© Tarasenko Ya. The software complex for attack the linguistic stegosystem // Ukrainian Scientific Journal of Information Security, 2018, vol. 24, issue 1, p. 56-61.

DOI: 10.18372/2225-5036.24.12388

ПРОГРАМНИЙ КОМПЛЕКС ПРОВЕДЕННЯ АТАКИ НА ЛІНГВІСТИЧНУ СТЕГОСИСТЕМУ

Ярослав Тарасенко

Черкаський державний технологічний університет, Україна

ТАРАСЕНКО Ярослав Володимирович

Рік та місце народження: 1993 рік, м. Черкаси, Україна.
Освіта: Черкаський державний технологічний університет, 2016 рік;
Посада: аспірант кафедри інформаційної безпеки та комп'ютерної інженерії.
Наукові інтереси: комп'ютерна лінгвістична стеганографія, математична та прикладна лінгвістика, інформаційна безпека, комп'ютерні системи.
Публікації: 7 наукових публікацій.
E-mail: yaroslav.tarasenko93@gmail.com

Анотація. *У статті проводиться огляд поширених методів і алгоритмів текстового стегоаналізу і реферування текстів та основаних на них автоматизованих комп'ютерних систем. Описуються їх недоліки для виконання задачі проведення атаки на лінгвістичну стегосистему шляхом семантичного стиснення з метою видалення прихованого повідомлення. Описується розробка та попереднє тестування програмного комплексу по виявленню та нейтралізації загроз, спричинених комп'ютерною лінгвістичною стеганографією. Розроблений програмний комплекс складається з 9 взаємопов'язаних модулів, 25 систем та 6 підсистем. В роботі описано їх взаємодію та наведено алгоритм роботи системи. Попередні результати тестування доводять ефективність проведення атаки з використанням запропонованої програми. Стегоповідомлення видаляється на 97%, коефіцієнт стиснення в середньому дорівнює 2,5 для окремих речень та 4,6 для суцільного тексту. Таким чином доводиться ефективність розробленого авторського методу стиснення текстової інформації для лінгвістичної стеганографії.*

Ключові слова*: стегоатака, програмний комплекс, лінгвістична стеганографія, протидія методам стеганографії, семантичне стиснення, стегоаналіз, текстова стеганографія, автоматизований стегоаналіз, лінгвістичні системи стегоаналізу, видалення стегоповідомлення.*

Вступ

На сьогоднішній день комунікація займає визначальне місце в житті кожної людини. Основуючись на твердженні, описаному в [1], комунікація є областю, що стрімко розвивається як в технічному чи політичному, так і в економічному і соціальному плані. Такий стан речей породжує відповідні негативні ефекти. Зі зростанням об'єму інформації, якою обмінюється все більше представників найрізноманітніших галузей зростає імовірність витоку секретної комерційної чи державної інформації на рівні з ризиком несанкціонованого обміну даними, що несуть в собі загрозу як у приватному так і у державному значенні. Це зумовлено використанням методів лінгвістичної стеганографії, які описані разом зі шляхами боротьби з ними у [2]. Таким чином, зростання кількості інформації зумовлює ефект інформаційного перевантаження, що на ряду з твердженням, що текстова стеганографія, хоч і є найбільш складною через наявність меншої кількості надлишкових даних у тексті, порівняно зі звуком чи зображенням [3, с. 3014], проте це робить її найбільш криптостійкою і це є підґрунтям для приховування інформації саме у тексті. Оскільки, найбільш поширені комп'ютерні системи та програмні комплекси, в тому числі у відкритому доступі на вітчизняному ринку направлені на виявлення слідів використання стеганографічних засобів у зображеннях, в меншій мірі в аудіо файлах, виникає потреба у створенні дієвих комплексів для проведення стегоаналізу, та атак на лінгвістичну стегосистему на основі цього аналізу, направленого на видалення прихованого стегоповідомлення у тексті зі збереженням семантичної структури та основної ідеї початкового тексту. Існує також значна необхідність у відповідних системах фільтрації, що могли б розпізнати неосмислений текст та проводити атаку на стегосистему, реалізовану з використанням штучно згенерованих відповідними автоматизованими засобами неосмислених текстів. Існуючі ж комп'ютерні системи проведення математичних атак потребують доопрацювання та адаптації з метою використання при атаці на лінгвістичну текстову стегосистему.

