

МЕТОД ВИЯВЛЕННЯ І ЗБЕРЕЖЕННЯ НЕПОДІЛЬНОЇ СЕМАНТИЧНОЇ СКЛАДОВОЇ АНГЛОМОВНОГО МАНІПУЛЯТИВНОГО ТЕКСТУ ПРОПАГАНДНОГО ДИСКУРСУ НА ОСНОВІ КВАНТОВОГО ФОРМАЛІЗМУ

У статті проведено дослідження особливостей використання квантового формалізму в задачах комп'ютерного аналізу та обробки маніпулятивних текстів. Розроблений метод виявлення і збереження неподільної семантичної складової маніпулятивного тексту пропагандного дискурсу дає можливість підвищити ефективність виявлення і збереження неподільної семантичної складової порівняно з методами семантичного стиснення, що базуються на інформаційній ентропії тексту, що дає змогу виокремити та зберегти ключові індивідуальні семантичні конструкції, притаманні психолінгвістичному портрету зловмисника для подальшого впровадження в них зворотного впливу. Змодельовано роботу системи визначення реми маніпулятивного англomовного тексту та проведено порівняння з результатами роботи систем семантичного стиснення, що доводить можливість подолання межі семантичного стиснення. Попередні експериментальні дані дають змогу стверджувати про високу ефективність розробленого методу. Окреслено перспективи використання подібного методу в процесі забезпечення інформаційної безпеки.

Ключові слова: визначення теми, визначення реми, протидія пропаганді, протидія маніпулятивним текстам, психолінгвістичний портрет, квантово-семантичний аналіз.

Вступ. Згідно з чинною Доктриною інформаційної безпеки України [1] забезпечення національних інтересів держави в інформаційній сфері, зокрема захист від руйнівних інформаційно-психологічних впливів, є пріоритетним та актуальним завданням. Допускається, що організація зворотного впливу на пропагандиста є тим необхідним засобом, що дасть можливість підвищити ефективність протидії злочинному інформаційному впливу на свідомість громадян та нівелювати наслідки. Для вбудовування націленого на зловмисника контейнера використовується його психолінгвістичний портрет, отриманий завдяки квантово-семантичному психолінгвістичному аналізу [2]. Однак для вбудовування необхідного контейнера знання про психолінгвістичний портрет пропагандиста є необхідною, проте недостатньою умовою. Для підвищення ефективності сприйняття та зменшення шансу можливої протидії зворотному впливу необхідно виокремити та зберегти неподільну семантичну складову маніпулятивного тексту для подальшого комп'ютеризованого форму-

вання зворотного тексту відповідно до психолінгвістичного портрету пропагандиста [3].

Ключовою у цьому контексті виступає саме потреба в методиках виявлення неподільної семантичної складової, які було б легко формалізувати і, відповідно, реалізувати за допомогою інформаційних систем.

Більшість сучасних методів виокремлення базової семантичної складової [4–6] ґрунтуються на семантичному стисненні тексту. В роботі [7] визначається межа семантичного стиснення, однак досягти її без значних втрат семантики вкрай важко. В дослідженні маніпулятивного тексту пропагандного дискурсу важливим є збереження притаманних зловмиснику маркерів маніпулятивного характеру, а той факт, що текст англomовний, зумовлює необхідність врахування сталого порядку слів. В такому разі використовується методика актуального членування речення для виділення теми та реми як окремого речення, так і всього тексту. В роботах [8–10] описується застосування цього підходу. Особливе значення має робота [9], де розглядається ідеологічне пред-

ставлення результатів аналізу, що доцільно при дослідженні текстів маніпулятивного характеру. Однак подібні методи потребують доопрацювання та формалізації. Таким чином, невирішеною залишається задача врахування невизначеності теми/реми при семантичному стисненні, в результаті чого може бути забезпечено захист реми від видалення, а отже, і збереження неподільної семантичної складової англомовного тексту.

Мета дослідження полягає у підвищенні ефективності виявлення та збереження неподільної семантичної складової англомовного тексту, що дасть можливість ефективніше протидіяти маніпулятивним текстам пропагандного дискурсу шляхом здійснення зворотного впливу на пропагандиста за рахунок його індивідуального психолінгвістичного портрету.

В роботі ставляться наступні **задачі дослідження**:

1. Визначення теми та монореми.
2. Визначення дирими.
3. Кореляція реми з пропагандним дискурсом.
4. Збереження реми.

