POLITECHNIKA LUBELSKA

WYDZIAŁ ELEKTROTECHNIKI I INFORMATYKI

INSTYTUT INFORMATYKI

ZAKŁAD OCHRONY INFORMACJI

mgr inż. Grzegorz Kozieł

ZMODYFIKOWANE METODY CYFROWEGO PRZETWARZANIA SYGNAŁÓW DŹWIĘKOWYCH W STEGANOGRAFII KOMPUTEROWEJ

Praca doktorska

Promotor: Prof. dr hab. inż. Volodymyr Harbarchuk

LUBLIN 2010

Spis treści

Wykaz skrótów i symboli

AD - średnia bezwzględna różnica wartości pomiędzy sygnałami,

AF - przezroczystość znaku wodnego,

BER - współczynnik ilości błędnie odczytanych bitów (ang. bit error rate),

bs - rozmiar bloku (w próbkach),

D - relatywna entropia,

DFT – dyskretna transformata Fouriera,

DSP - przetwarzanie sygnałów cyfrowych (ang. digital signal processing),

F_{dif} - maksymalna odległość wykorzystywanego prążka od największego,

f_1, f_2 - częstotliwości, na których występują wybrane do modyfikacji prążki widma,

f_{max} - częstotliwość na jakiej występuje prążek o największej wartości,

FT – transformata Fouriera,

H - entropia,

HAS - system słyszenia człowieka (ang. human auditory system),

IDFT - odwrotna dyskretna transformata Fouriera,

LMSE - Laplasjan błędu średniokwadratowego,

LSB - metoda najmniej znaczących bitów (ang. Least Significant Bit),

MD - maksymalna różnica wartości pomiędzy sygnałami,

MSE - błąd średniokwadratowy,

N - liczba próbek,

NAD - znormalizowana średnia bezwzględna różnica wartości pomiędzy sygnałami,

NC - znormalizowana korelacja krzyżowa,

NCQ - jakość korelacji,

NMR - odległość szumu od jego progu słyszalności (ang. noise-to-mask ratio),

NMSE - znormalizowany błąd średniokwadratowy,

p_1,p_2 - prążki widma przeznaczone do dołączenia informacji,

PSNR - szczytowa odległość sygnału od szumu,

Q - miara przezroczystości znaku wodnego,

R - różnica wartości pomiędzy prążkami,

R_p - różnica wartości pomiędzy prążkami, określona jako proporcja do W_{max},

S - sygnał oryginalny,

S' - sygnał zmodyfikowany,

f_p - częstotliwość próbkowania,

SMR - odległość pomiędzy sygnałem maskującym a maskowanym (ang. signal-to-mask ratio),

SNR - odległość sygnału od szumu (ang. signal-to-noise ratio),

TSM - operacje modyfikujące przebieg czasowy sygnału (ang. time scale modifications),

w_1, w_2 - wartości wybranych do modyfikacji prążków,

W_g - górna granica przedziału nieużywanego,

W_i_dopuszczalna – maksymalna wartość jaką może przyjąć i-ty prążek po modyfikacji,

W_i_docelowa – wartość i-tego prążka po modyfikacji,

W_{max} - wartość największego prążka we fragmencie sygnału,

β - liczba określająca maksymalną różnicę pomiędzy wyliczoną a ostateczną wartością modyfikowanego prążka,

γ - współczynnik określający rozmiar przedziału nieużywanych wartości prążków,

θ_{max} - maksymalny dzielnik wykorzystywany podczas zmniejszania prążka.

1. Wstęp

Łatwość w dostępie do szerokopasmowego łącza telekomunikacyjnego przenosi wiele dziedzin życia w świat komputerów i Internetu, który dla wielu z nas stał się niezbędnym narzędziem. Wykorzystujemy go nie tylko do komunikacji czy rozrywki, ale również do innych ważnych dla nas celów. Jednak wielu użytkowników sieci ma nieuczciwe zamiary. Zmusza to nas do korzystania z zabezpieczeń. Najczęściej polega to na zabezpieczeniu przesyłanych danych przed niepowołanym dostępem. Obecnie dużą popularnością cieszą się metody kryptograficzne[7]. Ich zadaniem jest przekształcenie informacji do postaci niezrozumiałej dla nieuprawnionych odbiorców, tak by wiadomość mogła odczytać tylko osoba posiadająca klucz deszyfrujący[46]. Gwałtowny wzrost mocy obliczeniowej komputerów i rozwój metod kryptoanalizy pozwala jednak coraz szybciej i efektywniej łamać zabezpieczenia. Niezbędne są więc alternatywne metody ochrony informacji. Doskonałe uzupełnienie stanowi tu steganografia.

Steganografia, jest to nauka zajmująca się ochroną cennej informacji poprzez jej ukrywanie w innej nie mającej wartości.
Wskutek tego osoba postronna nie jest w stanie wykryć obecności przekazu. Steganografia istniała już w czasach starożytnych, czego potwierdzenie możemy znaleźć nawet w mitologii[7,41]. Pojawienie się techniki cyfrowej, komputerów oraz efektywnej techniki łączności otworzyło nowe możliwości rozwoju tej nauki. Do ukrywania danych zaczęto wykorzystywać sygnał w postaci cyfrowej. Od tego momentu możemy mówić o wyodrębnieniu się steganologii cyfrowej oraz komputerowej.

Skuteczne ukrycie informacji znakomicie wypełnia lukę istniejącą w dotychczasowych systemach zabezpieczeń. Duża moc obliczeniowa nie daje przewagi w odnajdywaniu informacji ukrytej za pomocą metod steganograficznych. Jeśli adwersarz nie będzie świadomy faktu istnienia ukrytej informacji, lub nie będzie znał miejsca jej ukrycia, to nie będzie w stanie jej odczytać. Najczęściej w celu zwiększenia stopnia bezpieczeństwa algorytmy łączą techniki steganograficzne z kryptograficznymi. Informacja najpierw jest szyfrowana a dopiero później ukrywana. W ten sposób uzyskujemy pewność, że nawet przypadkowe wykrycie nie umożliwi osobie postronnej odczytania przekazu. Ponadto szyfrowanie powoduje, że ciąg binarny

wprowadzanych danych wykazuje dużą losowość zbliżoną do ciągu pseudolosowego, co znacznie zwiększa poziom bezpieczeństwa i utrudnia wykrycie.

Krytyczny przegląd dostępnej literatury wykazał, że obecnie do celów steganografii z powodzeniem wykorzystywanych jest wiele metod elektrotechniki teoretycznej, na przykład cyfrowego przetwarzania sygnałów (DSP) [4,14,31,41,64,78]. Jednak obecnie steganografia komputerowa nie posiada efektywnych podstaw matematycznych oraz dobrego opisu teoretycznego.

Przydatność metod często zależy od rodzaju nośnika informacji. Wiele publikacji prezentuje algorytmy ukrywające informację w dziedzinie transformacji Fouriera dające bardzo dobre wyniki w steganografii obrazu [11,31,78,69]. Jednak w przypadku, gdy nośnikiem jest dźwięk okazuje się, że dotychczas opracowane metody bazujące na transformacie Fouriera wprowadzają słyszalne zniekształcenia. Wynika to z różnic we właściwościach percepcji ludzkiego wzroku i słuchu. Wzrok nie jest wrażliwy na zmiany częstotliwości występujące w obrazie natomiast słuch jest bardzo czuły w tej dziedzinie. Modyfikacje współczynników transformaty Fouriera wykonywane podczas ukrywania informacji wpływają na zmianę częstotliwości w sygnale wynikowym. Dlatego też pomimo niezaprzeczalnych zalet dotychczas opracowane metody oparte o transformatę Fouriera nie zyskały popularności w steganografii sygnałów dźwiękowych. Przekształcenie to wykorzystywane jest jednak w metodach znakowania wodnego utworów[34,82], gdzie wprowadzane zmiany są o wiele bardziej przezroczyste ze względu na małą pojemność steganograficzną.

Okazuje się jednak, że słuch ludzki ma również swoje niedoskonałości i ograniczenia. Nie jest w stanie zarejestrować wszystkich docierających do ucha dźwięków. Między innymi nie pozwala na to zjawisko maskowania. W jego efekcie dźwięki cichsze są "zagłuszane" przez dźwięki głośniejsze o podobnej częstotliwości[40]. Dlatego rozpoczęte zostały badania mające na celu określenie możliwości wykorzystania tego zjawiska do opracowania **modyfikacji** metody steganograficznej bazującej na transformacie Fouriera, która nie będzie wprowadzała słyszalnych zniekształceń do sygnału dźwiękowego.

1.1. Cel i teza rozprawy

Celem pracy jest przeanalizowanie metod cyfrowego przetwarzania sygnałów pod kątem ich przydatności w steganografii komputerowej, wykorzystującej dźwięk jako kontener oraz opracowanie na tej podstawie nowej metody steganograficznej.

Zalety metod bazujących na teorii sygnałów oraz praktyka wskazują na ich przydatność steganograficzną.

Pozwala to na postawienie tezy rozprawy:

Możliwe jest wykorzystanie współczesnych metod cyfrowego przetwarzania sygnałów oraz ich modyfikacji do realizacji efektywnych technologii steganografii komputerowej wykorzystującej dźwięk jako kontener.

Postawiona teza zostanie udowodniona przez:

1. ocenę przydatności metod cyfrowego przetwarzania sygnałów do wykorzystania w steganografii i wybór jednej z nich jako przekształcenia bazowego w nowej metodzie steganograficznej,
2. teoretyczne opracowanie nowej metody w oparciu o transformatę Fouriera i maskowanie,
3. praktyczną weryfikację funkcjonalności i ograniczeń nowej metody,
4. porównanie tej metody z tradycyjną metodą opartą o transformatę Fouriera,
5. porównanie z innymi metodami steganograficznymi i dostępnym oprogramowaniem steganograficznym używającym dźwięku jako nośnika ukrytej informacji.

1.2. Zakres pracy

Rozprawę doktorską podzielono na wstęp i trzy rozdziały.

W rozdziale pierwszym w sposób syntetyczny przedstawiono znane z literatury podstawowe pojęcia dotyczące ochrony informacji i steganografii sygnałów dźwiękowych. Zawiera on analizę obecnego dorobku steganografii komputerowej wykorzystującej dźwięk jako nośnik ukrytej informacji oraz wybranych metod przetwarzania sygnałów pod kątem oceny ich przydatności jako przekształcenia bazowego nowej metody steganograficznej. Omówione zostały aspekty ludzkiego

słuchu z uwzględnieniem jego wad możliwych do wykorzystania w procesie ukrywania wprowadzanych zmian. Przedstawione zostało również zjawisko maskowania wykorzystywane w opracowanym algorytmie steganograficznym. Zakończenie tej części rozprawy zawiera konkretyzację celu i zadań pracy badawczej.

Drugi rozdział rozprawy zawiera teoretyczny opis nowej metody steganograficznej opracowanej na podstawie autorskich badań problemu.

Zawarta tu została zasada działania nowego algorytmu bazującego na transformacie Fouriera, uwzględniającego możliwości ukrycia wprowadzanych zmian poprzez wykorzystanie zjawiska maskowania częstotliwościowego.

Przedstawiony został proces przygotowywania nośnika do ukrycia informacji, jego podział na bloki, algorytm ukrywania dodatkowej informacji a także proces scalania przetworzonych bloków. Rozdział zawiera również wyniki praktycznej weryfikacji proponowanego algorytmu steganograficznego, czyli określenie dopuszczalnej zmienności parametrów opracowanej metody oraz składników klucza steganograficznego.

Trzecia część rozprawy dotyczy praktycznej weryfikacji jakości nowego algorytmu ukrywania danych, uzyskane przez porównanie z innymi aplikacjami steganograficznymi. Na wstępie scharakteryzowano aplikacje porównawcze i przedstawiono algorytm ich działania. Następnie porównano nośniki uzyskane za pomocą analizowanych aplikacji pod kątem oceny jakości ich procesu steganograficznego w zakresie pojemności steganograficznej, stopnia zniekształceń nośnika wprowadzanych podczas ukrywania danych, odporności dołączonych danych na uszkodzenia i zniszczenie podczas wykonywania przekształceń nośnika zawierającego ukrytą informację.

Podsumowanie rozprawy zawiera wnioski wynikające z przeprowadzonych badań oraz propozycje dalszych prac badawczych i rozwoju opracowanej metody steganograficznej.

Praca zawiera 139 stron, 54 rysunki, 12 tabel i 93 pozycje literatury.

2. Analiza i ocena możliwości wykorzystania metod elektrotechniki teoretycznej w steganografii

2.1. Postawienie problemu badawczego

Dotychczasowy rozwój steganografii zaowocował powstaniem dużej liczby metod steganograficzych. W wielu z nich wyniki uzyskane zostały na drodze eksperymentu, ze względu na brak dobrych podstaw teoretycznych i brak opisu matematycznego. Istnieje jednak wiele innych dziedzin posiadających dobrze poznane i opisane metody, które można wykorzystać do ukrywania informacji. Jedną z takich dziedzin jest elektrotechnika teoretyczna.

Problem badawczy polega na tym że:

a) istnieje potencjalnie bardzo efektywny i perspektywiczny kierunek steganografii komputerowej, lecz brakuje w nim dobrze opracowanych podstaw teoretycznych,
b) istnieje dobrze rozwinięta elektrotechnika teoretyczna zawierająca kierunek cyfrowego przetwarzania sygnałów,

Konieczne jest wykonanie badań pozwalających na ocenę i wybór najbardziej skutecznych metod cyfrowego przetwarzania sygnałów do celów steganografii komputerowej wykorzystującej dźwięk jako kontener.

2.2. Główne problemy ochrony informacji

Nauka ochrony informacji składa się obecnie z następujących gałęzi:

1. kryptologii,
2. kryptologii kwantowej,
3. steganologii,
4. metod kombinowanych 1,2,3.

Każda z nich dzieli się na dwie przeciwstawne dziedziny: jedną zajmującą się ochroną informacji oraz drugą przeciwstawną jej zajmującą się łamaniem istniejących zabezpieczeń. W przypadku kryptologii będą to: kryptografia mająca na celu przekształcenie cennej informacji w taki sposób, by stała się niezrozumiała dla osób postronnych oraz kryptoanaliza dążąca do umożliwienia odczytania zaszyfrowanych danych osobie postronnej.

Kryptografia kwantowa jest rozwinięciem kryptologii tradycyjnej wykorzystującym mechanizmy mechaniki kwantowej do celów kryptografii.

Steganologia obejmuje dwie dziedziny – steganografię zajmującą się ochroną informacji oraz przeciwstawną do niej steganoanalizę. Pomimo tego że tak jak w przypadku kryptologii celem tej nauki jest ochrona informacji, to sposób realizacji jest zupełnie inny. Steganologia nie dokonuje przekształceń informacji lecz zajmuje się ukrywaniem jej istnienia, tak by osoby postronne nie były w stanie stwierdzić jej istnienia[31]. W ostatnich latach głównym kierunkiem rozwoju steganologii jest steganologia komputerowa.

Nieustający "wyścig" pomiędzy twórcami zabezpieczeń a łamiącymi je wymusza tworzenie nowych metod pozwalających chronić informację. Konieczna jest również weryfikacja skuteczności tworzonych metod. Wykonywanie tego na drodze eksperymentu nie zawsze pozwala na osiągnięcie wiarygodnych wyników. Niezbędne staje się wykorzystanie metod matematycznych. Jednak pojawia się tu problem, ze względu na to, że często brak jest podstaw matematycznych opracowanych algorytmów. Problem ten szczególnie wyraźnie uwidacznia się w steganografii. Do jego rozwiązania może się przyczynić wykorzystywanie przekształceń pochodzących z innych dziedzin nauki, mających dobrze opracowane podstawy teoretyczne.

Obecnie najczęściej wykorzystywana jest postać cyfrowa sygnału dźwiękowego a on sam przekazywany i przetwarzany jest w tej właśnie postaci. Występuje tu zjawisko uniezależnienia danych od nośnika[41]. Konieczne jest więc opracowywanie metod ochrony powiązanych bezpośrednio z danymi cyfrowymi. Dlatego też rozważania przedstawione w dalszej części pracy dotyczyły będą cyfrowego zapisu sygnału dźwiękowego.

2.2.1. Podstawowe definicje ochrony informacji

Aby prowadzić rozważania niezbędne jest precyzyjne określenie czym będziemy się zajmować. Zacznijmy od definicji tego, co podlega ochronie czyli – informacji.

Informacja - jest niematerialnym zasobem, będącym produktem intelektualnej, produkcyjnej, społecznej działalności człowieka przy współdziałaniu człowieka z otaczającym środowiskiem. Zasób ten może występować, być mierzony, przekazywany, przechowywany, przetwarzany, eksploatowany i wykorzystywany w formie faktów, danych, wiedzy, obrazów i ich kombinacji.

Nie każdą informację chronimy. Ochronie podlegają tylko informacje, których udostępnienie innym mogłoby przynieść nam wymierne straty. Należy tu również zaznaczyć, że wartość informacji często zmienia się w czasie. Najprostszym tego przykładem mogą być kursy walut. Jeśli wiemy jak zmienią się w ciągu najbliższej godziny, to możemy za pomocą tej informacji osiągnąć korzyści materialne. Jednak warunkiem jest utrzymanie tego w tajemnicy. Jeśli więc zastosujemy zabezpieczenie, które będzie działało przez godzinę to informacja będzie skutecznie chroniona. Nie ma znaczenia to, że po upływie tego czasu będzie możliwe odczytanie informacji przez osoby postronne, ponieważ straci ona już swoją wartość.

Wartość informacji definiujemy jako wskaźnik, miarę znaczenia konkretnej ilości informacji do podjęcia decyzji w konkretnej sytuacji z określeniem następstw tych decyzji w zależności od danej informacji.

Im cenniejsza jest informacja tym większe jest prawdopodobieństwo ataków na nią, a ich moc jest wprost proporcjonalna do wartości tej informacji. Tak więc siła ochrony informacji powinna być adekwatna do mocy ataków.

Ochrona informacji z definicji jest całokształtem działań aktywnych, biernych lub aktywno-biernych dotyczących pewnej informacji, ukierunkowanych na uniemożliwienie dostępu do tej informacji osobom nie posiadającym do tego prawa (pozwolenia), zastosowanie specjalnych środków w celu uniemożliwienia dostępu, zniszczenia, uszkodzenia, zniekształcenia i innych celowych zmian w tej informacji, ukrycie miejsca położenia informacji lub w ogóle utajnienia faktu istnienia danej informacji.

Celem ochrony informacji jest zagwarantowanie niezmienności i nienaruszalności informacji w określonym przedziale czasu.

Oczywiste jest, że chroniona musi być tylko informacja, do której mogą uzyskać dostęp osoby postronne. Informacja niedostępna nie wymaga ochrony. W przypadku gdy mamy do czynienia z komunikacją pomiędzy dwoma oddzielnymi środowiskami informacja musi zostać pomiędzy nimi przekazana. Wykorzystywane są do tego *kanały łączności*, czyli środki łączności pomiędzy komunikującymi się stronami.

Informacja przekazywana poprzez kanał łączności jest narażona na *atak*, który jest każdym bezprawnym działaniem, ukierunkowanym na pozyskanie, przechwycenie, zniekształcenie, zniszczenie lub zmianę adresata przekazywanej informacji.

Definicje pojęć przedstawionych w tym rozdziale pochodzą z monografii [31].

2.2.2. Struktura steganologii

Steganologia jest nauką zajmującą się ukrywaniem informacji. Obejmuje ona dwie dziedziny: steganografię i stegoanalizę[31].

Steganografia zajmuje się ukrywaniem cennej informacji w innych danych nie mających wartości, tak by osoba postronna nie była w stanie wykryć istnienia dodatkowych danych. Dokonuje się tego poprzez wprowadzenie do oryginalnego nośnika niewielkich zmian, których ludzkie zmysły nie są w stanie wykryć. Istotne jest również zachowanie poprawnej wartości wszystkich parametrów, które mogą zostać zbadane podczas analizy sygnału[64].

Steganoanaliza zajmuje się bezprawnym wykrywaniem ukrytych informacji oraz ich odczytywaniem lub niszczeniem. Możliwe jest to dzięki wnikliwej analizie nośników i badaniu ich parametrów. W przypadku gdy znacznie odbiegają one od spodziewanej wartości sygnał poddawany jest szczegółowej analizie ze względu na duże prawdopodobieństwo istnienia w nim ukrytych danych.

Ze względu na obszary zastosowań i związane z nimi różne wymagania, metody steganograficzne można podzielić na kilka grup[31]:

Anonimowa komunikacja – polega na ukryciu faktu komunikacji a co za tym idzie tożsamości komunikujących się. Nadawca przygotowuje nośnik zawierający ukrytą informację a następnie w sposób anonimowy umieszcza go w publicznym miejscu, skąd każdy może go pobrać. Najczęściej do tego celu wykorzystywany jest Internet. Odbiorca również w sposób anonimowy pobiera zasób i odczytuje z niego dane. Najistotniejszym wymaganiem jest tu zachowanie przezroczystości, czyli braku zmian mogących wskazywać na istnienie dołączonej wiadomości. Przygotowany nośnik musi być jak najbardziej podobny do oryginału, tak by zmysły oraz analiza komputerowa nie były w stanie wykryć żadnych nieprawidłowości.

Silne znakowanie wodne – to sposób na oznakowanie danych cyfrowych, najczęściej stosowane w multimediach w celu udowodnienia praw autorskich. Polega na trwałym dołączeniu do oryginalnego utworu znaku identyfikującego. Wymogi stawiane znakom wodnym to odporność i przezroczystość. Znak wodny nie może ulec zniszczeniu podczas przetwarzania sygnału. Dopóki sygnał posiada wartość czyli do momentu jego znacznego uszkodzenia znak wodny musi dać się odczytać. Ponadto wprowadzenie znaku wodnego nie może zmniejszać wartości utworu poprzez wprowadzenie do niego dodatkowych zakłóceń[41,92,93].

Słabe znakowanie wodne – jest to odmiana znakowania wodnego mająca na celu potwierdzenie oryginalności nośnika, aby odbiorca mógł mieć pewność, że oznakowane dane nie były modyfikowane po ich oznakowaniu. Wiąże się to z koniecznością stosowania znaków ulegających zniszczeniu podczas każdej modyfikacji sygnału[41,92,93].

Ochrona przed powielaniem – to odmienny rodzaj steganografii służący uniemożliwieniu wykonywania nielegalnych kopii nośników optycznych. Można tego dokonać przez umieszczenie ukrytych danych, bez których nośnik jest bezużyteczny, w miejscach, które nie są kopiowane lub zorganizować zapis, tak by spowodować błędne działanie urządzenie kopiującego, co w efekcie prowadzi do uszkodzenia kopii[52].

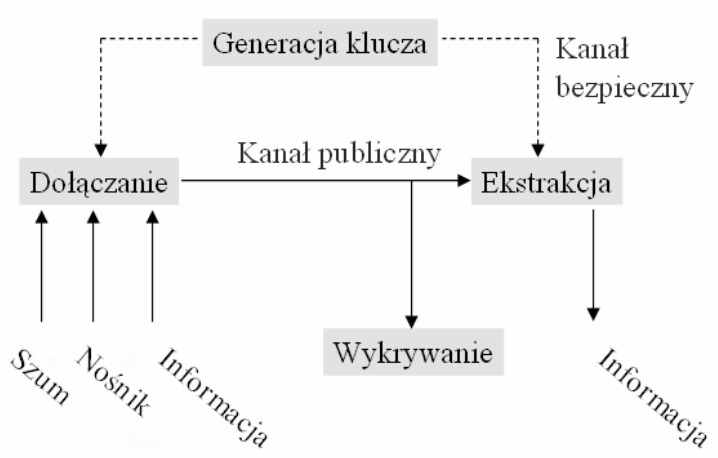
2.2.3. Model systemu steganograficznego

Pojęcia związane ze steganologią zdefiniowane zostały na konferencji First Information Hidding Workshop w 1996 roku w Cambridge[67]. Do najważniejszych z nich należą:

- *kontener (nośnik)* – dane, które będą wykorzystane do ukrycia w nich dodatkowej informacji,
- *dołączony* – coś ukrytego w czymś innym, czyli dane dołączone do kontenera za pomocą metod steganograficznych,
- *klucz steganograficzny (stegoklucz)* – dodatkowe tajne dane niezbędne do ukrycia informacji w kontenerze. Ten sam klucz jest niezbędny do odczytania ukrytej informacji (w steganografii asymetrycznej niezbędny jest powiązany z nim klucz),
- *stegokontener* – nośnik zawierający ukrytą informację powstały w wyniku połączenia kontenera i dołączonej informacji,
- *stegoanalityk* – przeciwnik, osoba próbująca w nieautoryzowany sposób pozyskać, przechwycić, zniekształcić lub zniszczyć informację zawartą w stegokontenerze, ktoś przed kim tę informację ochraniamy,
- *dołączanie* – proces ukrywania informacji w kontenerze za pomocą metody steganograficznej,
- *ekstrakcja* – proces wydobywania i odczytywania dołączonej informacji ze stegokontenera,

- *atak* – ogół działań podejmowanych przez stegoanalityka prowadzących do złamania stegosystemu, czyli wykrycia a następnie odczytania lub uszkodzenia ukrytej informacji. Wyróżniamy dwa typy ataków:
  - *pasywny* – stegoanalityk w żaden sposób nie ingeruje w przesyłaną wiadomość, jedynie próbuje złamać system steganograficzny za pomocą obserwacji danych,
  - *aktywny* – w którym w celu złamania systemu steganograficznego atakujący ingeruje w strumień danych wprowadzają do niego zmiany.

    Model systemu steganograficznego zaprezentowany został na rysunku 2.1.



Rysunek 2.1. Schemat systemu steganograficznego[30]

Zanim będzie możliwa komunikacja, strony muszą uzgodnić klucz steganograficzny, z którego będą korzystać. Musi on zostać dostarczony w sposób bezpieczny, tak by przeciwnik nie miał możliwości zdobycia go. Gwarantuje on bowiem bezpieczeństwo systemu. Mając uzgodniony klucz nadawca dołącza za jego pomoca informację do kontenera. Często jest ona wcześniej szyfrowana. Niekiedy jednocześnie z informacją dołączany jest dodatkowy szum, który ma za zadanie dodatkowo utrudnić detekcję właściwej informacji. Przygotowany stegokontener jest przesyłany kanałem publicznym do odbiorcy, który za pomocą klucza odczytuje z niego dołączoną informację. Działanie takiego systemu polega na "grze" pomiędzy trzema graczami. Dwaj z nich porozumiewają się przesyłając kanałem publicznym informacje ukryte w innym nośniku. Trzeci z nich próbuje wykryć, która część przesyłanych danych zawiera utajnioną informację. Jeśli mu się to uda przystępuje do wyodrębniania wiadomości z nośnika[10,30,36].

2.2.4. Bezpieczeństwo systemu steganograficznego

Ze względu na główny cel steganografii, jakim jest ukrycie informacji, bezpieczeństwo systemu steganograficznego ocenia się w aspekcie wykrywalności dołączonych danych[31].

System bezwarunkowo bezpieczny to taki, dla którego nie występuje możliwość wykrycia istnienia dołączonej wiadomości przy następujących założeniach[4,28,31,78]:

- bezpieczeństwo systemu zależy tylko od klucza steganograficznego, który jest znany jedynie stronom komunikującym się,
- algorytmy dołączania i ekstrakcji oraz kontener, stegokontener i ukryta informacja są znane stegoanalitykowi,
- stegoanalityk dysponuje nieograniczonymi zasobami, nieograniczoną mocą obliczeniową i dostępem do kanału łączności oraz w nieograniczonym czasie wykonuje różnego typu ataki.

W opisaniu bezpieczeństwa systemu steganograficznego pomocna jest entropia. *Entropia* (*H*) to miara oczekiwanej ilości informacji zawartej w przypadkowym sygnale ζ pochodzącym ze źródła informacji[3,31].

Aby podać definicję bezpieczeństwa systemu należy założyć, że nośnik wiadomości jest wybierany losowo jako zmienna losowa *C* o rozkładzie prawdopodobieństwa P_c, dołączanie poufnej wiadomości do nośnika może być rozumiane jako funkcja określona na zbiorze *C*, a P_s będzie rozkładem prawdopodobieństwa zbioru stegoobiektów tworzonych przez system steganograficzny. Po spełnieniu powyższych założeń możemy powiedzieć, że system steganograficzny jest ε-bezpieczny w przypadku ataku pasywnego, jeśli relatywna entropia *D* spełnia warunek:

$$D(P_C \parallel P_S) \geq \varepsilon \qquad (2.1),$$

gdzie: *D* – relatywna entropia określona dla dwóch rozkładów prawdopodobieństwa P_C i P_S zdefiniowanych na zbiorze *Q* według równania:

$$D(P_C \parallel P_S) \leq \sum_{q \in Q} P_C(q) \log_2 \frac{P_C(q)}{PS(q)} \qquad (2.2).$$

System jest bezwarunkowo bezpieczny, jeżeli ε=0. Relatywna entropia wynosi zero, gdy oba rozkłady prawdopodobieństwa są równe.

Zwiększenie skuteczności ochrony informacji możliwe jest przez jednoczesne zastosowanie metod steganografii komputerowej i kryptografii.