Аналіз досліджень та постановка завдання

Найбільш вдалим способом видалення стегоповідомлення у тексті є його семантичне стиснення. Перш за все, для виконання задачі проведення атаки на лінгвістичну стегосистему шляхом стиснення з метою видалення прихованого повідомлення необхідно провести стегоаналіз тексту. Одним із поши-

рених алгоритмів автоматизованого стегоаналізу є [5], де представлено статистичний алгоритм стегоаналізу текстів. Розробка базується на твердженні, що в стеготексті менше високочастотних слів, ніж у звичайному. В роботі [6] описано статистичний алгоритм стегоаналізу шляхом визначення статистичної характеристики кореляцій між службовими словами. В роботі [7] описується статистичний алгоритм, що оснований на дослідженні розподілу слів у текстовому сегменті. У роботі [8] запропоновано метод, що базується на використанні інформаційної ентропії в ролі статичної змінної. Усі згадані методи мають свої недоліки, що значно знижують ефективність використання їх з метою виконання поставленої задачі. Таким чином, при стегоаналізі описані системи характеризуються лише урахуванням статистики розподілу слів та незвичної кореляції. Проведення дослідження текстів невеликого об'єму сприяє підвищенню ефективності використання методів стеганографії та зниженню ефективності систем стегоаналізу на практиці. Недоліком методів є вузьконаправленість стегоаналізу та відсутність дій, призначених протидіяти засобам стеганографії та знищувати стегоповідомлення. Звідси витікає, що комп'ютерні системи, які використовують ці алгоритми для проведення атак на текстову стегосистему наслідують їх вразливості та недоліки. Оскільки більшість систем основуються на дослідженні статистично-частотних характеристик, це залишає невирішеним питання дослідження осмисленості тексту та прогнозування мети написання тексту. Більш ефективно в даному випадку використовувати автоматизований дискурсний аналіз. В свою чергу алгоритми та системи стиснення тексту також не позбавлені недоліків, що перешкоджають ефективному видаленню стегоповідомлення. Метод, описаний в роботі [9] направлений на виділення основного смислу тексту шляхом побудови вектору частот термінів у тексті. В методі [10] передбачається ймовірність того факту, що поточне речення буде частиною підсумку, що дає стан попереднього речення. Алгоритм [11] основується на методі [9], при цьому відмінністю є те, що подібність окремих речень обчислюється за центроїдою окремих векторів речення. У методі [12] скорочений текст є абстракцією вхідного тексту. Істотним недоліком цих методів і систем на їх основі є відсутність врахування можливості існування стегоповідомлення, що дозволяє зберегти значний його відсоток при використанні онтологічного підходу стеганографії, а отже повідомлення можливо відновити після стиснення початкового тексту. Таким чином, можна зробити висновок, що дуже затребуваними на даний час є автоматизовані системи проведення атак на лінгвістичну стегосистему шляхом семантичного стиснення зі збереженням початкової семантичної структури тексту та його основної думки. Цим зумовлена актуальність дослідження. Метою роботи є створення програмного комплексу на базі авторського методу стиснення текстової інформації для лінгвістичної стеганографії, що дозволяє проводити автоматизовану атаку на лінгвістичну стегосистему шляхом семантичного стиснення, яка базується на результатах стегоаналізу, що проводиться за допомогою дискурсного аналізу та елементів інтенсіональної логіки.

Основна частина дослідження

Для вирішення поставленої задачі було створено програмний комплекс, що реалізує авторський метод стиснення текстової інформації для лінгвістичної стеганографії та призначений експериментально підтвердити ефективність методу. Програмний комплекс має модульну архітектуру та об'єднаний в єдину систему. Основним його завданням є виявлення наявності слідів модифікації тексту засобами лінгвістичної стеганографії та внесення змін шляхом стиснення і модифікації тексту без втрати початкової семантичної структури та смислового навантаження для видалення можливого стегоповідомлення навіть без точних даних про його наявність. Модулі, системи та підсистеми комплексу, а також їхні зв'язки можна простежити на схематичному зображенні його архітектури (рис. 1).