Для виявлення і збереження неподільної семантичної складової пропонується використання квантового формалізму в задачах комп'ютерної лінгвістики [11]. Закони формальної семантики [12], синонімічний розподіл та метод контент-аналізу семантичної частки [13] використовуються для визначення теми та монореми. Метод визначення координат семантичної частки при відомому психолінгвістичному портреті пропагандиста [14], правила побудови англомовного речення та тексту і метод контент-аналізу семантичної частки [13] використовуються для визначення дирим. В свою чергу, власне квантовий формалізм [11] разом із методом квантово-семантичного психолінгвістичного аналізу англомовного тексту пропагандного дискурсу [2] та контент-критеріями семантичної категорії психолінгвістичного портрету [15] застосовується для кореляції реми з пропагандним дискурсом. Збереження реми відбувається із використанням методу контент-аналізу семантичної частки [13].

Виклад основного матеріалу. З метою пояснення базових механізмів, що лежать в основі процесу виявлення і збереження непо-

дільної семантичної складової тексту, необхідно надати визначення.

Темою тексту при здійсненні інформаційного впливу виступає функціональний блок тексту, який вже є відомим об'єкту здійснення інформаційного впливу.

Ремою тексту при здійсненні інформаційного впливу виступає функціональний блок тексту, який ще є невідомим об'єкту здійснення інформаційного впливу.

Диремою тексту виступає функціональний блок, в якому тема пов'язана з попередніми темами тексту (може виступати як темою, так і ремою).

Функціональний блок може бути як темою, так і ремою залежно від суб'єкта здійснення інформаційного впливу та при повторі попередньої вже відомої інформації.

Визначення теми та реми неможливе без контексту. Звідси випливає, що тему та рему можна вважати ядром семантики, а оточуючі лексико-граматичні структури речення – вказівниками. Виходячи з визначень, вказівники на тему можна вилучити без шкоди для індивідуальної семантики тексту пропагандного дискурсу. Водночас постає задача визначення теми та реми та, оскільки тема та рема перебувають у квантовій невизначеності, то стабілізації системи тема-рема.

Виокремлення теми, монореми та дирими. Виходячи з [13], при дослідженні розподілу синонімів можливо визначити тему тексту та підтеми, однак, необхідно визначити лише множину підтем Z (1):

$$Z = O_{0_n}^m - O'_{0_n}, \quad (1)$$

де $O_{0_n}^m$ – множина початкових синонімічних об'єктів O_{0_n} , розподілених по m групах; O'_{0_n} – головна тема тексту, що визначається за (2):

$$O'_{0_n} = \sum_{k=1}^m O_{0_n}^k. \quad (2)$$

Основна тема тексту видаляється, оскільки відноситься до усього тексту загалом і не матиме відповідної реми в цьому випадку.

Таким чином, кожна з підмножин z множини підтем Z буде складатися з синонімічного ряду $\sum_{i=1}^m z_i$, що належить одній із

m синонімічних груп. Зі сформульованих на початку статті визначень випливає, що початковий елемент z_0 кожної з m синонімічних груп, або множина z у випадку, коли $|z|=1$, буде ремою в усіх випадках, окрім зумовлених квантовою невизначеністю стану системи теми/реми згідно з психолінгвістичним портретом зловмисника. Решта елементів будуть ремою лише у випадках, зумовлених квантовою невизначеністю системи. Крім того, в результаті синонімічного зв'язку елементів рема, пов'язана з такою темою, буде диремою і безпосередньо потребує розгляду.

Зв'язний осмислений текст T завжди є послідовністю функціональних блоків, що розкривають суть попереднього вислову (3):

$$T = \sum_{i=1}^n (U + R)_i, \quad (3)$$

де U – тема; R – рема.

В такому випадку рема є моноремою та визначається шляхом пошуку нової, доданої інформації (4), адже можна стверджувати, що тема кожного наступного вислову є ремою попереднього, за винятком дирем:

$$R_n = T - U_n - \sum_{i=1}^{n-1} (U + R)_i. \quad (4)$$

В свою чергу, дирему можна знайти, знаючи тему (елемент множини z) і знайшовши вказівник на ядро семантики, яким є рема. Ядро треба шукати, базуючись на категоріальних ознаках семантичної одиниці із переходом на найвищий ступінь ієрархії [14].