2.2.5. Ocena odporności systemu steganograficznego

Odporność dołączonej informacji to miara pewności prawidłowego odczytu ukrytej informacji po wykonaniu sekwencji określonych przekształceń sygnału stegokontenera[27,61].

Ma ona szczególne znaczenie zwłaszcza w przypadku znakowania wodnego utworów. Dołączona informacja musi być w stanie przetrwać kilkakrotne wykonanie powszechnych operacji obróbki dźwięku. Dołączona informacja jest traktowana jako odporna do momentu jej usunięcia lub uszkodzenia w takim stopniu, że jednoznaczny odczyt jest niemożliwy. Do oceny odporności danych ukrytych w sygnale dźwiękowym wykorzystuje się procedurę testową składającą się z trzech kroków:

1. do każdego badanego sygnału dźwiękowego dołączana jest informacja, z maksymalną mocą, która nie narusza jej przezroczystości, czyli nie powoduje zmian mogących wskazywać na istnienie dołączonej informacji,
2. na utworzonym stegokontenerze wykonywany jest zestaw stosownych operacji przekształcających sygnał dźwiękowy,
3. z przekształconego stegokontenera odczytywana jest dołączona informacja i mierzony jest stopień jej uszkodzenia[27,61].

W [64] autorzy traktują dołączoną informację jako odporną tylko wtedy, gdy całą informację uda się odczytać bezbłędnie. Jednak obecnie odchodzi się od tak rygorystycznej oceny, podając współczynnik ilości błędnie odczytanych bitów (ang. BER – bit error rate). Powodem tego jest wprowadzenie kodów korekcji błędów umożliwiających poprawne odczytanie dołączonej wiadomości, nawet jeśli wystąpiły błędy odczytu[27,32,61,89].

BER obliczany jest według wzoru:

$$BER = \frac{100}{l}\sum_{n=0}^{l-1}\begin{cases}1, W'_n = W_n\\ 0, W'_n \neq W_n\end{cases} \qquad (2.3),$$

gdzie: l-długość dołączanej informacji, W_n – wartość n-tego bitu ukrywanej informacji, W_n'– wartość n-tego bitu odczytanej informacji.

Aby procedura testowa była wiarygodna musi zostać powtórzona wielokrotnie z użyciem losowych sekwencji dołączanych danych oraz różnych sygnałów kontenera. Aby uzyskać wiarygodne wyniki należy poddać sygnał różnorodnym przekształceniom edytorskim. Według [65] podzielone one zostały na 9 grup:

Dynamiczne – Są to wszelkie operacje modyfikujące amplitudę sygnału. W najprostszym przypadku będzie to zwykła zmiana głośności. Bardziej złożone przypadki polegają na losowym wyciszaniu i podgłaśnianiu fragmentów utworu oraz na innych nieliniowych modyfikacjach głośności.

Filtrowanie – Operacje polegające na modyfikowaniu określonych pasm częstotliwości, zarówno przez pojedyncze filtry jak i ich zespoły.

Kształtowanie efektu przestrzennego – Sygnał rozchodzący się w pomieszczeniu zawsze napotka przeszkodę, od której się odbije. Docierający do słuchacza odbity sygnał jest opóźniony i ma mniejszą amplitudę niż sygnał oryginalny. Do operacji kształtowania efektu przestrzennego zaliczymy więc sekwencje przekształceń opóźniające i zmniejszające amplitudę sygnału. Najlepszym przykładem takiej operacji jest dodanie echa.

Konwersja – Niejednokrotnie sygnał zapisywany jest w innym formacie lub na innym nośniku. Podczas każdego z tych przekształceń wprowadzane są zmiany do sygnału. Do operacji konwersji zaliczymy zmianę formatu zapisu z cyfrowego na analogowy i odwrotnie oraz zmianę liczby kanałów.

Kompresja – Są to wszelkie operacje mające na celu zmniejszenie rozmiaru danych niezbędnych do reprezentowania sygnału dźwiękowego. Jest to kłopotliwa operacja ze względu na fakt, że większość algorytmów stosuje kompresję stratną. Bazuje bowiem ona na redukowaniu tego co jest niesłyszalne w dźwięku, czyli wykorzystuje te same właściwości co metody steganograficzne, które modyfikują niesłyszalne dla człowieka parametry sygnału.

Zaszumienie – proces dołączania szumu do sygnału. Szum może być dodawany przez dołączenie go poprzez operację połączenia lub generowany przez maszyny przesyłające i przetwarzające dźwięk.

Modulacja – dodawanie różnych efektów dźwiękowych modyfikujących brzmienie utworu takich jak chorus czy efekt wibracji. Najczęściej są to przekształcenia wpływające na barwę dźwięku, polegające na podwyższeniu lub obniżeniu tonacji lub też nałożeniu na siebie zmodyfikowanych kopii utworu.

Modyfikacja zależności czasowych – wszystkie operacje modyfikujące zależności czasowe w utworze. Najbardziej istotna jest tu zmiana tempa nagrania.

Przestawianie próbek – do tego rodzaju przekształceń zaliczamy operacje modyfikujące ułożenie próbek w utworze. Może to być zarówno losowe wstawianie

dodatkowych próbek lub fragmentów do utworu, wycinanie ich lub też zamiana kolejności identycznych fragmentów dźwięku.

Dołączona informacja powinna przetrwać wszystkie te operacje, ponieważ nigdy nie wiadomo, na jakie przekształcenia będzie narażony stegokontener.

Należy wziąć jednak pod uwagę fakt, że stegoanalityk będzie próbował usunąć lub uszkodzić dołączoną informację za pomocą wyspecjalizowanych narzędzi. Należy więc rozszerzyć obszar testów o dodatkowe przekształcenia[27,61,64].

2.2.6. Ocena przezroczystości

Dołączenie dodatkowych danych zawsze niesie za sobą zmiany w sygnale oryginalnym. Jednak zmiany te nie mogą być zbyt duże. Sugerowałoby to istnienie wewnętrznej informacji i mogłoby stanowić wskazówkę dla stegoanalityka. Ponadto zbyt wyraźne zmiany mogą pogorszyć jakość dźwięku, wprowadzając słyszalne zakłócenia. Szczególnie istotne jest to w przypadku znakowania wodnego utworów audio. Wytwórnie fonograficzne przywiązują ogromną wagę do zachowania doskonałej jakości znakowanego sygnału[13].Minimalne wymagania stawiane znakom wodnym stanowią, że znak wodny nie może być wykrywalny zmysłami a odległość sygnału od szumu wprowadzanego podczas znakowania musi być większa od 36dB[92]. Wymagania te zmusiły steganografów do określenia miar zniekształceń wprowadzanych podczas znakowania oraz określenia stopnia wprowadzanych zniekształceń. Użyto do tego celu miar obiektywnych[27,61,89]:

- błędu średniokwadratowego zdefiniowanego wzorem:

$$MSE = \frac{1}{N}\sum_{n}(S_n - S'_n)^2 \qquad (2.4),$$

- znormalizowanego błędu średniokwadratowego opisanego równaniem:

$$NMSE = \sum_{n}(S_n - S'_n)^2 / \sum_{n} S_n^{\,2} \qquad (2.5),$$

- laplasjana błędu średniokwadratowego zdefiniowanego wzorem:

$$LMSE = \sum_{n}(\nabla^2 S_n - \nabla^2 S'_n)^2 / \sum_{n}(\nabla^2 S_n)^2 \qquad (2.6),$$

- odległości sygnału od szumu według:

$$SNR = 10\log(\sum_{n} S_n^2 / \sum_{n}(S_n - S'_n)^2) \qquad (2.7),$$

- szczytowej wartości odległości sygnału od szumu zdefiniowanego wzorem:

$$PSNR = 10\log(R^2 / MSE) \quad (2.8),$$

- normy LP opisanej równaniem:

$$LP = \left(\frac{1}{N} \sum_n | S_n - S'_n |^2 \right)^{1/p} \quad (2.9),$$

- maksymalnej różnicy pomiędzy sygnałem oryginalnym i zmodyfikowanym wg.:

$$MD = \max | S_n - S'_n | \quad (2.10),$$

- średniej bezwzględnej różnicy pomiędzy sygnałami opisanej równaniem:

$$AD = \frac{1}{N} \sum_n | S_n - S'_n | \quad (2.11),$$

- znormalizowanej średniej bezwzględnej różnicy pomiędzy sygnałami wg.:

$$NAD = \sum_n | S_n - S'_n | / \sum_n | S_n | \quad (2.12),$$

- przezroczystości znaku wodnego opisanej równaniem:

$$AF = 1 - \sum_n (S_n - S'_n)^2 / \sum_n S_n^2 \quad (2.13).$$

W powyższych wzorach przyjęto oznaczenia: S_n – wartość n-tej próbki sygnału oryginalnego, R- rozdzielczość próbkowania (liczba możliwych wartości), S_n'- wartość n-tej próbki sygnału z dołączonym znakiem wodnym, N – liczba próbek w sygnale.

Wszystkie przedstawione miary służą do mierzenia różnic występujących pomiędzy sygnałem oryginalnym i zmodyfikowanym. W sposób obiektywny określają zniekształcenia wprowadzane podczas znakowania sygnału. Jednak nie do końca przekłada się to na postrzegalność wprowadzanych zmian przez człowieka, ponieważ zależy ona od indywidualnych predyspozycji każdego z nas[27,61,64]. W celu określenia zniekształceń postrzeganych przez człowieka wprowadzono miary korelacji. Są one bowiem o wiele bardziej zbliżone do modelu ludzkiego słuchu. Stosowanymi miarami korelacji są:

- znormalizowana korelacja krzyżowa obliczona według wzoru:

$$NC = \sum_n S_n S'_n / \sum_n S_n^2 \quad (2.14),$$

- jakość korelacji określona równaniem:

$$NCQ = \sum_n S_n S'_n / \sum_n S_n \quad (2.15).$$

Często również używana jest pięciostopniowa miara przezroczystości znaku wodnego wyliczana według wzoru:

$$Q = \frac{5}{1 + NC \cdot SNR} \qquad (2.16),$$

gdzie *NC* oznacza stałą normalizacji[27,61].

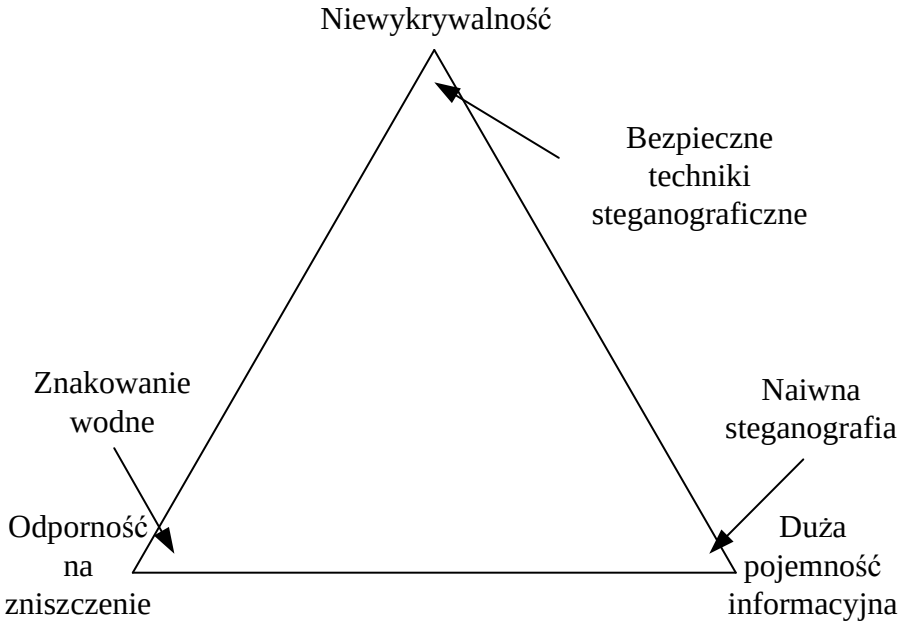
2.2.7. Problemy steganologii komputerowej

Steganologia jest nauką znaną od setek lat. Początkowo stosowana była do przekazywania wiadomości, później pojawiły się inne zastosowania. Możemy znaleźć dokumenty świadczące o tym, że w czasach starożytnych wykorzystywano steganografię do przekazywania tajnej korespondencji. Znane są przypadki tatuowania na skórze głowy, pisania na tabliczkach glinianych pokrywanych później woskiem czy też używania atramentu sympatycznego[33,64]. Znani drukarze znakowali swoje wyroby umieszczając na papierze znaki wodne[64]. Niewiele później po wprowadzeniu znaków wodnych, pojawiły się próby ich fałszowania, oparte zwykle na procesach chemicznych. Pomimo tego uważane były za wiarygodny dowód potwierdzający autentyczność i niejednokrotnie służyły za dowód w procesach sądowych. Taka sytuacja ma miejsce także w dzisiejszych czasach. W ostatnich latach rozwój technik cyfrowych uniezależnił dane od nośnika. Tradycyjne techniki steganograficzne stały się niewystarczające. Potrzebne są metody znakowania danych a nie nośnika. Funkcję tą pełnią cyfrowe znaki wodne. Kolejną dziedziną, w której znajduje zastosowanie cyfrowa postać steganografii jest ochrona danych przesyłanych przez publiczne medium. Mogło by się wydawać, że na tym polu doskonale się sprawdzają techniki kryptograficzne. Jednak steganografia umożliwia zastosowanie dodatkowej funkcji – ukrywa fakt istnienia przekazu. Bardzo dobrze prezentuje to zastosowanie przykład więźniów przedstawiony w [74]. Dwoje więźniów osadzonych jest w dwóch odległych celach. Komunikować się ze sobą mogą tylko za pośrednictwem strażnika, który przekazuje wiadomości jednocześnie je kontrolując. Więźniowie chcą ułożyć plan ucieczki. Muszą dokonać tego komunikując się przy pomocy strażnika a jednocześnie zachowując wszystko w tajemnicy. Dlatego też muszą zawierać poufne treści w niewinnie wyglądających wiadomościach. Jak widać na tym przykładzie, najważniejszą cechą technik steganograficznych wykorzystywanych do poufnej komunikacji jest wysoki poziom niewykrywalności ukrytych informacji.

Steganologia jest dziedziną, w której trwa nieustający wyścig pomiędzy steganografią i steganoanalizą. Ochrona informacji jest dla nas coraz ważniejsza i szukamy nowych metod. Z drugiej strony pozyskanie poufnej informacji przynosi duże korzyści, więc rozwijane są techniki wykrywania i odczytywania ukrytych danych. Każdy chciałby móc korzystać z idealnej metody posiadającej trzy ważne cechy:

- dużą pojemność,
- niewykrywalność,
- odporność na uszkodzenia i zniszczenie.

Okazuje się jednak, że nie jest możliwe stworzenie takiej metody. Doskonale obrazuje to trójkąt sprzeczności wymagań przedstawiony na rys. 2.2.



Rysunek 2.2. Trójkąt sprzeczności wymagań[44]

W komunikacji istotna jest wysoka niewykrywalność i duża pojemność informacyjna. Jednak oczywiste jest, że im więcej danych będziemy dołączać tym większe zniekształcenia wprowadzimy do nośnika. Tak więc te dwa warunki stoją ze sobą w sprzeczności.

Znak wodny musi być odporny na uszkodzenia i zniszczenie, ale również niewykrywalny. Nie może bowiem uszkadzać znakowanych danych. Odporność wzrasta wraz ze zwiększaniem wielkości wprowadzanych zmian i stopniem rozproszenia w całym znakowanym zbiorze. Jednak zwiększając wartość zmian wprowadzamy większe zakłócenia, czyli działamy na niekorzyść niewykrywalności.

Podobnie wygląda sytuacja dla pary pojemności informacyjnej i odporności. Żeby zwiększyć odporność należy zastosować dodatkowe kody korekcji błędów, zwiększyć wielkość wprowadzanych zmian, rozproszyć ukrywaną informację w całym kontenerze, czy też wielokrotnie ją zapisać. Wszystkie te operacje wymagają dodatkowego miejsca na zapis, czyli zmniejszają pojemność steganograficzną.

Niemożliwe jest więc skonstruowanie idealnej metody. Pojawia się tu problem doboru odpowiednich parametrów. Zadaniem steganografa jest dobór metody najlepszej do danego zastosowania[18].

Zadaniem stegoanalityka jest odszukanie ukrytego przekazu. Najprościej można tego dokonać przez porównanie zmodyfikowanego nośnika z oryginałem. Jednak sytuacje, w których stegoanalityk dysponuje oryginałem należą do rzadkości. Najczęściej więc przeprowadzane są analizy statystyczne i na ich podstawie określane jest prawdopodobieństwo istnienia ukrytej informacji wewnątrz nośnika.

2.2.8. Przegląd najważniejszych metod steganograficznych opracowanych dla sygnału dźwiękowego

Metody steganograficzne ze względu na sposób działania możemy podzielić na sześć grup[31]:

1. metody substytucji - polegające na zastępowaniu nadmiarowych danych nośnika ukrywaną informacją,
2. metody transformacyjne – polegające na przetransformowaniu sygnału kontenera do dziedziny częstotliwości i dołączanie informacji poprzez modyfikację uzyskanych współczynników transformaty,
3. metody rozproszonego widma – wykorzystujące całe pasmo częstotliwości do rozproszenia ukrywanych danych, które dodatkowo rozpraszane są w całym nośniku,
4. metody statystyczne – ukrywające dane poprzez modyfikację statystycznych cech nośnika,
5. metody zniekształceniowe – działające poprzez wprowadzanie zniekształceń do sygnału kontenera; aby odczytać dołączoną informację konieczne jest porównanie z oryginałem,
6. metody generacji nośnika – tworzące stegokontener na podstawie ukrywanej informacji.

Wśród popularnych technik podlegających ciągłemu rozwojowi możemy wyróżnić kilka głównych grup[29] przedstawionych w kolejnych podrozdziałach.

2.2.8.1 Metody najmniej znaczących bitów

Metoda modyfikacji amplitudy znana jest jako metoda najmniej znaczących bitów (LSB). Jest bardzo znana i popularna zarówno w komunikacji jak i znakowaniu wodnym[18,39,54,76,85]. Do ukrywania informacji wykorzystuje najmniej znaczące bity próbek dźwięku, które nie przenoszą wartościowej informacji, lecz zawierają jedynie szum kwantyzacji. Modyfikacja wartości tych bitów nie wpływa na zmianę parametrów dźwięku i zazwyczaj jest niesłyszalna dla ludzkiego ucha. Niestety metoda ta nie wykazuje również odporności na obróbkę dźwięku. Większość popularnych przekształceń bezpowrotnie niszczy informację.

Dodawanie informacji przez podmienienie najmniej znaczących bitów we wszystkich próbkach pozwala na uzyskanie dużej pojemności steganograficznej. Może to jednak prowadzić do łatwego odczytania dołączonej informacji oraz wystąpienia słyszalnego szumu. Aby tego uniknąć można do ukrycia informacji wykorzystać tylko niektóre próbki. Do ich wyznaczania należy użyć odpowiedniego algorytmu, który powinien losowo wybierać miejsca umieszczenia bitów danych, uwzględniając jednocześnie parametry dźwięku w celu osiągnięcia jak najlepszych efektów przy jak najmniejszych zniekształceniach sygnału [6,22,39]. Wpływ zakłóceń może być minimalizowany poprzez dobranie odpowiedniego nośnika. Na przykład hałas powodowany przez publiczność podczas nagrań koncertowych jest bardzo dobrym sygnałem maskującym wprowadzane zakłócenia[41]. Inną możliwością prezentowaną w [18] jest kształtowanie dołączanej sekwencji, tak by dopasować ją do sygnału kontenera w celu uzyskania jak najbardziej zbliżonej charakterystyki. Pozwoli to na zredukowanie poziomu wprowadzanych zakłóceń.

Główną wadą metod opierających się na modyfikacji amplitudy jest bardzo mała odporność na operacje wykonywane na dźwięku. Najczęściej już zwykła zmiana formatu zapisu bezpowrotnie niszczy ukrytą informację. Aby zwiększyć odporność metody można stosować korekcję błędów lub powielać ukrywane dane. Odbywa się to kosztem zmniejszenia pojemności steganograficznej.

Niektórzy autorzy proponują zastosowanie metody LSB w dziedzinie transformaty. Najpierw na sygnale kontenera wykonywana jest transformata. Następnie do uzyskanych współczynników dołączane są dodatkowe dane metodą LSB.

Po przeprowadzeniu odwrotnej transformaty uzyskiwany jest sygnał wynikowy zawierający ukrytą informację. W tej metodzie mogą zostać wykorzystane transformaty: Fouriera, Cosinusowa lub falkowa.

Aby uzyskać większą pojemność steganograficzną oraz zmniejszyć poziom wprowadzanych zakłóceń metoda LSB jest łączona z techniką rozpraszania błędów, uwzględnia psychoakustyczny model słuchu[16,17] oraz wykorzystuje efekt maskowania do ukrycia wprowadzanych zakłóceń[1]. Bardzo dużą pojemność steganograficzną udało się uzyskać autorom [1,15,21] przez użycie techniki LSB w dziedzinie transformaty falkowej przy wykorzystaniu maskowania.

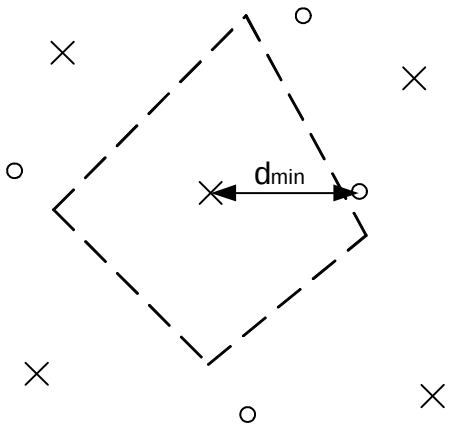
2.2.8.2 Wykorzystanie szumu

Każdy sygnał zawiera pewną ilość szumu. Może on zostać wykorzystany do ukrycia danych. Najprostszym sposobem jest dołączenie do sygnału własnego szumu zawierającego ukryte informacje lub zmodyfikowanie istniejącego [9,63].

Podczas próbkowania sygnału powstają zniekształcenia. Aby je zminimalizować i zachować jak najwierniejszy przebieg sygnału cyfrowego dodajemy sygnał rozpraszający (ang. dither). Jest to odpowiednio kształtowany sygnał o niewielkiej mocy. Dzięki temu możliwe jest znaczne wyeliminowanie zniekształceń kosztem wprowadzenia niewielkiej ilości szumu[41].

Aby dołączyć do sygnału dodatkowe dane wystarczy zmodyfikować sygnał rozpraszający zgodnie z przyjętymi założeniami a następnie przy jego pomocy spróbkować oryginalne nagranie[18]. Technika ta znana jest również pod nazwą modulacji indeksu kwantyzacji (ang. quantization index modulation)[12]. W praktyce realizowana jest przy użyciu dwóch różnych kwantyzerów. Każdy z nich musi dawać inne wartości próbek wynikowych. Wówczas wartość uzyskana za pomocą jednego z nich będzie odpowiadała zakodowanej wartości jeden, a wartość uzyskana z drugiego pozwoli na zakodowanie wartości zero[12]. Rysunek 2.3 obrazuje zasadę działania metody. Punkty oznaczone jako "x" reprezentują wartości uzyskiwane za pomocą jednego kwantyzera, a oznaczone jako "o" są wartościami uzyskanymi z drugiego z nich. Jeśli bit dołączanej informacji ma wartość zero to jako wartość wyjściową przyjmujemy wartość reprezentowaną przez najbliższe "x", a jeśli chcemy zakodować binarną jedynkę na wyjściu pojawi się wartość najbliższego "o". Odległość d_{min} jest miarą odporności metody (im większa odległość tym większa odporność metody).

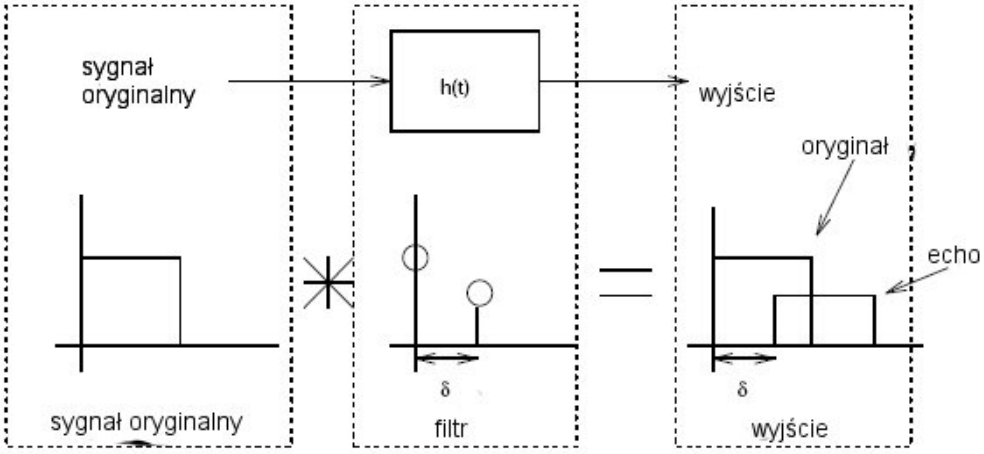
Rozmiar komórki kwantyzacji jest miarą zniekształceń wprowadzanych podczas procesu kwantyzacji.



Rysunek 2.3. Schemat modulacji indeksu kwantyzacji[85]

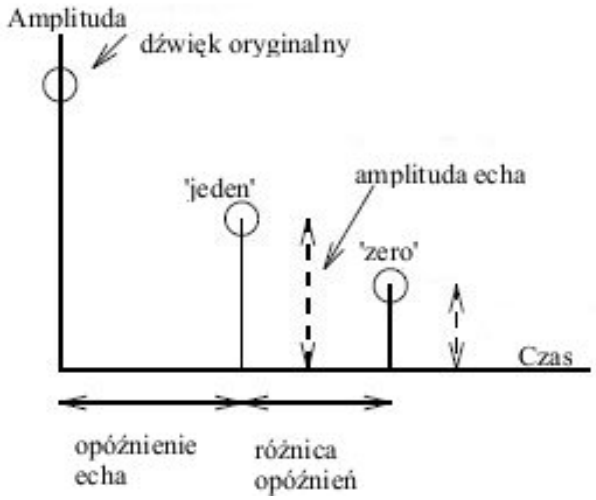
2.2.8.3 Techniki wykorzystujące echo

Słuch nie jest w stanie wychwycić faktu istnienia echa sygnału o sile nie przekraczającej 0.4 amplitudy dźwięku oryginalnego, które występuje w ciągu 2 milisekund po sygnale lub też bezpośrednio przed nim[8,25,26,51]. Można więc wykorzystać to zjawisko do ukrycia dodatkowej informacji. Ukrywanie danych odbywa się za pomocą sterowalnego filtru opóźniającego, którego zadaniem jest dodanie do oryginalnego sygnału jego kopii opóźnionej o zadany czas co zaprezentowane zostało na rys 2.4.



Rysunek 2.4. Proces dodawania echa do sygnału [26]

Aby uzyskać możliwość zakodowania symboli wystarczy przypisać im określone wartości opóźnienia echa np. dla jedynki może to być opóźnienie 1 ms a dla zera 2 ms (rys. 2.5) [18,23,39,64].



Rysunek 2.5. Sygnały echa odpowiadające wartościom logicznym zero i jeden [8]

W pracach [24,68] określone zostały warunki percepcji dodanego sygnału echa. Jako graniczne parametry niesłyszalności przyjęto opóźnienie w granicach 1-2 ms oraz amplitudę nie przekraczającą 0,3-0,5 amplitudy sygnału oryginalnego. Zakodowanie danych binarnych w nośniku wymaga podzielenia go na bloki, w których realizowane jest niezależne przesunięcie echa o zadaną wartość odpowiadającą kodowanej wartości. W rzeczywistości najłatwiej jest utworzyć dwie kopie sygnału z dodanym echem. W pierwszej stosujemy przesunięcie odpowiadające binarnej wartości jeden a w drugim przesunięcie odpowiadające zeru. Następnie obydwa sygnały dzielimy na jednakowe bloki danych, które pobieramy odpowiednio z pierwszego lub drugiego sygnału składając trzeci, który będzie nośnikiem ukrytej informacji. O wiele bardziej problematyczna okazała się zmiana opóźnienia, która powodowała słyszalny efekt „terkotania". Aby go uniknąć należy stosować zmiany opóźnień o co najwyżej 0.05 ms z częstością nie przekraczającą 10 ms. Dobre efekty możliwe są również do uzyskania poprzez płynne zmienianie opóźnienia[23].

Detekcja ukrytego sygnału możliwa jest poprzez obliczenie autokorelacji fragmentu sygnału z opóźnieniami odpowiadającymi wartościom logicznym "1" i "0" [8]. Możliwe jest również wykorzystanie charakterystyki amplitudowej filtra wstawiającego echo do obliczenia cepstrum sygnału. W wyniku otrzymujemy prążki na pozycji k oraz na pozycjach równych wielokrotności k, gdzie k oznacza opóźnienie echa wyrażone w ilości próbek. Położenia prążka o wartości maksymalnej oznacza zastosowane opóźnienie echa i pozwala odczytać zakodowaną wartość. W [24,68] przedstawiono korzystny wpływ wybielania widma sygnału na obniżenie stopy błędów.

