такі моделі використовують алгоритми машинного навчання, статистичні методи та підходи до інтелектуального аналізу даних для виявлення прихованих взаємозв'язків та закономірностей у базах даних про кримінальні правопорушення [2].

Наприклад, певні райони чи вулиці можуть демонструвати вищі показники крадіжок або пограбувань у певний час доби, на що впливають такі фактори, як щільність населення, соціально-економічні умови чи наявність потенційних цілей для скоєння злочину [3]. Картографування цих

закономірностей та розуміння їх основних факторів дозволяє правоохоронним органам ефективніше розподіляти ресурси, посилювати патрулювання в зонах підвищеного ризику та впроваджувати цільові стратегії запобігання.

Застосування прогностичної аналітики в поліцейській діяльності не лише покращує здатність прогнозувати потенційну злочинну діяльність, але й уможлиблює прийняття рішень на основі даних та проактивних заходів для запобігання злочинам до їх скоєння [4]. Випереджаючи кримінальні елементи на крок, правоохоронні органи можуть визначати пріоритети своїх зусиль, оптимізувати розподіл ресурсів та сприяти створенню безпечнішого середовища для громад [5].

Таким чином, дослідження закономірностей злочинної діяльності та розробка ефективних методів їх виявлення є актуальним науково-практичним завданням, вирішення якого сприятиме підвищенню ефективності роботи правоохоронних органів та забезпеченню громадської безпеки.

Аналіз останніх джерел

Останні дослідження у сфері кримінології зосереджені на комплексному вивченні чинників та умов, що впливають на динаміку злочинності в конкретних регіонах та громадах. Особлива увага приділяється дослідженню методів запобігання злочинам на локальному рівні, враховуючи специфіку кримінального середовища окремих міст та районів [6]. А. Лісовська-Керепка (A. Lisowska-Kierepka) представила комплексний огляд існуючих методологій, які використовують для аналізу просторових явищ на основі доступних джерел даних. Автор представила рішення, адаптоване для аналізу даних про злочинність у Польщі [7]. Х. Кім та Е. Сон (H. Kim, E. Seong) провели масштабний аналіз злочинності у Південній Кореї. Дослідники розробили просторову економетричну модель для вивчення зв'язку між забудованим середовищем та трьома видами злочинів (крадіжки, насильство та сексуальні напади). Отримані результати підтвердили обмежену ефективність традиційних методів запобігання злочинності та виявили тісний зв'язок між соціально-економічними факторами та рівнем злочинності в країні [8].

М. А. Хайдер та П. Іамтракул (M. A. Haider, P. Iamtr) розробили модель логістичної регресії для аналізу віктимізації вуличної злочинності, а також модель аналізу гарячих точок з використанням контурних карт та карти просторової автокореляції. Автори визначили вразливі місця в житлових та ділових районах міста Читтагонг, Бангладеш, де мешканці часто стають жертвами вуличних злочинів [9]. Дослідники Д. Лю (D. Liu) та співавтори застосували байєсівську просторово-часову модель для аналізу закономірностей злочинності в місті Чанчунь (Китай), що забезпечило можливість виявити значущий вплив щільності доріг, комерційних, рекреаційних, житлових та промислових зон на рівень злочинності, а також визначити райони з підвищеним ризиком злочинності, які потребують додаткової уваги правоохоронних органів [10]. П. Кадена-Урзуа (P. Cadena-Urzuá) та співавтори використали просторову панельну модель з фіксованими ефектами для дослідження впливу соціально-економічних та демографічних факторів на злочини, що відбуваються в комунах (муніципалітетах) Столичного регіону Чилі. Дослідники ідентифікували просторові та часові ефекти, які можуть виявити закономірності злочинності [11].

Незважаючи на наявність низки наукових досліджень з питань виявлення неочевидних патернів у даних про злочинність, існує гостра потреба у проведенні подальших розвідок. Такі дослідження можуть виявити суттєві територіальні відмінності у географічних, демографічних, соціально-економічних та культурних особливостях різних регіонів, що безпосередньо впливають на характер, структуру та динаміку злочинності та потребують розробки специфічних підходів до її попередження та протидії скоєнню правопорушень. Для України ця сфера є новою та потребує особливої уваги.