Таким чином, можна позначити два різновиди ядра семантики: ядро семантики, що відповідає темі D' , та ядро семантики, що відповідає ремі D'' . Необхідно знайти D'' , знаючи z та проаналізувавши вказівник на ядро семантики. Оскільки вказівник на ядро семантики може бути не одним словом, а словосполученням чи групою словосполучень, то можливо позначити групу вказівника на ядро семантики, що відповідає темі d' , та групу вказівника на ядро семантики, що відповідає ремі d'' . В такому випадку d'' визначатиметься за формулою (5):

$$d'' = \{B_m^j \mid z_n \notin B_m^j\}, \quad (5)$$

де B_m^j – найвище категоріальне значення з j категоріальних значень, яке відноситься до m -ї групи.

Відповідно до сталого порядку слів в англійській мові та висновків, отриманих в [13], що визначають осмислену лінгвістичну інтерпретацію теми тексту, група вказівника на ядро семантики, що відповідає ремі, буде виглядати (6):

$$d'' = B_{noun}^j + B_{verb}^j + B_{adj}^j, \quad (6)$$

де B_{noun}^j , B_{verb}^j , B_{adj}^j – найвищі категоріальні значення синонімічної конструкції, відповідно, іменника, дієслова та прикметника в послідовній хронології, пов'язаних з синонімічною групою теми z .

В такому випадку дирема буде визначатись як (7):

$$D'' = z_n \cdot d''_n. \quad (7)$$

В свою чергу, група вказівника на ядро семантики, що відповідає темі d' , буде визначатись аналогічно (8):

$$d' = B_{noun}^j + B_{verb}^j + B_{adj}^j, \quad (8)$$

при виконанні умови (9):

$$d'' = \{B_m^j \mid z_n \in B_m^j\}. \quad (9)$$

Неподільна семантична складова тексту загалом буде сумою всіх монорем та дирем, що в ньому наявні. Однак маніпулятивний текст пропагандного дискурсу має особливості, пов'язані з психолінгвістичним портретом зловмисника, від якого залежатиме стан системи тема-рема, оскільки, як вже зазначалося, система тема-рема знаходиться в квантовій невизначеності, а отже, рема може бути одночасно темою і ремою.

Кореляція реми з пропагандним дискурсом. Основний засіб визначення реми в пропагандному дискурсі – квантовий формалізм у зв'язку з використанням методів квантово-семантичного психолінгвістичного аналізу [2]. Однак, перш ніж перейти безпосередньо до ідентифікації реми в системі тема-рема, необхідно співвіднести її з психолінгвістичним портретом пропагандиста, що й буде впливати на процедуру квантування.

В [15] описується класифікація складників семантичної категорії психолінгвістичного портрету, де найвища категорія буде відповідати формулам (6), (8). Звідси випливає, що досліджуваний об'єкт тексту, який є підкатегорією III роду, і найвища категорія, визначена в ході аналізу теми та реми, вказують на відповідну характеристику психолінгвістичного портрету автора, що робить психолінгвістичний вплив (відбиток особистості, демографічний профіль чи детермінація психологічного портрету відповідно до методів впливу) відповідно до загального логічного рівня досліджуваного об'єкта. Таким чином, логічний рівень L^h буде розкладатися на п'ять підпросторів значень (10):

$$L^h = L_1 \oplus L_2 \oplus L_3 \oplus L_4 \oplus L_5, \quad (10)$$

де h – характеристика психолінгвістичного портрету, на яку впливає логічний рівень; L_1 – логічний рівень переконань та цінностей; L_2 – логічний рівень особистої своєрідності; L_3 – логічний рівень оточення; L_4 – логічний рівень поведінки; L_5 – логічний рівень можливостей, при виконанні умови (11):

$$\forall L, B^3 \in B^2 \in B^1 \in B^j \in L, \quad (11)$$

де B^3 – підкатегорія III роду (досліджуваний об'єкт тексту); B^2 – підкатегорія II роду; B^1 – підкатегорія I роду; B^j – найвища категорія.

Оскільки кожен із п'яти логічних рівнів складається з множини B^j , кожна з яких, в свою чергу, містить від одного до трьох елементів, які можна зіставити з частинами мови в наступній відповідності: іменник – означення предмета, дієслово – дія, прикметник – прикмети, то справедливе твердження (12):

$$\begin{cases} L \in d' \\ L \in d'' \end{cases} \mid \forall L \in L^h. \quad (12)$$

У процесі верифікації реми варто використати квантовий формалізм.

