УДК 004.056.5:004.8 DOI: https://doi.org/10.32515/2414-3820.2025.55.309-313

В.А. Резніченко, викл., **А.Я. Клюй,** студ.
Центральноукраїнський національний технічний університет, м. Кропивницький, Україна
e-mail: upsbilly@meta.ua, nastya.klui@gmail.com

Алгоритми фільтрації контенту в комп'ютерних системах: аналіз ефективності, виклики впровадження та перспективи розвитку

У статті представлено порівняльний огляд сучасних методів фільтрації контенту, що використовуються в комп'ютерних системах. Основну увагу приділено аналізу ефективності, ресурсоємності, масштабованості та практичної реалізації зазначених підходів. Оцінено сильні й слабкі сторони найпоширеніших методів фільтрації, а також їхню актуальність в умовах зростання обсягів інформації та цифрових загроз.
фільтрація контенту, машинне навчання, адаптивні системи

Постановка проблеми. За даними звіту CYBER DIGEST (2024) [3], понад 60% користувачів стикаються з онлайн-загрозами. Це стає особливо гострою проблемою на фоні стрімкого збільшення обсягів даних. Системи фільтрації повинні не лише швидко виявляти шкідливий контент, а й водночас не блокувати корисну інформацію. Отже, виникає запитання: як створити алгоритми, які поєднуватимуть точність, ефективність і невисоку ресурсоємність? Одним із потенційних напрямів вирішення є застосування інтелектуальних підходів – зокрема, алгоритмів машинного навчання (ML) та методів обробки природної мови (NLP). Актуальність цієї теми пояснюється її міждисциплінарним характером: вона охоплює аспекти кібербезпеки, аналізу даних і розробки адаптивних програмних рішень. Саме тому фільтрація контенту розглядається як прикладне завдання та водночас перспективний напрям досліджень у сфері комп'ютерних наук.

Аналіз останніх досліджень і публікацій. Проблема алгоритмів фільтрації отримала помітне висвітлення в наукових джерелах. Існує чимало робіт, у яких розглядаються різні методи вирішення цього завдання – від найпростіших до високотехнологічних.

Так, Джон Росс Андерсон у своїй праці «Інженерія безпеки: посібник із побудови надійних розподілених систем» [1] описує метод використання чорних списків. Автор підкреслює, що це досить простий і швидкий спосіб фільтрації, однак його ефективність сильно залежить від регулярного оновлення списку. Тобто, якщо не підтримувати актуальність даних, система швидко втрачає здатність блокувати нові загрози.

У роботах Д.О. Бухаленкова та Т.М. Заболотної [2] акцент зроблено на методі фільтрації за ключовими словами. Такий підхід дозволяє обробляти великі обсяги інформації у короткий час. Проте дослідники наголошують, що високий рівень хибнопозитивних спрацювань (наприклад, коли нешкідливий контент блокується через випадкове співпадіння ключових слів) залишається серйозною проблемою.

Більш сучасні підходи пропонують застосовувати машинне навчання. Відомий дослідник Том Мітчелл у книзі «Машинне навчання» [5] відзначає, що такі алгоритми здатні самостійно навчатися та підлаштовуватися під нові загрози. Це, безумовно, робить їх ефективними у динамічному цифровому середовищі. Однак, як зазначає науковець, такі рішення потребують значних обчислювальних ресурсів, що може бути недоступним для деяких організацій.

Цікавою є також модель фільтрації на основі контекстного аналізу, яку запропонували Картік Дінакар, Генрі Ліберман та Розалінд В. Пікард. Вони стверджують, що такий підхід дозволяє глибше розуміти зміст тексту та зменшити кількість хибних блокувань. Проте, як і у випадку з ML, він вимагає багато ресурсів та чутливий до неоднозначностей мови [4].

Загалом, аналіз наявних публікацій показує, що жоден із підходів не є універсальним. У кожного є як переваги, так і недоліки. Це підкреслює необхідність подальших досліджень, спрямованих на створення більш ефективних та збалансованих рішень для фільтрації, які б одночасно забезпечували точність, адаптивність і оптимальне використання ресурсів. На сьогодні залишаються невирішеними такі проблеми, як зниження кількості хибнопозитивних спрацювань без шкоди для продуктивності систем, а також удосконалення фільтрів для роботи в умовах обмежених обчислювальних потужностей. Саме ці питання й визначають головний акцент дослідження, представленого в цій статті.

Постановка завдання. Метою цієї статті є аналіз та порівняння сучасних алгоритмів фільтрації контенту з точки зору комп'ютерної науки. Особливу увагу приділено оцінці таких критеріїв, як ефективність, ресурсоємність, масштабованість та складність впровадження в умовах динамічного інформаційного середовища. Завдання полягає не лише у виявленні сильних і слабких сторін кожного підходу, але й у вивченні практичних сфер їх застосування та розгляді прикладів використання в реальних системах. Для систематизації результатів створено порівняльну таблицю, що узагальнює ключові характеристики методів фільтрації. Вона ґрунтується на аналізі наукових джерел і слугує інструментом для обґрунтованого вибору найбільш придатного рішення залежно від потреб конкретної системи.