Метою роботи є розробка інтелектуальної асоціативної моделі для виявлення прихованих закономірностей та тенденцій скоєння кримінальних правопорушень (крадіжок, пограбувань та незаконного заволодіння транспортними засобами) у Тернопільській області. Дослідження спрямоване на формування аналітичного підґрунтя для прийняття правоохоронними органами обґрунтованих рішень щодо ефективного розподілу ресурсів, планування превентивних заходів, оптимізації маршрутів патрулювання та посилення охорони найбільш вразливих локацій у найбільш ризиковані періоди часу, що сприятиме підвищенню ефективності запобігання та протидії злочинності в регіоні.

Виклад основного матеріалу

У нашій роботі представлено нову методологію виявлення шаблонів та взаємозв'язків у великих масивах даних про злочини. Ми застосували аналіз асоціативних правил до набору даних про різні види злочинів (крадіжки, пограбування, незаконне заволодіння транспортним засобом), який містить інформацію про скоєні кримінальні правопорушення у Тернопільській області з 2012 по 2023 рік. Запропонована методологія відповідає робочому процесу, представленому на рис. 1. Спочатку дані проходять етапи попередньої обробки, після чого застосовується алгоритм пошуку асоціативних правил, що генерує набір асоціативних правил [12].

Для виявлення неочевидних, але важливих закономірностей серед багатьох різноманітних факторів, які враховуються судами при винесенні кримінальних вироків, ми використали RapidMiner Studio – візуальний конструктор робочих процесів, оснащений інструментами для прогнозу аналітики, та машинного навчання [13]. На рис. 2 представлено оператори процесу, що реалізує алгоритми пошуку асоціативних правил.

За допомогою алгоритму пошуку асоціативних правил згенеровано часті набори елементів та асоціативні правила. Побудована модель асоціативних правил спрощує розуміння таких закономірностей щодо факторів, що характеризують злочинну діяльність:

- Крадіжки зазвичай вчиняються групою осіб незалежно від освітлення місця скоєння злочину.
- Більшість крадіжок було скоєно в Тернополі.
- Найчастіше групові злочини вчиняються у другій половині тижня (четвер, п'ятниця та субота).
- Крадіжка часто вчиняється одночасно з незаконним заволодінням транспортним засобом.
- Злочинні угруповання активізують свою діяльність у квітні.
- Незаконне заволодіння транспортним засобом в основному вчиняється групою осіб в темний час доби.

