

*Приурочено до 125-ї річниці створення
Національного технічного університету України
“Київський політехнічний інститут імені Ігоря Сікорського”*

МАТЕРІАЛИ

II Міжнародної науково-практичної конференції “КІБЕРБЕЗПЕКА ДЕРЖАВНИХ ІНСТИТУЦІЙ ТА ПОДОЛАННЯ КРИЗОВИХ СТАНІВ”

(Том 1)

25 травня 2023 року

Київ – Вроцлав

УДК 621

Матеріали II Міжнародної науково-практичної конференції “Кібербезпека державних інституцій та подолання кризових станів” : в 2 т. Київ : ІСЗЗІ КПІ ім. Ігоря Сікорського, 2023. Т. 1. 398 с.

Матеріали II Міжнародної науково-практичної конференції “Кібербезпека державних інституцій та подолання кризових станів”, яка присвячена 125-річчю Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, видані в 2-х томах. На конференції обговорено актуальні проблеми та узагальнено отримані наукові результати у сферах інформаційної та кібербезпеки, кіберзахисту, інформаційних і інформаційно-комунікаційних технологій, а також питання діяльності органів державної влади в умовах кризи. В матеріалах опубліковано тези доповідей, що є фундаментом для подальшого розвитку освітньої, наукової та професійної діяльності учасників конференції в справі наближення Перемоги України.

РЕЦЕНЗЕНТИ:

Олександр ПУЧКОВ	к.філос.н., професор
Сергій КОНЮШОК	к.т.н., доцент
Вадим РОМАНЕНКО	к.т.н., доцент
Дмитро МОГИЛЕВИЧ	д.т.н., професор
Олена УВАРКІНА	д.філос.н., професор
Ігор СУБАЧ	д.т.н., доцент
Сергій ІВАНЧЕНКО	д.т.н., професор
Ярослав ЗІНЧЕНКО	к.т.н., с.н.с.

Рекомендовано до друку Вченою радою ІСЗЗІ КПІ ім. Ігоря Сікорського (протокол № 12 від 23.05.2023)

Святослав ПАРШУКОВ	
Формування проєктної документації на основі використання системи pattern instance notation	143
Артем СОБОЛЄВ	
Інноваційні підходи зберігання даних з відкритих джерел інформації	144
Анатолій ГЛАДУН	
Застосування онтологій в задачах багатокритеріального співставлення складних інформаційних об'єктів	146
Тимофій АНДРІЯШИК; Сергій МІГІН	
Система захисту поштового сервера	147
Нікіта АНДРОЦУК; Олександр УСПЕНСЬКИЙ	
Програмне забезпечення віддаленого мережевого моніторингу	148
Єлизавета АНДРОЦУК; Олександр ШАПОВАЛ	
Аналіз та моніторинг мережевого трафіку комп'ютерної мережі	150
Юрій БАРАНОВ; Ігор ЯКОВІВ	
Програмний засіб автоматизованого збору даних від хостових IDS	151
Євгеній БЕРДНИК; Василь ЦУРКАН	
Специфікація вимог до програмного комплексу аналізування ризиків інформаційної безпеки за стандартом PCI DSS	152
Кирил БЕРЕЗОВСЬКИЙ; Дмитро ШАРАДКІН	
Програмний застосунок розпізнавання користувачів за клавіатурним почерком	153
Владислав БОЙКО; Сергій ОБРАШКО	
Система захисту вебсервера	155
Ілона БЕРБЕР; Олександр ШАПОВАЛ	
Програмний модуль реєстрації дій користувачів Internet Of Thing	156
Владислав БОРИСОВ; Олександр УСПЕНСЬКИЙ	
Програмне забезпечення стеганографічного захисту інформації	157
Валерія ГВОЗДІНСЬКА	
Спосіб структурування діяльності червоної команди протягом кібернавчань	159
Дмитро ДАШКЕВИЧ; Артем ЖИЛІН	
Специфікація вимог до програмного комплексу моделювання загроз безпеці інформаційно-комунікаційних систем	160
Віра ГИРДА; Дмитро ЛАНДЕ	
Визначення фейків за допомогою Language Network	161

ВИЗНАЧЕННЯ ФЕЙКІВ ЗА ДОПОМОГОЮ LANGUAGE NETWORK

Анотація. У тезах викладено метод, за допомогою якого в подальшому буде здійснено визначення фейкової інформації, що ґрунтується на побудові Language Network.

