УДК 343.9:159.9.075
ДМИТРО ЮРІЙОВИЧ УЗЛОВ
Кандидат технічних наук,
начальник Управління інформаційно-аналітичної підтримки ГУ НП в Харківській області, полковник поліції
ВОЛОДИМИР МИХАЛОВИЧ СТРУКОВ
Кандидат технічних наук, доцент, завідувач кафедри інформаційних технологій факультету № 4 Харківського національного університету внутрішніх справ

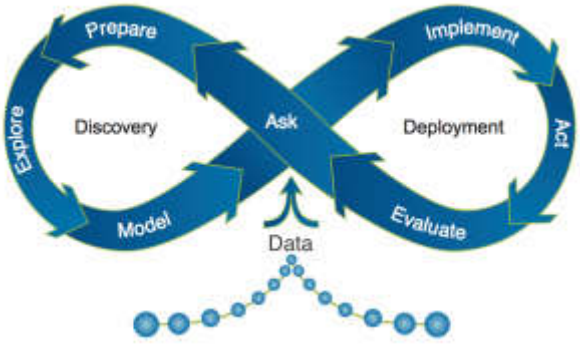
ВИКОРИСТАННЯ МЕТОДІВ І ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ В КРИМІНАЛЬНОМУ АНАЛІЗІ

Основним завданням кримінального аналізу є пошук неочевидних рішень в умовах неповної інформації, що є основною ознакою інтелектуальності. Кримінальний аналіз базується на великих масивах кримінально значимої інформації, яка часто є потоками слабоструктурованих і неструктурованих текстових масивів. Аналіз цієї інформації здійснюється правоохоронцями на основі дедуктивних і індуктивних методів і асоціативного мислення, отже, може бути автоматизований з використанням методів і технологій штучного інтелекту.

Інтелектуальний аналіз кримінальних даних включає в себе такі обов'язкові етапи:

- колекція даних і їх підготовка (препроцесінг);
- класифікація, кластеризація і семантичний розбір підготовлених даних;
- пошук прихованих закономірностей;
- прогнозування;
- візуалізація результатів.

У разі безперервного аналізу в режимі реального часу, інтелектуальний аналіз переходить в циклічний процес з додаванням етапу коригування і уточнення (додавання або зменшення параметрів) вихідних даних.



Мал. 1. Модель Data Mining цикла.

1. Колекція даних і їх підготовка (препроцесінг).

На даному етапі відбувається визначення джерел інформації, витяг кримінально значимої інформації, вибір обов'язкових атрибутів, компресія.

Використовуються методи: антиціпаційного алгоритму (схеми передбачення), асоціативні правила, нейронні алгоритми навчання.

На практиці: очищення даних інформаційно-пошукових систем від сміття і перетворення їх в знання, формування метаданих з існуючих баз даних, формування метапошукових запитів в зовнішнє середовище.

2. Класифікація, кластеризація і семантичний розбір.

Дані групуються або за класами (в разі можливості їх визначити заздалегідь), або по кластерам (в разі неможливості їх визначити заздалегідь). Визначається смислове значення даних (семантичний розбір) і будується функція приналежності до того чи іншого класу або кластеру.

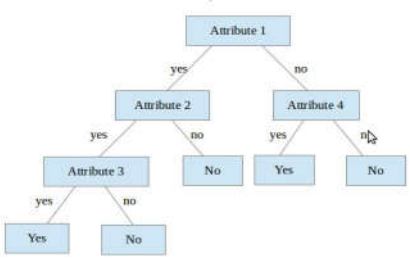
Використовувані методи: Байєсівський алгоритм, регресійний аналіз, K - means, C- means, Fuzzy Logic, латентно-семантичний аналіз, імовірнісний аналіз.

На практиці: угруповання подій за класами і групами; побудова поведінкових профілів подій, об'єктів і суб'єктів; угруповання осіб за поведінковим профілем; виявлення групової та серійної злочинності.

3. Пошук прихованих закономірностей.

Виявляється кореляція між різними процесами і множинами даних. Визначаються взаємодії множин, будуються перетини і відповідності.

Використовуються методи: дерева рішень, кластерний аналіз, нечітка логіка, асоціативна логіка, нейронні мережі, апріорний алгоритм.



Мал. 2 Дерево рішень.

На практиці: групування подій за схожими (в т.ч. заздалегідь невідомими) ознаками; визначення множини осіб, відповідній певній групі (кластеру) подій; modus operandi; визначення тенденцій і закономірностей прояву злочинності за локалізацією, часовими та іншими можливими залежностями.

4. Прогнозування.

Визначення ймовірностей і можливостей настання подій в певному місці, в певний час, певним чином.

Використовуються методи: дерева рішень, динамічні ряди, алгоритми нечіткої логіки, теорії ймовірності.

На практиці: визначення місць, часу, обставин, при яких ймовірність скоєння кримінальних подій висока; виявлення циклічних процесів на певній локалізації; виявлення осіб, схильних до вчинення протиправних дій; виявлення ланцюжків обставин, що сприяють вчиненню злочинів.

5. Візуалізація результатів.

Представлення результатів у вигляді, зручному для сприйняття і інтуітивно зрозумілому. Підтримка прийняття рішення на основі відображення множини станів різних індикаторів.

Використовуються методи: crime mapping, OLAP, Visual Mining.

На практиці: візуалізація на електронній карті місць концентрацій різноманітних подій; візуалізація зв’язків подій, осіб, об’єктів, процесів в просторі та часі.

Мал.3 Приклади візуалізації

Список використаних джерел

1. Investigative Data Mining for Security and Criminal Detectio. Jesus Mena - Butterworth Heinemann is an imprint of Elsevier Science.Copyright © 2003, Elsevier Science (USA).

2. Whitepaper, “Oracle’s Integration Hub For Justice And Public Safety”, Oracle Corp. 2004, available at:http://www.oracle.com/industries/government/IntegrationHub_Justice.pdf

3. Прикладний кримінальний аналіз на базі інформаційно-аналітичної системи «RICAS»: методичні рекомендації щодо аналітичної діяльності та кримінального аналізу на базі інформаційно-аналітичної системи «RICAS». – Харків : «Юрайт», 2018.- 92 с.