В [11] квантовий формалізм виступає в ролі інструменту опису просторів станів концептів. Для вирішення задачі, поставленої в статті, цей інструмент найкраще підходить з двох причин: складники семантичної категорії психолінгвістичного портрету можна

вважати концептами в рамках класифікації семантичної категорії; квантовий формалізм лежить в основі квантово-семантичного психолінгвістичного аналізу. Згідно з [11] для отримання імовірності того, що ядро семантики є ремою у разі дослідження дирем, вектор початкового стану семантичної одиниці необхідно записати у вигляді деякої множини базисних станів системи тема-рема (два базисні стани: тема і рема). Вибір базового стану визначається пропагандистом і процедурою квантування. Тоді для відповідника у квантовому формалізмі – оператора (13) необхідно:

$$\hat{A} |\varphi_i\rangle = L^h, \quad (13)$$

де \hat{A} – оператор; $|\varphi_i\rangle$ – базисний стан.

Визначений за допомогою квантово-семантичного психолінгвістичного аналізу [2] психолінгвістичний портрет також вплине на визначення реми. Таким чином, враховуючи хвильові особливості тексту відповідно до сприйняття семантики автором та логічні рівні я-концепції особистості нарівні з характеристикою психолінгвістичного портрету пропагандиста, функція (14) буде визначати вагу кожного логічного рівня для реального стану системи:

$$F : V_{\langle e.t \rangle, S_2^*, W^{L^h}, L^{S_2^* \times W^{L^h}}} \quad (14)$$

де $V_{\langle e.t \rangle}$ – предикативна константа; S_2^* – індивідні семантичні елементи; W^{L^h} – світ сприйняття пропагандистом дійсності відповідно до притаманних йому рівнів я-концепції особистості. Екстремум функції буде вказувати на переважання відповідних логічних рівнів.

В такому випадку, знаючи вагу кожного логічного рівня для реального стану системи, можна обчислити імовірність $|\psi^D\rangle$ одного з базисних станів ядра семантики за формулою (15) і визначити чи ядро семантики є ремою, чи ремою:

$$|\psi^D\rangle = \sum_i F : V_{\langle e.t \rangle, S_2^*, W^{L^h}, L^{S_2^* \times W^{L^h}}} |\varphi_i\rangle, \quad (15)$$

де $\{|\varphi_i\rangle\}$ – множина базисних станів системи.

Збереження реми. Усе вищезазначене є узагальненою моделлю визначення реми, однак, виходячи з визначень, тема та рема по-

стійно взаємозамінюються в ході розвитку семантичної лінії тексту. А отже, ядро семантики, що було ремою, в наступній структурній одиниці тексту буде темою, як інформація, що вже є відомою об'єкту здійснення впливу. Таким чином, розгляд тексту повинен здійснюватися послідовно, що виключає дослідження тексту частинами і не в хронологічній послідовності та вимагає введення додаткової темпоральної характеристики з метою збереження реми на усьому часовому проміжку розвитку семантичної лінії тексту.

Виходячи з формул [13], що визначають множини семантичних часток усіх п'яти рівнів я-концепції, можна вирахувати множину ядер семантики за усіма логічними рівнями (16) з урахуванням імовірності того, що ядро семантики є диремою та з огляду на часову характеристику:

$$M_L = \{B_L^{m_i} \in U(P(D^n)) \mid \max_{[a;b]} F[B_L^{m_i}]\}, \quad (16)$$

де $U(P(D^n))$ – множина тем з високою імовірністю того, що вони є ремою; $F[B_L^{m_i}]$ – функція, яка описується кривою, проведеною за множиною величин ентропії категоріальних значень $B_L^{m_i}$ за усіма логічними рівнями на відрізьку від a до b за часовою лінією розвитку семантики тексту відповідно до законів логічного і послідовного розвитку семантики тексту.

Таким чином, множина збережених рем R^* може бути визначена за формулою (17):

$$R^* = \{D'' - D' - d'' + R_n \mid D'' \in M_L\}. \quad (17)$$

Інакше кажучи, необхідно видалити усі теми, що не є ремою на момент послідовного дослідження тексту, а також власне вказівник на ядро семантики. В такому разі сума рем та дирем становитиме неподільну семантичну складову англомовного тексту пропагандного дискурсу, що несе в собі маркер психолінгвістичного портрету пропагандиста та вже адаптована до квантово-семантичної обробки.

