УДК 004.35

Б.Д. Будз, Ю.П. Дзюбінський
Національний університет «Львівська політехніка»
кафедра «Захист інформації»
79013, м. Львів, вул. С. Бандери, 12

ПРИХОВУВАННЯ ІНФОРМАЦІЇ В СЛУЖБОВИХ АТРИБУТАХ ФАЙЛІВ ТЕКСТОВОГО РЕДАКТОРА MS WORD

Розглянуті методи приховування інформації в мультимедійні, графічні і текстові файли. Досліджено службова область файлу текстового редактора MS Word, з метою визначення пропускної здатності каналу витоку інформації. Розроблено програмні макроси, для аналізу текстового файлу на предмет наявності прихованих даних.

Methods of hiding information in multimedia, graphic and text files have been studied in this work. MS Word test editor settings options have been analyzed with the aim to determine the capacity of canal to information leakage. Macro programs have been designed in order to analyze test files on the evidence of hidden data.

Актуальність. На сьогодні, більшість інформації обробляється в автоматизованих системах, так званих персональних комп'ютерах. В основному, дані накопичуються, обробляються і передаються в текстових файлах різних форматів. Найпоширенішим в Україні, та й у світі, програмним продуктом, який дозволяє обробляти інформацію в текстовому форматі є текстовий редактор MS Word пакету MS Office від компанії Microsoft. Даний редактор використовується більшістю приватних і державних установ для формування текстів різноманітних наказів, звітів, листів, тощо, які можуть бути як таємними, так і відкритими для загалу. Процес створення і циркуляція конфіденційних текстових файлів в установах контролюється спеціальними службами і спеціальними програмами, які унеможливлюють несанкціоноване потрапляння конфіденційних даних до сторонніх осіб. Натомість циркуляція файлів з вмістом неконфіденційного характеру, практично не контролюється і працівники можуть безперешкодно виносити на різноманітних носіях і надсилати на ззовні файли, вміст яких не містить конфіденційних даних.

Зазвичай основне, на що звертається увага при оцінці конфіденційності вмісту файлу, є дані, які безпосередньо можна проглянути за допомогою відповідних редакторів, при цьому, службові дані, які зазвичай можна переглянути в пункті меню «Властивості», переважно не перевіряються. Дані, які розміщуються в пункті меню «Властивості» містять службову інформацію, яку можна редагувати користувачем і спеціальними програмними засобами, яка є факультативною і не впливає на зміст інформативної частини вмісту документу. Ця область може використовуватись для розміщення конфіденційних даних, з подальшим переміщенням файлів за межі організацій, що може призвести до несанкціонованого витоку конфіденційних даних. Постає потреба у створенні спеціальних програмних продуктів, для забезпечення можливості перевірки службових полів в автоматичному режимі, з метою недопущення витоку конфіденційних даних.

Мета і задачі дослідження. Метою роботи є аналіз пропускної здатності технічного каналу, який дозволяє здійснювати приховування інформації в областях службових записів текстового редактора MS Word. На основі аналізу, створити програмні макроси, для аналізу і виявлення прихованих записів в службових областях.

Комп'ютерні методи приховування інформації у файлах даних. Найпоширенішими комп'ютерними файлами є: мультимедійні (аудіо, відео), графічні і текстові. Кожен тип файлів має свою структуру, яка і диктує алгоритми приховування даних, використовуючи ті чи інші властивості. Наведемо приклади приховування інформації у файлах, які були перелічені вище.

1. *Приховування інформації в аудіофайлах*. Найпростішим методом приховування інформації, є заміщення найменш значущих бітів (усіх або деяких) на біти прихованого повідомлення. Алгоритм, запропонований в роботі [1], задовольняє більшості з вимог, що пред'являються, він впроваджується в аудіосигнали (послідовність 8 - або 16-бітових відліків) шляхом незначної зміни амплітуди кожного відліку. Для виявлення даних не вимагається початкового аудіосигналу. Метод, який був запропонований В. Бендером, Н. Морімото та ін. пропонує використовувати слабку чутливість системи слуху людини до незначних змін фази сигналу. Метод вбудовування інформації за рахунок зміни часу затримки ехо-сигналу дозволяє записувати конфіденційні дані в сигнал прикриття, змінюючи параметри відлуння сигналу. Метод маскування сигналу використовує не лише особливості будови аудіосигналів, але і системи слуху людини, в цьому методі слабке, але чутне звукове коливання стає нечутним за наявності іншого більш гучного (сигнал маскування).