Aby zminimalizować stopę błędów odczytu należy zwiększyć odległość pomiędzy sygnałami echa kodującymi zera i jedynki lub wykorzystać różne wartości opóźnienia echa podczas kodowania tej samej wartości[23]. Możliwe jest również wykorzystanie *pre-echa*. Jest to sygnał echa dodawany przed właściwym dźwiękiem. Wówczas wartość logiczna zero jest kodowana poprzez dodanie echa wyprzedzającego sygnał, a logiczne jedynka kodowana jest poprzez dodawanie echa po sygnale. Wykorzystywane jest tu zjawisko maskowania sygnałów cichych przez głośniejsze następujące tuż po nich. Należy zadbać o to, aby odstęp pomiędzy sygnałami i ich wartości spełniały warunki maskowania.

Autorzy [42] prezentują wykorzystanie echa polarnego. Pozwala to na odejście od kodowania informacji w funkcji odległości echa od sygnału. Kodowanie odbywa się poprzez dołączenie echa o różnej polaryzacji. Echo o dodatniej polaryzacji wykorzystywane jest do kodowania jedynki logicznej, zaś echo o polaryzacji ujemnej odpowiada zeru logicznemu. Wykorzystanie echa polarnego zwiększa odporność metody ale jednocześnie wprowadza zmiany barwy dźwięku.

Aby zredukować to niekorzystne zjawisko autorzy [73] zaproponowali wykorzystanie echa bipolarnego. Zniekształcenie powodowane przez dodane echo redukowane jest poprzez dodanie bezpośrednio przed lub za nim słabszego sygnału echa o przeciwnej polaryzacji.

Pojemność steganograficzna metody zależy od wielkości bloków, na które będziemy dzielić sygnał, odstępów pomiędzy blokami oraz nadmiarowości danych.

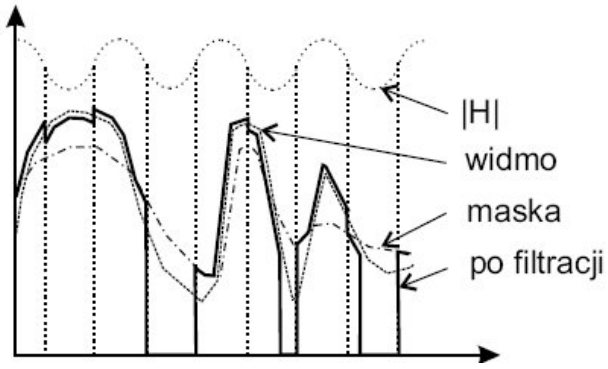
2.2.8.4 Metoda filtracji subpasmowej z wykorzystaniem maskowania

Informację cyfrową można ukrywać modyfikując cepstrum sygnału poprzez dokonywanie zmian wartości średnich w określonych przedziałach lub dodawanie określonej funkcji do wartości cepstrum. Modyfikacja cepstrum może również stanowić operację wspomagającą detekcję echa sygnału, zwiększając niezawodność metody. Polega to na obliczeniu odwrotnej dyskretnej transformaty Fouriera zlogarytmowanego widma sygnału i wykorzystaniu korelacji tego widma i odpowiednich funkcji bazowych[23]:

$$C(k_m) = \sum_{i=0}^{N-1} \hat{X}_i \exp(j2\pi k_m i / N) \qquad (2.17),$$

gdzie: X_i jest próbką zlogarytmowanego widma amplitudy sygnału, k_m – m-ty współczynnik odwrotnej dyskretnej transformaty Fouriera.

W celu zwiększenia wartości korelacji $C(k_m)$ zwiększamy próbki zlogarytmowanego widma w punktach i, dla których $cos(2\pi k_m i/N)>0$ oraz zmniejszamy, gdy $cos(2\pi k_m i/N)<0$. Zakres zmian nie może przekroczyć progu maskowania, aby nie wystąpiły słyszalne zakłócenia. Próg maskowania obliczany jest na zasadzie stosowanej w koderach MPEG-Audio. Modyfikacji podlega jedynie widmo amplitudy. Niezmienne natomiast pozostaje widmo fazowe. Filtracja wykonywana jest w podpasmach wyznaczanych przez znak funkcji $cos(2\pi k_m i/N)$ co obrazuje rysunek 2.6. Próbki widma zwiększa się o M_i jeśli przekraczały próg maskowania lub też do wartości M_i jeśli znajdowały się pod progiem. Po dokonaniu modyfikacji widma amplitudy należy dołączyć oryginalne widmo fazy a następnie wykonać odwrotną transformatę Fouriera.



Rysunek 2.6. Podział na podpasma i modyfikacja widma [23]

Badania zaprezentowane w [23] pokazują, że przedstawiona metoda ukrywania danych daje dobre wyniki zarówno jako metoda samodzielna jak też jako towarzysząca metodzie dodawania echa. Dodatkową zaletą tej metody jest o wiele mniejsza moc sygnału transmisji. Wadą jest możliwość wprowadzania słyszalnych zakłóceń w postaci „terkotania". Spowodowane jest to skokową zmianą wartości widma amplitudy. W celu zmniejszenia tego efektu należy zmodyfikować metodę, tak by na krańcach bloków danych nie wprowadzać zmian. Trzeba je wprowadzać stopniowo wewnątrz bloku, aż do osiągnięcia oczekiwanej wartości a następnie płynnie powrócić do stanu oryginalnego na końcu bloku. Aby było to możliwe konieczne będzie skorzystanie z wzoru:

$$x_n = \sum_{i=0}^{N-1} X_i(n) e^{j\phi_i} e^{j2\pi in/N} \qquad (2.18),$$

który posłuży do wyznaczenia wartości kolejnych próbek zmodyfikowanego sygnału.

$X_i(n)$ zmienia się w obrębie bloku od wartości oryginalnej do docelowej i z powrotem. Przy zastosowaniu bloków o rozmiarze 512 próbek i czasów przejść równych 60 próbek zakłócenia zostały praktycznie wyeliminowane.

2.2.8.5 Metody rozproszonego widma

Jednym z warunków skuteczności ukrycia informacji w sygnale jest jej mała moc. Rozproszenie informacji w całym spektrum widma wykorzystywanego sygnału kontenera pozwala na jednoznaczne dołączenie danych nawet jeśli moc ukrywanego sygnału jest mniejsza od mocy szumu[41]. Ponadto ukrycie informacji w pasmach o wysokiej częstotliwości ma minimalny wpływ na sygnał kontenera. Wykorzystanie pasm o niskiej częstotliwości pozwala uzyskać dużą odporność na uszkodzenia. Rozproszenie informacji we wszystkich pasmach pozwala uzyskać kompromis pomiędzy odpornością i przezroczystością. Tak dołączona informacja wnosi niewielkie zmiany do sygnału i jest odporna na uszkodzenia i usunięcie, ponieważ trudno jest oczyścić cały sygnał nie uszkadzając go w znacznym stopniu. Sposób rozpraszania informacji w szerokim paśmie częstotliwości wywodzi się z telekomunikacji, gdzie jest szeroko stosowany w łączności radiowej. Realizacja w steganografii polega na wymnożeniu sygnału dołączanej informacji z drugim pseudolosowym sygnałem o większej przepływności bitowej[8,18,19,22,43,72,85]. Powoduje to rozproszenie widma sygnału, który następnie łączy się z sygnałem kontenera. W odbiorniku sygnał szerokopasmowy wymnażany jest operacją EX-OR z identyczną sekwencją pseudolosową w wyniku czego następuje kompresja widma i możliwe jest odczytanie dołączonej wiadomości. Do odczytania informacji niezbędny jest klucz, którym jest używana sekwencja pseudolosowa. Aby utrzymać niski poziom zakłóceń moc sygnału znaku wodnego nie powinna przekroczyć 0.5% mocy sygnału kontenera[8].

Metoda rozproszonego widma jest jedną z lepszych metod steganograficznych dzięki wysokiej odporności na wykrycie, uszkodzenie i usunięcie oraz dużej pojemności steganograficznej [41,75].

2.2.8.6 Techniki wykorzystujące kodowanie fazy

Niewrażliwość słuchu na zmianę fazy sygnału jest często wykorzystywana przez algorytmy steganograficzne[8,18,39,70]. Wyróżniamy tu dwie grupy metod:

- kodowania fazy,
- modulacji fazy.

Zasada działania metody kodowania fazy opiera się na podziale oryginalnego sygnału dźwiękowego na fragmenty o określonej wielkości. Dane dołączane są tylko w pierwszym fragmencie poprzez modyfikację widma fazy sygnału. Stanowi to jego wadę ze względu na znaczne ograniczenie pojemności steganograficznej. Ponadto łatwe jest usunięcie dołączonych danych ze strumienia dźwiękowego, ponieważ są one umieszczone w konkretnym fragmencie sygnału, który można usunąć lub uszkodzić.

Dołączanie do nośnika audio za pomocą metod modulacji fazy polega na niezależnej modyfikacji fazy każdego z podpasm sygnału. Aby zmiana ta była niesłyszalna należy wprowadzać jedynie niewielkie zmiany spełniające zależność:

$$\left|\Delta\phi(z)/\Delta z\right| < 30° \qquad (2.19),$$

gdzie $\phi(z)-$ oznacza fazę sygnału, z – jest oznaczeniem skali Barka. Każdy stopień w tej skali jest równoważny częstotliwości krytycznej.

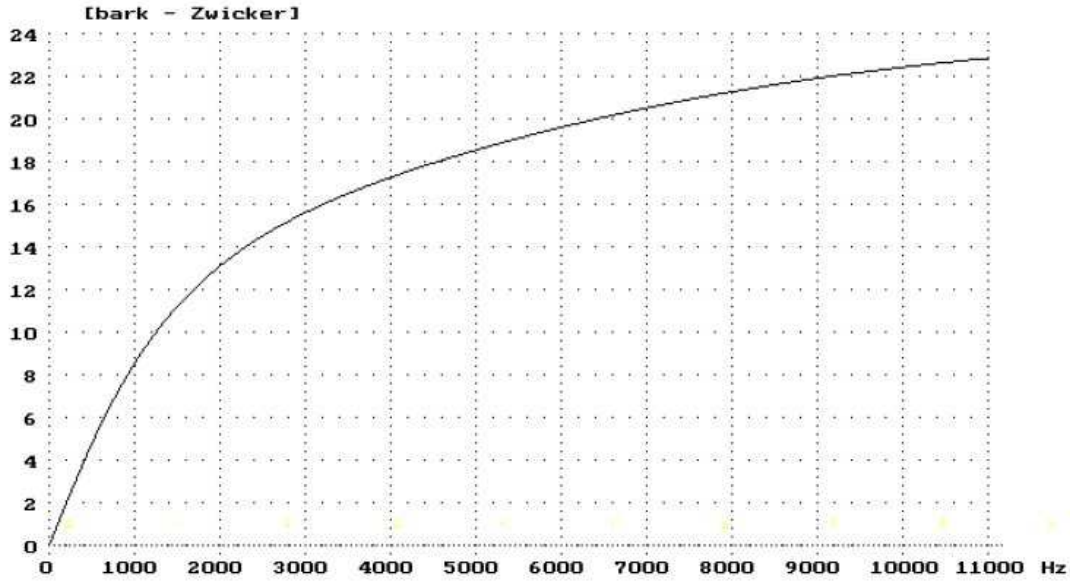
Skala barkowa jest ściśle powiązana z pojęciem **pasma krytycznego**, wynikającego z badań nad percepcją głośności szumu wąskopasmowego (Zwicker) lub zjawisk maskowania tonu prostego przez taki szum (Schröder). Całe pasmo słyszenia zostało podzielone na 24 pasma krytyczne. Możliwe stało się określenie zależności pomiędzy wysokością tonu w barkach a częstotliwością w hercach. Aby przejść na opis częstotliwości wyrażonej w hercach według. Zwickera należy skorzystać ze wzoru:

$$b = 13\arctan(0.00076f) + 3{,}5\arctan((f/7500)^2) \qquad (2.20),$$

gdzie f oznacza częstotliwość wyrażoną w hercach. Zależność pomiędzy skalą częstotliwości a skalą barkową prezentuje rysunek 2.7[20,90].

Skala barkowa według Schrödera obliczana jest według zależności:

$$b = 7 \cdot \operatorname{arcsin} h\left(\frac{f}{0.65}\right) \qquad (2.21).$$



Rysunek 2.7. Zależność pomiędzy liniową skalą częstotliwości a skalą barkową Zwickera [90]

Zmianę fazy wykonujemy tylko raz w każdym z określonych wcześniej bloków danych. Użycie dużych bloków daje gwarancję niewielkich zmian fazy w dziedzinie czasu. Ukrytą informację zapisujemy za pomocą zmian fazy, tak by zmiana o jeden stopień w skali Barka przenosiła informację o jednym bicie dołączanym. Zwiększenie zakresu zmiany kodującej bit pozwala uzyskać większą odporność dołączonych danych, lecz dzieje się to kosztem zmniejszenia pojemności steganograficznej.

Odczytanie dołączonej informacji możliwe jest poprzez porównanie stegokontenera z oryginałem. Należy zwrócić uwagę na poprawne rozpoznanie i dopasowanie bloków danych. Niewielkie zmiany sygnału nie wpływają negatywnie na odczyt danych.

Metoda przedstawiona w [39] również ukrywa informację w przesunięciach fazy sygnału. Operacja ukrycia wykonywana jest poprzez zastępowanie fragmentów sygnału ich odpowiednikami o przesuniętej fazie. Wówczas każdej wartości binarnej odpowiada inne przesunięcie fazy. Dla każdego z fragmentów obliczana jest transformata Fouriera, która pozwala na uzyskanie macierzy faz, która jest odpowiednio modyfikowana, aby odpowiadała ukrywanej wartości binarnej. Umieszczenie obok siebie fragmentów o przesuniętej fazie byłoby niekorzystne ze względu na wrażliwość słuchu na względne przesunięcie fazy. Dlatego też pomiędzy fragmentami pozostawiane są odstępy, w których macierz faz jest modyfikowana, tak by zachować ciągłość fazy sygnału.

Do odczytania ukrytej informacji niezbędne jest posiadanie informacji o położeniu fragmentów będących nośnikami ukrytych danych oraz o ich rozmiarze. Metoda ta umożliwia uzyskanie pojemności steganograficznej rzędu 32 bitów na sekundę.

2.2.8.7 Techniki bazujące na modyfikacji histogramu

Autorzy [80,81] proponują ukrywanie informacji poprzez modyfikację histogramu sygnału. Działanie metody składa się z pięciu etapów.

1. Ze znakowanego sygnału *F*, wybieramy próbki o amplitudzie z zakresu $B=[-\lambda A, \lambda A]$ (λ-nieujemne, *A*- średnia wartość modułu amplitudy w utworze), na podstawie których tworzony jest histogram H_M obrazujący liczbę próbek posiadających amplitudę o określonej wartości. Wielkości przedziałów *M* dobierane są tak, by uzyskać wystarczającą ich liczbę do ukrycia informacji.
2. Poprzez zakwalifikowanie próbek do poszczególnych przedziałów otrzymujemy porcje oznaczone jako π. Następnie próbki wewnątrz każdego z przedziałów przetwarzane są dyskretną transformatą falkową.
3. Tworzona jest pseudolosowa sekwencja pełniąca rolę znaku wodnego, która następnie dołączana jest do histogramu.
4. Za pomocą odwrotnej transformaty falkowej uzyskiwana jest pierwotna postać histogramu, która podlega następnie rozgrupowaniu w celu uzyskania oznakowanego sygnału *F'*.
5. Klucz jest zapamiętywany w detektorze w celu późniejszego wykorzystania.

Ze względu na małą pojemność steganograficzną i dosyć dużą odporność na uszkodzenia metody te stosowane są przeważnie do znakowania wodnego sygnałów.

Wstawianie znaku wodnego rozpoczyna się od jego utworzenia, czyli wygenerowania pseudolosowej sekwencji $W=\{w(i)|i=1,\ldots,P\}$, która zostanie ukryta w utworze *F*. Następnie wybierany jest przedział amplitudy $B=[-\lambda A, \lambda A]$, który będzie podstawą do utworzenia histogramu H_M. Jako że podczas różnych przetworzeń sygnału wartości próbek zmieniają się, należy odnieść się do średniej z modułu amplitudy próbek obliczonej według wzoru:

$$A=\frac{1}{N}\sum_{i=1}^{N}F(i) \qquad (2.22).$$

Autorzy [80,81] sugerują, że najlepsze wyniki można osiągnąć przy wykorzystaniu zakresu $\lambda <0.5A, 2A>$. Do zakodowania jednego bitu potrzebne są trzy przedziały histogramu. Należy więc dobrać rozmiar przedziałów tak, by ich liczba była

wystarczająca do umieszczenia znaku wodnego. Jeśli *P* jest rozmiarem dołączanych danych to liczba przedziałów histogramu *L* nie może być mniejsza niż *3P*.

Zakodowanie pojedynczego bitu polega na odpowiedniej zmianie proporcji pomiędzy liczbami próbek w każdym z przedziałów. Oznaczmy liczbę próbek w przedziałach przeznaczonych do zakodowania jednego bitu jako *a,b* i *c.* Aby ukryć w nich bit danych należy zmodyfikować ich rozmiary zgodnie ze wzorem:

$$\begin{cases} \frac{2b}{a+c} \geq T \; dla \; w(i)=1 \\ \frac{a+c}{2b} \geq T \; dla \; w(i)=0 \end{cases} \qquad (2.23),$$

gdzie *T* jest ustalonym progiem.

Modyfikacja rozmiaru przedziału polega na zmianie amplitudy próbek, tak by znalazły się w sąsiednim przedziale histogramu. Operację tą zawsze wykonujemy jednocześnie na trzech fragmentach histogramu, które będą przechowywać jeden bit informacji. Aby uzyskać przekształcenie odporne na operacje modyfikujące przebieg czasowy utworu (TSM) oraz kompresję mp3 próg *T* musi mieć wartość większą niż 1,1.

W praktyce autorzy [80,81] przedstawiają wyniki przy *T=1.4, λ=2.4* i *P=40.* Ponadto zakładają możliwość poprawnego zidentyfikowania znaku wodnego przy 15% błędów odczytu. Zastosowanie takich wartości współczynników pozwala na uzyskanie odporności na TSM w zakresie od -10% do +10%, zmianę częstotliwości próbkowania, filtrację dolnoprzepustową, dodanie szumu, losowe wstawianie niewielkich fragmentów sygnału, zmianę głośności zakresie +-20%, kompresję stratną mp3 i efekt jitter.

1.2.8.7 Wykorzystanie znaczących fragmentów sygnału

Bardzo dużym zagrożeniem dla wszystkich technik steganograficznych są operacje TSM[51]. Niewiele metod jest w stanie dołączyć informację tak, by przetrwała wydłużenie lub skrócenie utworu o kilka procent. Daje to więc możliwość łatwego usunięcia dołączonej informacji zwłaszcza, że słuch nie jest zbyt wrażliwy na TSM. Obecnie wykorzystywane algorytmy TSM zostały zaprojektowane tak, by zachować jak najlepszą jakość przeskalowanego dźwięku. Możliwe jest to poprzez zróżnicowanie operacji wykonywanych na poszczególnych fragmentach sygnału. Dźwięk nie jest bowiem "rozciągany" równomiernie. Najpierw jest on analizowany w celu określenia fragmentów, na które słuch jest szczególnie wrażliwy. Te fragmenty nie są modyfikowane lub podlegają zniekształceniu w minimalnym stopniu. Natomiast pozostałe części sygnału są skalowane w takim stopniu, aby możliwe było osiągnięcie

pożądanej długości całego utworu. W [58] autorzy zauważyli, że stwarza to możliwość dołączenia znaku wodnego odpornego na TSM. Zaproponowali oni wykorzystanie do ukrywania informacji wycinków utworu niepodlegających modyfikacji podczas TSM. Aby odnaleźć takie regiony należy przeprowadzić filtrowanie pasmowo przepustowe sygnału pozostawiając pasmo d3 (o zakresie 3kHz – 5kHz). Jest to zakres częstotliwości bardzo dobrze słyszanych przez człowieka i nie modyfikowany przez algorytmy TSM. Ponadto dźwięk większości instrumentów zawiera tą składową. Maksima występujące w analizowanym podpaśmie wskazują miejsca, które w minimalnym stopniu podlegają modyfikacji podczas TSM. Lee i Xue zaproponowali wykorzystanie tych miejsc do dołączenia znaku wodnego.

Maksymalny rozmiar dołączanej sekwencji nie powodujący słyszalnych zakłóceń, dla regionu o rozmiarze 4096 próbek, został określony na 64 bity. Przed dołączeniem ukrywana sekwencja podlega zmodyfikowaniu algorytmem binarnego kluczowania fazy (ang. binary phase shift keying – BPSK), który zamienia wartości {0,1} na {-1,1} dając w efekcie sekwencję *W'* zgodnie z zależnością:

$$W'=\{w'(i)|w'(i)=1-2\cdot w(i),\ w'(i)\in\{+1,-1\},\ 1\leq i\leq 64\} \qquad (2.24).$$

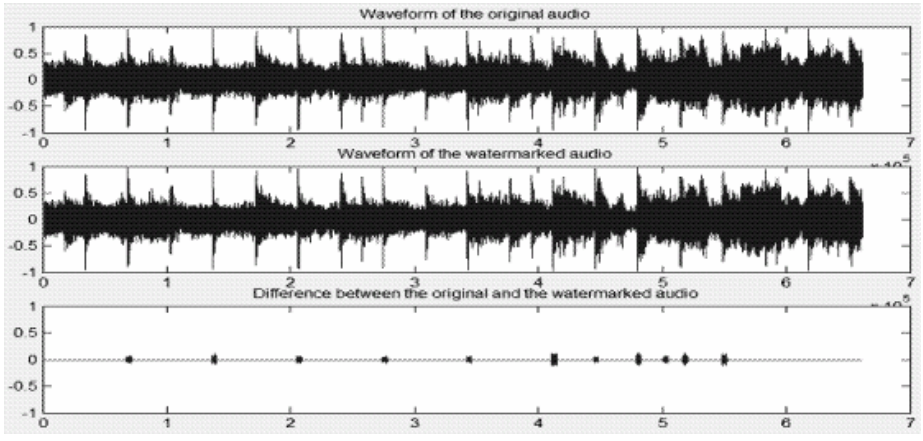
Następnie do regionu dołączania dodawane są kolejno bity *W'* poprzez zamianę współczynników wykonywanej transformaty Fouriera.

Do powrotu w dziedzinę czasu wykorzystywana jest odwrotna transformata Fouriera. Sygnał oznakowany za pomocą przedstawionej metody oraz wprowadzane zakłócenia przedstawione zostały na rysunku 2.8.

(a)

(b)

(c)



Rysunek 2.8. Sygnał oznakowany metodą [58]: (a) sygnał oryginalny, (b) sygnał oznakowany, (c) różnica sygnałów [58]

Testy przeprowadzone przez twórców algorytmu wykazały, że jest on odporny na kompresję mp3 o przepustowości 32kb/s, filtrowanie dolnoprzepustowe, dodanie

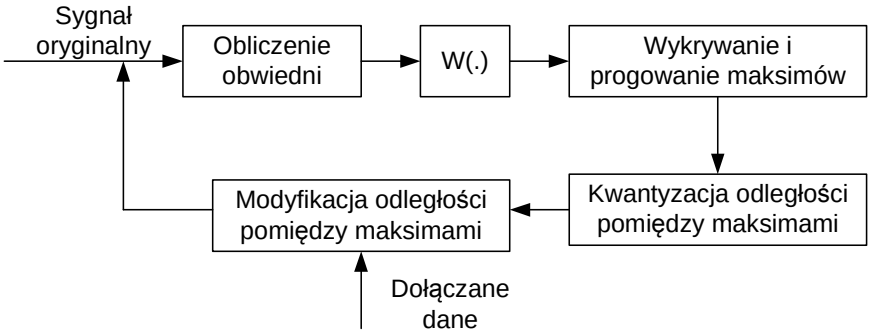
szumu, zmianę częstotliwości próbkowania, dodanie echa, losowe wstawianie dodatkowych fragmentów sygnału o wielkości nie przekraczającej 10000 próbek, odszumianie i TSM do 16%.

W [59] autorzy proponują bardzo podobną metodę dołączania znaków wodnych, w której miejscami dołączania znaku są krawędzie sygnału lub maksima lokalne energii sygnału.

Istnieją również algorytmy przeznaczone dla konkretnego rodzaju dźwięku modyfikujące zależności czasowe występujące w nich. Przykładem może być metoda zaprezentowana w [53], ukrywająca dane w przesunięciach rytmu utworów muzycznych.

1.2.8.8 Modyfikacja odległości pomiędzy znaczącymi punktami sygnału

Mansour i Tewfik w [62] proponują dołączanie informacji do sygnału dźwiękowego poprzez modyfikację odległości pomiędzy znaczącymi punktami sygnału. Położenie tych punktów ustalane jest za pomocą transformaty falkowej. Ekstrema lokalne współczynników falkowych wskazują położenie minimów i maksimów pochodnej sygnału. Tak więc wskazują one położenie krawędzi sygnału. Krawędziami sygnału nazywamy przejścia od ciszy do dźwięku oraz miejsca zmiany częstotliwości podstawowej. Proponowany algorytm opiera się na idei modyfikacji odległości pomiędzy maksimami współczynników falkowych. Schemat algorytmu przedstawiony został na rysunku 2.9.



Rysunek 2.9. Schemat algorytmu[62]

Działanie algorytmu składa się z pięciu kroków:

1. obliczenia obwiedni sygnału poprzez filtrację dolnoprzepustową,
2. przeprowadzenia nieortogonalnej dekompozycji falkowej obwiedni, której współczynniki $C(\tau)$ będą użyte w dalszych krokach,

3. znormalizowania mocy wektora współczynników $C(\tau)$ i wyszukanie w nim maksimów lokalnych,
4. progowania ekstremów w celu pozostawienia tylko największych,
5. modyfikacji odległości pomiędzy wyselekcjonowanymi ekstremami w celu ukrycia informacji.

Algorytm ten opiera się na zasadzie działania algorytmów TSM, które w niewielkim stopniu modyfikują znaczące miejsca w sygnale. Fragmenty sygnału wyszukane za pomocą dekompozycji falkowej są analizowane pod kątem określenia istotności wykrytej zmiany. Do ukrycia informacji wykorzystywane są tylko najbardziej znaczące krawędzie, które nie ulegną modyfikacji podczas przetworzeń sygnału. Odległości pomiędzy poszczególnymi maksimami są kwantowane poprzez określenie przedziałów odległości odpowiadających zakodowanej wartości binarnej jeden oraz zero. Następnie są one modyfikowane zgodnie z ciągiem dołączanych danych tak, aby każda z odległości pomiędzy analizowanymi maksimami znajdowała się w centralnej części przedziału. Pozwoli to uzyskać zwiększoną odporność na modyfikacje wprowadzane w skali czasu, które mogą nierównomiernie "rozciągać" lub skracać sygnał.

Testy przeprowadzone przez twórców algorytmu wykazały, że jest on odporny na kompresję mp3 o przepustowości 32kb/s, filtrowanie dolnoprzepustowe, zmianę częstotliwości próbkowania, odszumianie i TSM do 10%.

1.2.8.9 Metody operujące w dziedzinie transformaty

Metody transformacyjne opierają się na przekształceniu tradycyjnego zapisu sygnału w dziedzinę zastosowanej transformaty. Ukrywanie danych polega na modyfikacji uzyskanych współczynników transformaty, które są następnie poddawane transformacie odwrotnej w celu ponownego uzyskania sygnału w funkcji czasu zawierającego dołączoną informację[41,78]. Autorzy [2,5,15,21,57] proponują algorytmy wykorzystujące transformatę falkową ukrywające informację w wybranych współczynnikach transformaty lub w najmniej znaczących bitach wszystkich współczynników. [5] prezentuje algorytm znakowania wodnego oparty o transformatę falkową fragmentów sygnału, cechujący się dobrą odpornością uzyskaną kosztem znacznego ograniczenia pojemności steganograficznej. W [57] informacja ukrywana jest poprzez modyfikację cech statystycznych w dziedzinie transformaty falkowej. [79] prezentuje algorytm odporny na przekształcenie sygnału do postaci analogowej.

Metody wykorzystujące transformaty Fouriera, kosinusową (DCT), oraz inne podobne przekształcenia zaprezentowane zostały w [34,37,56,60]. Algorytm przedstawiony w [37] wykorzystuje dołączanie wąskopasmowego szumu do wybranego pasma częstotliwości za pomocą DCT. [35] proponuje dołączenie sygnału synchronizacji do dołączanego w dziedzinie DCT znaku wodnego oraz uwzględnienie właściwości ludzkiego słuchu w celu zwiększenia odporności na przekształcenia oraz przezroczystości wprowadzanego znaku wodnego. Prezentowane metody oparte o DCT wrażliwe są jednak na kompresję mp3. Rozwiązanie tego problemu przedstawione jest w [60]. Bazuje ono na modyfikacji energii widma transformaty kosinusowej we fragmentach sygnału.