Результати досліджень. В роботі розроблено метод виявлення і збереження неподільної семантичної складової англомовного маніпулятивного тексту пропагандного дискурсу, що базується на квантовому формалізмі для формалізації й адаптації методів визна-

чення реми тексту до квантово-семантичного дослідження, що дає змогу більш ефективно і точно виявляти ключові семантичні елементи маніпулятивного тексту та зберігати їх для здійснення зворотного впливу на зловмисника, шляхом квантово-семантичної лінгвістичної стеганографії.

Водночас визначення теми проводиться на основі синонімічного розподілу, що дає змогу із застосуванням законів формальної семантики та методу контент-аналізу семантичної частки визначити монорему, яка є одним із елементів неподільної семантичної складової наряду з диремою.

Дирема визначається шляхом застосування методу визначення координат семантичної частки при відомому психолінгвістичному портреті пропагандиста на основі правил побудови англомовного речення та тексту із використанням методу контент-аналізу семантичної частки, що дає змогу сформулювати загальну модель визначення реми з відривом від загальної структури тексту в часовому проміжку розвитку семантики.

Кореляція реми з пропагандним дискурсом здійснюється на основі квантового формалізму із застосуванням методу квантово-семантичного психолінгвістичного аналізу англомовного тексту пропагандного дискурсу шляхом врахування контент-критеріїв семантичної категорії психолінгвістичного портрету для стабілізації системи тема-рема, яка перебуває у квантовій невизначеності, спричиненій психолінгвістичним портретом суб'єкта здійснення впливу.

Збереження реми відбувається шляхом використання методу контент-аналізу семантичної частки для отримання можливості подальшого здійснення маніпуляцій з неподільною семантичною складовою, ґрунтуючись на часовій характеристиці розвитку семантичної лінії тексту.

Обговорення результатів. Моделювання функціонування розробленого методу за попередніми даними порівняно з експериментальними даними роботи аналогів дає змогу прогнозувати ефективність. Так, на рисунку 1 зображено графік залежності об'єму залишкового тексту від об'єму початкового.

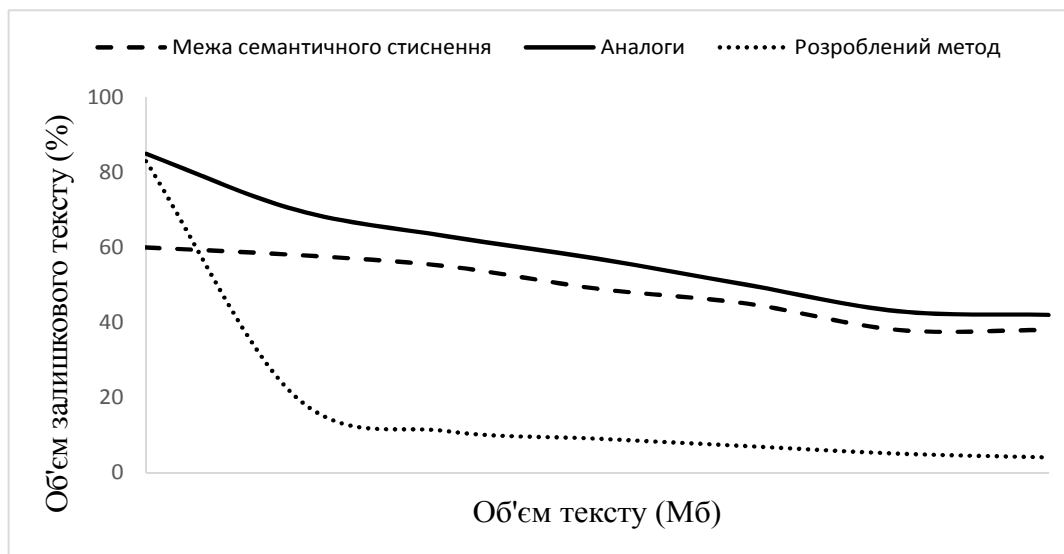


Рисунок 1 – Залежність об'єму залишкового тексту від об'єму початкового

Як видно з графіка, виділення семантичної складової за допомогою семантичного стиснення передбачає наявність межі семантичного стиснення, подолання якої несе в собі загрозу втрати значної частини семантики тексту. Запропонований метод позбавлений цього недоліку і долає межу семантичного стиснення, в результаті чого отримується залишкова неподільна семантична складова, яка передбачає

наявність значно меншого відсотка початкового тексту порівняно з аналогами.