Таким чином, програмний комплекс являє собою набір функціонально взаємопов'язаних структурних елементів: модуль налаштування системи (МНС), модуль взаємодії з користувачем (МВК), модуль підготовки вхідних даних (МПВД), модуль лінгвістичного аналізу тексту (МЛАТ), система морфологічного аналізу (СМА), система визначення імовірності використання морфологічних засобів стеганографії (СВІВМЗС), система синтаксичного аналізу (ССА), система визначення імовірності використання синтаксичних засобів стеганографії (СВІВСЗС), модуль оцінки осмисленості тексту та виділення основної думки (МООТВОД), система текстуального дискурсного аналізу (СТДА), система виділення теми та мікротем (СВТМ), система аналізу семантичної цілісності і завершеності (САСЦЗ), підсистема оцінки комунікативної цілісності (ПОКЦ), підсистема оцінки смислової цілісності (ПОСЦ), підсистема оцінки структурної цілісності (ПОСТЦ), система дослідження з використанням інтенсіональної логіки (СДІЛ), система оцінки осмисленості виразів (СООВ), система оцінки осмисленості тексту (СООТ), система визначення мети (СВМ), система інтертекстуального дискурсного аналізу (СІДА), модуль стиснення тексту (МСТ), система обчислення ентропії (СОЕ), система визначення надлишковості (СВН), система внутрішнього стиснення (СВС), підсистема видалення (ПВ), підсистема узагальнення (ПУ), підсистема заміни (ПЗ), система скорочення тексту частинами (ССТЧ), система заміни синонімів (СЗС), система стиснення неосмисленого тексту (ССНТ), модуль прийняття рішень (МПР), база даних (БД), модуль модифікації тексту (ММТ), система видалення інтервалів (СВІ), система виправлення морфологічних помилок (СВМП), система виправлення синтаксичних помилок (СВСП), модуль формування результуючого тексту (МФРТ), система формування морфологічно-синтаксичної структури (СФМСС), система формування семантичної структури (СФСС), система опису модифікацій (СОМ).

Робота з програмою здійснюється безпосередньо через модуль налаштувань системи.