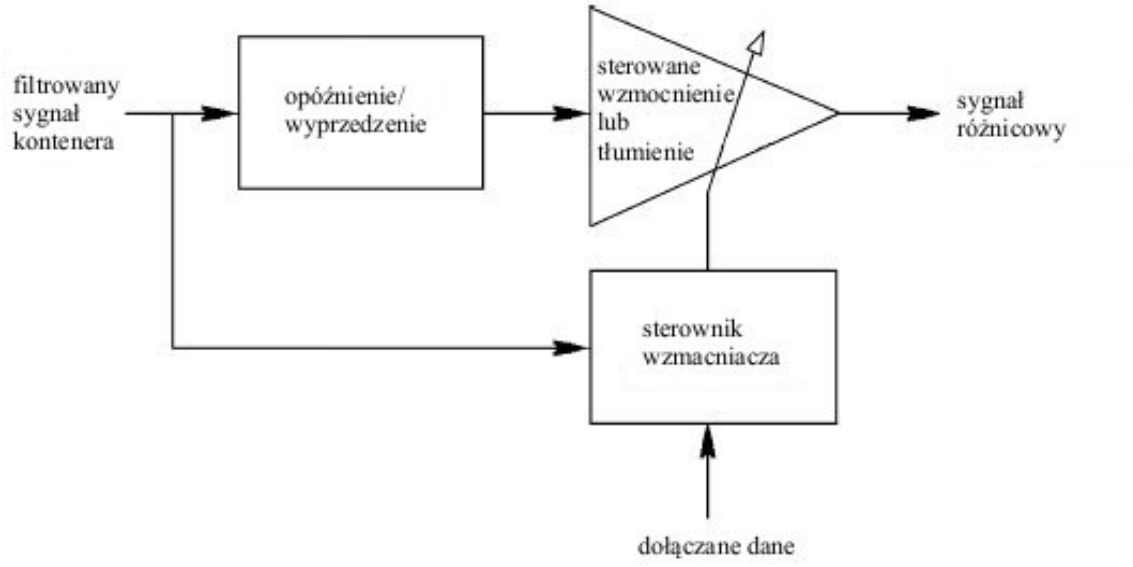
Opracowanie [77] proponuje rozwiązanie wykorzystujące zalety połączenia metod wykorzystujących przekształcenia falkowe i DCT.

Algorytm opisany w [56] ukrywa informacje poprzez wprowadzenie modyfikacji w wybranych współczynnikach cepstrum sygnału za pomocą DFT. To samo przekształcenie zostało zastosowane w [34], przy uwzględnieniu rozkładu energii sygnału w czasie, co pozwoliło uzyskać odporności na usuwanie i wstawianie niewielkich fragmentów sygnału. [82] prezentuje metodę dołączającą znak wodny poprzez modyfikację znaczącego prążka widma. Niestety metoda ta charakteryzuje się niezbyt wysoką odpornością na kompresję.

Metody transformacyjne wykazują dużą odporność na kompresję sygnału. Niestety nie są odporne na modyfikację w skali czasu. Autorzy [77] prezentują metodę odporną na desynchronizację powodowaną wstawianiem losowych fragmentów sygnału oraz zwiększoną odpornością na kompresję. Algorytm dzieli sygnał kontenera na segmenty. Każdy segment jest dzielony na dwie części. Pierwsza jest wykorzystywana do dołączenia kodu synchronizacji pozwalającego na zidentyfikowanie miejsca dołączenia informacji. Do tego celu jest wykorzystywany kod Barkera. Druga część segmentu wykorzystywana jest do dołączenia informacji. Aby zwiększyć odporność procedura ta wykonywana jest dwuetapowo. Najpierw na przetwarzanym fragmencie wykonywana jest transformata falkowa. Następnie na uzyskanych współczynnikach transformaty falkowej odpowiadających niskim częstotliwościom wykonywana jest transformata kosinusowa, w której współczynnikach ukrywana jest dołączana informacja.

1.2.8.10 Techniki wykorzystujące autokorelację

Metoda autokorelacji należy do rodziny metod modyfikacji parametrów statystycznych sygnału. Pierwszym etapem jest odpowiednie przygotowanie sygnału kontenera. Uzyskuje się go poprzez zastosowanie odpowiedniego filtru, którego zadaniem jest utworzenie sygnału zawierającego minimalną ilość zakłóceń i posiadającego jak najlepsze właściwości ukrywające. Tak zmodyfikowany sygnał jest wykorzystywany do uzyskania sygnału różnicowego. Powstaje on w wyniku przepuszczenia przez opóźniający filtr sterowalny o zmiennym wzmocnieniu. Wielkość opóźnienia lub wyprzedzanie sygnału oryginalnego przez uzyskany odpowiada opóźnieniu autokorelacji, które powstanie w sygnale wynikowym. Wzmocnienie modyfikowanego sygnału zależne jest od właściwości kontenera oraz ukrywanych danych. Schemat układu zaprezentowany został na rysunku 2.10.



Rysunek 2.10. Generator sygnału różnicowego [66]

Autokorelacja przetworzonego sygnału może zostać opisana za pomocą wzoru:

$$R(t,\tau) = \int_{t-T}^{t} s(x)s(x-\tau)dx \qquad (2.25),$$

gdzie: s(t) - sygnał kontenera, R(t,τ) - autokorelacja, τ - opóźnienie autokorelacji, T- standardowa różnica czasu pomiędzy sygnałem oryginalnym a skorelowanym, t- czas

Po dodaniu sygnału różnicowego e(t) do sygnału kontenera funkcja korelacji przyjmuje postać:

$$R_m(t,\tau) = \int_{t-T}^{t} (s(x)+e(x))(s(x-\tau)+e(x-\tau))dx$$
$$= R(t,\tau) + \int_{t-T}^{t} (s(x)e(x-\tau)+e(x)s(x-\tau)+e(x)e(x-\tau))dx \quad (2.26).$$

Właściwy wybór sygnału różnicowego pozwala uzyskać różne opóźnienia autokorelacji w sygnale wynikowym, co daje możliwość zakodowania w nich dodatkowej informacji [66]. Technika ta bazuje na słowach kodowych reprezentowanych przez odpowiednie zmiany opóźnień autokorelacji. Wśród metod reprezentacji słowa kodowego można wyróżnić:

- kodowanie proste – każdy symbol posiada przypisany jedną wartość opóźnienia
- wielowartościowe mapowanie symboli – jednemu symbolowi przypisujemy określony zbiór wartości opóźnień
- kodowanie Manchester – sygnał dzielony jest na dwie równe części, dla każdej z nich obliczane są współczynniki autokorelacji, informacja ukrywana jest poprzez modyfikację różnicy pomiędzy współczynnikami
- zmienność opóźnień – poszczególne elementy sygnału posiadają swoje indywidualne opóźnienia, które można zamienić na wartości za pomocą tajnej tablicy klucza[14,75].

Pojemność steganograficzna zależy od zastosowanego sposobu kodowania, oraz częstotliwości zmian autokorelacji. Zazwyczaj pojemność dostępna do wykorzystania jest niewielka ale za to dołączone dane charakteryzują się wysoką odpornością na uszkodzenia. Według [75] dodanie białego szumu od sile 36 dB oraz zmiana tempa o 10% nie uszkadza dołączonej informacji. Ze względu na swoje właściwości najczęściej metoda jest wykorzystywana do znakowania wodnego.

1.2.8.11 Pozostałe metody

Oprócz zaprezentowanych już rozwiązań istnieje jeszcze wiele innych, już nie tak popularnych ze względu na ich ograniczenia, małą przydatność praktyczną czy też z innych powodów. Często nie są one rozwijane, więc w dostępnej literaturze istnieją tylko pojedyncze wzmianki na ten temat. Należą do nich:

- Algorytmy ukrywające informację podczas kompresji poprzez modyfikowane współczynników algorytmu kompresującego. Popularne jest wykorzystywanie

powszechnego formatu mp3. Informacja dołączona w ten sposób nie jest jednak odporna na zmianę formatu zapisu danych[45,87].

- Metody kombinowane polegające na transformacji dźwięku do postaci obrazu, na którym przeprowadzany jest właściwy proces ukrywania informacji. Po jego przeprowadzeniu sygnał ponownie transformowany jest do postaci dźwięku. Tego typu rozwiązanie prezentowane jest w [71].
- Metody zniekształceniowe bazujące na wprowadzaniu zakłóceń w losowych miejscach nośnika. Odczyt ukrytej informacji możliwy jest poprzez porównanie oryginalnego nośnika ze zmienionym. Szczególnym przypadkiem jest wykorzystanie procesu naprawiania uszkodzeń do dołączania dodatkowych danych[55]. Wadą tego systemu jest konieczność zapewnienia bezpiecznego kanału służącego do przesyłania oryginalnych nośników.
- Metody statystyczne pozwalające na ukrycie jednego bitu informacji w bloku danych. Wymagają one podzielenia nośnika na bloki, w których umieszczane będą kolejne bity wiadomości. Ukrywanie polega na odpowiedniej modyfikacji danych w bloku, tak aby wybrana funkcja testująca zwróciła wartość bliską zeru lub znacznie różniącą się od zera. Kluczem systemu opartego na tej metodzie jest znajomość funkcji testującej oraz podziału na bloki.
- Metody generacji nośnika różniące się od przedstawionych powyżej koncepcją systemu działania. Nie korzystają bowiem z istniejących nośników informacji. Pobierają one dane, które mają zostać ukryte, analizują je a następnie generują dla nich nośnik w taki sposób, aby wyglądał na zwykły plik z danymi. Szczególny nacisk przy generowaniu nośnika kładziony jest na prawidłową postać wszystkich charakterystyk. Zwiększa to prawdopodobieństwo, że dany nośnik nie zostanie zakwalifikowany jako zawierający ukryte dane [4,31,41,47,48].
- Algorytmy wykorzystujące określony rodzaj dźwięku jako kontener. Przykładem może być opisana w [83] metoda wykorzystująca syntezę mowy do ukrywania informacji w sygnale mowy.
- Inną możliwością jest umieszczenie danych w szumie, tak aby jego charakterystyka była jak najbardziej naturalna – zbliżona do szumu białego. Tak przygotowany szum dołączany jest do nagrania. Wprowadza to jednak zakłócenia do sygnału. Jeśli będą one słyszalne obniżą jakość dźwięku a ponadto mogą skłonić odbiorcę do przeprowadzenia operacji odszumiania, która uszkodzi lub usunie dołączone

dane. Jeśli chcemy uniknąć wprowadzania słyszalnych zakłóceń musimy znacznie zmniejszyć moc sygnału szumu, tak aby został on zamaskowany przez oryginalny dźwięk. Efektem tego będzie bardzo niska energia sygnału a co za tym idzie niewielka odporność na uszkodzenia oraz trudności w odczytaniu ukrytej informacji[47].

2.3. Metody cyfrowego przetwarzania sygnałów stosowane w dźwięku

Dziedzina cyfrowego przetwarzania sygnałów zawiera wiele dobrze opracowanych teoretycznie metod operujących na cyfrowej postaci sygnału dźwiękowego. W niniejszym rozdziale został przedstawiony krótki przegląd najpopularniejszych metod przetwarzania sygnałów wraz z ich porównaniem i oceną przydatności w steganografii. Na tej podstawie wybrana została jedna z metod, którą wykorzystano jako przekształcenie bazowe w nowej metodzie steganograficznej.

2.3.1. Filtracja sygnałów

Filtracja sygnałów to ogół działań mających na celu przepuszczanie lub blokowanie sygnałów o określonym zakresie częstotliwości lub zawierających określone harmoniczne. Ze względu na szeroki zakres tego zagadnienia w niniejszej pracy została zaprezentowana tylko operacja splotu, za pomocą której realizowana jest filtracja sygnałów dyskretnych przy użyciu komputera.

Splot opisuje operację filtracji sygnału (*x(k)*) przez sygnał (*h(k)*) według wzoru:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \qquad (2.27).$$

Obliczanie polega na odwróceniu drugiego z sygnałów ze względu na *k*, przesunięciu go o *n* próbek, wymnożeniu go z pierwszym sygnałem a następnie zsumowaniu wszystkich iloczynów próbek.

Równanie 2.27 opisuje mechanizm filtracji sygnałów dyskretnych. Sygnał *x(k)* jest sygnałem filtrowanym, *h(k)* jest sygnałem filtrującym będącym najczęściej odpowiedzią impulsową filtra, przez który przechodzi sygnał. W wyniku filtracji w sygnale *x(k)* redukowane lub usuwane są wybrane składowe częstotliwościowe[84].

Użycie różnych sygnałów filtrujących podczas filtrowania różnych fragmentów sygnału umożliwia sterowanie procesem filtracji. Pozwala to na realizację filtracji adaptacyjnej a także na wykorzystanie omawianego przekształcenia do innych celów w tym do steganograficznego ukrywania danych w przetwarzanym sygnale.

2.3.2. Korelacja sygnałów

Funkcję korelacji dla stacjonarnych dyskretnych sygnałów losowych *x(n)* i *y(n)* definiuje się następująco:

$$R_{xy}(m) = E[x(n)y(n-m)] \tag{2.28}.$$

We wzorze 2.28 symbolem *E* oznaczono wartość oczekiwaną po zbiorze wszystkich realizacji procesu dla ustalonego *n*.

Dla sygnałów niestacjonarnych funkcja korelacji jest wartością oczekiwaną (najbardziej prawdopodobną) iloczynu *x(n)* i *y(n-m)*, zależną od wyboru *n* i *m*[84].

W zastosowaniach praktycznych korelację często wykorzystuje się do określania miary podobieństwa pomiędzy sygnałami.

2.3.3. Transformacja Fouriera

Transformacja Fouriera jest transformacją całkową z dziedziny czasu w dziedzinę częstotliwości. Transformata jest wynikiem transformacji Fouriera, która rozkłada funkcję na szereg funkcji okresowych tak, że uzyskany wynik określa jakie częstotliwości składają się na pierwotną funkcję[88]. Transformacja Fouriera jest podstawowym narzędziem analizy częstotliwościowej sygnałów. Prosta i odwrotna transformacja Fouriera zdefiniowane są wzorami:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \tag{2.29},$$

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{-j2\pi ft}df \tag{2.30}.$$

X(f) jest zespolonym *widmem Fouriera* sygnału *x(t)* i zawiera informację o częstotliwościach składowych sygnału (*f* – częstotliwość wyrażona w hercach).

Wykonując transformację Fouriera dokonujemy rozłożenia sygnału na jego składowe o różnych częstotliwościach. Możliwe jest również wykonanie operacji odwrotnej i zsyntetyzowanie oryginalnej reprezentacji postaci sygnału w funkcji czasu. Uzyskujemy ją wykorzystując odwrotną transformację Fouriera (2.30).

W przypadku sygnałów cyfrowych korzystamy z transformacji Fouriera dla sygnału dyskretnego określonej parą równań:

$$X_{\delta}(e^{j\Omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\Omega n} \tag{2.31},$$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\Omega}) e^{j\Omega n} d\Omega \qquad (2.32),$$

gdzie Ω jest ciągłą pulsacją unormowaną względem częstotliwości próbkowania. zgodnie z wzorem:

$$\Omega = \omega \Delta t = 2\pi f \,/\, f_p \qquad (2.33).$$

Równania (2.31) i (2.32) są prawdziwe przy założeniu, że sygnał został spróbkowany zgodnie z twierdzeniem Nyquista z częstotliwością f_p dwukrotnie większą od maksymalnej częstotliwości występującej w sygnale.

W realizacji praktycznych zadań analizy fourierowskiej nie jest możliwe wykonanie sumowania w nieskończonych granicach ze względu na to, że analizie poddawany jest zawsze sygnał o określonej długości. Dlatego też granice te są zawsze skończone. W związku z tym wzory określające transformację Fouriera dla sygnału o długości *N* próbek przyjmą postać:

$$X^{(N)}(e^{j\Omega}) = \sum_{n=0}^{N-1} x(n) e^{-j\Omega n} \qquad (2.34),$$

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X^{(N)}(e^{j\Omega}) e^{j\Omega n} d\Omega \qquad (2.35).$$

Jest to równoznaczne z wyznaczaniem transformaty Fouriera iloczynu analizowanego sygnału i okna prostokątnego, które wycina fragment sygnału do analizy. W wyniku uzyskujemy splot widma sygnału i okna. Pulsacja Ω przyjmuje wartości z przedziału (-π,π). Ze względu na ograniczenia praktycznej implementacji algorytmu konieczne jest wybranie skończonej liczby wartości Ω, dla których obliczane będą wartości widma. Im więcej *prążków* (próbek) widma wyznaczamy tym dokładniejszy wynik uzyskamy. Najczęściej oblicza się tyle prążków widma ile próbek zawiera sygnał. Dodatkowo dla uproszczenia zakłada się, że sygnał jest sygnałem okresowym o okresie *N*. Wówczas równania (2.34) i (2.35) przyjmują postać:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn}, k = 0,1,2,...,N-1 \qquad (2.36),$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j\frac{2\pi}{N}kn}, n = 0,1,2,...,N-1 \qquad (2.37).$$

Transformacja o postaci (2.36) i (2.37) jest powszechnie nazywana w literaturze dyskretną transformacją Fouriera (DFT)[84,88].

2.3.4. Transformacja kosinusowa

W systemach perceptualnego kodowania dźwięku często wykorzystywana jest zmodyfikowana dyskretna transformata kosinusowa (ang. modified discrete cosinus transform) będąca odmianą transformacji kosinusowej. Przejście do dziedziny widmowej realizowane jest według wzoru:

$$X_j(n) = \sum_{k=0}^{N-1} w(k) x_j(k) \cos\left[\frac{\pi}{2N}\left(2k+1+\frac{N}{2}\right)(2n+1)\right] \qquad (2.38),$$

gdzie: $X_j(n)$ - wartości próbek w dziedzinie częstotliwości, *W(k)* – funkcja okna, *N*- długość bloku w dziedzinie czasu, *N/2* – długość bloku w dziedzinie częstotliwości.

Odwrotną transformatę kosinusową oblicza się z zależności:

$$y_j(m) = \sum_{n=0}^{N-1} X_j(n) \cos\left[\frac{\pi}{2N}\left(2k+1+\frac{N}{2}\right)(2n+1)\right] \qquad (2.39),$$

gdzie: $X_j(n)$ - wartości próbek w dziedzinie częstotliwości, *N*- długość bloku w dziedzinie czasu, *N/2* – długość bloku w dziedzinie częstotliwości[20].

2.3.5. Transformacja Z

Przekształcenie Z jest jednym z najczęściej stosowanych przekształceń w cyfrowym przetwarzaniu sygnałów. Transformatę dyskretnego ciągu *x(n)* można wyznaczyć korzystając z definicji:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) \cdot (r \cdot e^{j\omega})^{-n} \qquad (2.40),$$

gdzie *z* jest zmienną zespoloną.

Wyrażenie zmiennej zespolonej w postaci biegunowej ($z = re^{j\omega}$) pozwala na przedstawienie wzoru 2.40 w postaci:

$$X(r \cdot e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n) \cdot (r * e^{j\omega})^{-n} \qquad (2.41).$$

Z równania 2.41 wynika, że transformatę Z ciągu *x(n)* można interpretować jako transformatę Fouriera ciągu *x(n)* pomnożonego przez ciąg wykładniczy. Dla *|z|=1* transformata Z jest równa transformacie Fouriera.

Odwrotna transformata Z wyliczana jest według wzoru:

$$x(n) = \frac{1}{2\pi j} \oint_C X(z) \cdot z^{n-1} dz \qquad (2.42),$$

gdzie *C* jest krzywą zamkniętą obieganą w kierunku dodatnim w obszarze zbieżności *X(z)* i obejmującą początek układu współrzędnych na płaszczyźnie zmiennej *z* [20].

2.3.6. Transformacja Gabora

Transformacja Gabora to rozłożenie sygnału cyfrowego na kombinację liniową przesuniętych w czasie, modulowanych funkcji Gaussa. Uzyskanie pełnego zestawu współczynników Gabora możliwe jest przy użyciu wzoru:

$$C_{m,n} = \sum_{i=0}^{\infty} x(i) \cdot \gamma \cdot (i - m\Delta M) \cdot \exp\left(-\frac{j2\pi n\Delta N i}{L}\right) \qquad (2.43),$$

gdzie: *x(i)* – analizowany sygnał, *γ(i)* – okno analizy, funkcja bi-ortogonalna do okna syntezy *h(i)*, przy czym:

$$h(i) = (\pi\sigma^2)^{0,25} \cdot \exp\left(\frac{-i - 0,5 \cdot (L-1)^2}{2\sigma^2}\right) \qquad (2.44).$$

W zależności 2.44 zastosowano oznaczenia: *L* – długość okna, *ΔM*- przesunięcie w dziedzinie czasu, *ΔN* – przesunięcie w dziedzinie częstotliwości, *m* – indeks czasu, σ- dyspersja funkcji Gaussowskiej zastosowanej do filtracji.

Funkcja okna jest dana w postaci gaussoidy, jak to wynika z równania 2.44. Współczynniki $C_{m,n}$ obliczane są według wzoru 2.43 poprzez przesuwanie zastosowanego okna w dziedzinie czasu oraz częstotliwości.

Analizowany sygnał *x(i)* może być przedstawiony jako suma:

$$x(i) = \sum_{m=0}^{\infty} \sum_{n=0}^{N-1} C_{m,n} \cdot h(i - m\Delta M) \cdot \exp\left(-\frac{j2\pi\Delta N i}{2\sigma^2}\right) \qquad (2.45).$$

Transformacja Gabora jest przydatnym narzędziem w procesie przetwarzania i analizy sygnałów fonicznych. Jednak jej zastosowanie jest ograniczone ze względu na trudności z obliczaniem współczynników $C_{m,n}$ oraz stałą szerokość okna analizy ograniczającą jej precyzję w dziedzinie czasu i częstotliwości[20].

2.3.7. Transformacja falkowa

Transformacja falkowa jest najnowszą z transformacji, charakteryzującą się dużą rozdzielczością zarówno w dziedzinie czasu jak i częstotliwości. Stanowi ona reprezentację sygnału w postaci superpozycji pojedynczych składowych, będących pakietami przebiegów czasowych nazywanych falkami. Falki powstają przez przeskalowanie tego samego podstawowego kształtu zwanego falką macierzystą.

Stanowią one zestaw funkcji analizujących stosowanych do dekompozycji sygnału. Falki analizujące tworzone są z falki macierzystej g(t) poprzez jej przeskalowanie i przesunięcie w dziedzinie czasu według wzoru:

$$g_{b,a}(t) = \frac{1}{a} \cdot g\left(\frac{t-b}{a}\right) \tag{2.46},$$

gdzie: $g_{b,a}(t)$ – funkcja analizująca, $g(t)$ – funkcja macierzysta, a- współczynnik rozszerzenia, b- wartość przesunięcia czasowego.

Dyskretna transformata falkowa DWT (ang. discrete wavelet transform) może zostać obliczona według wzoru:

$$DWT(a,n) = \frac{1}{\sqrt{a}} \cdot \sum_k h\left(\frac{k}{a} - n\right) \cdot x(k) \tag{2.47},$$

gdzie:k – indeks czasu, $h(k)$ – funkcja prototypowa, $x(k)$ – analizowany sygnał cyfrowy.

Ze względu na własność liniowości transformaty falkowej, transformata odwrotna jest sumą elementarnych funkcji skalowanych przy użyciu wag równych współczynnikom otrzymanym w wyniku transformacji prostej[20].

2.3.8. Analiza cepstralna

Znane są dwie definicje cepstrum widma mocy C(τ):

$$C(\tau) = |F\{\log S_{AA}(f)\}|^2 \tag{2.48},$$

$$C(\tau) = F^{-1}\{\log S_{AA}(f)\} \tag{2.49}.$$

We wzorach 2.48 i 2.49 $S_{AA}(f)$ reprezentuje dwuwstęgowe widmo mocy sygnału $f(t)$, czyli $S_{AA}(f)=|F\{f(t)|^2$.

Porównując definicje 2.48 i 2.49 można zauważyć, że jedna z nich wykorzystuje transformację Fouriera a druga transformację odwrotną. Jednak w wyniku obydwu przekształceń otrzymujemy reprezentację cepstralną przedstawiającą sygnał w dziedzinie czasu τ. Jest to dziedzina odpowiadająca odwrotności częstotliwości.

W przypadku sygnału cyfrowego cepstrum może być obliczone jako zestaw współczynników C_r na podstawie równania:

$$C_r = \sum_{i=1}^{m} \ln A_i \cos\left(r\frac{i\pi}{m}\right),\ m = \frac{l_{pr}}{f_p} f_c \tag{2.50},$$

gdzie: r - rząd współczynnika cepstralnego, l_{pr} – liczba próbek w ramce, i – numer kolejnej próbki widma, A_i – amplituda i-tej próbki, f_p – częstotliwość próbkowania, f_c – maksymalna częstotliwość uwzględniona w analizie cepstralnej[20].

2.3.9. Ocena możliwości wykorzystania wybranych metod cyfrowego przetwarzania sygnałów w steganografii sygnałów dźwiękowych

Różnorodność istniejących metod przetwarzania sygnałów daje duże możliwości obróbki i przekształcania sygnałów. Jednak nie każde przekształcenie pozwala na osiągnięcie oczekiwanych rezultatów. Problemem jaki pojawił się na początku pracy badawczej był wybór odpowiedniej metody przetwarzania sygnałów do wykorzystania w projektowanym nowym algorytmie steganograficznym. Ze względu na często występującą trudność dokonania analitycznej oceny metody, przeprowadzona została krytyczna analiza literatury pod kątem porównania wybranych metod cyfrowego przetwarzania sygnałów i wstępnej oceny możliwości ich wykorzystania w steganografii. Wyniki przeprowadzonej analizy prezentuje tabela 2.1

Jak widać w tabeli 2.1 wiele z przedstawionych metod znalazło już zastosowanie w steganografii, poza tym każda z tych metod umożliwia osiągnięcie innych właściwości bazującego na niej algorytmu steganograficznego. Ze względu na opracowywanie coraz skuteczniejszych metod steganoanalizy istniejące metody tracą swoją wartość, konieczne jest opracowywanie nowych, najlepiej bazujących na niewykorzystywanych do tej pory przekształceniach.

Celem autora było opracowanie metody steganograficznej wykazującej odporność na popularne przekształcenia dźwięku. Dlatego też wybrane przekształcenie powinno wprowadzać zmiany odporne na uszkodzenia podczas tych przekształceń. Drugim istotnym kryterium wyboru było istnienie dobrego opisu teoretycznego przekształcenia. Istotna była również odwracalność przekształcenia, precyzyjność, niski poziom wprowadzanych szumów oraz mała popularność w wybranym obszarze zastosowań. Zgodnie z powyższymi założeniami i przeprowadzoną analizą literatury wybór ograniczył się do transformaty Fouriera oraz transformaty falkowej. Jako przekształcenie bazowe wybrana została transformata Fouriera, ze względu na małą popularność wykorzystania w tajnej komunikacji. Analiza literatury wykazała, że przekształcenie to jest stosowane w metodach znakowania wodnego o małej pojemności steganograficznej. Zastosowanie w metodach komunikacji wymagających większej pojemności steganograficznej jest niewielkie w przeciwieństwie do transformaty falkowej.