До того ж, ефективність використання розробленого методу обумовлена і низькою імовірністю виникнення похибки, яка полягає у виокремленні другорядної інформації, що має надлишковість і не є неподільною семантичною складовою (рисунок 2). Це позитивно вплине на вбудовування зворотного впливу.

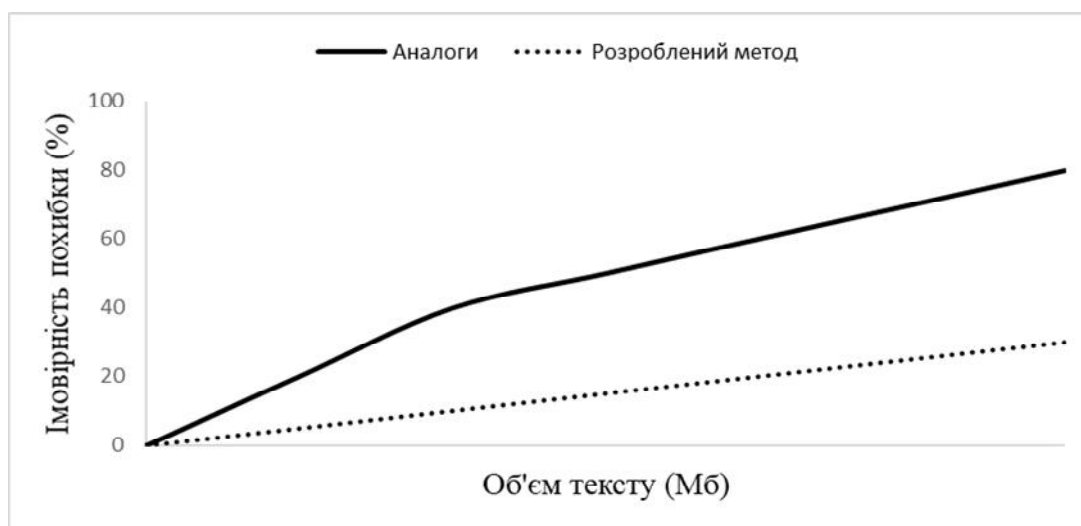


Рисунок 2 – Залежність імовірності виникнення похибки від об'єму початкового тексту

Недолік запропонованого методу полягає у значних ресурсах, які необхідні на проведення обчислень, а відповідно і час обчислень буде збільшуватись прямо пропорційно до зменшення виділених системі ресурсів.

Однак описані переваги та ефективність нівелюють цей недолік.

Висновки. Розроблений метод виявлення і збереження неподільної семантичної складової маніпулятивного тексту пропагандного

дискурсу дає можливість підвищити ефективність виявлення і збереження неподільної семантичної складової за рахунок використання квантового формалізму порівняно з методами семантичного стиснення, що базуються на інформаційній ентропії тексту. Метод дає змогу виокремити та зберегти ключові індивідуальні семантичні конструкції, притаманні психолінгвістичному портрету зловмисника, які неможливо змінити чи видалити без втрати базової семантики, для подальшого впровадження в них зворотного впливу.

Порівняння з аналогічними методами показує можливість подолання межі семантичного стиснення без втрати семантики, що значно зменшує результуючу множину семантичних категорій та підвищує ефективність виділення неподільної семантичної складової за рахунок зменшення кількості надлишкових і неточних семантичних складників.

Практичне значення має також формалізація методів визначення реми, що завдяки застосованому квантовому формалізму адаптована до використання в квантово-семантичному аналізі та в квантово-семантичній лінгвістичній стеганографії для протидії інформаційній пропаганді.

Результати, описані в статті, дадуть можливість провести вбудовування індивідуальної семантичної лінії пропагандиста в неподільну семантичну складову тексту з метою підвищення ефективності зворотного прихованого інформаційного впливу на зловмисника.

