BADANIE METODY STEGANOGRAFICZNEJ NA PODSTAWIE PRZEKSZTAŁCENIA NAJMNIEJ ZNACZĄCYCH BITÓW

Cel: Zbadanie metody steganograficznej wbudowywania/wyodrębniania tajnych informacji za pomocą pliku elektronicznego kontenera opartego na transformacji najmniej znaczących bitów (LSB), zdobycie praktycznych umiejętności implementacji tej metody.

Zadania:

1. Umacnianie teoretycznej wiedzy z zakresu steganograficznej transformacji informacji, modelowania steganosystemów, klasyfikacji i istoty metod cyfrowej steganografii.

2. Zbadanie algorytmu wbudowywania/wyodrębniania tajnych informacji na podstawie metody LSB (Least Significant Bit), zdobycie doświadczenia w praktycznej implementacji metody.

3. Opracowanie aplikacji do realizacji algorytmu wbudowywania/wyodrębniania tajnych informacji za pomocą pliku elektronicznego kontenera opartego na metodzie LSB.

4. Zapoznanie się z metodami oceny odporności steganograficznej metody LSB.

5. Wyniki przeprowadzonej pracy laboratoryjnej należy przedstawić w formie opisu opracowanej aplikacji, metodyki przeprowadzania eksperymentów z użyciem aplikacji oraz wyników eksperymentu.

12.1.1. Podstawowe definicje, klasyfikacja i istota metod steganograficznych.

Definicja 1. System steganograficzny - zbiór środków i metod wykorzystywanych do tworzenia ukrytego kanału transmisji (lub przechowywania) informacji. Ukryty kanał jest organizowany na bazie i wewnątrz kanału otwartego, wykorzystując cechy percepcji informacji. "Ukrycie" kanału transmisji tajnych informacji odróżnia steganografię od kryptografii: w pierwszym przypadku tajne jest samo istnienie kanału (transmisji informacji).

Definicja 2. Abstrakcyjny system steganograficzny zazwyczaj jest definiowany jako pewien zbiór odwzorowań jednej przestrzeni (zbioru możliwych komunikatów M) w inną przestrzeń (zbioru możliwych komunikatów steganograficznych S), i odwrotnie. Podstawowe składniki systemu steganograficznego:

- Kontener C (plik-kontener lub elektroniczny dokument dowolnego formatu), w którym znajduje się (osadzana, ukrywana) tajna wiadomość M; to właśnie kontener jest wspomnianym ukrytym kanałem;

- Tajna wiadomość M, osadzana w kontenerze w celu transmisji lub przechowywania (na przykład w celu udowodnienia lub ochrony praw autorskich do dokumentu-kontenera [2, 53–56]; może to dotyczyć niewidocznych cyfrowych znaków wodnych (CZV));

- Klucze lub informacje kluczowe, K systemu, wykonujące tę samą funkcję, co klucze kryptograficzne; kluczy może być kilka, zgodnie z tym nowoczesne steganosystemy charakteryzuje się jako wielokluczowe: jeden klucz identyfikuje metodę wbudowywania/wyodrębniania tajnej informacji, inny - wybór elementów (na przykład bitów) kontenera do jego modyfikacji podczas osadzania tajnej informacji, trzeci - do wstępnego (przed wbudowaniem) przetwarzania tajnej informacji (na przykład na podstawie kodowania odpornego na zakłócenia, kompresji lub szyfrowania) itd.;

- Kontener z wbudowaną wiadomością, czyli steganokontener, S, który jest przesyłany przez otwarty kanał, również będący ważnym składnikiem analizowanego systemu; steganokontener będzie również nazywany steganokomunikatem;

• dla kompletności wymienimy również podmioty systemu: nadawcę i odbiorcę.

W zależności od formatu dokumentu-kontenera cyfrową (lub komputerową) steganografię dzieli się na klasy:

• audiosteganografia;

• wideo steganografia;

• graficzna steganografia;

• tekstowa steganografia;

• i inne.

