INFORMATION PLATFORM "CENTER FOR INNOVATIVE THINKING"
UKRAINIAN INSTITUTE OF SCIENTIFIC STRATEGIES
EUROPEAN UNION RESEARCH DEPARTMENT
SCIENTIFIC AND PUBLISHING CENTER "PROGRESS"

SCIENCE, TECHNOLOGY AND CULTURE: INTERACTION, EVOLUTION AND PROGRESS

PROCEEDINGS OF THE INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE

DECEMBER 21-23, 2025
COPENHAGEN, DENMARK

INFORMATION PLATFORM "CENTER FOR INNOVATIVE THINKING"
UKRAINIAN INSTITUTE OF SCIENTIFIC STRATEGIES
EUROPEAN UNION RESEARCH DEPARTMENT
SCIENTIFIC AND PUBLISHING CENTER "PROGRESS"

SCIENCE, TECHNOLOGY AND CULTURE: INTERACTION, EVOLUTION AND PROGRESS

PROCEEDINGS OF THE INTERNATIONAL SCIENTIFIC
AND PRACTICAL CONFERENCE

December 21-23, 2025

Copenhagen, Denmark

This edition was approved for publication on January 11, 2026.

Published in A4 format online on website:
https://naukainfo.com/conference?id=84

Publisher: Sole proprietor Soloviov O. V. Certificate of registration in the State Register of Publishers, Manufacturers, and Distributors of Publishing Products series DK № 8227, dated April 23, 2025.

Copenhagen, Denmark
2026

UDC 001.3-048.35:0/9](06)

Proceedings of the International scientific and practical conference “Science, Technology and Culture: Interaction, Evolution and Progress” (December 21-23, 2025) / Publisher website: www.naukainfo.com. – Copenhagen, Denmark, 2026. – 161 p.

ISBN 978-617-8680-29-9
https://doi.org/10.64828/conf-84-2025

The recommended citation for this publication is:

Shevchenko T. G. Research into the specifics of the development of performing arts in Ukraine under martial law // Science, Technology and Culture: Interaction, Evolution and Progress : proceedings of the International scientific and practical conference (December 21-23, 2025). – Copenhagen, Denmark : naukainfo.com, 2026. - Pp. 15-21. - URL: https://naukainfo.com/conference?id=84

Editor
Soloviov O. V.

M.Sc.Ed., M.P.A., Hon. PhD, Academic Advisor,
Head of the European Union Research Department,
Ukrainian Institute of Scientific Strategies

The collection of scientific articles is a scientific and practical publication that includes research papers by students, postgraduate students, Candidates and Doctors of Sciences, researchers, and practitioners from Ukraine, Europe, neighboring countries, and beyond. The articles reflect studies of processes and changes in the structure of modern science. This collection is intended for students, postgraduate and doctoral candidates, educators, researchers, practitioners, and all those interested in current trends in the development of modern science.

E-mail: journal@naukainfo.com

Publisher website: https://www.naukainfo.com

УДК 004.056

Ланде Дмитро Володимирович
д.т.н., професор, зав. каф. ІБ
ORCID ID: 0000-0003-3945-1178
Косигін Олександр Сергійович
магістрант каф. ІБ
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
м. Київ, Україна

ЗАСТОСУВАННЯ СТИЛОМЕТРІЇ В OSINT ДЛЯ ВИРІШЕННЯ ЗАДАЧ КІБЕРНЕТИЧНОЇ ТА ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Анотація. У роботі досліджено ефективність застосування методів стилометрії в рамках розвідки з відкритих джерел для вирішення задач атрибуції кіберзагроз. Запропоновано архітектуру системи ідентифікації авторів текстів та програмного коду, що базується на аналізі мікро-синтаксичних ознак та «технічного шуму». Експериментально доведено, що використання алгоритму One-Class SVM у поєднанні з символьними n-грамами дозволяє з високою точністю виявляти аномалії авторського стилю

Ключові слова: OSINT, кібербезпека, стилометрія, атрибуція кіберзагроз, машинне навчання, One-Class SVM, аналіз програмного коду.

Вступ

В умовах глобалізації кіберзлочинності атрибуція атак стає критично важливим завданням. Фахівцям з кіберзахисту стає дедалі складніше ідентифікувати зловмисників, оскільки традиційні технічні індикатори (IP-адреси, хеш-суми файлів, цифрові підписи) можуть фальсифікуватись або приховуються засобами анонімізації (Tor, VPN).

Одним із перспективних інструментів, що дозволяє вирішити цю проблему в рамках розвідки з відкритих джерел (OSINT) , є метод кількісного аналізу авторського стилю – стилометрія. Відомо, що кожен автор має унікальний «стилістичний відбиток» (ідіолект), який проявляється у несвідомому використанні специфічних мовних конструкцій, пунктуації та патернів кодування [1].

Актуальність задачі

Серед актуальних проблем атрибуції кіберзагроз та аналізу даних в OSINT можна виділити такі [2]:

1. Зміна семантичного домену. Стандартні NLP-моделі (наприклад, BERT) часто «вивчають тему», а не автора. Коли зловмисник змінює контекст спілкування, точність ідентифікації критично знижується.
2. Обфускація програмного коду. Розробники шкідливого ПЗ застосовують пакувальники та обфускатори. Лексичні методи аналізу дають високий рівень помилкових спрацювань (False Positive).
3. Компрометація облікових записів. Необхідність виявлення ситуацій, коли легітимний акаунт захоплюється зловмисником, вимагає аналізу поведінкових аномалій, а не лише факту входу в систему.

У цій роботі пропонується архітектура системи для автоматизованої атрибуції текстів та програмного коду, яка базується на аналізі мікро-синтаксичних ознак та «технічного шуму» (n-грами, службові слова, відступи в коді).