Summary. The theses outline the method by which fake information will be determined in the future, which is based on the construction of the Language Network.

Ключові слова: фейки, Language network, компатифікований граф горизонтальної видимості, вага слова.

Побудова Language Network є методом аналізу текстових документів та інформаційних потоків в кіберпросторі. Кожен текст з текстового корпусу можна подати у вигляді складної мережі, де вузлами є слова та словосполучення, тобто окремі фрагменти Language Network, які дозволяють виявити структурні елементи тексту, що забезпечують зв'язність між словами.

Першим кроком у створенні Language Network є побудова корпусу текстових документів - масиву текстових документів, який відповідає конкретній тематиці.

На наступному етапі відбувається процес розбиття тексту на елементарні одиниці, тобто – лєми. Цей процес дозволяє працювати зі словами як з окремими сутностями та використовувати їхню важливість для подальшого аналізу.

Далі, за допомогою статичного вагового показника TF-IDF визначається важливість кожного слова в текстовому документі. Цей показник враховує частоту вживання слова в документі, а також його значимість у всьому корпусі текстових документів.

Для уникнення ситуації, коли інформаційно-важливі слова або словосполучення в документах одного корпусу мають низький ваговий показник TF, застосовують глобальний ваговий показник GTF.

Далі будується компатифікований граф горизонтальної видимості (КГТВ). Використання КГТВ породжує нову мережу слів, відповідно до якої здійснюється перевизначення ваги слів.

Перерахунок ваги здійснюється за допомогою використання алгоритму NITS, що дозволяє визначити авторство або посередництво для кожного вузла КГТВ.

В процесі аналізу текстових документів слід прибрати слова, які не несуть жодного смислового навантаження і є інформаційно неважливими.

Для визначення вагових значень вузлів використовується матриця суміжності графа, яка формується шляхом конкатенації стовбців та рядків – зваженої компатифікації графа горизонтальної видимості. Отримана матриця визначає орієнтований зважений граф, де вершини відповідають унікальним словам у тексті, а вагове значення ребра, що з'єднує одну вершину з іншою, визначається кількістю появи слова в тексті.

Для кожного текстового документу в корпусі будується компатифікований граф горизонтальної видимості та матриця суміжності.

Щоб визначити фейків за допомогою Language Networks необхідно порівняти матриці суміжностей з текстового корпусу та зробити лексичний та семантичний аналіз. Під час візуального перегляду порівняти дані з різних документів корпусу та визначити в них аномалії.

Щоб здійснити порівняння матриць суміжностей, представимо матрицю A як різницю матриць зважених мереж, що відповідає мережам слів цього корпусу, норма якої оцінюється як міра розбіжності. У своїй роботі використовуватимемо норму Фробеніуса $\|A\|_F$, яка дорівнює кореню квадратному із суми всіх елементів матриці

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$$

Розмірність двох матриць, які порівнюються, має співпадати. Оскільки в різних матрицях суміжностей різна кількість вузлів, мережі слів, що порівнюються можуть бути доповнені словами, що входять до загального складу.

Порівняння текстів за допомогою міри Фробеніуса матриць зважених мереж, які отримані з різних текстових документів корпусу дозволяє виявити семантичну близькість та ступінь подібності текстів.

Висновки. Language Network – метод який здатний здійснювати аналіз та порівняння текстових документів. Використання алгоритмів, які лежать в основі Language Network дозволяє побудувати нові мережі слів, відповідно до яких буде здійснено порівняння, але для ефективного визначення фейкової інформації даний метод слід використовувати разом з іншими методами та підходами.