Definicja 3. Systemem steganograficznym ∑ nazwiemy zbiór komunikatów M, kontenerów C, kluczy K, steganokomunikatów (wypełnionych kontenerów) S i transformacji (prostych F i odwrotnych F–1), które je łączą:

∑ = (M, C, K, S, F, F–1).

Jak widać, istotą badanego systemu jest tajne przechowywanie lub przesyłanie jednej informacji w innej informacji, która jest jawna.

W związku z tym, przy projektowaniu systemu steganograficznego należy uwzględnić następujące podstawowe zasady:

• właściwości kontenera muszą być modyfikowane w taki sposób, aby zmiany nie mogły być wykryte wizualnie; to wymaganie określa jakość ukrycia osadzonej wiadomości: aby zapewnić swobodną transmisję steganokomunikatu przez kanał komunikacyjny, nie powinien on w żaden sposób przyciągać uwagi atakującego;

• przeciwnik (intruz) posiada pełną wiedzę na temat systemu steganograficznego i szczegółów jego implementacji; jedyną informacją, która pozostaje mu nieznana, jest klucz, za pomocą którego tylko jego posiadacz może potwierdzić obecność i zawartość ukrytej wiadomości;

• jeśli przeciwnik w jakiś sposób dowiaduje się o istnieniu ukrytej wiadomości, nie powinno to umożliwić mu wydobyć podobnych wiadomości, dopóki klucz pozostaje tajny;

• potencjalny przeciwnik powinien być pozbawiony jakichkolwiek technicznych i innych przewag w rozpoznawaniu lub ujawnianiu treści tajnych wiadomości.

12.1.2. Metoda Najmniej Znaczącego Bitu (NZB) i cechy jej realizacji

Większość badań w dziedzinie steganografii poświęcona jest wykorzystaniu obrazów jako steganokontenerów (tekst można również traktować jako obraz). Jest to podyktowane następującymi powodami:

• relatywnie dużą ilością cyfrowego reprezentowania obrazów, co umożliwia osadzanie dużych ilości danych;

• z góry znanym rozmiarem kontenera, brakiem ograniczeń narzucanych przez wymagania czasu rzeczywistego;

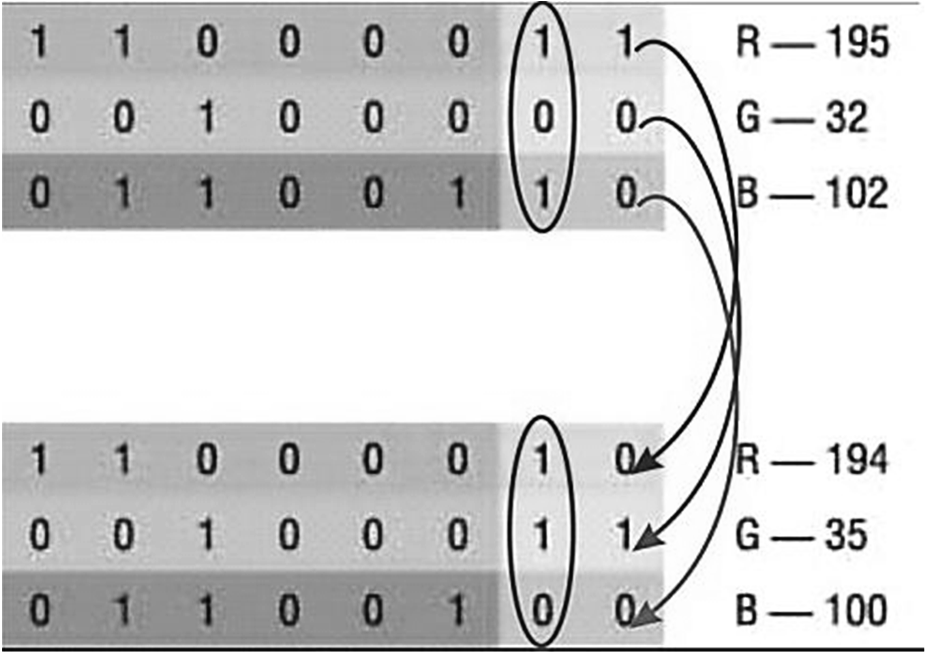
• obecnością w większości realnych obrazów obszarów teksturalnych, posiadających szumową strukturę i dobrze nadających się do osadzania informacji;