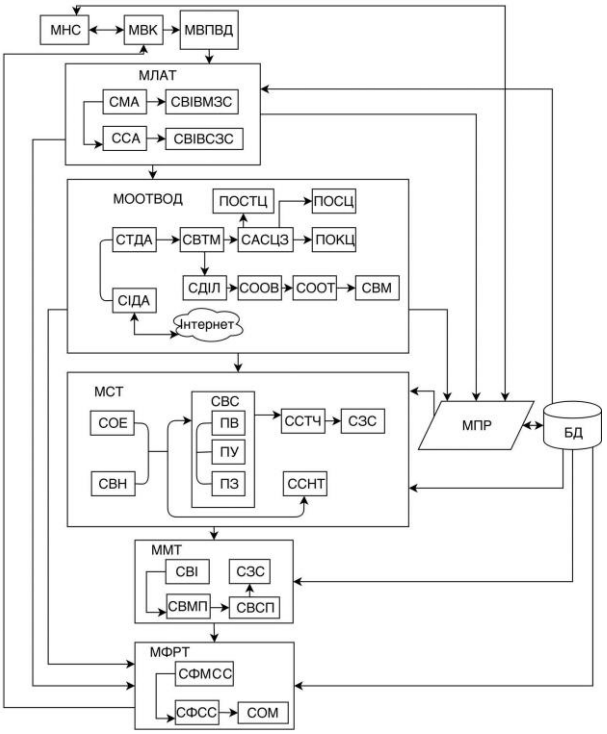

Рис. 1. Архітектура програмного комплексу

Він зчитує початкову інформацію, після чого відправляє її на обробку модулю підготовки вхідних даних, а також виводить оператору результати роботи програми, що направляє йому модуль формування результуючого тексту. Отримані налаштування і текст, який потребує семантичного стиснення передається модулю підготовки вхідних даних, який здійснює первинну обробку тексту. Після валідації їх оброблює модуль лінгвістичного аналізу тексту. Проводиться морфологічне дослідження, що визначає імовірність використання морфологічних засобів стеганографії. Результати передаються до системи синтаксичного аналізу для визначення імовірності використання синтаксичних засобів стеганографії. Крім того, результати роботи систем морфологічного і синтаксичного аналізу визначають морфологічно-синтаксичну структуру початкового тексту для формування результуючого, тому ці дані отримує модуль формування результуючого тексту та модуль прийняття рішень для побудови дерева відповідності та адекватного реагування на події. Дані морфологічно-синтаксичного дослідження є основою для роботи модуля оцінки осмисленості та виділення основної думки, який відповідає за два рівноправні напрями дослідження, що проводять системи текстуального (визначення семантики тексту) та інтертекстуального (потребує зв'язку з мережею Інтернет) дискурсного аналізу. В мережу відправляється запит, автоматично побудований на основі досліджуваного тексту, а з мережі завантажуються та аналізуються отримані результати. Дослідження морфологічної та синтаксичної структури має на увазі використання ряду словників, що зберігаються в базі даних. Система текстуального дискурсного аналізу є основою для послідовного виконання системи виділення теми та мікротем, системи аналізу семантичної цілісності і завершеності (складається з трьох підсистем оцінки: комунікативної, структурної, смислової цілісності) та системи дослідження з використанням інтенсіональної логіки, яка є основою для послідовного виконання трьох систем: оцінки осмисленості виразів, оцінки осмисленості тексту, визначення мети. Результати дослідження семантичної структури передаються до модуля формування результуючого тексту, а результати дискурсного дослідження до модуля прийняття рішень. Наступним виконується модуль стиснення тексту, який отримує дані з модуля прийняття рішень. На системі оцінки ентропії та системі визначення надлишковості основується два паралельні напрями: система внутрішнього стиснення та система стиснення неосмисленого тексту. Система внутрішнього стиснення складається з трьох паралельних підсистем: видалення, узагальнення і заміни та забезпечує вхідними даними дві інші системи: стиснення тексту частинами та заміни синонімів. Модуль модифікації тексту завантажує з бази даних відповідні словники для послідовного виконання чотирьох систем: видалення інтервалів, виправлення морфологічних помилок, виправлення синтаксичних помилок та заміни синонімів, після чого передає керування модулю формування результуючого тексту, який завантажує відповідну інформацію з бази даних, та виконує три системи: формування морфологічно-синтаксичної структури, формування семантичної структури та опису модифікацій, основою для виконання яких є результат роботи систем морфологічного і синтаксичного аналізу та передає дані до модуля взаємодії з користувачем. Модуль прийняття рішень у двосторонній формі взаємодіє з базою даних, куди записується інформація, отримана при виконанні модуля лінгвістичного аналізу тексту і модуля оцінки осмисленості тексту та виділення основної думки, звідки потім завантажується інформація, необхідна для прийняття рішень. В той час, як база даних забезпечує роботу чотирьох модулів: лінгвістичного аналізу, стиснення, модифікації та формування результуючого тексту, модуль прийняття рішень пов'язаний з модулем налаштувань системи. Із опису взаємодії модулів, систем та підсистем програмного комплексу витікає алгоритм його роботи, що можна зобразити за допомогою блок-схеми (рис. 2). Наведена принципова блок-схема відображає загальну концепцію роботи алгоритму.

Модульна структура за допомогою відповідних процедур та функцій забезпечує виконання того чи іншого етапу дослідження, перетворення або виводу результатів. Перш за все система отримує налаштування Parametr. Якщо вони не задані користувачем, то автоматично обираються стандартні налаштування stParametr. Після валідації, текст проходить обробку та перетворюється на масив даних, елементами якого є окремі лексичні одиниці чи слова. Після цього завантажується функція get_morphAn(), що відповідає за проведення морфологічного аналізу. Циклічно перевіряється кожне слово до останнього елементу масиву. Параметр partText[i]=1 означає наявність високої імовірності використання засобів стеганографії у досліджуваній частині тексту. Ці дані записуються до бази даних для подальшого використання функцією get_eval(),

що проводить обробку даних і прийняття рішень. За синтаксичне дослідження відповідає функція get_parse(), що циклічно викликається для аналізу кожної синтаксичної одиниці тексту до його закінчення. У випадку наявності ризику використання синтаксичних засобів стеганографії інформація записується до бази даних. Після цього керування системою передається функції get_eval(), яка після виконання передає управління функції get_disc(), що здійснює дискурсний аналіз тексту.