2. *Приховування інформації у відеофайлах***.** Найбільш популярними стандартами кодування відео є MPEG-2 і MPEG-4. Наведемо методи приховування інформації у відео, що стискується за стандартом MPEG-2. Алгоритм стискування MPEG базується на гібридній схемі кодування [2]. Ця схема об'єднує міжкадрове і внутрішньо кадрове кодування послідовності відеоданих. В методі вбудовування інформації на рівні коефіцієнтів [3], здійснюється додавання псевдовипадкового масиву до DC-коефіцієнтів відео, стислого за стандартом MPEG. В процесі вбудовування безпосередньо беруть участь тільки значення яскравості в I-кадрах. Метод вбудовування інформації на рівні бітової площини базується на методі приховування інформації в найменш значущий біт нерухомих зображень. Цей метод відрізняється високою пропускною здатністю і невеликою обчислювальною складністю для даних, стиснених за стандартом MPEG. Метод приховування інформації завдяки енергетичній різниці між коефіцієнтами поєднує в собі переваги методів, що працюють з початковим і стислим відео. В його основі лежить диференціальне вбудовування енергії даних.

3. *Приховування інформації в графічних файлах*. Більшість досліджень в сфері приховування інформації (стеганографії) присвячена використанню в якості стегоконтейнерів зображення. Це обумовлено наступними причинами:

- існуванням практично значущого завдання захисту фотографій, картин, відео від незаконного тиражування і поширення;
- відносно великим об'ємом цифрового представлення зображень, що дозволяє впроваджувати інформацію великого об'єму або підвищувати відновлення впровадження;
- заздалегідь відомим розміром контейнера, відсутністю обмежень, що накладаються вимогами реального часу;
- наявністю в більшості реальних зображень областей текстур, що мають шумову структуру і добре відповідних для вбудовування інформації;
- слабкою чутливістю людського ока до незначних змін кольорів зображення, його яскравості, контрастності, змісту в нім шуму, спотворенням поблизу контурів;
- добре розробленими останнім часом методами цифрової обробки зображень.

Щодо зображень, то найбільш поширеними є приховування даних в просторовій області і в області перетворення.

4. *Приховування інформації в текстових файлах*. Для приховування інформації в текстових файлах, використовуються синтаксичні, лексичні методи і мімікрія. Щодо синтаксичних методів, то тут конфіденційна інформація найчастіше кодується шляхом зміни кількості пропусків, використовуючи невидимі символи, регістру букв, шляхом зміни міжрядкових інтервалів, табуляцій, тощо. Лексичні методи базуються на лексичній структурі тексту, коли в одне слово, кодується, наприклад, два біти інформації, таким чином слова завдовжки в 4 і 8 символів можуть означати комбінацію біт "00", завдовжки в 5 і 9 - "01", 6 і 10 - "10", 7 і 11 букв - "11". Слова, коротше 4 і довше 11 букв, можна вставляти де завгодно для лексичної і граматичної зв'язки слів в пропозиції - програма-декодер буде просто ігнорувати їх. Мімікрія генерує осмислений текст, використовуючи синтаксис, описаний в Context Free Grammar (CFG), і вбудовує інформацію, вибираючи з CFG певні фрази і слова.

Аналіз службових полів опції «Властивості» текстового документі MS Word. Опція «Властивості» документу MS Word розділяються на дві групи: постійні та змінні.

Постійних властивостей є дев`ять. Вікно постійних властивостей документу MS Word наведено на рис .1. Кількість змінних полів власник документу може встановлювати сам, користуючись переліком, в який входить 27 пунктів, або створювати свої. Вікно змінних властивостей документу MS Word наведено на рис 2.