• słabą wrażliwością ludzkiego oka na niewielkie zmiany kolorów obrazu, jego jasności, kontrastu, zawartości szumu oraz zniekształceń wokół konturów;

• dobrze opracowanymi w ostatnim czasie metodami cyfrowej obróbki obrazów.

Metoda NZB opiera się na ograniczonych zdolnościach ludzkiego wzroku lub słuchu, w wyniku czego ludzie mają trudności w rozróżnieniu niewielkich wariacji koloru lub dźwięku. Przyjrzyjmy się temu na przykładzie 24-bitowego obrazu rastrowego RGB. Jak wiadomo, każdy piksel jest kodowany za pomocą trzech bajtów. Każdy bajt określa intensywność koloru czerwonego (Red), zielonego (Green) i niebieskiego (Blue). Zestawienie intensywności koloru w każdym z trzech kanałów określa odcień piksela.

Wyobraźmy sobie piksel jako trzy bajty w postaci bitowej, jak pokazano na rysunku 1.



Rys.1. Przykład ilustrujący zasadę implementacji metody LSB

Najmłodsze bity (wyróżnione na jasny kolor, po prawej stronie) mają niewielki "udział" w obrazie w porównaniu do starszych.

Zastąpienie jednego lub nawet kilku najmłodszych bitów będzie prawie niezauważalne dla ludzkiego oka, ponieważ człowiek faktycznie może odróżnić około stu pięćdziesięciu odcieni kolorów.

Jednym z najprostszych i najbardziej zrozumiałych formatów do rozwiązania naszych zadań jest format BMP (Bitmapa) - jeden z formatów reprezentacji grafiki rastrowej. Obraz jest reprezentowany jako macierz pikseli, gdzie każdy punkt charakteryzuje się trzema parametrami: współrzędnymi x i y oraz kodem koloru w oparciu o model RGB. Wszystkie operacje graficznego wejścia-wyjścia na ekranie monitora (drukarka i niektóre inne urządzenia) ostatecznie są wykonywane w tym formacie. Dla pracy z tym formatem w systemie Windows istnieje wiele specjalnych funkcji i struktur API, które pomagają wykonywać wszystkie niezbędne operacje na dość wysokim poziomie logicznym.

Kontenery oparte na formacie BMP można podzielić na dwie klasy: czyste i zaszumione. W pierwszych jest widoczna zależność między młodszymi a pozostałymi bitami elementów koloru, a także między samymi młodszymi bitami. Osadzenie komunikatu w takim kontenerze narusza te zależności, co łatwo wykryć przez analityka. Jeśli jednak obraz jest zaszumiony (na przykład uzyskany ze skanera lub aparatu fotograficznego), to trudniej jest określić osadzone w nim komunikaty. W związku z tym jako kontenery do metody LSB zaleca się używanie plików, które nie zostały pierwotnie utworzone na komputerze.

Innym z formatów rastrowych używanych w steganografii jest format PNG (Portable Network Graphics). Pod względem jakości wyświetlania kolorów ten format przewyższa formaty JPEG (Joint Photographic Experts Group) i GIF (Graphics Interchange Format), ale rozmiar pliku będzie o 30–40% większy.