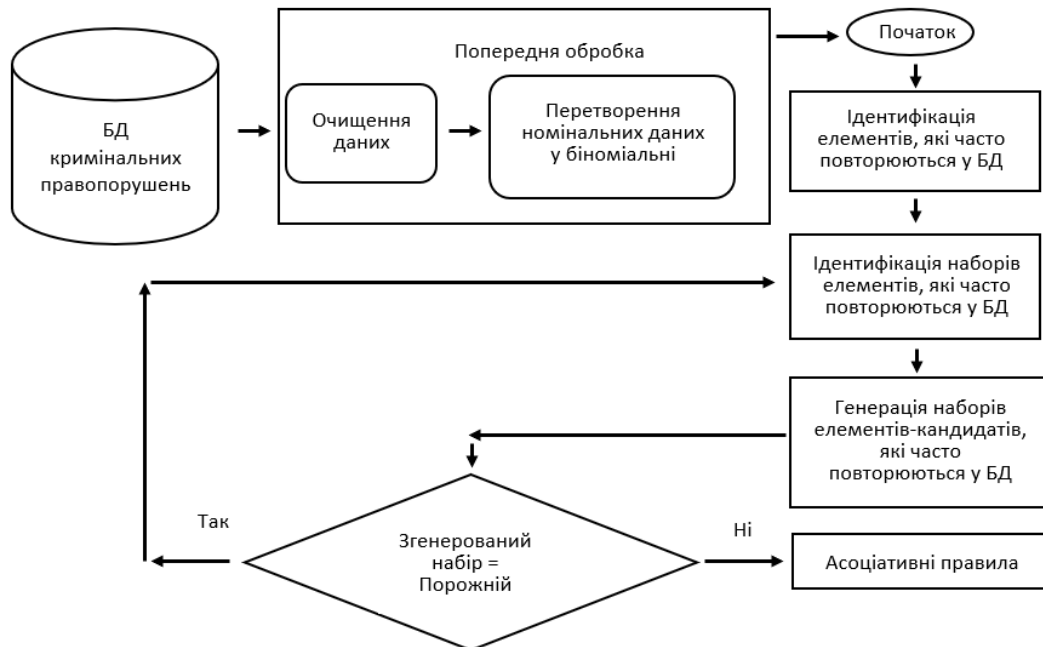


Рис. 1. Схема робочого процесу для запропонованого підходу

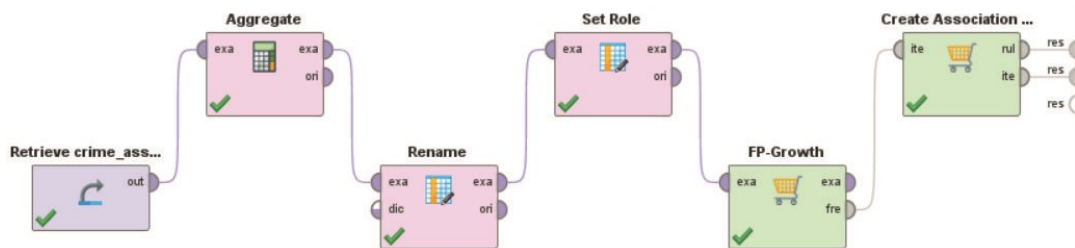


Рис. 2. Процес виявлення асоціативних правил

Рис. 3 ілюструє правила асоціацій, отримані в результаті аналізу бази даних кримінальних правопорушень, зокрема зображує правила, пов'язані зі злочинами, скоєними групою осіб.

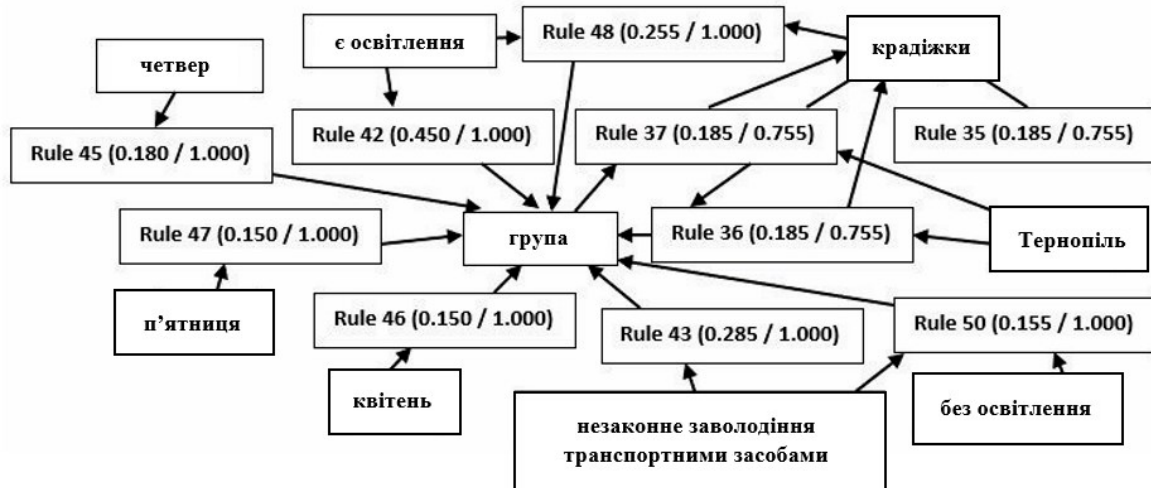


Рис. 3. Мережева діаграма правил асоціацій для злочинів, скоєних групою осіб

Отримані результати підтвердили оцінки інших авторів: більшість злочинів відбувається в густонаселених регіонах (містах). Визначення сприятливого часу для злочинів та зосередження на конкретних умовах є ключовими застосуваннями. Зокрема, результати показали тісний зв'язок між крадіжками та незаконним заволодінням транспортними засобами, причому останнє часто здійснюється групами в темний час доби. Тому поліція може очікувати збільшення випадків незаконного заволодіння транспортними засобами вночі та запобігати їм, збільшуючи патрулювання вулиць та паркувальних майданчиків у ці години. Згідно з виявленими правилами, певні злочини, такі як незаконне заволодіння транспортними засобами, частіше вчиняються групами в темний час доби. Поліція може врахувати цю інформацію та зосередити більше ресурсів на нічному патрулюванні в районах з високим ризиком таких інцидентів. Маючи глибше розуміння закономірностей злочинності та детермінант, отриманих з аналізу правил асоціації, поліція може оптимізувати розподіл обмежених ресурсів, зосередившись на конкретних місцях, періодах та видах злочинів, для яких ризик є найвищим. Однак побудована модель асоціативних правил може надати додаткові можливості для розуміння місцевої злочинності. Наприклад, які види злочинів найімовірніше можуть статися в певному регіоні, які дні чи місяці найбільше сприяють злочинній діяльності? Крім того, запропонована модель може надати правоохоронним органам додаткову інформацію про діяльність організованої злочинності в конкретному регіоні.