Архітектура рішення

Загальну схему алгоритму стилометричного аналізу для підтримки OSINT-розслідувань подано на рис. 1.

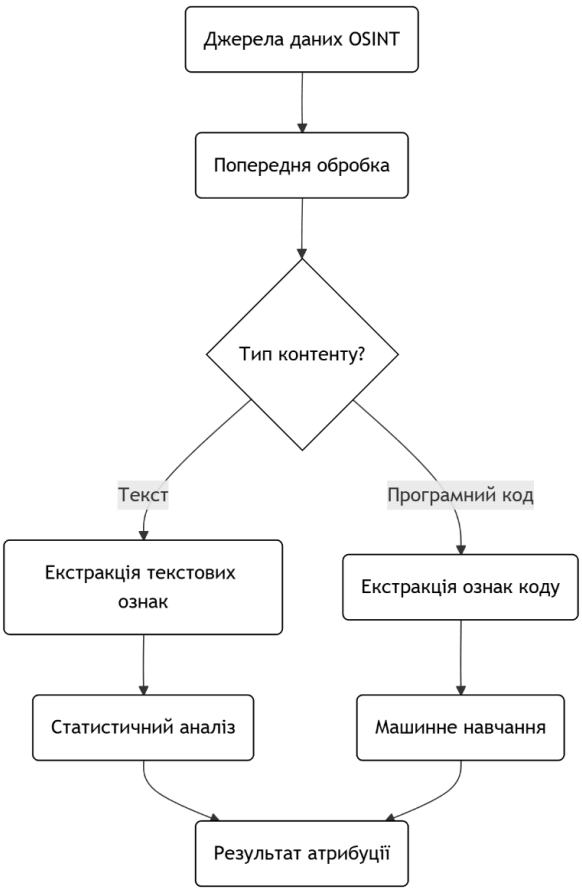

Рис. 1. Алгоритм стилометричної ідентифікації та профілювання суб'єктів у кіберпросторі

Методика, що пропонується, включає такі етапи:

1. Збір даних. Акумуляція даних з джерел OSINT (GitHub, форуми, соцмережі).
2. Розгалуження аналізу. Залежно від типу контенту застосовуються різні вектори ознак:
   - Для тексту*:* символьні n-грами (3-grams) та частоти функціональних слів.
   - Для коду: структурні вузли абстрактного синтаксичного дерева (AST) та «технічний шум».
3. Аналітичне ядро. Використання методу Burrows' Delta для текстової атрибуції та алгоритмів ML для класифікації коду.
4. Поведінкова біометрія. Використання One-Class SVM для детекції аномалій.

Результати досліджень

Ключовим етапом роботи стала перевірка ефективності алгоритмів навчання без учителя для задачі виявлення підміни автора. Порівнювалися два методи векторизації: Default (NLP підхід на основі слів) та Stylometry (підхід на основі символьних n-грам).

Таблиця 1.

Порівняння ефективності алгоритмів детекції аномалій

Algorithm	**Method**	**AUC-ROC**	**True**	**False**
Isolation Forest	Stylometry	0.389257	14/50	157/350
Isolation Forest	Default	0.237314	0/50	87/350
Isolation Forest	Stylometry	0.770000	25/50	56/350
Isolation Forest	Default	0.636629	18/50	50/350
One-Class SVM	Stylometry	0.805486	37/50	140/350
One-Class SVM	Default	0.680286	38/50	211/350
SGD Linear SVM	Stylometry	0.729371	50/50	349/350
SGD Linear SVM	Default	0.598857	50/50	342/350

Аналіз даних таблиці показав:

1. Метод Stylometry (n-грами) стабільно перевершує стандартний NLP за показником AUC-ROC. Це підтверджує, що структура слів несе більше інформації про особу автора, ніж словниковий запас.
2. Найкращий результат надав алгоритм One-Class SVM зі стилометричними ознаками (0.805). Він забезпечує оптимальний баланс між виявленням атак та кількістю хибних спрацювань.
3. Алгоритм SGD Linear SVM, хоч і виявив усі атаки, продемонстрував неприйнятно високий рівень хибних тривог (близький до 100%), що робить його непридатним для реальних систем.

Висновки

Інтеграція методів обчислювальної стилометрії в інструментарій OSINT дозволяє значно підвищити надійність атрибуції кіберзагроз. Експериментально показано, що методи, орієнтовані на аналіз формальних ознак, є більш ефективними та ресурсоощадними, ніж важкі нейромережеві моделі. Комбінація One-Class SVM та символьних n-грам є рішенням, що рекомендується для побудови систем захисту від атак типу Account Takeover.

СПИСОК ВИКОИРСТАНИХ ДЖЕРЕЛ:

1. Stamatatos E. A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology*. 2009. Vol. 60, № 3. P. 538–556. URL: https://doi.org/10.1002/asi.21001
2. Afroz S., Brennan M., Greenstadt R. Adversarial Stylometry: Circumventing Authorship Recognition to Preserve Privacy and Anonymity. *ACM Transactions on Information and System Security*. 2012. Vol. 15, № 3. URL: https://doi.org/10.1145/2382448.2382450
3. Caliskan A. et al. De-anonymizing Programmers via Code Stylometry. *Usenix Security Symposium*. 2015. P. 255–271. URL: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/caliskan-islam
4. Burrows J. F. ‘Delta’: a Measure of Stylometric Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*. 2002. Vol. 17, № 3. P. 267–287. URL: https://academic.oup.com/dsh/article/17/3/267/985972
5. Koppel M., Schler J. Authorship Verification as a One-Class Classification Problem. *Proceedings of the 21st International Conference on Machine Learning*. 2004. URL: https://dl.acm.org/doi/10.1145/1015330.1015448