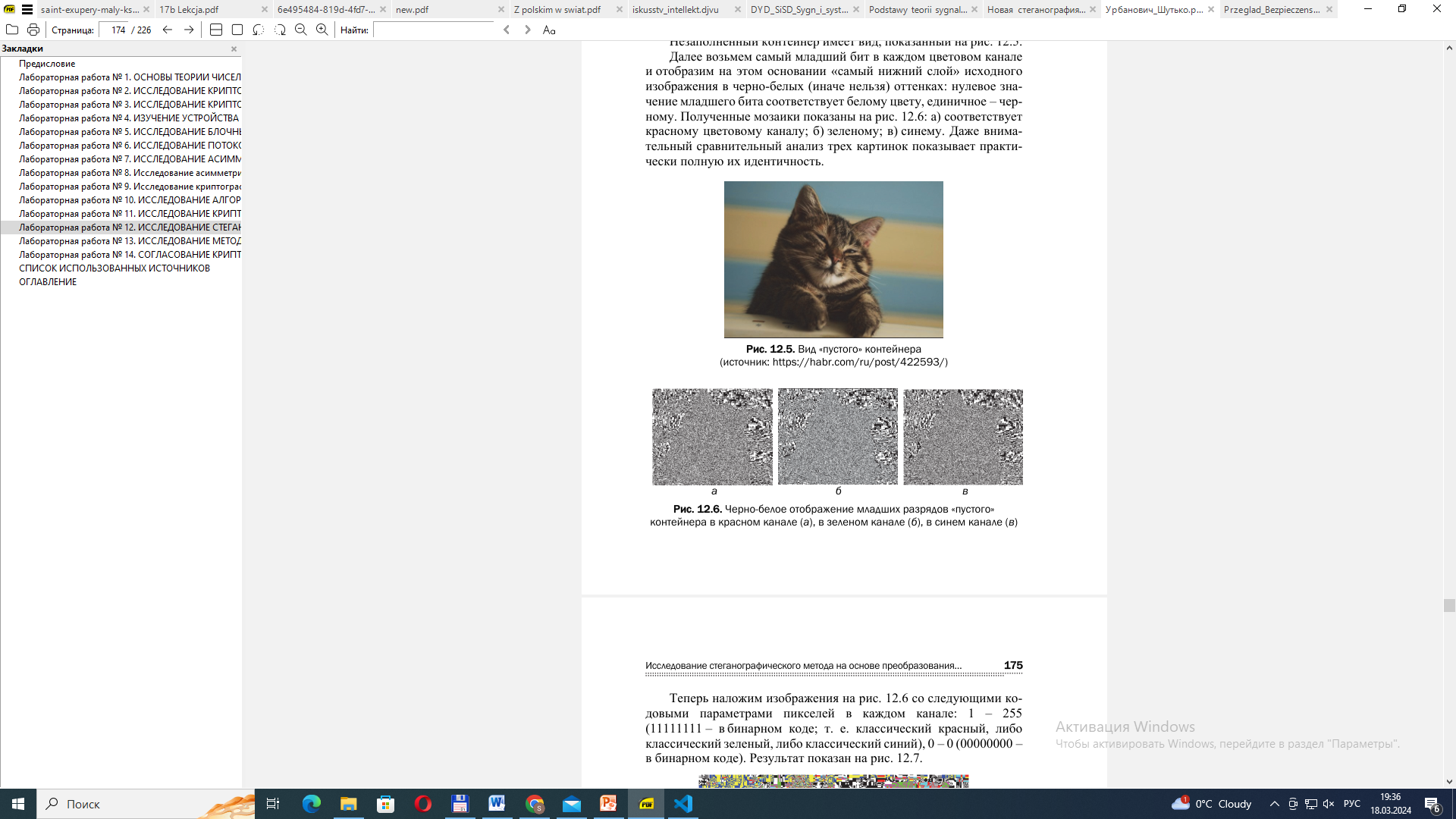
Przedstawione powyżej dane tabelaryczne i ilustracyjne, oraz doświadczenie specjalistów, pokazują, że nawet przy modyfikacji 3–4 najmłodszych bitów stan steganokontenera graficznego nie budzi podejrzeń u ekspertów podczas wizualnej kontroli. Na podstawie takiej oceny należy porównywać wielkość osadzonej wiadomości VM z wielkością VC używanego kontenera. Na przykład, jeśli rozmiar obrazu to 500×500 = 250 000 pikseli, to uwzględniając używany model 3-kolorowy, mamy 750 000 jednostek kolorowych współrzędnych. Jeśli planujemy modyfikować tylko najmłodsze bity wszystkich kanałów kolorów macierzy, to maksymalny rozmiar osadzonej wiadomości (VMmax) nie powinien przekraczać 750 000 bitów.