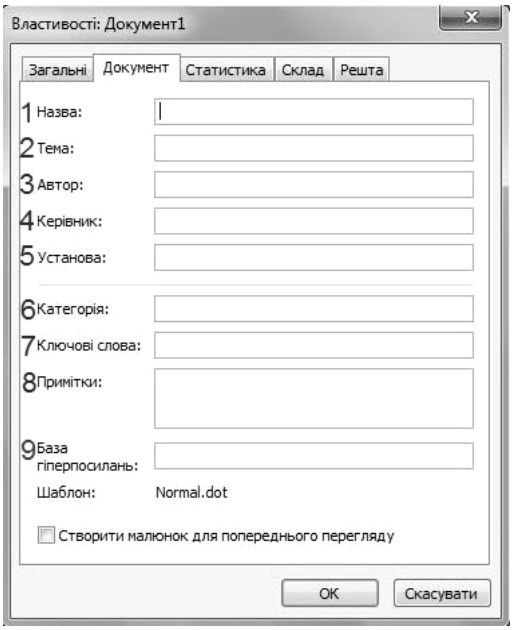

Рис. 1 Вікно постійних властивостей документу MS Word

Рис. 2 Вікно змінних властивостей документу MS Word

Перші вісім постійних властивостей вміщують 32767 символів тексту, включаючи пробіли та розділові знаки, останнє поле властивостей, яким є поле «база гіперпосилання» вміщує лише 2048 символів. Якщо прийняти до уваги, що один символ тексту займає один байт, то можна зробити висновок, що пропускна здатність такого каналу є обмеженою і становить 264184 байт, або ≈ 258 кілобайт.

Кожне поле змінних властивостей файлу вміщує 255 символів, або $\approx 0{,}25$ кілобайт що є набагато менше ніж місткість полів постійних властивостей. Але для прихованої передачі інформації змінні властивості підходять краще, тому, що їх кількість не є обмеженою.

Створення програмних макросів. MS Office пропонує два способи створення макропрограм: безпосередньо введення тексту (процедури VBA) в редакторові VBA або застосування вбудованого засобу запису. Перший варіант розрахований на досвідчених користувачів, програмістів або інтеграторів і дозволяє створювати повнофункціональні надбудови, функції і модулі. Другий, навпаки, надзвичайно простий і призначений для початківців, який не вимагає знання принципів програмування і синтаксису команд VBA.

Написання макросу для внесення даних в поля властивостей файлу («Введення»). Як контейнер для приховування інформації будемо використовувати поля змінних властивостей. Далі позбуваємось переносів рядків і розривів сторінок. Після того ми ділимо текст на блоки по 256 символів, враховуючи останній пробіл і створюємо нові поля властивостей, їх кількість буде такою, яка потрібна, щоб вмістити текст. Результат роботи макросу наведений на рис. 3.

Написання макросу для виведення даних з полів властивостей файлу («Виведення»). Макрос дозволить вивести інформацію з полів властивостей файлу. Інформація буде виводитись після тексту документу. Для цього створено цикл, що дозволить послідовно вилучати дані з полів змінних

властивостей. Наступним кроком буде повернення початкового форматування тексту. Для цього необхідно відшукати перенос рядка і замінити його пробілом, відшукати подвійний пробіл і замінити його переносом рядка, віднайти потрійний пробіл і замінити його розривом сторінки. Результат роботи макросу наведений на рис. 4.

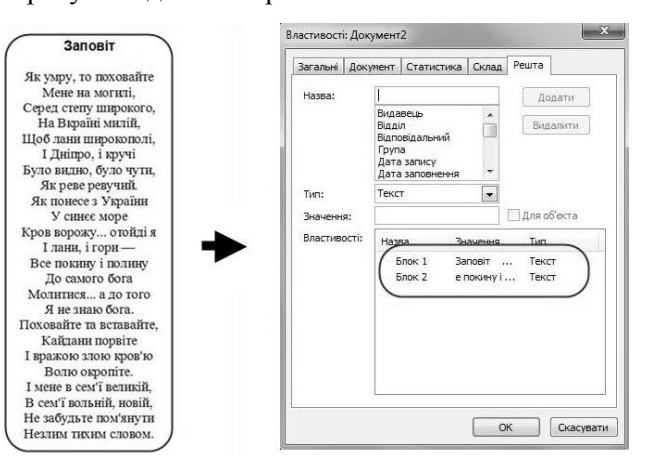

Рис. 3. Результат роботи макросу «Введення».