Tabela 2.1. Porównanie wybranych metod cyfrowego przetwarzania sygnałów [opracowanie na podstawie literatury]

Metoda DSP	Zalety	Wady
Modyfikacja szumu	• duża pojemność informacyjna • brak wpływu na istotną informację przenoszoną przez sygnał • minimalne modyfikacje nośnika	• duża liczba metod steganograficznych opartych o LSB • wprowadzanie zmian statystyk nośnika i łatwa wykrywalność • bardzo mała odporność dołączonych danych na uszkodzenia
Modulacja	• możliwość modulacji pseudoszumu na podstawie ukrywanej wiadomości dodawanego do nośnika • możliwość modyfikacji bezpośrednio amplitudy sygnału • składnik wielu metod steganograficznych	• wprowadzanie znacznego poziomu zakłóceń do nośnika • konieczność znajomości sygnału oryginalnego przy bezpośredniej modulacji amplitudy
Zmiana częstotliwości próbkowania	• możliwość wykorzystania decymacji i interpolacji lub samej interpolacji do wprowadzenia nowych skwantowanych według klucza próbek, • lepsza skuteczność niż modyfikacji szumu	• brak odporności na uszkodzenia, każda modyfikacja uszkadza ukryte dane • może powodować pogorszenie jakości dźwięku, zwłaszcza w przypadku zmniejszania częstotliwości próbkowania
Kwantyzacja	• możliwość wyboru kierunku zaokrąglenia oraz wartości według ustalonych kluczem progów kwantyzacji	• ograniczone możliwości kwantyzacji podczas kompresji • mała odporność dołączonych danych na uszkodzenia • wykorzystywana w wielu metodach steganograficznych
Transformacje DFT i DCT	• rozpraszanie informacji na obszarze całego nośnika • możliwość rozproszenia informacji w szerokim zakresie częstotliwości • odporność na wiele przekształceń sygnału, kompresję, filtrowanie, dołączanie szumu • możliwość ukrycia danych przez modulację amplitudy i fazy • metoda dobrze poznana, precyzyjna i posiadająca dobrze	• duże modyfikacje amplitudy widma wprowadzają słyszalne zakłócenia

	opisane podstawy matematyczne • niski poziom szumów przetwarzania	
Filtrowanie	• możliwość doboru współczynników filtra wg klucza • możliwość modyfikacji parametrów filtra w czasie ukrywania na podstawie wiadomości, • rozproszenie dołączanej informacji	• konieczna znajomość oryginalnego nośnika w przypadku modyfikacji parametrów filtra podczas ukrywania • trudność doboru odpowiednich parametrów pozwalających uniknąć słyszalnych zmian w sygnale
Kompresja	• wykorzystanie gotowych algorytmów kompresji • możliwość ukrywania „w locie” - podczas kompresji sygnału	• bardzo dużo znanych metod steganograficznych wykorzystujących zmodyfikowane algorytmy kompresji do ukrywania informacji • ograniczenie do określonego formatu danych, • mała odporność na uszkodzenia, zmiana formatu przeważnie bezpowrotnie niszczy ukrytą informację
Transformacja falkowa	• rozpraszanie informacji w całym nośniku • duża odporność dołączonych danych na uszkodzenia • niski poziom szumów przetwarzania • możliwość osiągnięcia dużej pojemności steganograficznej	• duża popularność metody w steganografii dźwięku • duże modyfikacje amplitudy widma wprowadzają słyszalne zakłócenia
Dodanie echa	• wysoka odporność na wiele przekształceń dźwięku • niewielka złożoność obliczeniowa procesu ukrywania	• mała pojemność steganograficzna • łatwość wykrycia i usunięcia dołączonej informacji • może wprowadzać słyszalne zniekształcenia do sygnału
Cepstrum	• rozpraszanie informacji na obszarze całego nośnika • odporność na wiele przekształceń sygnału i kompresję • precyzyjna o dobrze opisanych podstawach matematycznych	• możliwość generowania słyszalnych zakłóceń • trudności przy odczytywaniu ukrytych danych
Transformata Gabora	• brak istniejących metod steganograficznych wykorzystujących to przetworzenie	• ograniczona precyzja w dziedzinie czasu i częstotliwości • trudności w wyznaczaniu współczynników $C_{m,n}$

2.4. Aspekty wykorzystania transformaty Fouriera jako przekształcenia bazowego nowej metody steganograficznej

Na podstawie analizy przedstawionej w rozdziale 2.3 przekształcenie Fouriera zostało wybrane jako podstawa do opracowania nowego algorytmu steganograficznego wykorzystującego dźwięk jako kontener. W niniejszym rozdziale przedstawione zostaną istniejące rozwiązania bazujące na transformacie Fouriera. Poruszony zostanie również problem generowania słyszalnych zakłóceń. Opisany zostanie proces słyszenia wraz z jego niedoskonałościami i możliwością ich wykorzystania do ukrycia wprowadzanych zmian poprzez wykorzystanie zjawiska maskowania.

2.4.1. Zastosowanie transformaty Fouriera w steganografii dźwięku

W wyniku transformaty Fouriera uzyskujemy reprezentację częstotliwościową analizowanego sygnału. Modyfikacja wartości poszczególnych prążków lub też fazy sygnału umożliwia wprowadzenie trwałych zmian w sygnale. Może to zostać wykorzystane do ukrycia dodatkowej informacji. Istniejące techniki steganograficzne bazują na zmianie wartości poszczególnych prążków modułu widma lub modyfikują fazę sygnału. Możemy wyróżnić wśród nich kilka grup.

- Techniki najmniej znaczących bitów – metody polegające na zamianie najmniej znaczących bitów modułu widma sygnału na bity ukrywanej informacji. Działają na identycznej zasadzie jak tradycyjne metody najmniej znaczących bitów. Różnica polega jedynie na wykorzystaniu reprezentacji sygnału w dziedzinie częstotliwości a nie czasu. W wyniku wykorzystania widma częstotliwościowego uzyskujemy jednak większą odporność na uszkodzenia ukrytych danych.
- Techniki modyfikujące wartości prążków modułu widma mające na celu uzyskanie wysokiej odporności na uszkodzenia. Wprowadzają one znaczące modyfikacje wartości niewielkiej ilości prążków. Duża wartość zmiany gwarantuje, że nie zostanie ona usunięta w wyniku niewielkich modyfikacji nośnika. Niewielka liczba modyfikowanych prążków sprawia, że zmiana w skali całego sygnału jest niewielka i trudna do wykrycia. Metody te wykazują dużą odporność na uszkodzenia i zniszczenie dołączonej informacji. Odbywa się to jednak kosztem znacznego zmniejszenia pojemności steganograficznej i możliwości wprowadzenia słyszalnych zakłóceń. Do zapisu informacji może zostać wykorzystany tylko jeden

prążek. Odczyt możliwy jest wówczas poprzez porównanie widma stegokontenera z oryginalnym. Metoda ta jednak nie zyskała popularności ze względu na konieczność zapewnienia bezpiecznego kanału do wymiany oryginalnych nośników. Wady tej pozbawione są metody wykorzystujące większą liczbę prążków do zapisu informacji. Najczęściej używane są dwa prążki. W przypadku wykorzystania dwóch częstotliwości modyfikacji podlega stosunek ich wartości. Jeśli większy udział ma częstotliwość f1 to jest to równoznaczne z zakodowaniem jedynki a jeśli f2 to ukryte jest zero W niektórych metodach wykorzystywane są trzy prążki widma. Osiągnięcie stosunku wartości f1<f2<f3 jest interpretowane jako dołączenie zera, f1>f2>f3 jako dołączenie jedynki binarnej. Nie spełnienie żadnej z dwóch powyższych nierówności oznacza, że w danym fragmencie nie została dołączona żadna informacja. Takie podejście pozwala na zwiększenie pewności odczytu oraz zmniejszenie poziomu wprowadzanych zniekształceń poprzez zapewnienie możliwości uniknięcia wprowadzania modyfikacji we fragmentach, które zostałyby w znacznym stopniu zniekształcone.

- Techniki modyfikujące prążki widma o największej energii zostały przedstawione w [19]. Zaproponowano tam wykorzystanie *k* największych współczynników transformaty do ukrycia znaku wodnego wprowadzanego za pomocą DFT. Zaletą tego podejścia jest uzyskanie dużej odporności na uszkodzenia oraz zniszczenie. Aby usunąć tak dołączone dane trzeba w znacznym stopniu uszkodzić sygnał co spowoduje utratę jego wartości. Jednak modyfikacja największych współczynników transformaty wprowadza znaczne zmiany do sygnału. O ile oko ludzkie ma niewielką wrażliwość na zmiany częstotliwości i można zaimplementować proponowaną metodę do znakowania obrazów to w dźwięku nie znajduje ona zastosowania ze względu na dużą wrażliwość ludzkiego słuchu na zmiany częstotliwości. Do zidentyfikowania dołączonego znaku wodnego niezbędne jest porównanie z oryginałem.
- Metody wykorzystujące znaczące punkty sygnału. W [62] przedstawiono metodę opracowaną przez Mansour i Tewfik. Autorzy wskazują, że jest ona w stanie uzyskać bardzo wysoką odporność na różne przekształcenia sygnału. Osiągają tak wysoką odporność przez wielokrotne umieszczenie tej samej sekwencji w utworze. Odbywa się to kosztem znacznego zmniejszenia pojemności steganograficznej, co czyni metodę przydatną jedynie do znakowania wodnego utworów.

- Techniki zmiany fazy sygnału ukrywające informację poprzez wprowadzanie modyfikacji w przebiegu fazy sygnału. Metody te również charakteryzują się wysoką odpornością na uszkodzenia oraz mają małą pojemność steganograficzną ze względu na konieczność modyfikacji fazy w całym przetwarzanym fragmencie.

W metodach wykorzystujących cały przetwarzany sygnał do ukrycia bitu informacji konieczne jest podzielenie sygnału na bloki. Każdy z bloków przenosi wówczas jeden bit informacji. Najpierw blok przekształcany jest za pomocą DFT do reprezentacji częstotliwościowej i określane są wartości wybranych prążków widma. Następnie wykonywana jest modyfikacja tych wartości zgodnie z przyjętym algorytmem. Tak przygotowane widmo transformowane jest z powrotem za pomocą IDFT do postaci przebiegu czasowego. Przetworzone bloki łączone są w jeden sygnał.

Metody bazujące na transformacie Fouriera są z powodzeniem wykorzystywane w steganografii obrazu. W przypadku sygnałów audio ich zastosowanie w technikach anonimowej komunikacji ma marginalne znaczenie ze względu na wprowadzanie słyszalnych zakłóceń przy dołączaniu dużych ilości informacji. Sygnał dołączany za pomocą zaprezentowanych metod musiałby mieć bardzo małą moc, co spowodowałoby jego nieodporność na zakłócenia. Wykorzystanie częstotliwości większych niż 16kHz pozwala uniknąć słyszalnych zakłóceń, jednak to pasmo częstotliwości jest pomijane podczas kompresji stratnej czy zmniejszeniu częstotliwości próbkowania. Ponadto zakres częstotliwości dźwięku zawiera się najczęściej w przedziale od 20Hz do 10kHz. Pojawienie się wyższych częstotliwości jest bardzo łatwe do zauważenia podczas analizy częstotliwościowej sygnału. W efekcie obecność dołączonej informacji jest łatwa do wykrycia.

2.4.2. Selektywność układu słuchowego

Znajomość własności ludzkiego słuchu (HAS) jest również wykorzystywana w dziedzinie steganografii. Istotne jest tu bowiem wykrycie niedoskonałości słuchu, które pozwolą na niezauważalną modyfikację sygnału dźwiękowego.

Słuch ma relatywnie niewielką wrażliwość na zmiany fazy sygnału stacjonarnego. Ponadto zauważamy silną tendencję do maskowania sygnałów występujących w niewielkim odstępie czasu. Przykładem może tu służyć zupełne zagłuszanie dźwięków cichych przez głośne niezależnie od ich brzmienia.

Postrzeganie dźwięków przez ludzki aparat słuchu jest oparte na jego analizie w uchu środkowym. Reprezentacja sygnału nie jest tu jednak ciągła w dziedzinie

częstotliwości, lecz bazuje na określonej ilości pasm zwanych pasmami krytycznymi. Narząd słuchu jest obecnie modelowany jako zestaw filtrów pasmowo-przepustowych o częściowo nakładających się pasmach o szerokości około 100 Hz w zakresie 500 Hz – 5kHz. W zakresie wyższych częstotliwości ucho jest o wiele mniej wrażliwe, więc szerokości pasm przepustowych filtrów modelujących znacznie się zwiększają. Można więc zauważyć, że człowiek nie jest w stanie rozróżnić większości dźwięków o zbliżonych częstotliwościach.

Najczęściej do celów steganografii wykorzystywane jest maskowanie w dziedzinie częstotliwości lub też w dziedzinie czasu poprzez umieszczenie dźwięków na tyle blisko, by jeden z nich został zagłuszony przez drugi. Zjawiska te pozwalają na dodanie dodatkowych dźwięków niesłyszalnych dla człowieka, które mogą być nośnikiem dodatkowej informacji.

Należy jednak pamiętać, że pomysł wykorzystania niedoskonałości słuchu wykorzystywany jest w algorytmach kodowania i kompresji stratnej[88]. Istnieje więc możliwość usunięcia lub uszkodzenia dołączonej informacji podczas przetwarzania stegokontenera lub zmiany formatu zapisu[40].

2.4.3. Możliwości maskowania zmian częstotliwości kontenera dźwiękowego

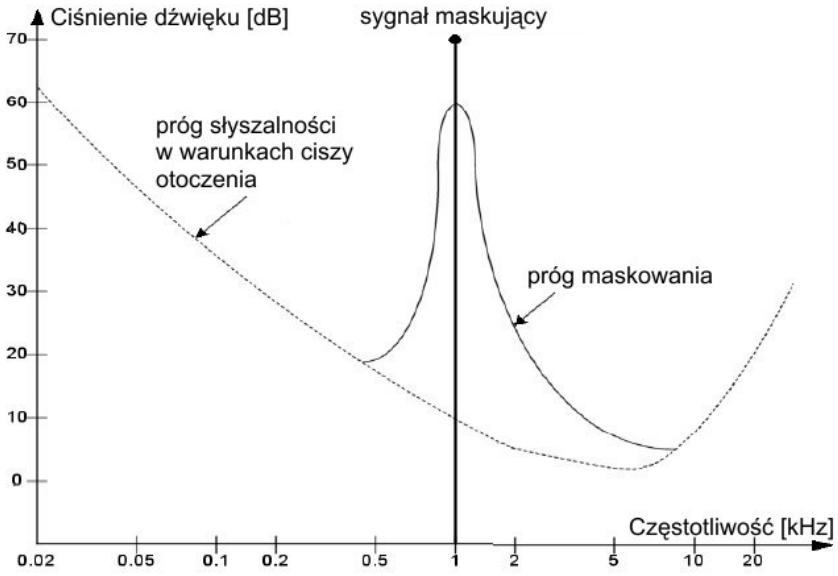
Maskowanie to zjawisko, które powoduje, że słuch nie jest w stanie zarejestrować pewnych dźwięków (maskowanych), ponieważ są one „zagłuszane przez inne dźwięki (maskujące)[38]. Możemy wyróżnić dwa rodzaje maskowania:

- nierównoczesne,
- równoczesne.

Maskowanie nierównoczesne (czasowe) polega na blokowaniu percepcji sygnału przez głośny sygnał następujący w odstępnie czasu nie większym niż 40 ms po maskowanym lub do 200 ms przed nim. Zjawisko to dotyczy maskowania dźwięków odpowiednio cichszych od maskera. Pojawienie się dźwięku głośniejszego zostanie zauważone w każdym momencie.

Maskowanie częstotliwościowe (równoczesne) polega na maskowaniu dźwięku cichszego przez równocześnie występujący dźwięk głośniejszy o zbliżonej częstotliwości. Warunkiem maskowania jest to, aby dźwięk maskowany znajdował się poniżej progu maskowania. Wartość progu maskowania zależy od częstotliwości oraz charakteru tonu maskowanego i maskującego (czy będzie to czysty dźwięk, czy też

wąskopasmowy szum). Zależność tę dla sygnału o częstotliwości 1kHz maskującego zmiany przedstawia rysunek 2.11

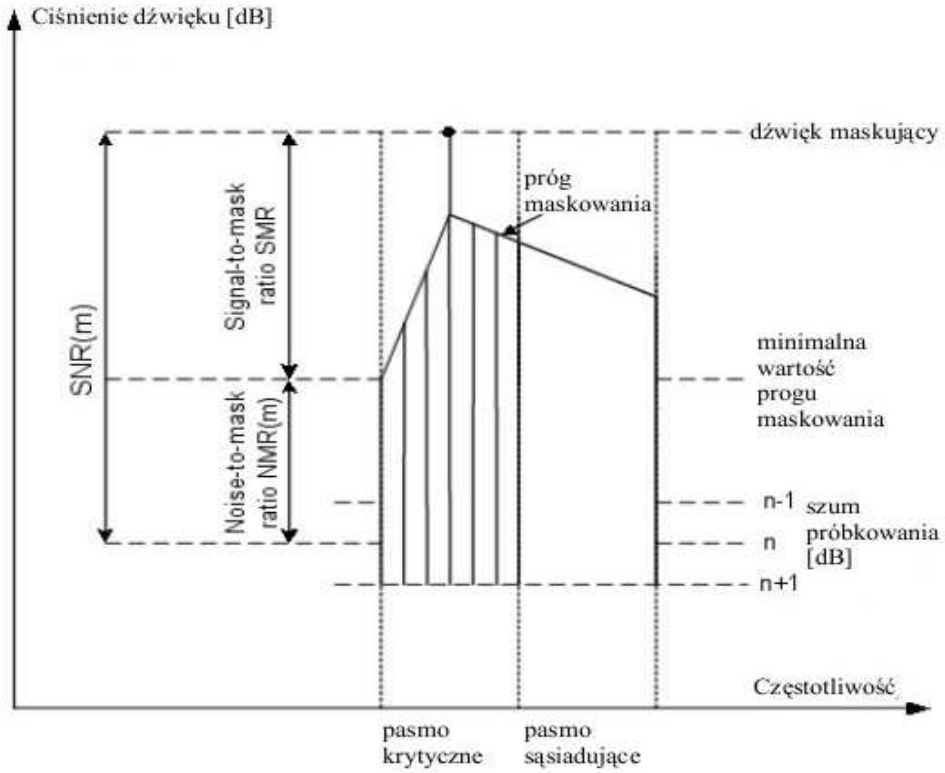


Rysunek 2.11. Próg maskowania równoczesnego dla 1kHz sinusoidalnego sygnału maskującego, przy maskowaniu „czystego” dźwięku [14]

O wiele trudniejsze jest maskowanie szumu, ponieważ słuch jest szczególnie wyczulony na występowanie tego typu zakłócenia. Tak więc próg słyszalności w przypadku maskowania szumu za pomocą czystego dźwięku będzie dużo niższy niż w przypadku maskowania czystego dźwięku szumem. Odległość pomiędzy sygnałem maskującym a maskowanym jest oznaczana symbolem SMR. Wartość maksymalną posiada ona na lewej granicy *pasma krytycznego* będącego elementarnym pasmem częstotliwości zawierającym w sobie moc akustyczną równą mocy akustycznej tonu prostego o częstotliwości f położonej w środku tego pasma, przy czym rozpatrywany ton prosty ma taką intensywność, że zagłuszany przez nieograniczone widmo szumów ciągłych znajduje się na granicy słyszalności. Wewnątrz tego pasma szum powodowany dołączaniem informacji jest słyszalny dopóki odległość sygnału od szumu SNR jest większa niż SMR. Jeśli przyjmiemy, że wszystkie zakłócenia zostały spowodowane dołączeniem dodatkowych danych i przez SNR(*m*) oznaczymy zakłócenia w *m*-tym paśmie krytycznym wówczas słyszalne zakłócenia mogą zostać określone jako odległość szumu od progu słyszalności tego zakłócenia NMR obliczane według wzoru:

$$NMR(m)=SMR - SNR(m) \qquad (2.51).$$

Zmienność wartości SMR i SNR w dziedzinie częstotliwości została zaprezentowana na rysunku 2.12.



Rysunek 2.12. Wartości SMR i SNR [14]

2.5. Konkretyzacja celu pracy

Na podstawie pozytywnej oceny przydatności transformaty Fouriera połączonej ze zjawiskiem maskowania cel badania skonkretyzowano w następujący sposób:

Skonkretyzowanym celem pracy jest opracowanie nowej metody budowy efektywnych stegokontenerów dźwiękowych steganografii komputerowej, bazującej na połączeniu transformaty Fouriera ze zjawiskiem maskowania częstotliwościowego.

Aby zrealizować cel pracy przyjęto następujący zakres prac:

1. porównanie metod DSP,
2. wybór metody stanowiącej podstawę opracowywanej metody steganograficznej,
3. opracowanie algorytmu ukrywania informacji,
4. wyznaczenie parametrów metody oraz określenie ich zakresów zmienności,
5. określenie składników klucza steganograficznego,
6. określenie jakości dołączania przez analizę wprowadzanych zniekształceń oraz odporności na uszkodzenia,
7. porównanie opracowanej metody z innymi znanymi algorytmami steganograficznymi pod kątem stopnia zniekształcenia nośnika, pojemności steganograficznej oraz odporności na uszkodzenia.

3. Podstawy teoretyczne oraz badania doświadczalne nowej metody steganograficznej

W rozdziale pierwszym przedstawiono analizę obecnego stanu technik steganograficznych wykorzystujących dźwięk jako kontener. Pozwoliło to na zauważenie istotnej luki, jaką stanowi brak efektywnych metod wykorzystujących możliwości oferowane przez transformatę Fouriera. Skuteczność działania tych metod w dziedzinie steganografii obrazu sprawiła, że temat ten stał się przedmiotem zainteresowania autora, co zaowocowało projektem nowej metody steganograficznej.

Rozdział drugi zawiera badania teoretyczne oraz opis opracowanego algorytmu steganograficznego, który bazuje na przekształceniu Fouriera. Dodatkowo, aby wprowadzone zmiany nie były słyszalne przez człowieka wykorzystane zostało zjawisko maskowania częstotliwościowego.

W celu osiągnięcia odporności na zmianę częstotliwości próbkowania oraz filtrowanie dane dołączane są w paśmie słyszalnym dźwięku poprzez zmianę wartości wybranych prążków transformaty Fouriera.

Przedstawiono również wyniki badań weryfikujące właściwości opracowanej metody i pozwalające na określenie dopuszczalnych zakresów zmienności parametrów a także wpływu wprowadzanych zmian na stegokontener.

3.1. Model nowego systemu steganograficznego

Opracowany model systemu steganograficznego pozwala na ukrycie informacji w nośniku będącym cyfrowym zapisem sygnału dźwiękowego. Jego schemat został zaprezentowany na rysunku 3.1.

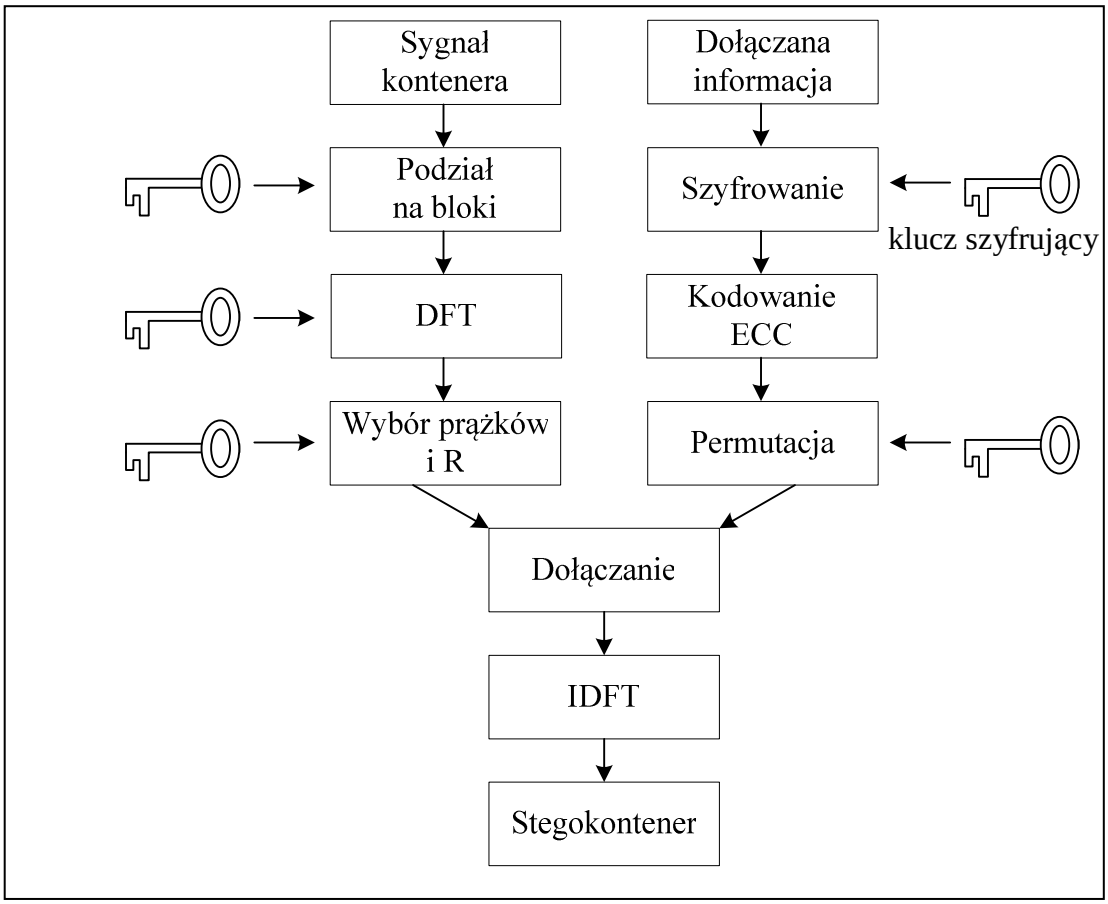
W początkowym etapie przetwarzanie odbywa się w dwóch niezależnych ścieżkach. W jednej z nich przetwarzana jest ukrywana informacja, natomiast w drugiej obróbce poddawany jest kontener.

Celem przetwarzania ukrywanej informacji jest przygotowanie jej do ukrycia i nadanie charakteru jak najbardziej zbliżonego do sekwencji losowej. Składa się ono z trzech etapów:

1. Szyfrowania mającego na celu dodatkowe ochronienie informacji przed odczytem, jak również zmianę charakteru dołączanego ciągu znaków, tak by nawet po wykryciu i odczytaniu nie można było określić, czy są to konkretne dane, czy losowy ciąg bitów. W ten sposób zwiększamy bezpieczeństwo dołączanych danych

metody i dodatkowo łączymy dwie techniki ochrony informacji w celu uzyskania jak najskuteczniejszej ochrony. Do zaszyfrowania niezbędna jest znajomość klucza jaki będzie stosowany. Etap ten nie wpływa jednak na poziom bezpieczeństwa steganograficznego metody, może więc zostać pominięty na etapie badań.

2. Kodowania ECC – wprowadzenie kodu korekcji błędów (ang. Error Correcting Code), pozwalającego na naprawienie pewnej ilości błędów powstających podczas przetwarzania stegokontenera.
3. Permutacji, której zadaniem jest dodatkowe wymieszanie danych. Do jej przeprowadzenia niezbędna jest znajomość macierzy permutacji, która jest kolejną częścią klucza steganograficznego. Zadaniem tego etapu jest wymieszanie bitów dołączanej informacji, tak by uzyskać ciąg jak najbardziej zbliżony do losowego. W ten sposób nawet po wyodrębnieniu ze stegokontenera ukryta informacja będzie przypominała przypadkowy zbiór bitów nie budzący podejrzeń.



Rysunek 3.1. Schemat opracowanego modelu stegosystemu [opracowanie własne]

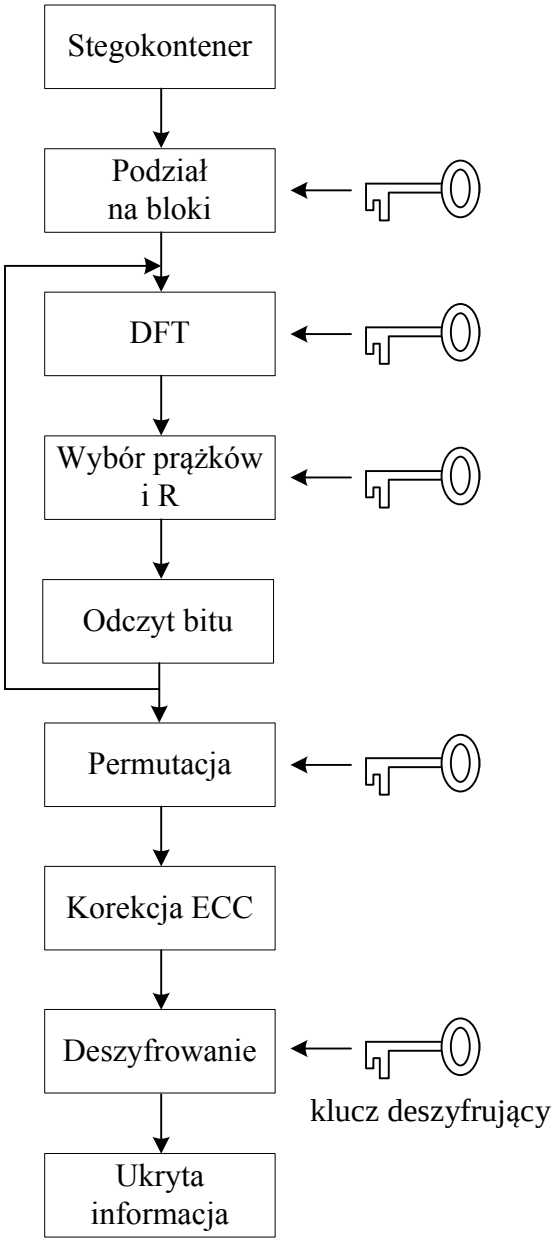
Obróbka kontenera składa się z trzech etapów:

1. Podziału na bloki - polegającego na podzieleniu sygnału na fragmenty, z których połowa będzie mogła być użyta do dołączenia informacji. Reszta służy do łączenia bloków przenoszących informację. Wielkość oraz położenie fragmentów definiuje klucz steganograficzny.
2. Dyskretnej transformaty Fouriera, której zadaniem jest zmiana reprezentacji czasowej sygnału na reprezentację częstotliwościową, czyli obliczenie widma.
3. Wyboru prążków i określenia *R*. W tym etapie wybierane są prążki widma, które będą przenosić informację. Wybór dokonywany jest niezależnie w każdym fragmencie. Do przeniesienia jednego bitu informacji wykorzystywane są dwa prążki. Wartość wprowadzanych zmian określana jest proporcjonalnie do wielkości największego prążka, tak by metoda mogła adaptować wartość wprowadzanych zmian do energii sygnału w danym fragmencie. Pozwoli to na uniknięcie wprowadzania słyszalnych zakłóceń w cichych fragmentach sygnału. Wybór prążków oraz wielkość zmian zależą również od klucza steganograficznego.