Następnie przyjrzymy się niektórym szczegółom technicznym i cechom implementacji metody LSB przy użyciu obrazu w formacie PNG jako kontenera.

Pusty kontener ma postać przedstawioną na rys. 2. Następnie weźmiemy najmłodszy bit w każdym kanale kolorowym i na jego podstawie przedstawimy "najniższą warstwę" oryginalnego obrazu w odcieniach czerni i bieli: wartość zerowa najmłodszego bitu odpowiada kolorowi białemu, a wartość jedynkowa - czarnemu. Otrzymane mozaiki przedstawione są na rys. 3: a) odpowiada czerwonemu kanalowi kolorowemu; b) zielonemu; c) niebieskiemu. Nawet dokładna analiza porównawcza trzech obrazków pokazuje ich prawie całkowitą tożsamość.



Rys. 2. Widok "pustego" kontenera



Rys. 3. Czarno-białe przedstawienie młodszych bitów "pustego" kontenera w kanale czerwonym (a), w kanale zielonym (b), w kanale niebieskim (c)

Teraz nałożymy obrazy na rys. 12.6 z następującymi parametrami kodowymi pikseli w każdym kanale: 1 - 255 (11111111 - w kodzie binarnym; czyli klasyczny czerwony, zielony lub niebieski), 0 - 0 (00000000 - w kodzie binarnym). Wynik pokazano na rys. 4.



Rys. 4. Kolorowe przedstawienie młodszych bitów "pustego" kontenera w kanale czerwonym

Praktyczne zadanie

1. Przed przystąpieniem do głównego zadania zaleca się zapoznanie się ze strukturą, interfejsem oraz funkcjonalnościami dostępnych i godnych uwagi aplikacji, w których zaimplementowano metodę LSB. Należą do nich:

- Stegano (https://github.com/cedricbonhomme/Stegano) - działa nie tylko z klasycznym LSB; posiada elastyczne ustawienia, może być również używany jako moduł Python; osadzanie jest realizowane za pomocą polecenia stegano-lsbhide --inputcover.jpg-fsecret.txt -eUTF-8 --outputstego.png, ekstrakcja – stegano-lsbreveal -istego.png -eUTF-8 -ooutput.txt; działa również z kontenerami PNG;

- LSB-Steganography (https://github.com/RobinDavid/LSB-Steganography) - aplikacja napisana w języku Python; działa z kontenerami PNG i BMP.

2. Opracuj własną aplikację, w której zostanie zaimplementowana metoda LSB. W tym celu:

- wybierz plik- kontener w porozumieniu z prowadzącym;

- zaimplementuj dwa warianty osadzanego/wydobywanego komunikatu:

− własne nazwisko, imię i drugie imię;

− tekstowa część raportu z jednego z wykonanych laboratoriów;

- zaimplementuj dwa wybrane przez siebie metody umieszczania bitowego strumienia osadzanego komunikatu w treści kontenera;

- stwórz macierze kolorów (analogicznie do rysunku 12.7), odzwierciedlające każdy używany do osadzania poziom młodszych bitów znaczących kontenera;

- przeprowadź wizualną analizę (z udziałem kolegów jako ekspertów) stegokontenerów o różnej wewnętrznej zawartości; wyciągnij wnioski na podstawie przeprowadzonej analizy.

3. Wyniki pracy należy udokumentować w formie raportu zgodnie z ustalonymi zasadami.