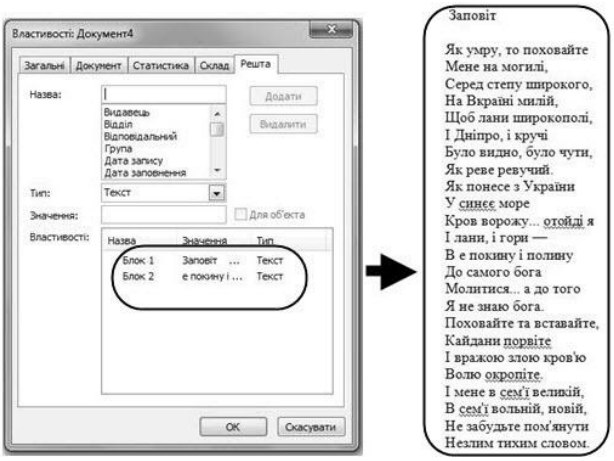

Рис. 4. Результат роботи макросу «Виведення».

Написання макросу для аналізу службових полів («Перевірка»). Тут інформація буде виводитись після тексту документу. Для цього створюємо цикл, що дозволить послідовно вилучати дані з полів постійних властивостей, а також цикл, що дозволить послідовно витягувати дані з полів змінних властивостей. Для покращення читабельності здійснюємо форматування тексту. Знаходимо перенос рядка і заміна його пробілом, заміна двох чи більше пробілів між словами на один пробіл, останнім кроком буде збереження звіту, що містить виведену інформацію для його подальшої перевірки, місце зберігання вказується в тексті макросу. Результат роботи макросу наведений на рис. 5.

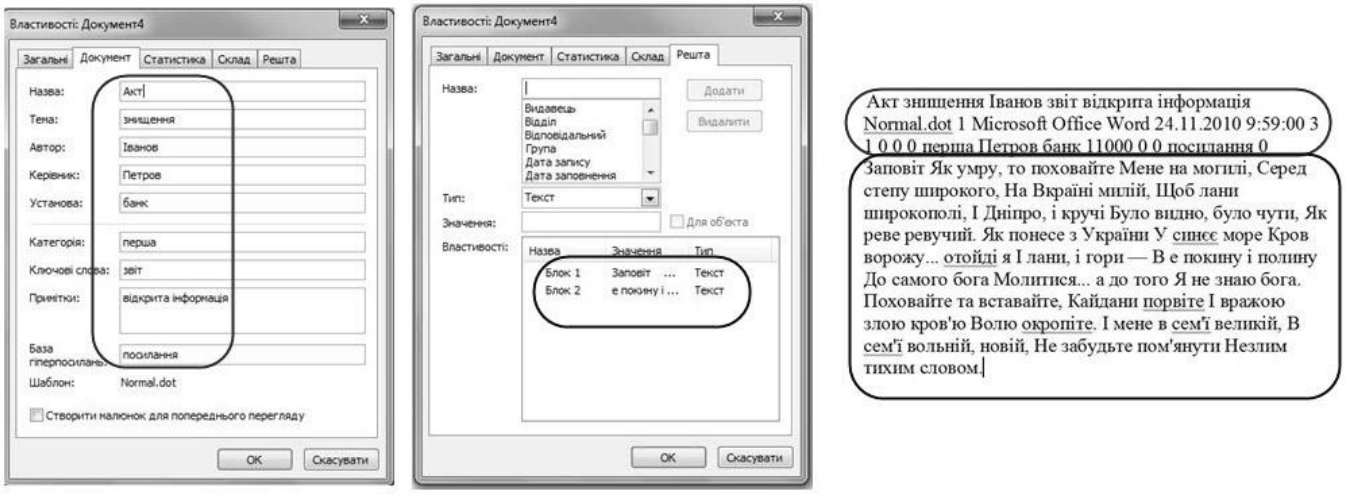

Рис. 5. Результат роботи макросу «Перевірка»

Написання макросу для видалення записів в службових полях («Видалення»). Для видалення записів в службових полях необхідно виділити текст документу і скопіювати його в новий пустий документ. Для видалення постійних властивостей чистого документу будемо реалізована наступна команда: ActiveDocument.BuiltInDocumentProperties(wdPropertyTitle) = "". Для збереження документу використовуємо команду: ActiveDocument.Save. Результат роботи макросу наведений на рис. 6.