Po przygotowaniu wiadomości oraz wykonaniu DFT, określeniu prążków oraz wartości wprowadzanych zmian, algorytm przechodzi do ukrywania bitu w przetwarzanym fragmencie. Odbywa się to poprzez modyfikację wartości wybranych prążków widma, tak by uzyskać pomiędzy nimi różnicę wartości, wyznaczoną przez algorytm dla przetwarzanego bloku.

Po wprowadzeniu zmian, na bloku wykonywana jest odwrotna transformata Fouriera, w wyniku której następuje powrót do dziedziny czasu a uzyskany sygnał jest już stegokontenerem zawierającym ukryty bit informacji. Przetworzony fragment wstawiany jest do sygnału w miejsce jego oryginalnej postaci.

Rysunek 3.2 prezentuje schemat drugiej części modelu stegosystemu przeznaczonej do odczytywania informacji ze stegokontenera. Jak widać na przedstawionym schemacie, do odczytania dołączonej informacji niezbędne są wszystkie składniki klucza. Cały proces przebiega w odwrotnej kolejności niż podczas ukrywania.

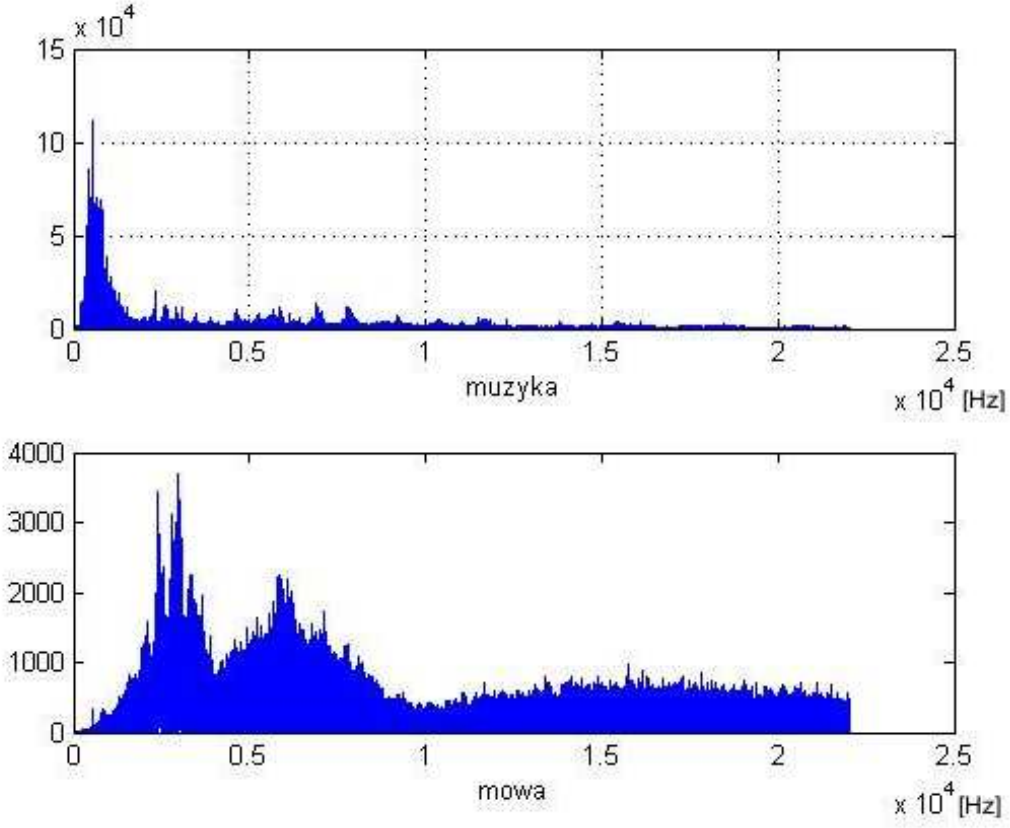


Rysunek 3.2. Schemat stegosystemu odczytującego ukrytą informację [opr. własne]

3.2. Metoda dołączania wiadomości do kontenera

Jak już wcześniej zostało napisane, kontenerem jest sygnał dźwiękowy. Omawiany algorytm ukrywa jeden bit informacji w przetwarzanym fragmencie. Należy więc podzielić kontener na mniejsze fragmenty (bloki) w celu uzyskania większej pojemności steganograficznej. Wyodrębniony blok transformowany jest do dziedziny częstotliwości za pomocą transformaty Fouriera. Następnie obliczany jest moduł uzyskanego widma. Na jego podstawie można określić, jaki jest udział ilościowy poszczególnych częstotliwości w badanym fragmencie.

Kolejnym etapem jest wybór prążków widma do modyfikacji. Aby dołączona informacja była odporna a zarazem niesłyszalna, muszą one spełniać dwa warunki: znajdować się w paśmie słyszalnym oraz w pobliżu odpowiednio silniejszego maskera. Aby je spełnić należałoby ograniczyć analizę do pasma słyszalnego a następnie odnaleźć w nim prążki maskowane. Jednak analizując rozkład częstotliwości w rzeczywistych nagraniach dźwiękowych można zauważyć, że częstotliwości o największej energii, które są najlepszymi maskerami znajdują się najczęściej w paśmie słyszalnym. Nawet wysoki poziom szumów występujący w nagraniu nie zmienia tego rozkładu. Obrazuje to rysunek 3.3, na którym przedstawione zostały widma przykładowych nagrań audio. Nagranie pierwsze jest dobrej jakości, natomiast drugie jest mocno zaszumione.



Rysunek 3.3. Wykres rozkładu częstotliwości w przykładowych nagraniach [opracowanie własne]

Przeprowadzone badania wykazały, że prawie wszystkie prążki widma o największej amplitudzie znajdują się w paśmie słyszalnym. Nie będą one więc usuwane podczas kompresji stratnej oraz innych przekształceń kontenera dźwiękowego. Dlatego też prążki te zdecydowano się wykorzystać do maskowania obecności

wprowadzanej informacji. Uwzględniając powyższe założenia możemy zacząć wyszukiwanie odpowiednich prążków widma do ukrycia informacji od znalezienia maskera. Odbywa się to poprzez znalezienie wartości maksymalnej w module widma przetwarzanego bloku. Wartość tą oznaczmy jako W_{max}. Następnie określane jest położenie prążka o tej wartości, oznaczanego jako p_{max}. Jego częstotliwość oznaczmy jako f_{max}.

Położenie prążków widma częstotliwościowego przeznaczonych do ukrycia informacji będzie determinowane przez trzy czynniki:

1. położenie w obszarze maskowanym,
2. maksymalną odległość prążków od f_{max} zdefiniowaną w kluczu steganograficznym,
3. rozdzielczość transformaty Fouriera.

Aby poprawnie wyznaczyć położenie prążków widma musimy uwzględnić wszystkie trzy aspekty. Przede wszystkim należy zadbać o to, by wyznaczone prążki znajdowały się w obszarze maskowanym przez wyznaczonego maskera. Tu oczywiście najkorzystniejsze jest położenie jak najbardziej zbliżone do f_{max}. Jednak ze względu na bezpieczeństwo istotne jest wykorzystanie prążków oddalonych w różnym stopniu od maskera. Nie mamy tu jednak nieograniczonej ilości prążków do wykorzystania, gdyż ograniczani jesteśmy przez transformatę Fouriera, która pozwala na uzyskanie określonej liczby prążków widma. Aby uniknąć wprowadzania modyfikacji poza pożądanym obszarem, dodano do klucza opis zakresu, w którym mogą być wykonywane modyfikacje. Realizowane jest to poprzez określenie maksymalnej wartości prążka przeznaczonego do modyfikacji oraz największej dopuszczalnej różnicy częstotliwości pomiędzy nim a maskerem. W tak określonym obszarze wyszukiwane są prążki spełniające postawione warunki. Jeśli nie uda się znaleźć prążków widma spełniających powyższe założenia to wykorzystywane są dwa dowolnie wybrane prążki.

Kolejnym etapem jest określenie oczekiwanej różnicy wartości (R) pomiędzy wykorzystywanymi prążkami. Aby zmiany nie były słyszalne nie może ona być zbyt duża. Należy więc dopasować ją do mocy sygnału. Nie wystarczy tu jednak określenie średniej mocy dla całego sygnału, ze względu na zmienny charakter dźwięku w czasie. Autor proponuje więc niezależne dopasowywanie mocy wprowadzanych zmian w każdym fragmencie. Pozwoli to na uniknięcie zakłóceń w cichszych fragmentach nagrania i zapewni odpowiednią moc dołączania we fragmentach głośnych. Konieczne

jest więc określenie sposobu określania mocy zmian. Ważne jest również, by odbiorca mógł ustalić, jaka wartość została użyta. Pozwoli mu to na odczyt ukrytej informacji.

Ze względu na losowy przebieg sygnału nie jest możliwe odgórne ustalenie wartości *R*. Konieczne jest ustalenie sposobu określania mocy zmian. Autor proponuje wykorzystanie wartości W_{max} jako poziomu odniesienia. Wielkość zmian byłaby określana w kluczu steganograficznym jako proporcja do W_{max} według wzoru:

$$R = W_{\max} \cdot R_p \tag{3.1}$$

Wartość przechowywaną w kluczu oznaczmy jako R_p i nazwijmy *siłą dołączania.*

Po określeniu wszystkich parametrów możemy przystąpić do ukrywania bitu *(b)* w przetwarzanym bloku. Polega ono na przetworzeniu sygnału w taki sposób, by spełnić zależność:

$$\begin{cases} |w_1 - w_2| \geq R, dla\ b = 1 \\ |w_1 - w_2| < \beta R, dla\ b = 0 \end{cases} \tag{3.2}$$

gdzie:

β jest wartością umieszczoną w kluczu oznaczającą maksymalny zakres wartości dodawanej do obliczonej wartości prążka widma (podawana jako wielkość procentowa), w_1,w_2 są wartościami prążków wybranych do ukrycia informacji.

Po spełnieniu powyższej nierówności, fragment transformowany jest z powrotem do dziedziny czasu za pomocą odwrotnej transformaty Fouriera (IDFT) i umieszczany w sygnale w miejsce oryginalnego.

Pełny przebieg algorytmu ukrywania bitu informacji *b*=1 we fragmencie sygnału zaprezentowany został poniżej.

1. Sygnał przetwarzany jest za pomocą DFT w wyniku czego uzyskujemy wektor wartości zespolonych Y_c,
2. wyliczamy wartość bezwzględną z wartości wektora $Y_r=|Y_c|$,
3. w Y_r wyszukujemy wartość maksymalną $W_{max}=max(Y_r)$,
4. obliczamy oczekiwaną różnicę $R= W_{max} \cdot R_p$,
5. określamy położenie f_{max} prążka widma o wartości maksymalnej p_{max} oraz prążków spełniających wymogi stawiane prążkom przenoszącym dołączoną informację,
6. Wybieramy prążki widma p_1 i p_2, przeznaczone do dołączenia dodatkowej informacji oraz określamy ich wartości w_1 i w_2,
7. Dla każdego z nich na podstawie klucza określamy maksymalną dopuszczalną wartość prążka: W_1*_dopuszczalna* i W_2*_dopuszczalna*

8. Jeśli $|w_2-w_1|\geq R$ to koniec algorytmu (prążki widma mają odpowiednią wartość).
9. Jeśli $|w_2-w_1|\leq R$ to sprawdzamy który prążek jest mniejszy, a który większy, oznaczamy odpowiednio ich wartości: w_m, w_w. Następnie obliczamy docelowe wartości prążków widma:
   - jeśli $wm/\theta max+R\leq ww$ to $wm=ww-R-rnd(\beta)$, ($rnd(\beta)$ – funkcja zwracająca liczbę losową z przedziału $<-\beta, \beta>$, θmax -maksymalna wartość przez jaką można podzielić wartość prążka widma podczas jego zmniejszania),
   - jeśli $wm/\theta_{max}+R>w_w$ to $w_m=w_m/\theta_{max}$, $w_w=w_m+R+rnd(\beta)$.
10. Na podstawie wyliczonych wartości aktualizujemy wektor Y_c,
11. Za pomocą IDFT zaktualizowany wektor Y_c zmieniamy na sygnał w funkcji czasu.

Gdy istnieje konieczność wprowadzenia zmian wartości prążków widma to najpierw modyfikacji poprzez zmniejszenie poddawany jest prążek o mniejszej wartości. Pozwala to na ograniczenie wzmocnienia drugiego prążka a często pozwala na zupełne uniknięcie wprowadzania modyfikacji. W ten sposób redukujemy energię drugiej wzmacnianej częstotliwości a co za tym idzie uzyskujemy zniekształcenie, które o wiele łatwiej będzie maskowane – większy z prążków będzie dodatkowo maskował zmiany wprowadzone w mniejszym. Wartość mniejszego prążka mogłaby zostać zredukowana do zera. Jednak ze względu na to, że w widmie sygnału praktycznie nie występują wartości zerowe autor zdecydował, że niekorzystne byłoby wprowadzenie ich ze względu na możliwość ich łatwej detekcji podczas analizy częstotliwościowej sygnału. Dlatego też wartość mniejszego prążka jest dzielona przez określony współczynnik θ, $\theta \in (1, \theta_{\max} >$, gdzie θ_{max} jest zawartą w kluczu graniczną wartością współczynnika θ.

3.2.1. Określenie rozmiaru bloku

Rozmiar bloku zastosowanego podczas ukrywania informacji ma istotny wpływ na działanie metody. Od niego zależy maksymalna rozdzielczość transformaty Fouriera, co przekłada się na liczbę uzyskanych prążków widma a także na odległości pomiędzy nimi. Jeśli będą one zbyt duże to nawet prążki widma najbliższe f_{max} mogą znaleźć się poza obszarem maskowania. Duże bloki pozwalają na uzyskanie dużej rozdzielczości transformaty Fouriera. Skutkuje to pojawieniem się większej ilości prążków widma w obszarze maskowanym oraz mniejszymi odległościami pomiędzy nimi wpływając

korzystnie na bezpieczeństwo steganograficzne metody. Korzystne byłoby tu zastosowanie bloków o jak największym rozmiarze.

Od rozmiaru bloku będzie również zależała pojemność steganograficzna. Jeden blok przenosi jeden bit ukrytej informacji. Im mniejsze będą bloki tym większą pojemność steganograficzną będziemy w stanie uzyskać.

Niezbędne jest więc określenie kompromisu pomiędzy przedstawionymi powyżej wymaganiami. Na rysunku 3.4 zaprezentowano wyniki badań, określające słyszalność wprowadzanych zakłóceń w zależności od rozmiaru bloku. W eksperymencie do ukrycia informacji wykorzystywano dwa kolejne prążki widma występujące po f_{max}. Ich położenie zależało więc od rozdzielczości transformaty Fouriera. Im mniejszy był rozmiar użytego bloku tym większa była odległość pomiędzy prążkami widma. Do celów eksperymentu funkcję określającą maksymalne dopuszczalne wartości zmodyfikowanych prążków zastąpiono wartością stałą, co umożliwiło uniezależnienie ograniczeń nakładanych na prążki od ich odległości od f_{max}.

Metodologia przeprowadzania testów odsłuchowych

Ze względu na brak możliwości dokonania numerycznej oceny słyszalności wprowadzanych zakłóceń, niezbędne było przeprowadzenie testów przy udziale ludzi. Do wszystkich badań wykorzystano słuchawki studyjne, dynamiczne zamknięte o parametrach:

- średnica wkładki: 40mm,
- pasmo przenoszenia: 20Hz – 20 kHz,
- moc maksymalna: 100mW,
- efektywność: 110dB dla częstotliwości 1kHz.

Każde z badań przeprowadzone zostało metodą podwójnie ślepego testu odsłuchowego. Uczestnicy testu otrzymywali zbiór losowo ponumerowanych katalogów zawierających po dwa utwory: oryginalny oraz przetworzony. Każdy z uczestników został poproszony o ocenienie słyszalności różnic pomiędzy nagraniami w pięciostopniowej skali:

1 – brak różnic,

2 - brak pewności czy różnice występują czy nie,

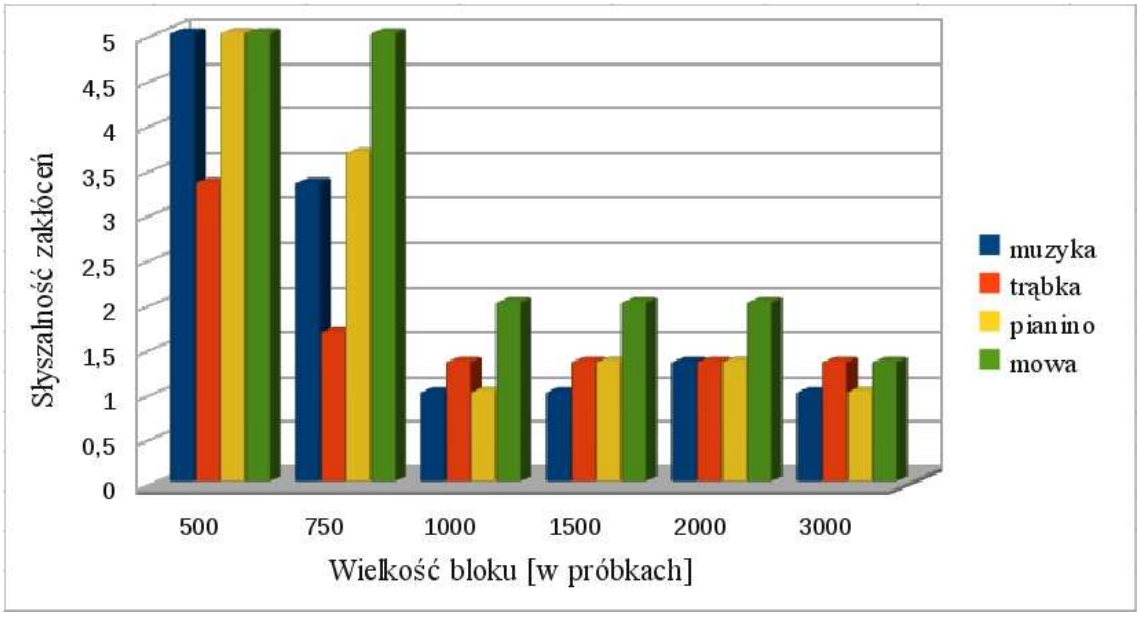
3 – występowanie bardzo słabych zakłóceń,

4 – słabe zakłócenia,

5 – wyraźnie słyszalne zakłócenia.

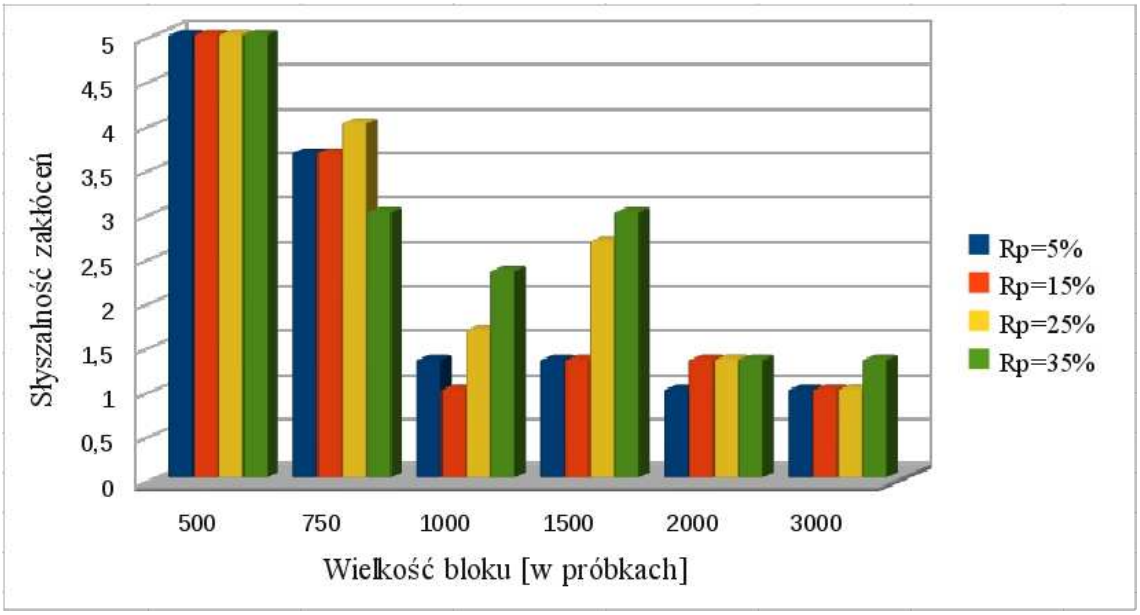
Aby umożliwić zweryfikowanie jakości otrzymanych wyników dodano do każdego z zestawów badawczych katalogi zawierające po dwa utwory oryginalne oraz kilka z przygotowanych wcześniej katalogów umieszczono w teście dwukrotnie.

W teście słyszalności zakłóceń w funkcji rozmiaru bloku otrzymane wyniki dodatkowo poddawane były weryfikacji. Odrzucono spośród nich te, w których tester określił poziom różnic pomiędzy dwoma oryginalnymi nośnikami na więcej niż 2, oraz te, w których rozbieżność w ocenie tego samego, umieszczonego wielokrotnie utworu, była większa niż 1.



Rysunek 3.4. Słyszalność zakłóceń w funkcji rozmiaru bloku [opracowanie własne]

Jak widać na rysunku 3.4 dla sygnałów próbkowanych z częstotliwością 44100Hz, w których informacja ukryta została z siłą dołączania R_p=15% użycie bloków o rozmiarze mniejszym niż 1000 próbek generowało słyszalne zakłócenia. Uwzględniając fakt, że wielu ankietowanych oceniało różnice pomiędzy dwoma identycznymi oryginalnymi fragmentami na poziomie "2", możemy uznać, że bloki o rozmiarze nie mniejszym niż 1000 próbek nadają się do ukrywania informacji i pozwalają na uniknięcie wprowadzania słyszalnych zakłóceń. Zastosowanie większych sił dołączania wpływa na zwiększenie poziomu zakłóceń i wymaga stosowania większych rozmiarów bloków, by osiągnąć możliwość zamaskowania wprowadzanych zmian. Wyniki testu odsłuchowego uwzględniającego wpływ siły dołączania na słyszalność zakłóceń przy zastosowaniu określonego rozmiaru bloku przedstawione zostały na rysunku 3.5.



Rysunek 3.5. Słyszalność zakłóceń w funkcji rozmiaru bloku z uwzględnieniem wpływu siły dołączania [opracowanie własne]

3.2.2. Dobór oczekiwanej różnicy wartości pomiędzy prążkami widma przenoszącymi informację

Podstawowym zagadnieniem podczas opracowywania metody było ustalenie sposobu kodowania dołączanych danych. Jako, że celem było uzyskanie metody wykazującej odporność na przekształcenia dźwięku, przeprowadzono badania określające zmiany różnych parametrów sygnału podczas poddawania go popularnym przekształceniom. Poszukiwano parametrów najmniej podatnych na zmiany wprowadzane podczas modyfikacji. Szczególnie interesujące okazały się wyniki uzyskane w dziedzinie częstotliwości. O ile wartości poszczególnych prążków widma ulegały znacznym zmianom to ich wzajemny stosunek wartości pozostawał zbliżony do oryginalnego, nieznacznie zmieniając się wraz ze zwiększaniem odległości pomiędzy prążkami widma poddawanymi badaniu. W tabeli 3.1 zaprezentowano zmiany wybranych parametrów podczas przekształcania przykładowego sygnału. Przedstawione wartości były obliczane według wzorów:

$$R_{max}=|(W_{max}-W'_{max})/W_{max}| \quad (3.3),$$

$$R_1=|(W_1-W'_1)/W_1| \quad (3.4),$$

$$R_2=|(W_2-W'_2)/W_2| \quad (3.5),$$

$$R_5=|(W_5-W'_5)/W_5| \quad (3.6),$$

$$RP_1=|W_1/W_{max}-W_1'/W'_{max}| \quad (3.7),$$

$$RP_{12}=|(W_1-W_2)/W_{max}-(W_1'-W_2')/W'_{max}| \quad (3.8),$$

$$RP_{34}=|(W_3-W_4)/W_{max}-(W_3'-W_4')/W'_{max}| \quad (3.9),$$

$$RP_{56}=|(W_5-W_6)/W_{max}-(W_5'-W_6')/W'_{max}| \quad (3.10),$$

$$RP_x=|(W_{21}-W_{32})/W_{max}-(W_{21}'-W_{32}')/W'_{max}| \quad (3.11),$$

gdzie: W_{max} – wartość największego prążka widma, W_n – wartość n-tego prążka widma po największym $n \in \{1,2,3,4,5,6,20,32\}$, znakiem ' oznaczono wartości uzyskane z sygnału przekształconego. Zmiana wartości oznaczonych wzorami 3.7 – 3.11 została odniesiona do wartości największego prążka widma, ze względu na wykorzystywane go w określaniu mocy wprowadzanych zmian, co zostanie opisane w dalszej części rozdziału. Przedstawione wyniki badań dotyczą przekształceń nagrania zapisanego w formacie "wav" o próbkowaniu 44,1kHz i 16 bitowej próbce.

Uzyskane wyniki skłoniły autora do podjęcia decyzji o ukrywaniu dodatkowych danych poprzez modyfikację wartości różnicy pomiędzy dwoma prążkami widma oraz do wykorzystania amplitudy największego prążka widma jako wartości odniesienia.

Informacja dołączana jest do fragmentu sygnału poprzez modyfikację jego widma częstotliwościowego w taki sposób, by osiągnąć założoną różnicę wartości *R* pomiędzy dwoma wybranymi prążkami widma f_1 i f_2. Istotne jest by *R* było znane zarówno nadawcy jak i odbiorcy. Ponadto wprowadzana zmiana wartości nie może być zbyt duża, gdyż spowodowałoby to pojawienie się słyszalnych zakłóceń. Zbyt mała wartość *R* może skutkować zmniejszeniem odporności metody. Ze względu na zmienność przebiegu sygnału niemożliwe okazało się dobranie jednej wartości *R* wspólnej dla całego sygnału. Przeprowadzony eksperyment wykazał wprowadzanie wyraźnych zakłóceń podczas wykonywania zmian o stałej wielkości równej średniej wartości *R* obliczonej dla R_p=10%. Zniekształceń nie zanotowano przy wykorzystywaniu niezmiennego *R* o mocy dopasowanej do najcichszego ze wszystkich fragmentów, jednak odbiło się to niekorzystnie na odporności dołączonych danych co prezentuje tabela 3.2. Ze względu na zmiany wartości prążków wprowadzane podczas przekształceń prawie wszystkie odczytane bity miały wartość "1". Przy zbyt małej oczekiwanej różnicy wartości *R* niemożliwe okazało się utrzymanie różnic pomiędzy prążkami przenoszącymi informację w zakresie odpowiadającym wartości "0". Ponadto w niektórych przypadkach, ze względu na bardzo małą różnicę wartości, dawały się nawet zauważyć błędy w odczycie spowodowane zaokrągleniami współczynników transformaty Fouriera.