Виклад основного матеріалу. На основі опрацювання сучасних наукових джерел [1-2, 4-5] узагальнено основні характеристики найбільш поширених методів фільтрації контенту (табл. 1).

Ця таблиця систематизує підходи за критеріями точності, адаптивності, ресурсоємності, сфер застосування та прикладами реалізації.

Таблиця 1 – Порівняльна характеристика методів фільтрації контенту за основними критеріями

Метод фільтрації контенту	*Чорні списки URL*	*Ключові слова*	*Машинне навчання*	*Контекстний аналіз*
Принцип роботи	Блокування доступу до сайтів за базою «чорних» адрес	Пошук та аналіз специфічних слів чи фраз у контенті	Навчання моделей на великих масивах даних для розпізнавання патернів у контенті	Семантичне та лінгвістичне аналізування змісту для точного визначення значення

Продовження таблиці 1

Переваги	Швидке та просте блокування. Мінімальні ресурси	Швидка реалізація. Просте масштабування	Висока точність. Самовнавчання. Адаптивність до нових загроз	Мінімізує хибнопозитивні блокування завдяки глибокому аналізу контенту
Недоліки	Низька точність. Потреба в регулярному оновленні списків. Не захищає від нових загроз	Високий ризик хибнопозитивних блокувань. Не враховує контекст	Високі обчислювальні витрати. Потреба у великих наборах тренувальних даних.	Ресурсоємний. Вимагає складних алгоритмів. Труднощі з багатомовністю
Сфера застосування	Оптимально для базового захисту невеликих організацій та початкового рівня фільтрації	Модерація контенту на форумах, соцмереж. Антиспам-фільтри для електронної пошти	Антиспам системи корпоративного рівня. Виявлення вторгнень у реальному часі. Динамічний моніторинг	Критично важливі системи. Фінансові установи. Інтелектуальний моніторинг коментарів
Ефективність	60 %	70 %	90 %	88 %
Вимоги до ресурсів	Низькі	Низькі–середні	Високі	Дуже високі
Складність впровадження	Низька	Низька–середня	Висока	Дуже висока
Приклади реалізації	Google Safe Browsing, DNS-based blacklists, Spamhaus DBL, OpenPhish	WordPress Akismet, Barracuda Email Security, SpamAssassin (keyword-based mode)	Google Perspective API, FastText NLP, Facebook DeepText, SpamAssassin з ML-модулями	GPT (OpenAI), BERT (Google), T5 (Google), ERNIE (Baidu), Hugging Face Transformers

Джерело: [1-2, 4-5]

Аналіз отриманих даних свідчить про те, що жоден із методів не є універсальним. Кожен має свої переваги та обмеження:

- *Чорні списки URL* ефективні для швидкого блокування відомих загроз із мінімальними витратами, але не здатні виявляти нові типи загроз і потребують постійного оновлення баз.

- *Фільтрація за ключовими словами* забезпечує оперативну обробку великих масивів тексту, проте водночас нерідко спричиняє хибні спрацювання через багатозначність слів у мові.

- *Машинне навчання* забезпечують високу точність і адаптивність, але вимагають значних обчислювальних ресурсів і складного налаштування.

- *Контекстний аналіз* дозволяє досягти глибшого розуміння змісту і мінімізує помилки блокування, однак є надзвичайно ресурсоємним і складним у реалізації.

З огляду на виявлені особливості, у дослідженні обґрунтовано доцільність побудови **багаторівневої системи фільтрації контенту**, яка поєднує переваги кожного з методів:

1) **Перший рівень –** чорні списки URL: базовий захист із мінімальним навантаженням, швидке блокування відомих джерел загроз.

2) **Другий рівень –** фільтрація за ключовими словами: швидка перевірка великого обсягу текстових даних на основі виявлення небажаних термінів.

3) **Третій рівень –** машинне навчання: адаптивний аналіз, який підвищує точність розпізнавання загроз.

4) **Четвертий рівень –** контекстний аналіз: фінальна перевірка у критично важливих сферах, де недопустимі помилкові рішення, таких як державні системи чи фінансові платформи.

Запропонована модель забезпечує збалансоване поєднання точності, адаптивності та обчислювальної ефективності. Такий підхід дозволяє надійно реагувати на сучасні інформаційні загрози в умовах зростання обсягів контенту та складності цифрового середовища.