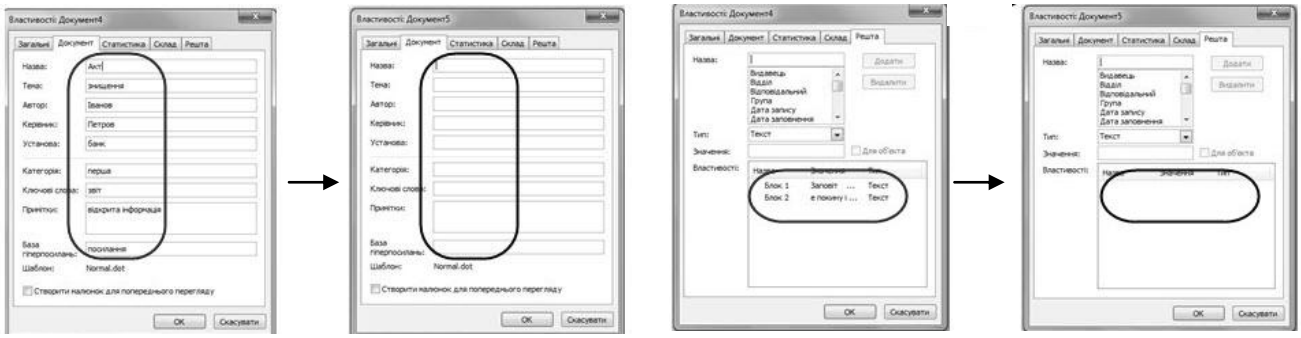

Рис. 6. Результат роботи макросу «Видалення»

Висновки. Питання прихованої передачі конфіденційної інформації в текстових файлах є актуальним на сьогодні, оскільки щодня генерується велика кількість таких файлів, які надсилаються за призначенням по незахищеним каналам зв'язку. В роботі було розглянуто методи прихованого передавання інформації в загальному, та безпосередньо в комп`ютерних файлах даних. Здійснено огляд методів приховування інформації у комп`ютерних файлах. Були розглянуті аудіо, відео, графічні та текстові файли. Також були розглянуті передумови для приховування інформації і особливості цих типів файлів, що можуть використовуватися для приховування інформації.

Було розглянуто службову область файлів текстового редактора MS Word і можливість використання її в якості контейнера для приховування інформації. Було визначена пропускна здатність такого каналу. Для аналізу службових полів були застосовані макропрограми написані в редакторі Visual Basic що є складовою частиною пакету програм Microsoft Office. Їхньою перевагою є відносна простота написання і те що для їхнього написання не потрібно використовувати програми сторонніх розробників, що часто бувають платними. Недоліком такого методу приховування є те, що розмір вихідного файлу змінюється в залежності від розміру інформації, що введена в службовій області файлу і також те що інформація для введення повинна бути текстовою.

Також можна стверджувати, що на відміну від технічних каналів, приховані канали функціонують до моменту їх виявлення, який може складати досить тривалий час. Пропускна здатність прихованих каналів є досить низькою. Для виявлення цих каналів необхідно здійснювати технічні і організаційні заходи.

Список використаної літератури

1. Bassia P., Pitas I., Robust audio watermarking in the time domain // Department of Informatics, University of Tressaloniki. 2. Koch E., Zhao J. Towards Robust and Hidden Image Copyright Labeling // IEEE Workshop on Nonlinear Signal and Image Processing. 1995. P. 123-132. 3. Wu T., Wu S. Selective encryption and watermarking of MPEG video // International Conference on Image Science, Systems, and Technology. 1997. 4. «Використання макросів в Microsoft Office» http://itc.ua/articles/ispolzovanie_makrosov_v_microsoft_office_28520. 4. «Цифрова стеганографія», В. Г. Грибунин, И. Н. Оков, И. В. Туринцев, 2002. 5. Стеганографічні методи захисту інформації», С. В. Ярмолик, Ю. Н. Листопад.