Після стиснення, змінній що відповідає за розділений текст присвоюється значення стисненого тексту, тобто aray(Text)=textCom, а керування системою передається процедурі get_change(). Після внесення відповідних модифікацій до тексту, змінній textCom присвоюється значення модифікованого тексту chTextCom, а керування передається процедурі формування остаточного тексту get_build().

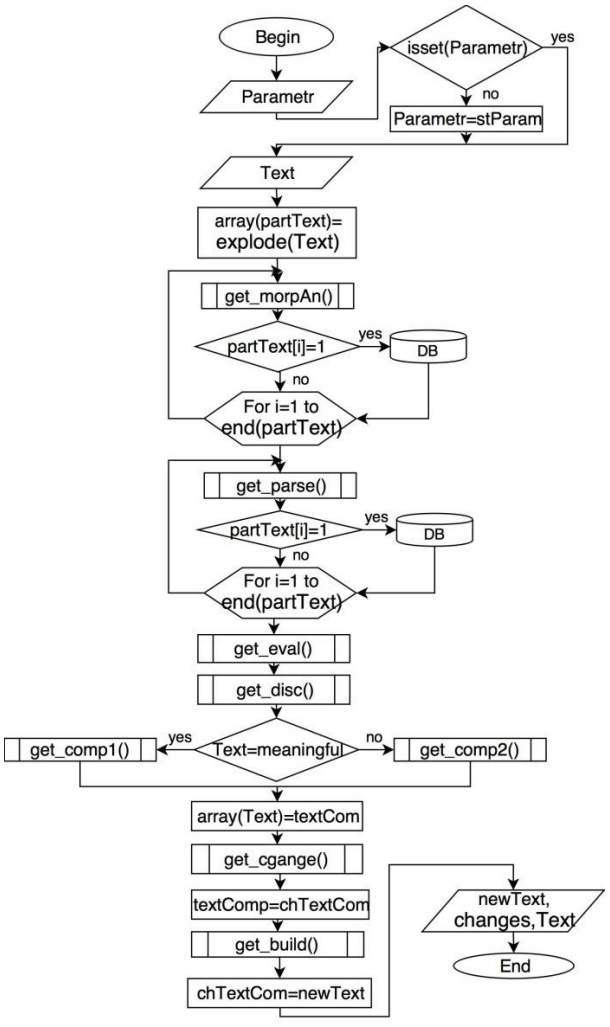

Рис.2. Принципова блок-схема алгоритму

Значення сформованого тексту присвоюється попередній змінній, таким чином chTextCom= =newText. Оператору системи на монітор чи носій інформації виводиться початковий текст Text, стиснений текст newText та описання всіх змін changes, після чого програма завершує роботу. Функціональні можливості розробленого програмного комплексу дозволяють стиснути будь-яку текстову інформацію для протидії методам комп'ютерної лінгвістичної стеганографії. Крім того, надається вибір початкових налаштувань системи.

Основною функцією програмного комплексу є семантичне стиснення тексту. При цьому наявність ознак осмисленості не є необхідною умовою стиснення. Програма ефективно проводить маніпуляції як зі звичайним текстом природньою мовою, так і зі штучно згенерованим неосмисленим текстом. Крім цього, надається можливість перегляду даних стегоаналізу тексту. Також система надає можливість модифікації та обробки текстових даних після проведення стиснення та вибір способу виведення результатів, а саме варіант запису на носій, виведення на екран чи збереження в одному з форматів .txt, .doc, .htm або .html. Ці ж формати підтримуються для завантаження початкових даних у систему в зв'язку з направленістю на отримання тексту, що передається чи зберігається у мережі Інтернет. Можливість первинного та вторинного налаштування забезпечує гнучкість системи в роботі з різними стилями тексту та форматами даних. Під первинним налаштуванням розуміється набір параметрів, що задається оператором перед проведенням стегоаналізу. Вторинне налаштування передбачає набір опцій та параметрів, що виявляються після стегоаналізу та впливають на процес стиснення та обробки даних. Для доведення ефективності та працездатності програмного комплексу було проведено попереднє тестування, матеріалом для якого стали 1000 сформованих текстів для перевірки, які поділялися за двома категоріями: метод приховування стегоповідомлення та об'єм тексту. Таким чином, попередні дані демонструють коефіцієнт стиснення, що в середньому дорівнює приблизно 2,5 для окремих речень та 4,6 для суцільного тексту. В той же час, стегоповідомлення було видалене в середньому на 97%, як видно з рисунку 3, де по осі X зображено об'єм текстового файлу, а по осі Y відсоток видаленого стегоповідомлення.