Список використаних джерел

- [1] Указ Президента України, Про рішення Ради національної безпеки і оборони України від 29 грудня 2016 р. "Про Доктрину інформаційної безпеки України". [Електронний ресурс]. Режим доступу: <https://zakon.rada.gov.ua/laws/show/47/2017#Text>. Дата звернення: Січ. 21, 2021.
- [2] Ya. Tarasenko, "The quantum-semantic psycholinguistic analysis method for the english-language text of propaganda discourse", *Advanced Information Systems*, vol. 3, no. 4, pp. 62-68, 2019.
- [3] Я. В. Тарасенко, "Метод комп'ютеризованого формування англомовного тексту відповідно до психолінгвістичного портрету пропагандиста", *Захист інформації*, т. 22, № 2, с. 66-73, 2020.
- [4] B. Hale, "Essence and definition by abstraction", *Synthese*, 2018. [Електронний ресурс]. Режим доступу: <https://link.springer.com/article/10.1007/s11229-018-1726-7>. Дата звернення: Лют. 2, 2021.
- [5] D. Ceglarek, "Semantic compression for text document processing", *Transactions on Computational Collective Intelligence XIV*, vol. 13, pp. 20-49, 2014.
- [6] Я. В. Тарасенко, О. Б. Півень, та І. М. Федотова-Півень, "Метод семантичного стиснення текстової інформації для протидії комп'ютерній лінгвістичній стеганографії", *Наука і техніка Повітряних Сил Збройних Сил України*, № 3 (32), с. 68-78, 2018.
- [7] Я. В. Тарасенко, та О. Б. Півень, "Особливості обчислення інформаційної ентропії тексту в умовах проведення атаки семантичним стисненням на лінгвістичну стегосистему", *Безпека інформації*, т. 24, № 2, с. 124-129, 2018.
- [8] M. Khedri, and S. F. Ebrahimi, "The essence of thematic structures in the academic translated texts", *Journal of Education and Practice*, no. 1, vol. 3, pp. 37-43, 2012.
- [9] L. Potter, "Ideological representations and theme-rheme analysis in English and Arabic news reports: a systemic functional approach", *Functional Linguistics*, 2016. [Електронний ресурс]. Режим доступу: <https://functionallinguistics.springeropen.com/articles/10.1186/s40554-016-0028-y>. Дата звернення: Лют. 3, 2021.
- [10] Ye. Yu. Savchenko, "Correlation of theme-rheme structure with dictum and modus in English and Ukrainian speech", *Science and Education a New Dimension. Philology*, no. 4 (27), iss. 107, pp. 15-18, 2016.
- [11] Р. А. Латыпов, "Квантовый формализм как инструмент описания пространств состояний концептов и сочетаний концептов", *Филологические науки. Вопросы теории и практики*, № 4 (58), ч. 1, с. 120-127, 2016.
- [12] И. А. Герасимова, *Формальная грамматика и интенциональная логика*. Москва, Россия: ИФ РАН, 2000.
- [13] Я. В. Тарасенко, "Метод контент-аналізу семантичної частки в текстах з ознаками психолінгвістичного впливу", *Системи*