Висновки. Результати аналізу свідчать, що ефективна фільтрація контенту в комп'ютерних системах потребує комплексного підходу. Жоден із досліджуваних методів – ані чорні списки, ані фільтрація за ключовими словами, ані машинне навчання чи контекстний аналіз – не є універсальним рішенням. Запропонована в роботі багаторівнева система дозволяє об'єднати сильні сторони різних підходів і забезпечити баланс між точністю, швидкодією та ефективним використанням ресурсів. Практична цінність моделі полягає у її гнучкості та здатності адаптуватися до змінних умов інформаційного середовища. Серед ключових проблем, які потребують подальшої уваги, можна виокремити: адаптацію алгоритмів до нових типів загроз, ефективне використання обчислювальних ресурсів, зменшення хибнопозитивних спрацювань без втрати точності. У майбутніх дослідженнях доцільно зосередитися на тіснішій інтеграції інтелектуальних рішень – зокрема, використанні моделей машинного навчання та методів аналізу природної мови. Це дозволить значно посилити інтелектуальні можливості систем фільтрації. Впровадження таких підходів сприятиме створенню масштабованих, надійних і точних систем, які відповідатимуть сучасним вимогам у сфері кібербезпеки.

Список літератури

1. Anderson R.J. Security Engineering: A Guide to Building Dependable Distributed Systems / R.J. Anderson. 3rd ed. Hoboken: John Wiley & Sons, 2020. 1232 p. URL: https://www.cl.cam.ac.uk/archive/rja14/book.html?utm_source=chatgpt.com (дата звернення 14.04.2025).
2. Бухаленков Д.О., Заболотня Т.М. Модифікований метод пошуку ключових слів та термінів у текстових даних. *Проблеми програмування*. 2024. №1. С. 12–22. (дата звернення 13.04.2025).
3. CYBER DIGEST: Огляд подій у сфері кібербезпеки. Січень, 2024. 44 с. URL: https://www.rnbo.gov.ua/files/2024/NATIONAL_CYBER_SCC/Cyber%20digest/Cyber%20digest_Jan_2024_UA.pdf (дата звернення 15.04.2025).

4. Dinakar K., Lieberman H., Picard R. Using Common Sense Reasoning to Detect and Respond to Cyberbullying. *ACM Transactions on Interactive Intelligent Systems*. 2022. Vol. 2, № 3. Article 18. P. 1-27.
5. Mitchell T.M. Machine Learning. New York: McGraw-Hill, 1997. 432 p. URL: https://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf?utm_source=chatgpt.com (дата звернення 15.04.2025).

References

1. Anderson, R.J. (2020). *Security Engineering: A Guide to Building Dependable Distributed Systems.* (3rd ed.). Hoboken: John Wiley & Sons. https://www.cl.cam.ac.uk/archive/rja14/book.html?utm_source=chatgpt.com [in English].
2. Bukhalenkov, D.O. & Zabolotnia, T.M. (2024). Modified Method for Searching Keywords and Terms in Text Data. *Problems of Programming*. № 1. 12–22 [in Ukrainian].
3. CYBER DIGEST: Review of Events in the Field of Cybersecurity. January 2024. 44 p. https://www.rnbo.gov.ua/files/2024/NATIONAL_CYBER_SCC/Cyber%20digest/Cyber%20digest_Jan_2024_UA.pdf [in Ukrainian].
4. Dinakar, K., Lieberman, H. & Picard, R. (2022). Using Common Sense Reasoning to Detect and Respond to Cyberbullying. *ACM Transactions on Interactive Intelligent Systems*, Vol. 2, No. 3. Article 18. P. 1-27. [in English].
5. Mitchell, T.M. (1997). *Machine Learning* . New York: McGraw-Hill [in English].

Vitalii Reznychenko, Lecturer**, Anastasiia Kliui,** 2nd-year Student
Central Ukrainian National Technical University, Kropyvnytskyi, Ukraine

Content Filtering Algorithms in Computer Systems: Analysis of Efficiency, Implementation Challenges, and Future Development Prospects

This study provides a structured analysis of modern methods used for content filtering in computer systems, with attention to their efficiency, computational demands, scalability, and implementation feasibility. The review outlines both the strengths and weaknesses of widely adopted strategies, while also assessing their applicability amid rising data volumes and emerging digital threats.

The research discusses four key filtering approaches: URL blacklists, keyword detection, machine learning solutions, and semantic context analysis. Drawing on recent academic literature, the article organizes the core attributes of each method into a comparative table based on accuracy, adaptability, complexity, and usage scope. The findings confirm that none of these methods is fully comprehensive, as each comes with its own trade-offs. As a result, the article supports the concept of a layered filtering framework that integrates the strengths of various methods. This structure is composed of four stages of content handling — from initial rule-based filtering to deeper semantic interpretation.

The developed model offers a well-balanced synergy between filtering precision, system responsiveness, and rational use of computational power. It remains adaptable to emerging threats within rapidly changing digital ecosystems. Future research directions may include reducing false positive detections, improving computational efficiency, and advancing the integration of AI-based solutions and linguistic analysis technologies to strengthen filtering intelligence. The implementation of such systems will contribute to the advancement of modern digital infrastructures that are scalable, intelligent, and secure, especially in the context of computer science.

smart content filtering, machine learning techniques, adaptive computing architectures

Одержано (Received) 18.04.2025 *Прорецензовано (Reviewed) 19.09.2025*
Прийнято до друку (Approved) 23.12.2025