Висновки

У дослідженні представлено новий підхід, що полягає у побудові асоціативних правил для аналізу бази даних кримінальних правопорушень, скоєних у Тернопільській області з 2012 по 2023 рік. Встановлено, що крадіжки часто вчиняються групами незалежно від умов освітлення, причому Тернопіль є гарячою точкою таких злочинів. Групова злочинна діяльність має тенденцію досягати піку в квітні та в другій половині тижня, особливо в четвер і п'ятницю. Цікаво, що крадіжки та незаконне заволодіння транспортними засобами виявилися тісно пов'язаними правопорушеннями, які часто відбуваються одночасно. Крім того, викрадення автомобілів зазвичай здійснювалося групами осіб у темний час доби.

Виявлені асоціативні правила дають цінне розуміння динаміки злочинної діяльності в регіоні, що дозволяє правоохоронним органам розробляти цільові стратегії запобігання та оптимізувати розподіл ресурсів. Однак важливо визнати специфічний для місцевості характер злочинних закономірностей, що вимагає подальших досліджень у різних регіонах для формування комплексного розуміння злочинної поведінки та факторів, що на неї впливають.

Література

1. Berezka, K. M., Kovalchuk, O. Ya., Banakh, S. V., Zlyvko, S. V., & Hrechaniuk, R. (2022). A binary logistic regression model for support decision making in criminal justice. *Folia Oeconomica Stetinensia*, 22(1), 1–17. <https://doi.org/10.2478/fofi-2022-0001>
2. Wang, Z., & Zhang, H. (2020). Construction, detection, and interpretation of crime patterns over space and time. *ISPRS International Journal of Geo-Information*, 9, 339. <https://doi.org/10.3390/ijgi9060339>
3. Modise, J. M. (2023). Key elements of predictive policing include crime analysis, crime mapping and geographies. *International Journal of Innovative Science and Research Technology*, 8(5), 3312–3326. <https://www.ijisrt.com/key-elements-of-predictive-policing-include-crime-analysis-crime-mapping-and-geographies>
4. Kovalchuk, O., Kasianchuk, M., Karpinski, M., & Shevchuk, R. (2023). Decision-making supporting models concerning the internal security of the state. *International Journal of Electronics and Telecommunications*, 69(2), 301–307. <https://doi.org/10.24425/ijet.2023.144365>
5. Ceccato, V., & Newton, A. (2025). *System thinking for sustainable crime prevention*. Routledge. <https://doi.org/10.4324/9781003281030>
6. Tavares, J. P., & Costa, A. C. (2021). Spatial modeling and analysis of the determinants of property crime in Portugal. *ISPRS International Journal of Geo-Information*, 10, 731. <https://doi.org/10.3390/ijgi10110731>
7. Lisowska-Kierepka, A. (2021). How to analyse spatial distribution of crime? Crime risk indicator in an attempt to design an original method of spatial crime analysis. *Cities*, 120, 103403. <https://doi.org/10.1016/j.cities.2021.103403>
8. Kim, H., & Seong, E. (2022). Pattern and explanation of inter-city crime variation in South Korea. *Sustainability*, 14(22), 15458. <https://doi.org/10.3390/su142215458>
9. Haider, M. A., & Iamtrakul, P. (2022). Analyzing street crime hotspots and their associated factors in Chittagong City, Bangladesh. *Sustainability*, 14, 9322. <https://doi.org/10.3390/su14159322>
10. Liu, D., Song, W., Xiu, C., & Xu, J. (2021). Understanding the spatiotemporal pattern of crimes in Changchun, China: A Bayesian modeling approach. *Sustainability*, 13(19), 10500. <https://doi.org/10.3390/su131910500>
11. Cadena-Urzuá, P. Á., Briz-Redón, Á., & Montes, F. (2022). Crime analysis of the Metropolitan Region of Santiago de Chile: A spatial panel data approach. *Social Sciences*, 11, 443. <https://doi.org/10.3390/socsci11100443>
12. Mathur, V. (2022). Association rule mining: Importance and steps. *Analytics Steps*. <https://www.analyticssteps.com/>
13. RapidMiner Documentation. (2023). *RapidMiner*. <https://docs.rapidminer.com>