- управління, навігації та зв'язку, т. 6, № 58, с. 92-96, 2019.
- [14] Я. В. Тарасенко, "Визначення координат семантичної частки в англomовному тексті при відомому психолінгвістичному портреті пропагандиста", *Захист інформації*, т. 21, № 3, с. 168-174, 2019.
- [15] Ya. Tarasenko, "Content-criteria of psycholinguistic portrait's semantic category for researching the group propaganda", *Ukrainian Scientific Journal of Information Security*, vol. 26, no. 1, pp. 5-13, 2020.
- ### References
- [1] *Decree of the President of Ukraine, On the decision of the National Security and Defense Council of Ukraine on December 29, 2016 "About the Doctrine of information security of Ukraine"*. [Online]. Available: <https://zakon.rada.gov.ua/laws/show/47/2017#Text>. Accessed on: Jan. 21, 2021 [in Ukrainian].
- [2] Ya. Tarasenko, "The quantum-semantic psycholinguistic analysis method for the english-language text of propaganda discourse", *Advanced Information Systems*, vol. 3, no. 4, pp. 62-68, 2019.
- [3] Ya. V. Tarasenko, "The method of English text's computerized formation in accordance with the propagandist's psycholinguistic portrait", *Zakhyst informatsii*, vol. 22, no. 2, pp. 66-73, 2020 [in Ukrainian].
- [4] B. Hale, "Essence and definition by abstraction", *Synthese*, 2018. [Online]. Available: <https://link.springer.com/article/10.1007/s11229-018-1726-7>. Accessed on: Feb. 2, 2021.
- [5] D. Ceglarek, "Semantic compression for text document processing", *Transactions on Computational Collective Intelligence XIV*, vol. 13, pp. 20-49, 2014.
- [6] Ya. V. Tarasenko, O. B. Piven, and I. M. Fedotova-Piven, "Method of textual information semantic compression for counteracting computer linguistic steganography", *Nauka i tekhnika Povitrianykh Syl Zbroinykh Syl Ukrainy*, no. 3 (32), pp. 68-78, 2018 [in Ukrainian].
- [7] Ya. V. Tarasenko, and O. B. Piven, "Features of calculating the information entropy of the text in case of attacking the linguistic stegosystem by semantic compression", *Bezpeka informatsii*, vol. 24, iss. 2, pp. 124-129, 2018 [in Ukrainian].
- [8] M. Khedri, and S. F. Ebrahimi, "The essence of thematic structures in the academic translated texts", *Journal of Education and Practice*, no. 1, vol. 3, pp. 37-43, 2012.
- [9] L. Potter, "Ideological representations and theme-rheme analysis in English and Arabic news reports: a systemic functional approach", *Functional Linguistics*, 2016. [Online]. Available: <https://functionallinguistics.springeropen.com/articles/10.1186/s40554-016-0028-y>. Accessed on: Feb. 3, 2021.
- [10] Ye. Yu. Savchenko, "Correlation of theme-rheme structure with dictum and modus in English and Ukrainian speech", *Science and Education a New Dimension. Philology*, no. 4 (27), iss. 107, pp. 15-18, 2016.
- [11] R. A. Latypov, "Quantum formalism as a tool to describe the state spaces of concepts and concept combinations", *Filologicheskiye nauki. Voprosy teorii i praktiki*, no. 4 (58), part 1, pp. 120-127, 2016 [in Russian].
- [12] I. A. Gerasimova, *Formal Grammar and Intensional Logic*. Moscow: IF RAN, 2000 [in Russian].
- [13] Ya. V. Tarasenko, "The content analysis method of semantic particle in texts with psycholinguistic influence signs", *Systemy upravlinnia, navihatsii ta zviazku*, vol. 6, no. 58, pp. 92-96, 2019 [in Ukrainian].
- [14] Ya. V. Tarasenko, "Determining the coordinates of semantic particle in English text with a known psycholinguist portrait of propagandist", *Zakhyst informatsii*, vol. 21, no. 3, pp. 168-174, 2019 [in Ukrainian].
- [15] Ya. Tarasenko, "Content-criteria of psycholinguistic portrait's semantic category for researching the group propaganda", *Ukrainian Scientific Journal of Information Security*, vol. 26, no. 1, pp. 5-13, 2020.

Ya. V. Tarasenko, Ph.D.
e-mail: yaroslav.tarasenko93@gmail.com
Cherkasy State Technological University
Shevchenko Blvd, 460, Cherkasy, 18006, Ukraine

**METHOD OF DETECTION AND PRESERVATION
OF INDIVISIBLE SEMANTIC COMPONENT OF ENGLISH MANIPULATIVE TEXT
OF PROPAGANDA DISCOURSE BASED ON QUANTUM FORMALISM**

The article studies the peculiarities of the use of quantum formalism in the problems of computer analysis and processing of manipulative texts. The developed method of detection and preservation of indivisible semantic component of the manipulative text of propaganda discourse makes it possible to increase the efficiency of detection and preservation of indivisible semantic component compared to semantic compression methods based on information entropy which leads to detecting and preserving the key individual semantic constructions inherent in the attacker's psycholinguistic portrait for further implementation of the reverse effect in them. The methods of quantum formalism and quantum-semantic research have been used to perform the tasks set in the article and to achieve the aim. The functioning of the system for determining the rheme of manipulative English text is modeled and the comparison with the results of the work of semantic compression systems is made, which proves the possibility of overcoming the semantic compression limit without losing semantics, which significantly reduces the resulting set of semantic categories and increases the efficiency of detecting the indivisible semantics components in comparison with the analogues by reducing the number of redundant and inaccurate semantic components.

Preliminary experimental data allow to assert the high efficiency of the developed method. The advantages and disadvantages of the developed method are outlined. The results described in the article will allow embedding the propagandist's individual semantic line into the text's indivisible semantic component in order to increase the effectiveness of the reverse hidden informational influence on the attacker.

Keywords: *theme definition, rheme definition, counteraction to propaganda, counteraction to manipulative texts, psycholinguistic portrait, quantum-semantic analysis.*

Стаття надійшла 04.02.2021

Прийнято 02.03.2021