Tabela 3.1. Zniekształcenia wprowadzane podczas przekształceń dźwięku [opracowanie własne]

Operacja/ zastosowany format zapisu	R_{max}	R_1	R_2	R_5	RP_1	RP_{12}	RP_{34}	RP_{56}	RP_x
próbkowanie 22 kHz	0,070	0,047	0,037	0,033	0,005	0,003	0,003	0,003	0,002
próbki 8bit	0,044	0,022	0,014	0,012	0,002	0,000	0,000	0,000	0,002
aac 128kbit/s	0,044	0,038	0,033	0,028	0,003	0,005	0,000	0,004	0,001
aac 64kbit/s	0,052	0,049	0,044	0,038	0,001	0,004	0,004	0,003	0,009
ape	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
filtrowanie dolnoprzepustowe 0-10kHz	0,079	0,041	0,027	0,018	0,000	0,000	0,000	0,000	0,000
mp3	0,043	0,025	0,021	0,018	0,000	0,000	0,007	0,001	0,000
ogg	0,015	0,017	0,016	0,014	0,003	0,002	0,003	0,001	0,000
wma quality=50%	0,021	0,020	0,021	0,020	0,000	0,001	0,004	0,003	0,004
wma quality=98%	0,003	0,003	0,002	0,002	0,000	0,000	0,001	0,000	0,000
podbicie basów 0-200Hz	0,673	0,293	0,187	0,088	0,011	0,002	0,007	0,018	0,006
narastanie poziomu	0,501	0,239	0,149	0,113	0,001	0,001	0,000	0,001	0,000

Tabela 3.2. Odporność badanej metody dla *R* o mocy dopasowanej do bloku o najmniejszej energii sygnału [opracowanie własne]

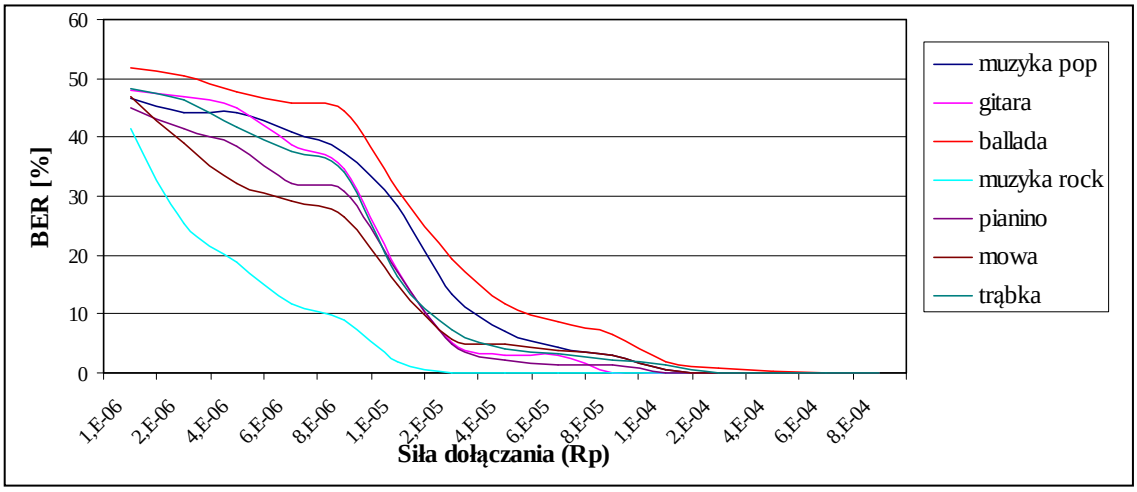
Operacja lub zastosowany format zapisu	Bity poprawnie odczytane [%]	Bity błędne [%]	Bity oznaczone jako niepewne [%]
ogg	50,4	49,4	0,2
aac 128kbit/s	48,6	51,2	0,2
aac 64kbit/s	49,3	50,5	0,2
wma quality=98%	53,1	46,8	0,1
wma quality=50%	50,2	49,7	0,1
ape	51,5	48,4	0,1
mp3 128kbit/s	50,5	49,5	0
wav 22050Hz	50,4	49,5	0,1
wav 8bit/próbka	49,1	50,7	0,2
filtr dolnoprzepustowy 0-10kHz	53,1	46,9	0

Przeprowadzone eksperymenty wykazały konieczność opracowania metody dopasowywania mocy dołączania do mocy sygnału. Ze względu na przyjętą zasadę ukrywania informacji oraz problemy przedstawione powyżej autor zdecydował się na ustalanie wartości *R* niezależnie dla każdego z bloków. Jest to rozwiązanie bardzo korzystne ze względu na możliwość ustalenia siły dołączania niezależnie dla każdego fragmentu, pozwalające na dopasowanie jej do mocy sygnału w danym punkcie. Pozwala to na wykorzystanie fragmentów sygnału o małej amplitudzie bez wprowadzania tam słyszalnych zakłóceń, jak również zapewnia odpowiednią siłę dołączania we fragmentach głośnych. Umieszczanie w kluczu informacji opisujących każdy blok z osobna spowodowałoby jego nadmierną wielkość. Niezbędne jest więc opracowanie metody obliczania *R* na podstawie charakterystyki sygnału oraz parametru zamieszczonego w kluczu. Parametrami mającymi bezpośredni wpływ na skuteczność maskowania są energia maskera i prążków widma maskowanych oraz ich odległości od maskera. Dlatego też postanowiono wykorzystać wartość prążka maskera (W_{max}) jako poziom odniesienia przy ustalaniu siły dołączania, która będzie wyrażona przez różnicę wartości (*R*) pomiędzy prążkami widma przenoszącymi informację.

Ze względu na odwrotnie proporcjonalną zależność pomiędzy poziomem wprowadzanych zniekształceń a odpornością dołączonej informacji niekorzystne byłoby odgórne ustalenie proporcji pomiędzy W_{max} i R. Proporcja ta powinna być dobierana przez urywającego w zależności od wymagań stawianych tworzonemu stegokontenerowi. Dlatego też w kluczu steganograficznym umieszczona została wartość R_p, która będzie określała proporcję pomiędzy W_{max} i R zgodnie ze wzorem 3.1.

Zastosowanie takiego rozwiązania pozwala na dopasowanie wartości wprowadzanych zmian do mocy maskera, niezależnie w każdym fragmencie, tak by nie przekroczyły one progu słyszalności. Zapewnia również odpowiednią siłę dołączania pozwalającą na uzyskanie dobrej odporności na uszkodzenia. Ponadto w ten sposób zapewniamy odbiorcy możliwość obliczenia R dla poszczególnych fragmentów na podstawie wspólnego R_p.

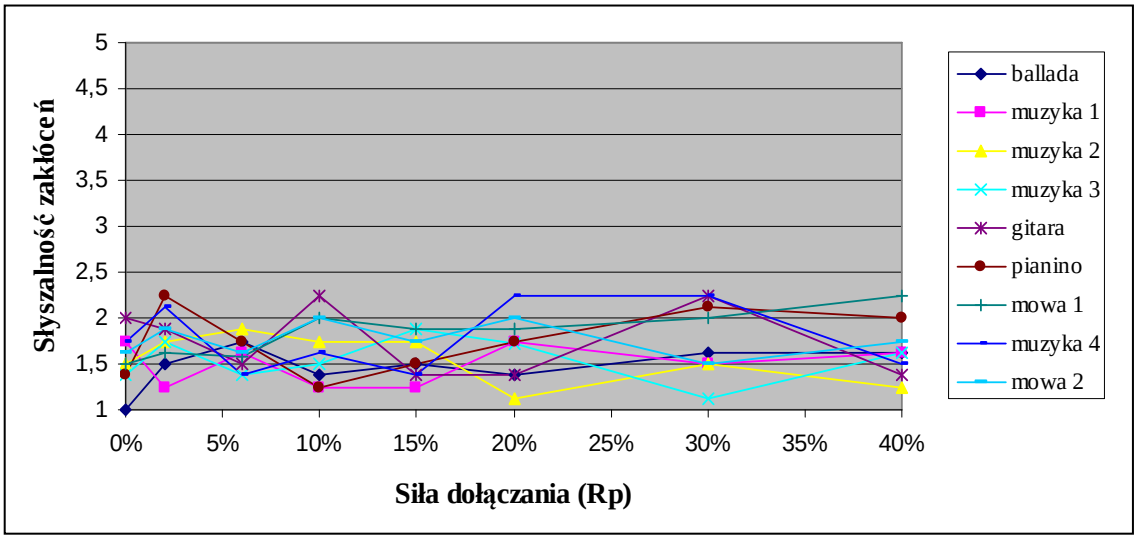
Ważnym zagadnieniem jest ustalenie dopuszczalnego zakresu zmienności R_p. Małe wartości będą korzystne ze względu na niewielki poziom wprowadzanych zniekształceń i trudność wykrycia istnienia dołączonych danych. Jednak odbędzie się to kosztem odporności na uszkodzenia. Zwiększenie wartości wprowadzanych zmian spowoduje wzrost odporności kosztem zwiększonego poziomu zakłóceń. Aby określić minimalną wartość R_p przeprowadzono eksperyment pozwalający na oszacowanie siły dołączania koniecznej do poprawnego odczytania ukrytych danych. Sygnał zawierający dodatkową informację zapisywany był w formacie "wav", a następnie podejmowano próbę odczytu dołączonej informacji oraz określenia ilości błędnie odczytanych bitów. Wyniki zostały przedstawione na rysunku 3.6.



Rysunek 3.6. Ilość błędnie odczytanych bitów w funkcji siły dołączania [opr. własne]

Jak widać na rysunku 3.6 dla małych sił dołączania występują błędy odczytu. Ich liczba zwiększa się w miarę zmniejszania wartości R_p. Spowodowane jest to zaokrągleniami podczas obliczania transformaty Fouriera oraz kwantyzacji sygnału. Dla zestawu badanych sygnałów dopiero zastosowanie siły dołączania $R_p=6\cdot10^{-4}$ pozwalało na bezbłędny odczyt ukrytych danych.

Zwiększanie siły dołączania nie ma negatywnego wpływu na poprawność odczytu, natomiast skutkuje zwiększaniem odporności dołączonych danych na uszkodzenia oraz wprowadza coraz większe zmiany do sygnału, co w pewnym momencie powoduje powstawanie słyszalnych zakłóceń. Górną granicę wartości siły dołączania można wyznaczyć jedynie za pomocą testów odsłuchowych. Za dopuszczalną siłę dołączania można uznać taką, przy której nie występują zniekształcenia dające się jednoznacznie zidentyfikować w teście odsłuchowym. Test został przeprowadzony dla nagrań, w których dokonano modyfikacji z siłą dołączania R_p zawierającą się w przedziale od 2% do 40%. Aby móc w wiarygodny sposób zinterpretować uzyskane wyniki dołączono do testu również oryginalne utwory (testerzy porównywali dwa oryginały). Ocena wprowadzanych zniekształceń była podawana w pięciostopniowej skali (1 - brak zakłóceń, 2 - nie mam pewności czy zakłócenia występują, 3 - bardzo słabe, 4 - słabe, 5 - wyraźne). Wyniki podane przez testerów dla porównania dwóch kopii oryginalnego utworu zostały umieszczone na wykresie jako zmodyfikowane z siłą 0%. Wyniki przeprowadzonego testu zaprezentowano na rysunku 3.7.



Rysunek 3.7. Słyszalność zakłóceń w funkcji siły dołączania [opracowanie własne]

Jak widać na wykresie przedstawionym na rysunku 3.7 testerom nie udało się jednoznacznie zidentyfikować utworów zawierających dołączone dane. Wyniki uzyskane dla nagrań zmodyfikowanych są zbliżone do wyników uzyskanych podczas porównywania dwóch identycznych sygnałów. W niektórych przypadkach oryginał wskazywany był przez słuchaczy jako bardziej "podejrzany" niż utwory zmodyfikowane. Ponadto skrajne wartości nieznacznie tylko przekraczają wartość dwa, która w przeprowadzonym teście oznacza brak pewności co do istnienia lub braku zakłóceń w sygnale. Można więc na tej podstawie wyciągnąć wniosek, że opracowana metoda w badanym przedziale nie wprowadza zniekształceń słyszalnych dla przeciętnego odbiorcy, co pozwala na stosowanie wartości $R_p \leq 40\%$.

3.2.3. Opracowanie kryteriów wyboru prążków widma

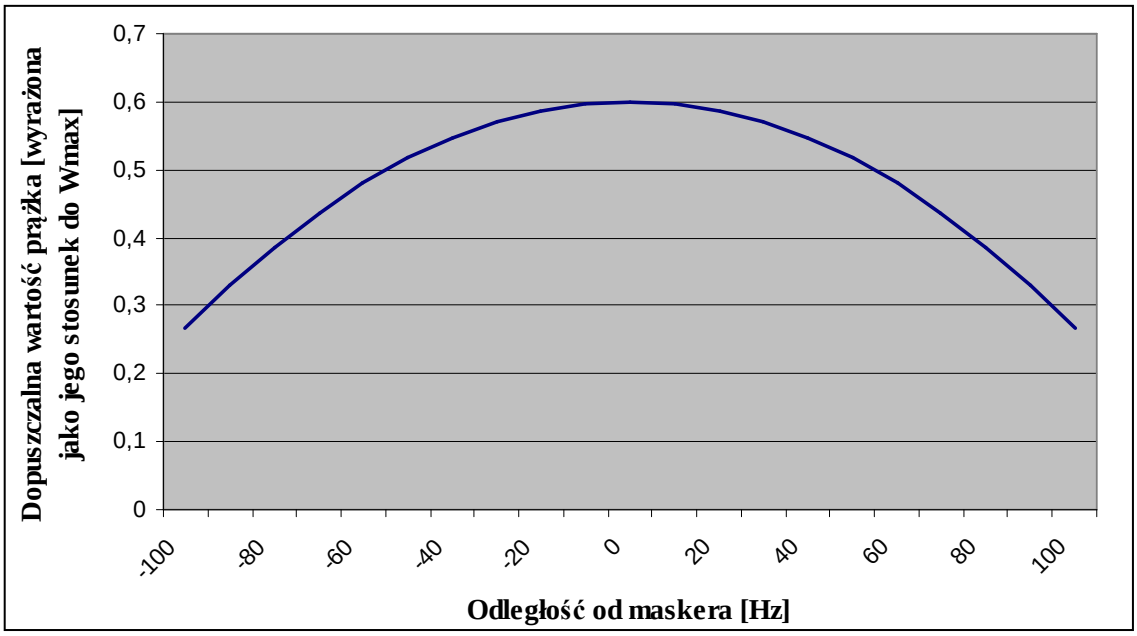
Aby spełniać założenia metody i umożliwić dołączenie informacji bez wprowadzania słyszalnych zakłóceń, modyfikowane prążki widma muszą być położone w sąsiedztwie prążka o większej energii i znajdować się w jego obszarze maskowania. W wyniku prowadzonych badań wyciągnięto wniosek, że funkcję maskera może skutecznie pełnić największy prążek w widmie częstotliwościowym. Dlatego też prążki przeznaczone do ukrycia dodatkowej informacji wybierane są spośród prążków położonych w jego sąsiedztwie. Nie jest to jednak warunkiem wystarczającym, gdyż jak to zostało przedstawione w rozdziale 2.4.3 nie będą słyszalne tylko te częstotliwości, które znajdują się poniżej krzywej maskowania. Niestety nie istnieje żadna metoda pozwalająca na określenie kształtu tej krzywej. Znany jest tylko jej przybliżony przebieg określony w wyniku testów odsłuchowych. Ponadto w rzeczywistym dźwięku nigdy nie występuje tylko jedna częstotliwość. Powoduje to występowanie bardzo złożonych zjawisk maskowania – każda z występujących częstotliwości wpływa na odbiór pozostałych[40]. Dlatego zdecydowano się opracować krzywą zbliżoną kształtem do progu maskowania, która będzie służyła do wyboru prążków przeznaczonych do ukrycia informacji. Kształt zastosowanej krzywej opisuje równanie:

$$W_g = (a - (f_{\max} - f)^2 / b) \cdot W_{\max} \qquad (3.12),$$

gdzie: a,b – współczynniki równania pobierane z klucza steganograficznego, f – częstotliwość.

Przykładowy przebieg krzywej opisanej równaniem 3.12 zilustrowano na rysunku 3.8.

Krzywa przedstawiona na rysunku 3.8 powstała po podstawieniu do równania 3.12 wartości a=0.6 oraz b=30000. Poprawność dobrania współczynników a i b potwierdzają wyniki przedstawione na rysunku 3.7, uzyskane z nagrań modyfikowanych z użyciem tej właśnie funkcji.



Rysunek 3.8. Krzywa aproksymująca przebieg progu maskowania [opr. własne]

Oprócz zapobiegania wprowadzaniu zmian w prążkach widma o zbyt dużej wartości krzywa opisana równaniem 3.12 wpływa na wybór wykorzystywanych prążków. Jest to dodatkowym atutem, gdyż bez znajomości jej przebiegu osoba próbująca odczytać dołączoną wiadomość nie jest w stanie określić, które prążki widma zostały wykorzystane do ukrycia dodatkowych danych.

Proces wyboru prążków widma przebiega według schematu:

1. znalezienie częstotliwości prążka widma o największej wartości (f_{max}),
2. utworzenie tablicy prążków widma sąsiadujących, znajdujących się w pobliżu f_{max}, znajdujących się w przedziale odległości (F_{dif1},F_{dif2}) od f_{max}; różnice częstotliwości (F_{dif1},F_{dif2}) zapisane są w kluczu steganograficznym; kolejność dodawania prążków widma do tablicy również jest zależna od klucza,
3. wybór dwóch pierwszych prążków widma z tablicy, znajdujących się pod krzywą opisaną równaniem 3.13 – czyli spełniających nierówność:

$$W_n \le (a - (f_{\max} - f_n)^2 / b) \cdot W_{\max} \qquad (3.13),$$

gdzie f_n jest częstotliwością wybranego prążka widma,

4. sprawdzenie ilości znalezionych prążków widma spełniających kryteria z punktu 3, jeśli znaleziono mniej niż dwa prążki wybieramy dodatkowe z tablicy.

Największy wpływ na wybór prążków widma mają parametry określone w punktach 2 i 3. Aby uniknąć wybierania prążków z jednego przedziału częstotliwości wybór uzależniono od częstotliwości mającej największy udział w widmie sygnału. Klucz steganograficzny definiuje przedziały odległości od niej, z których pobierane są prążki do modyfikacji. Ponadto prążki widma umieszczane są w tablicy w kolejności określonej przez klucz. Z tak przygotowanego zestawu wybierane są prążki spełniające nierówność 3.13 i znajdujące się jednocześnie w przedziale odległości (F_{dif1},F_{dif2}). Takie prążki w dalszej części niniejszej pracy będą nazywane "prążkami spełniającymi warunki". W przypadku, gdy nie uda się wybrać wystarczającej ilości prążków spełniających warunki, zestaw uzupełniany jest o dodatkowe prążki nie spełniające warunków. Uzupełnianie ma na celu umożliwienie wykorzystania wszystkich fragmentów sygnału, co istotnie wpływa na zwiększenie odporności dołączonych danych na przekształcenia stegokontenera. Wykorzystanie algorytmu omijającego niektóre fragmenty powodowało powstawanie przesunięć bitów odczytanych ze stegokontenera, który był poddawany przekształceniom dźwięku. Powodem tego była zmiana wartości poszczególnych prążków widma wprowadzana podczas modyfikacji nośnika. Po przeprowadzeniu tego procesu w niektórych fragmentach sygnału nie posiadających prążków widma spełniających warunki pojawiały się prążki, które były rozpoznawane przez algorytm jako przenoszące informację. Powodowało to odczytanie bitu informacji i wstawienie go do ciągu danych pomimo tego, że fragment nie był wykorzystywany w procesie ukrywania informacji. Dodatkowy bit przesuwał pozostałą część danych. Równie kłopotliwe były przypadki odwrotne. Fragment zawierający ukryty bit po przekształceniu rozpoznawany był jako niewykorzystany, co również powodowało przesunięcie pozostałych bitów informacji. Brak możliwości określenia pozycji dodatkowo wstawionych lub też pominiętych bitów niszczył dołączoną informację, gdyż prawidłowo odczytane bity umieszczone na niewłaściwych pozycjach nie pozwalały na odczyt dołączonej informacji. Aby umożliwić poprawny odczyt niezbędne byłoby wprowadzenie dodatkowych znaczników pozwalających na określenie poprawnej pozycji bitów. Ze względu na nieopłacalność zastosowania takiego rozwiązania zdecydowano się na wykorzystanie wszystkich fragmentów. Niewielka liczba fragmentów nie posiadających prążków widma spełniających warunki

nie wpływa w zauważalnym stopniu na parametry stegokontenera. Prowadzone badania wykazały, że w żadnym z badanych sygnałów nie wystąpiło więcej niż 1% takich fragmentów.

W przypadku, gdy algorytm natrafi na fragment nie posiadający wystarczającej ilości prążków widma spełniających przedstawione warunki, uzupełnia ich liczbę dodatkowymi prążkami nie spełniającymi warunków. Proces ten rozpoczyna się określeniem ilości brakujących prążków. Jeśli nie znaleziono żadnego prążka widma spełniającego warunki, wybierane są dwa pierwsze z tablicy prążków sąsiadujących (znajdujących się w przedziale odległości (F_{dif1},F_{dif2}) od f_{max}). Ze względu na zależną od klucza kolejność umieszczania prążków widma w tablicy oraz niewielką liczbę fragmentów, do których wybierane są dodatkowe prążki, zrezygnowano z umieszczania w kluczu dodatkowej informacji o prążkach do wybrania.

W przypadku, gdy znaleziony jest jeden prążek, niezbędne jest dołączenie drugiego. Wybierany jest wówczas prążek położony najbliżej znalezionego w kierunku f_{max}. Jeśli nie ma takiego prążka, to wybierany jest inny położony najbliżej prążek, z uwzględnieniem, że musi się znajdować w przedziale odległości (F_{dif1},F_{dif2}) od f_{max}.

3.2.4. Dobór sposobu modyfikacji wartości prążków widma

Po dokonaniu wyboru prążków można przystąpić do dołączenia bitu informacji. Dokonuje się tego poprzez modyfikację wartości wybranych prążków, tak by osiągnąć żądaną różnicę wartości pomiędzy prążkami. W przypadku ukrywania binarnej jedynki dążymy do uzyskania różnicy wartości pomiędzy prążkami nie mniejszej niż R-β. Ukrywanie binarnego zera polega na zmniejszeniu różnicy pomiędzy prążkami do wartości mniejszej lub równej β. Wartość β została wprowadzona ze względu na konieczność usunięcia słabego punktu metody, jakim byłaby stała różnica wartości pomiędzy wykorzystywanymi prążkami widma, łatwa do wychwycenia za pomocą analizy numerycznej. Dlatego w kluczu została umieszczona wartość β oznaczająca zakres przedziału <- β, β>, z którego losowana jest liczba dodawana do wyliczonej wartości jednego z prążków widma. Dzięki temu różnice wartości pomiędzy prążkami widma nie są takie same, co uniemożliwia łatwe zidentyfikowanie wykorzystanych prążków widma przy znajomości zastosowanej metody steganograficznej.

Zgodnie z założeniami opracowywany algorytm ma zapewniać wysoką odporność na przekształcenia dźwięku. Należy więc uwzględnić, że wykonując je wprowadzamy zmiany wartości prążków. Niezbędne jest więc zabezpieczenie danych

przed zniszczeniem przez to zjawisko. Dokonano tego przez określenie przedziału nieużywanych wartości prążków. Ze względu na wprowadzenie ograniczenia wartości prążków funkcją opisaną równaniem 3.12, przedział wartości niedozwolonych nie może być taki sam dla wszystkich prążków. Ponadto niezbędne jest uwzględnienie wartości prążków w przetwarzanym fragmencie – wielkość przerwy nie może być większa od wartości największego prążka, gdyż uniemożliwiłoby to ukrywanie danych. Nie może być również zbyt mała, gdyż nie spełni swojej roli. Niezbędne jest więc dopasowanie zakresu nieużywanych wartości do energii prążków w danym fragmencie. Rozwiązanie uzyskano poprzez użycie dwóch krzywych. Pierwszą z nich jest funkcja opisana równaniem 3.13, która wyznacza górną granicę nieużywanego przedziału. Drugą będzie krzywa o takim samym kształcie lecz obniżona o wartość W_p wyliczaną według wzoru:

$$W_p = \gamma \cdot W_{\max} \tag{3.14},$$

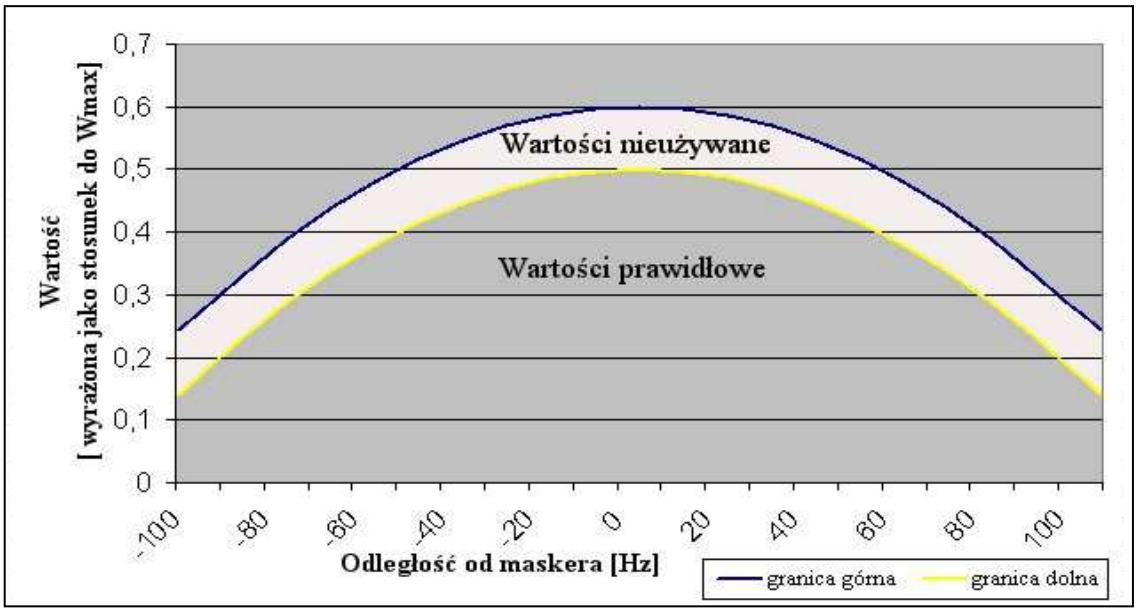
gdzie γ jest wartością zdefiniowaną w kluczu steganograficznym. Krzywa wyznaczająca dolną granicę będzie więc opisana wzorem:

$$W_d = (a - (f_{\max} - f)^2 / b - \gamma) \cdot W_{\max} \tag{3.15}.$$

Wartość funkcji W_g dla częstotliwości odpowiadającej przetwarzanemu prążkowi widma będzie nazywana odpowiadającą mu wartością górną, zaś W_d wartością dopuszczalną.

Uzależnienie wielkości przedziału nieużywanego od wartości największego prążka pozwala na dostosowanie go do energii przetwarzanego fragmentu sygnału. Kształt przedziału wynikający z ograniczających go funkcji zaprezentowany został na rysunku 3.9. Oznaczono go kolorem jasnoszarym. Podczas procesu ukrywania analizowane są wartości prążków, aby odpowiednio wyliczyć wartości docelowe, tak by żaden z prążków przenoszących informację nie znalazł się w przedziale nieużywanym.

Z powodu trudności w analitycznym wyznaczeniu optymalnej wielkości przedziału wartości nieużywanych, określono na drodze eksperymentu jej wartość na poziomie $0.15R_p$, pozwalającą na zachowanie dobrych parametrów metody. Dokładne ustalenie tej wartości pozostaje kwestią otwartą wymagającą dalszych badań.



Rysunek 3.9. Przedział wartości nieużywanych [opracowanie własne]

Przebieg procesu ukrywania zależny jest od wartości początkowych prążków widma oraz od ukrywanego bitu. Możemy wyróżnić następujące przypadki:

1. udało się znaleźć więcej niż jeden prążek widma o wartości mniejszej od W_g,
2. znaleziony został tylko jeden prążek widma o wartości mniejszej od W_g,
3. nie znaleziono żadnego prążka widma spełniającego ten warunek.

Proces ukrywania wartości binarnej "jeden" przebiega następująco:

1. W pierwszym przypadku pobierane są dwa pierwsze ze znalezionych prążków i następuje sprawdzenie czy da się ukryć jedynkę przy wykorzystaniu tej pary. Możliwe to będzie wtedy, gdy wartość dopuszczalna odpowiadająca większemu prążkowi będzie większa lub równa *R*. Jeśli okaże się, że niemożliwe jest wykorzystanie tych prążków, następuje poszukiwanie pary prążków pozwalających na uzyskanie różnicy *R*. Jeśli uda się taką parę znaleźć, wówczas wartości wszystkich prążków występujących w tablicy znalezionych prążków przed tymi prążkami są zwiększane do odpowiadających im wartości górnych. Jeśli natomiast nie uda się znaleźć odpowiedniej pary prążków, wówczas pozostawiany jest tylko jeden, natomiast wartości reszty prążków zwiększane są do odpowiadających im wartości górnych. Następnie przeprowadzany jest proces dołączania według pkt. 2.
2. Jeśli określono, że za pomocą wybranej pary prążków da się zakodować binarną jedynkę rozpoczyna się proces dołączania. Wyliczenie wartości docelowych prążków przebiega według algorytmu przedstawionego na rysunku 3.10.

3. W przypadku, gdy został znaleziony tylko jeden prążek spełniający warunki, jako drugi przyjmowany jest najbliższy mu prążek i jest on traktowany jako większy nie podlegający modyfikacji ze względu na wartość wykraczająca poza przyjęty próg maskowania. Schemat NS algorytmu wyliczenia wartości mniejszego prążka przedstawiony został na rysunku 3.11.
4. W sytuacji gdy nie udało się znaleźć żadnego prążka znajdującego się pod krzywą W_g, bit informacji dołączany jest na wybranych dwóch prążkach z sąsiedztwa największego prążka. Odbywa się to poprzez zmniejszenie wartości mniejszego o wartość z zakresu $< R\text{-}\beta, R+\beta >$. Proces obliczania wartości docelowych prążków przedstawia schemat NS zaprezentowany na rysunku 3.12.