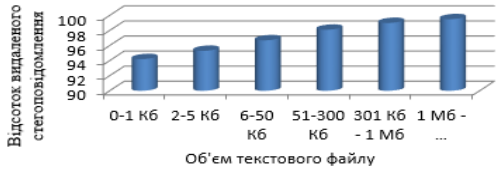

Рис. 3. Відсоток видаленого стегопоідомлення

Очікується, що подальше тестування уточнить отримані дані, проте статистика ефективності системи значно не зміниться. Очікуваний відсоток видаленого стегоповідомлення повинен знаходитись в межах 97 ±2 %, а коефіцієнти стиснення в межах 2,5 ± 0,3 для окремих речень та 4,6 ± 0,2 для тексту.

Висновок

Для виконання задачі проведення атаки на лінгвістичну стегосистему шляхом семантичного стиснення з метою видалення прихованого повідомлення було реалізовано програмний комплекс, що виявляє факт присутності слідів модифікації тексту засобами лінгвістичної стеганографії та вносить зміни у текст шляхом його стиснення і модифікації без втрати семантичної структури та смислового навантаження і видаляє можливе стегоповідомлення. Система складається із 9 взаємопов'язаних модулів, 25

систем та 6 підсистем. Попереднє тестування доводить ефективність та дієздатність програмного комплексу і можливість його використання на практиці. Головними перевагами системи є автоматизація процесу проведення атаки на текстову стегосистему, основану як на осмислених так і неосмислених текстах, широкий спектр дослідження при проведенні автоматизованого стегоаналізу та врахування можливості використання засобів стеганографії при проведенні реферування тексту. Недоліком програмного комплексу є значні затрати ресурсів комп'ютерної системи на проведення аналізу і обробки даних та вразливість системи до випадкового чи навмисного перевантаження і збільшення часу роботи при аналізі великих масивів даних. В подальших дослідженнях планується усунути виявлені недоліки та шляхом збільшення кількості та різноманітності тестових наборів отримати фінальні результати, що дадуть змогу зробити остаточні висновки щодо ефективності роботи системи, порівняти розроблену систему з існуючими аналогами та надати практичні рекомендації по її використанню.

Література

[1] K. Vineet, «The Changing World of Media & Communication», *Journal of Mass Communication & Journalism,*URL: https://www.omicsonline.org/open-access/the changing-world-of-media-and-communication21657912 .1000116.php ?aid=6473.

[2] І. Федотова-Півень, Я. Тарасенко, «Шляхи задоволення потреб сучасної кібербезпеки в рамках протидії методам комп'ютерної лінгвістичної стеганографії», *Безпека інформації*, №23(3), с. 190-196, 2017.

[3] R. Neha, J. Chaudhary, «Text Steganography Techniques: A Review», *International Journal of Engineering Trends and Technology,* Vol. 4, Issue 7, pp. 3013-3015, 2013.

[4] О. Бабина, «Лингвистическая стеганография: современные подходы. Часть 2», *Вестник ЮУрГУ. Серия: Лингвистика*, №4, с. 49-55, 2015.

[5] P. Meng, L. Hang, Z. Chen, Y. Hu, W. Yang, «STBS: A Statistical Algorithm for Steganalysis of Translation-Based Steganography», *12th International Conference «Information Hiding»*, Calgary, Canada, June 28-30, Vol. 6387, pp. 208-220, 2010.

[6] Z. Chen, L. Huang, Z. Yu, W. Yang, L. Li, X. Zheng, X. Zhao, «Linguistic Steganography Detection Using Statistical Characteristics of Correlations between Words», *10th International Workshop «Information Hiding»*, Santa Barbara, USA, May 19-21, Vol. 5284, pp. 224-235, 2008.

[7] Z. Chen, L. Huang, Z. Yu, L. Li, W. Yang, «A Statistical Algorithm for Linguistic Steganography Detection Based on Distribution of Words», *Third International Conference on Availability, Reliability and Security*, Barcelona, Spain, March 04-07, pp. 558 – 563, 2008.

[8] Z. Chen, L. Huang, Z. Yu, X. Zhao, X. Zhao, «Effective Linguistic Steganography Detection», *8th International Conference on Computer and Information Technology Workshops*, Sidney, Australia, July 08-11, pp. 224-229, 2008.

[9] Y. Gong, X. Liu, «Generic text summarization using relevance measure and latent semantic analysis», *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Informational Retrieval*, New Orleans, USA, September 9-12, pp. 19-25, 2001.

[10] J. Conroy, D. O'leary, «Text summarization via hidden markov models», *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Informational Retrieval*, New Orleans, USA, September 9-12, pp. 406-407, 2001.

[11] D. Radev, H. Jing, M. Stys, D. Tam, «Centroid-based summarization of multiple documents», *Information Processing & Management,* №40(6), pp. 919-938, 2004.

[12] R. Paulus, C. Xiong, R. Socher «A Deep Reinforced Model for Abstractive Summarization», May 2017, URL: https://arxiv.org/ab s/1705.04304.

УДК 003.26 (045)

Тарасенко Я. В. Программный комплекс проведения атаки на лингвистическую стегосистему

Аннотация. *В статье проводится обзор распространенных методов и алгоритмов текстового стегоанализа и реферирования текстов, как и основанных на них автоматизированных компьютерных систем. Описываются их недостатки для выполнения задачи проведения атаки на лингвистическую стегосистему путем семантического сжатия с целью удаления скрытого сообщения. Описывается разработка и предварительное тестирование программного комплекса по выявлению и нейтрализации угроз, вызванных компьютерной лингвистической стеганографией. Разработанный программный комплекс состоит из 9 взаимосвязанных модулей, 25 систем и 6 подсистем. В работе описано их взаимодействие и приведен алгоритм работы системы. Предварительные результаты тестирования доказывают эффективность проведения атаки с использованием предложенной программы. Стегосообщение удаляется на 97%, коэффициент сжатия в среднем равен 2,5 для отдельных предложений и 4,6 для сплошного текста. Таким образом, доказывается эффективность разработанного авторского метода сжатия текстовой информации для лингвистической стеганографии.*

Ключевые слова: *стегоатака, программный комплекс, лингвистическая стеганография, противодействие методам стеганографии, семантическое сжатие, стегоанализ, текстовая стеганография, автоматизированный стегоанализ, лингвистические системы стегоанализа, удаление стегосообщения.*

Tarasenko Ya. The software complex for attack the linguistic stegosystem

Abstract. *The article reviews the common methods and algorithms of textual steganalysis and text summarization, as well as automated computer systems based on them. It is described the disadvantages for accomplishing the task of the linguistic stegosystem attacking by semantic compression in order to remove a hidden message. The article describes the development and pre-testing of the software complex, based on the semantic compression of the text for attacking the linguistic stegosystem and taking into account the initial structure of this text. The development is aimed at detecting and neutralizing the threats caused by the use of computer linguistic steganography methods. Described program complex consists of 9 interconnected modules, 25 systems and 6 subsystems. Their interaction is highlighted and the system's algorithm is shown in the article. The pre-testing results on the basis of the material*

consisting of 1000 generated texts, divided into such two categories, as the stegomessage hiding method and the volume of text, prove the effectiveness of the attack using the proposed program. The stegomessage may be deleted by 97%, the average compression ratio is 2.5 for single sentences and 4.6 for solid texts. During the preliminary testing it was found the main advantages of the system, which include the automation the process of attacking a text-based stegosystem using both meaningful and meaningless texts and a wide range of research in stegoanalysis. Thus, the efficiency and practical capacity of the developed author's method, aimed at the textual data semantic compression for the purpose of counteracting the methods of linguistic steganography, has been proved. The further studies will be related to elimination the disadvantages detected during the preliminary testing and to confirming the preliminary conclusions of the work.

***Key words**: attacks on stegosystems, software complex, linguistic steganography, counteracting the methods of steganography, semantic compression, steganalysis, textual steganography, automated steganalysis, linguistic systems of steganalysis, removal of hidden message.*

Отримано *06 лютого 2018 року*, затверджено редколегією *14 березня 2018 року*