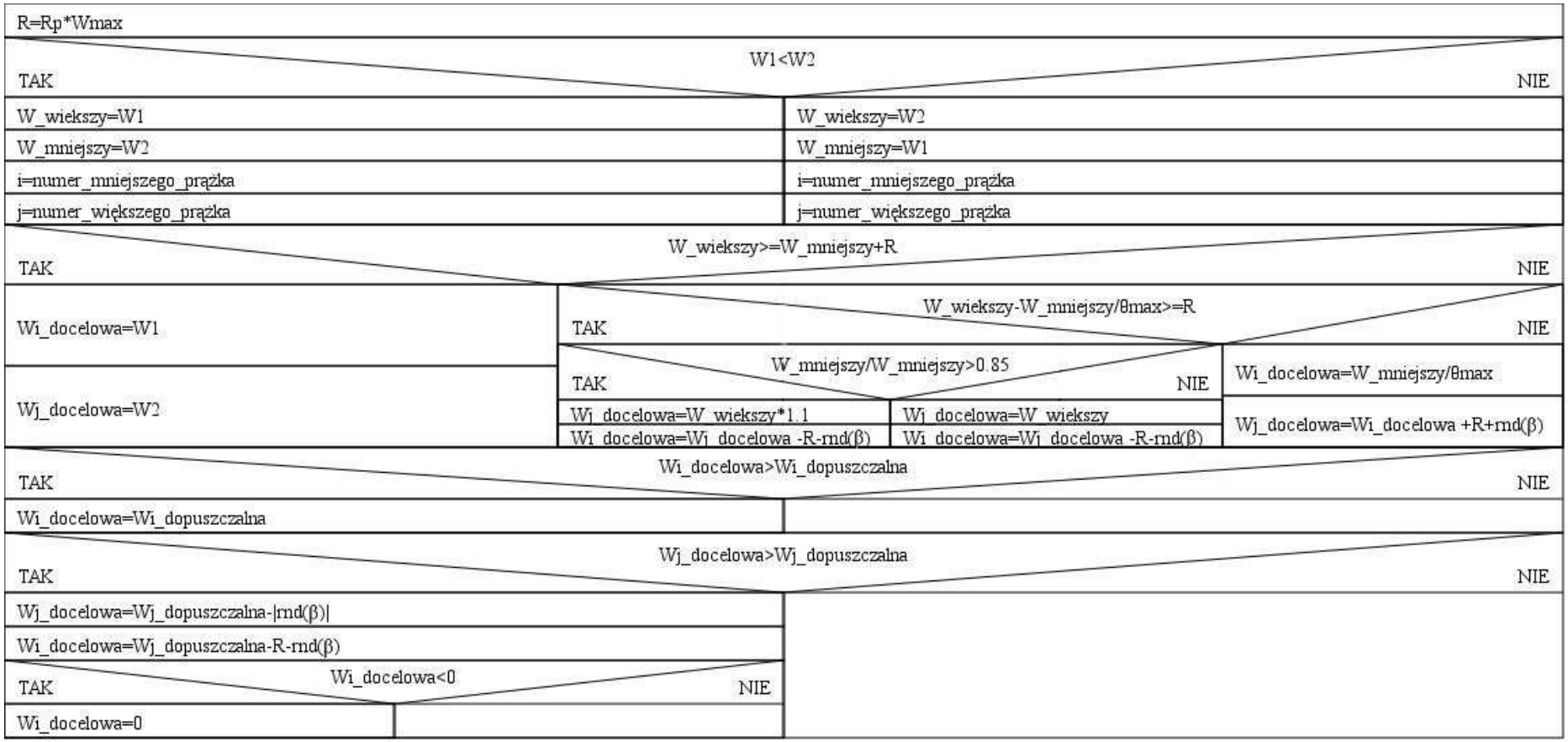
Nieco inaczej przebiega ukrywanie bitu informacji w przypadku, gdy celem jest ukrycie binarnego zera. W tym przypadku dążymy do uzyskania zbliżonych wartości wybranych prążków. Tutaj również staramy się unikać wprowadzania identycznych wartości ze względu na możliwość łatwego wykrycia za pomocą analizy komputerowej. Tak więc wartości prążków przenoszących wartość binarną "zero" dobieramy tak, by spełniały nierówność:

$$|W_1 - W_2| \leq \beta \qquad (3.16).$$

Jako że dążymy do usunięcia różnicy wartości pomiędzy przetwarzanymi prążkami, nie będzie konieczności sprawdzania czy ukrycie za pomocą pary wybranych prążków jest możliwe. Zawsze uda się znaleźć wartość docelową, mniejszą od wartości dopuszczalnych odpowiadających wybranym prążkom. Aby zminimalizować zmiany wprowadzane w sygnale, wartości docelowe prążków zbliżone są do średniej z ich oryginalnych wartości.

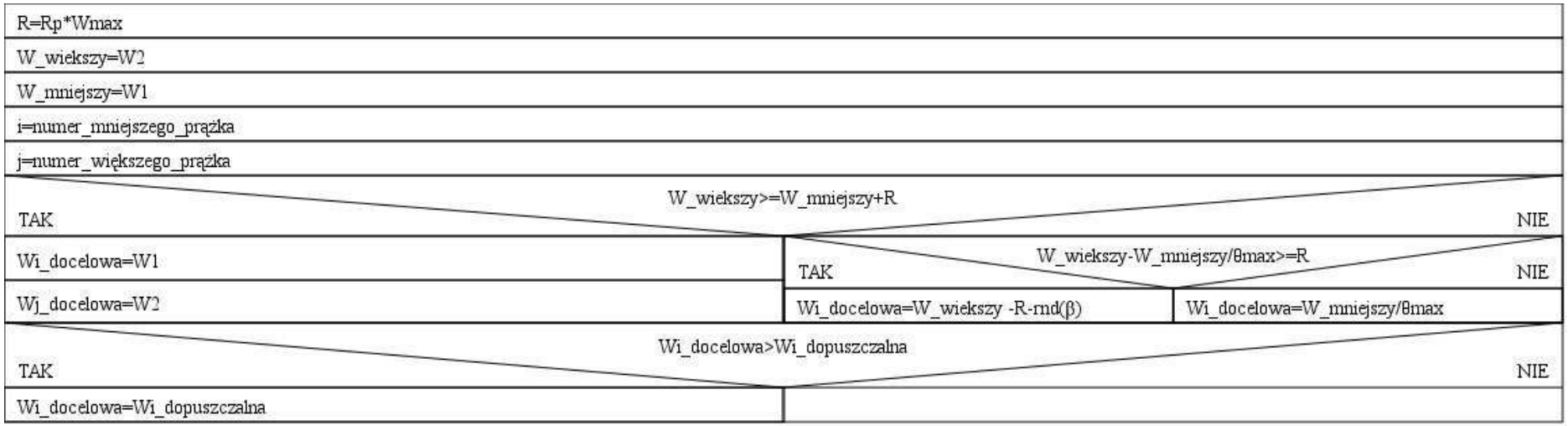
Jeśli udało się znaleźć więcej niż jeden prążek spełniający stawiane warunki, wybieramy dwa pierwsze prążki z tablicy znalezionych prążków a następnie modyfikujemy ich wartości według algorytmu przedstawionego na rysunku 3.13.

W przypadku, gdy istnieje tylko jeden prążek spełniający warunki lub nie ma żadnego, wówczas wybierane są dodatkowe prążki spośród pozostałych, tak by uzyskać parę, która zostanie wykorzystana do ukrycia informacji. Następnie sprawdzana jest wartość dopuszczalna dla każdego z wybranych prążków. Wartość docelowa prążków ustalana jest na poziomie mniejszej z dopuszczalnych wartości. Algorytm realizujący ten proces ilustruje schemat NS zaprezentowany na rysunku 3.14.

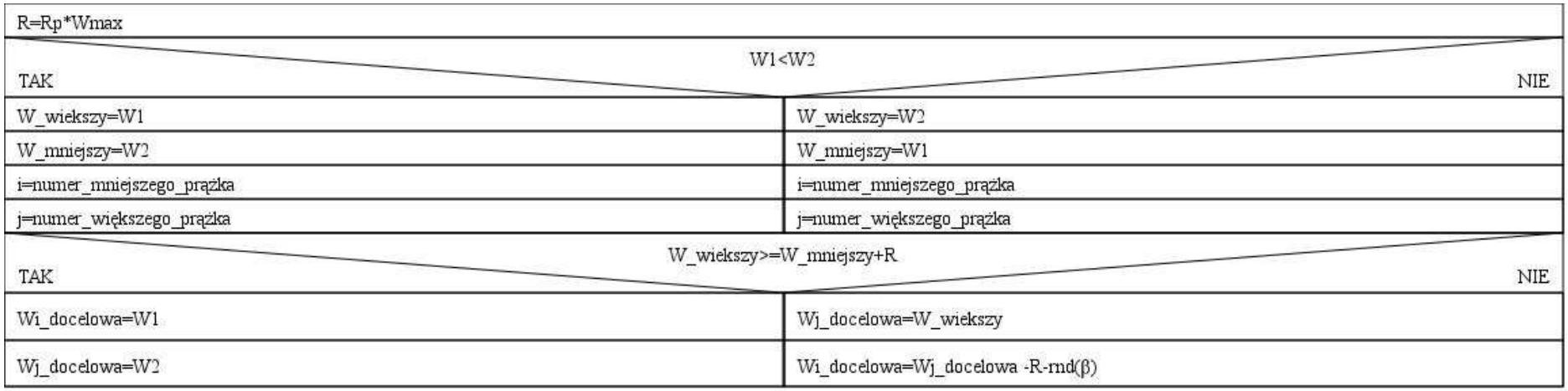


Rysunek 3.10. Dołączanie binarnej "jedynki" na dwóch prążkach widma spełniających warunki [opracowanie własne]

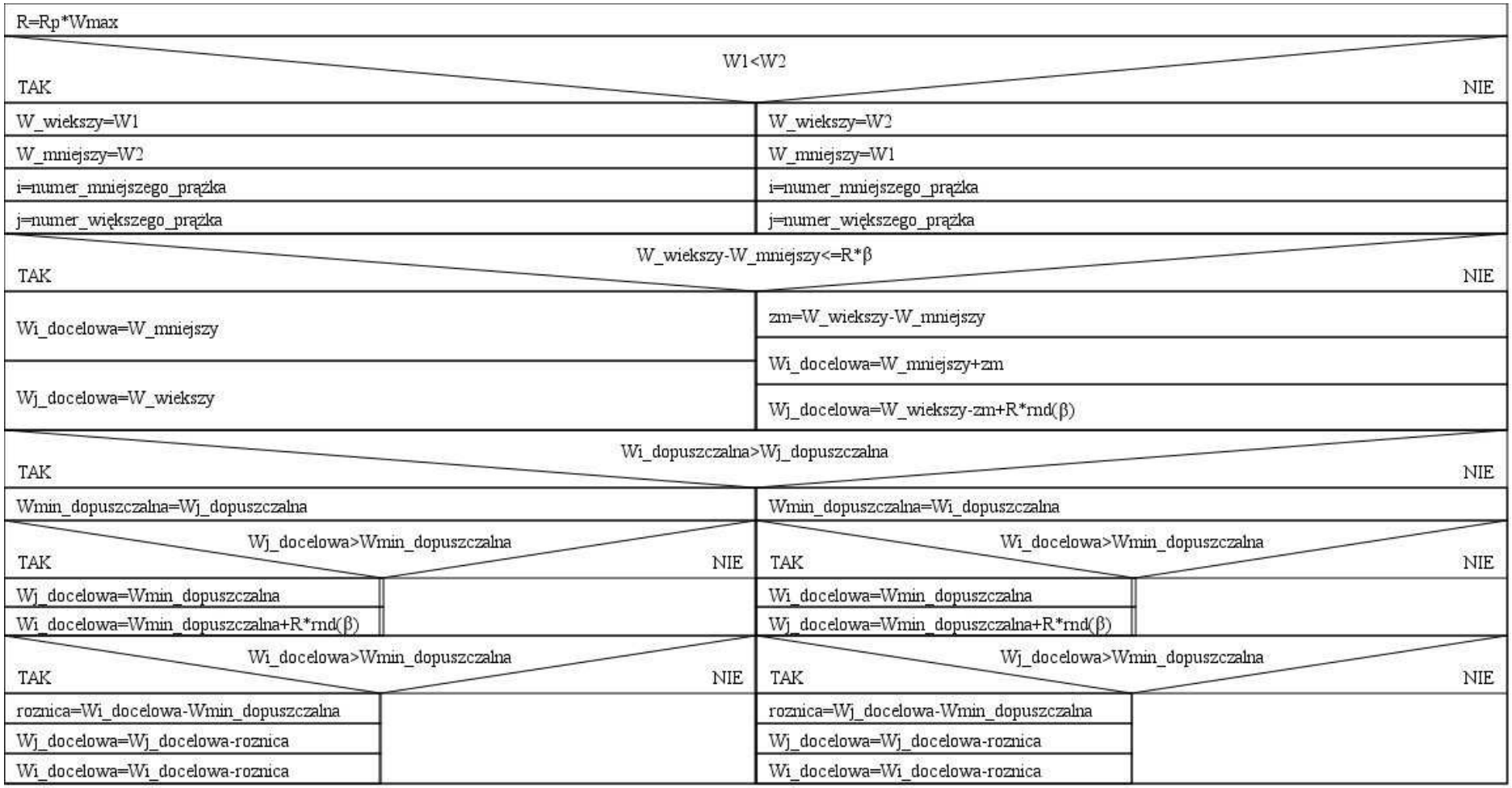
Na rysunku 3.10 zastosowano oznaczenia: *rnd(β)* – funkcja zwracająca wartość losową z przedziału <- *β*, *β* >, $W_i docelowa$, $W_j docelowa$ – wartości docelowe prążków *i* oraz *j*.



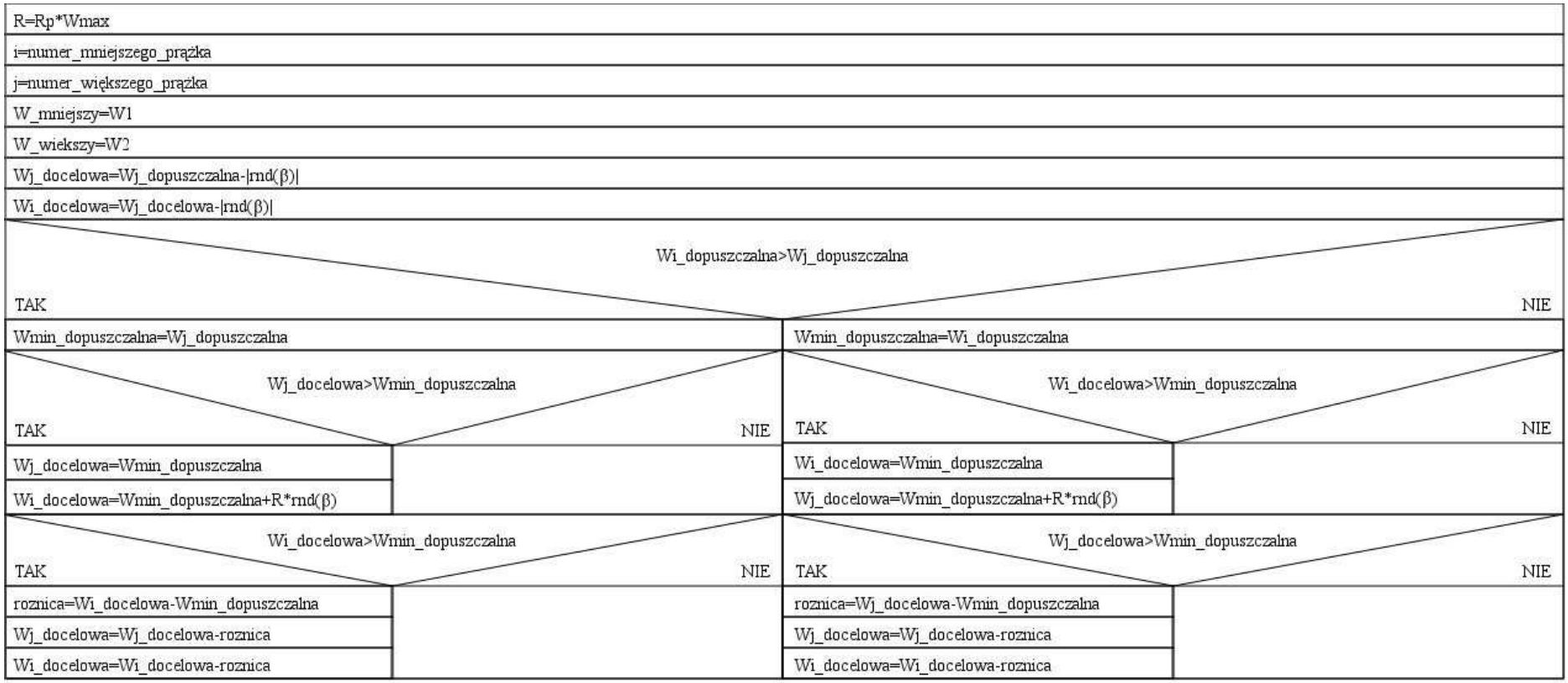
Rysunek 3.11. Dołączanie jedynki, gdy istnieje tylko jeden prążek spełniający warunki [opracowanie własne]



Rysunek 3.12. Dołączanie jedynki, przy braku prążków widma spełniających warunki [opracowanie własne]



Rysunek 3.13. Dołączanie bitu o wartości zero na dwóch prążkach widma spełniających warunki [opracowanie własne]



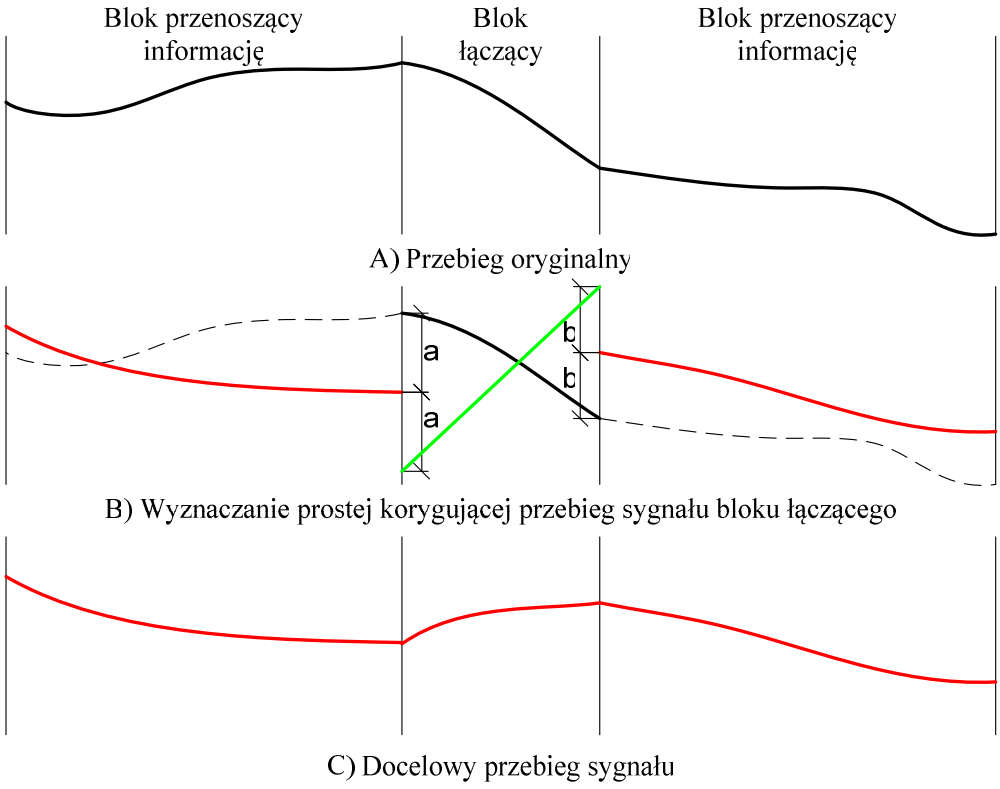
Rysunek 3.14. Dołączanie bitu o wartości zero, gdy znaleziono mniej niż dwa prążki widma spełniające warunki [opracowanie własne]

3.2.5. Określenie wpływu bloków łączących na redukcję poziomu zniekształceń powstających na łączeniach bloków

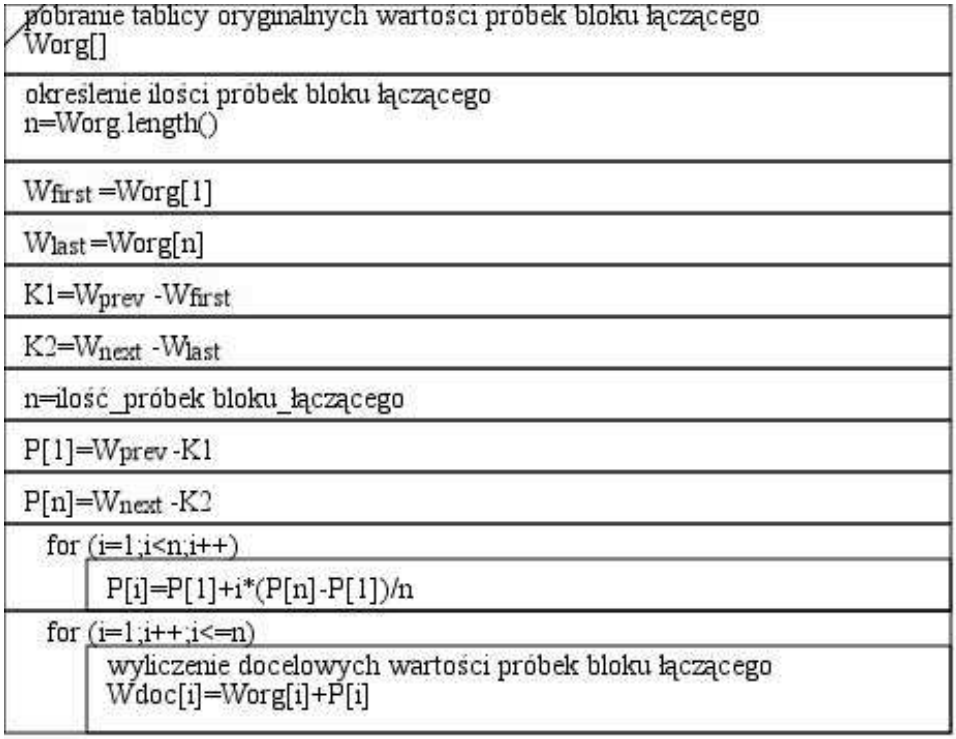
Ukrycie informacji w bloku powoduje zmianę przebiegu sygnału w czasie. Podczas łączenia bloków może to powodować powstawanie nieciągłości na ich granicy. W przypadku niektórych sygnałów powoduje to powstawanie słyszalnych zakłóceń. Aby ich uniknąć niezbędne jest dopasowanie amplitudy sygnału na krańcach łączonych bloków. Wprowadzanie modyfikacji wewnątrz bloku spowoduje jednak zniekształcenie sygnału, co niekorzystnie wpłynie na odporność dołączonych danych. Lepszym rozwiązaniem jest wprowadzenie bloków łączących. Pomiędzy blokami przenoszącymi informację umieszczone zostaną bloki, których zadaniem będzie zapewnienie odpowiednio ukształtowanego przejścia zmniejszającego zakłócenia, tak by stały się niesłyszalne.

Pierwszym etapem jest podział na bloki. Z całego sygnału wyodrębniane są na przemian bloki przeznaczone do ukrycia informacji oraz bloki łączące. Po dołączeniu danych do bloków przenoszących informację określane są wartości próbek sąsiadujących z blokiem łączącym. Oznaczmy je jako: W_{prev} – wartość ostatniej próbki bloku poprzedzającego łączący oraz W_{next} – wartość pierwszej próbki kolejnego bloku. Następnie określane są wartości skrajnych próbek bloku łączącego. Oznaczmy je jako: W_{first} – wartość pierwszej próbki oraz W_{last} – wartość ostatniej próbki bloku łączącego. Na podstawie powyższych danych określane są zmiany, które zostaną wprowadzone w bloku łączącym. Odbywa się to poprzez obliczenie różnicy pomiędzy skrajną próbką bloku przenoszącego informację a sąsiadującą z nią próbką bloku łączącego. Różnica ta obliczana jest na obu krańcach bloku łączącego. Otrzymane w ten sposób liczby po odjęciu od wartości skrajnych próbek sąsiadujących przedziałów (W_{prev} i W_{next}) wyznaczą punkty, które łączymy prostą. Będzie ona funkcją, której wartości zostaną dodane do próbek bloku łączącego. W ten sposób oryginalny przebieg sygnału w bloku łączącym zostanie zmodyfikowany, tak by próbki na jego krańcach miały takie same wartości jak sąsiadujące z nimi próbki bloków łączących informację. Graficznie przedstawiono ten proces na rysunku 3.15. Kolorem czarnym oznaczono oryginalny przebieg sygnału. Kolor czerwony reprezentuje docelową postać sygnału. Zieloną linią przedstawiono przebieg prostej, która po zsumowaniu z oryginalnym przebiegiem

sygnału da postać docelową. Schemat NS algorytmu modyfikacji bloku łączącego został zaprezentowany na rysunku 3.16.



Rysunek 3.15. Wyznaczenie kształtu sygnału w bloku łączącym [opracowanie własne]



Rysunek 3.16. Schemat NS algorytmu obliczania wartości docelowych próbek bloku łączącego [opracowanie własne]

W celu zweryfikowania zasadności stosowania bloków łączących przeprowadzono testy odsłuchowe. Badano zależność słyszalności zakłóceń w zależności od długości stosowanego bloku. W tabeli 3.3 przedstawiono otrzymane wyniki. Na ich podstawie można stwierdzić, że zastosowanie bloków łączących o długości większej niż 40 próbek pozwala zmniejszyć poziom zakłóceń spowodowanych nieciągłościami na łączeniach bloków do poziomu niesłyszalnego dla przeciętnego słuchacza.

Tabela 3.3. Słyszalność zakłóceń spowodowanych nieciągłościami na łączeniach bloków w funkcji długości bloku łączącego [opracowanie własne]

Rodzaj dźwięku i siła dołączania	Długość bloku łączącego (w próbkach)	Siła zakłóceń
Pop, R_p=40%	4	Słabe
Pop, R_p=40%	10	Bardzo słabe
Pop, R_p=40%	powyżej 40	Niezauważalne
Ballada, R_p=40%	4	Słabe
Ballada, R_p=40%	10	Bardzo słabe
Ballada, R_p=40%	20	Bardzo słabe
Ballada, R_p=40%	powyżej 40	Niezauważalne
Mowa, R_p=40%	4	Bardzo słabe
Mowa, R_p=40%	powyżej 10	Niezauważalne
Pianino, R_p=40%	2	Bardzo słabe
Pianino, R_p=40%	powyżej 4	Niezauważalne
Rock, R_p=40%	4	Słabe
Rock, R_p=40%	20	Bardzo słabe
Rock, R_p=40%	powyżej 40	Niezauważalne

3.3. Budowa klucza steganograficznego

Wszystkie informacje niezbędne do dołączenia a następnie odczytania ukrytych danych zawarte są w kluczu steganograficznym. Bez jego znajomości odczytanie tajnego przekazu nie jest możliwe. W opracowanej metodzie klucz steganograficzny zawiera:

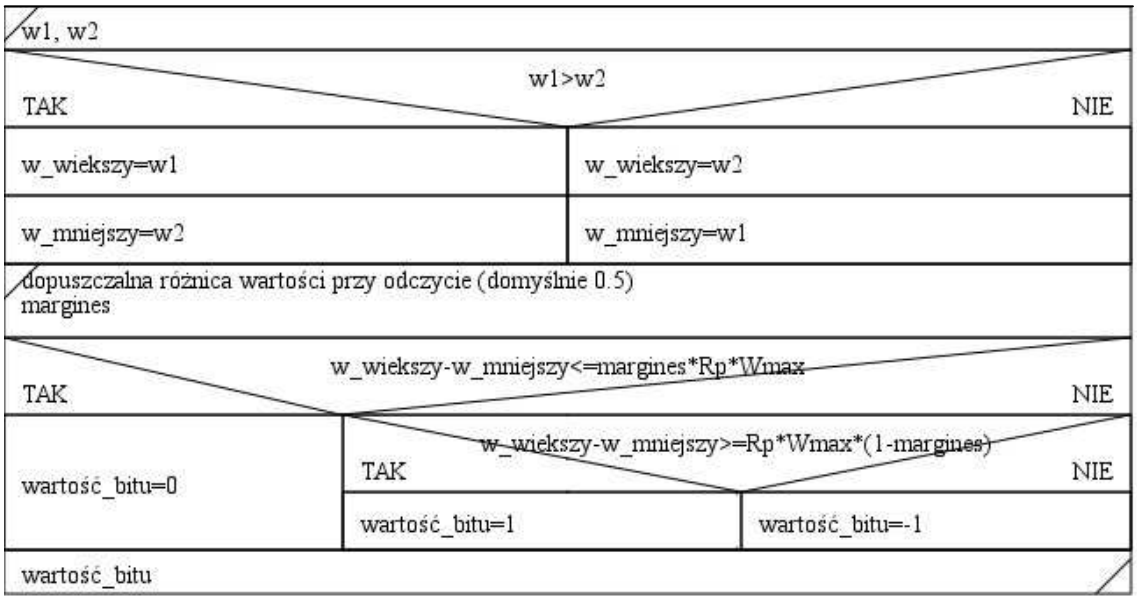
- rozmiar i położenie bloków przenoszących informację,
- rozmiar bloków łączących,
- siłę dołączania (R_p), określoną jako proporcja do wartości największego prążka,
- wartości współczynników a,b,γ służące do określenia granic przedziału nieużywanych wartości prążków widma (według wzorów 3.13 i 3.16),
- maksymalną odległość F_{dif2} prążków widma przenoszących informację od f_{max},
- minimalną odległość F_{dif1} prążków widma przenoszących informację od f_{max},
- kolejność umieszczania znalezionych prążków widma w tablicy, decydująca o kolejności wybierania ich jako prążków przenoszących informację,
- parametr β oznaczający zakres przedziału $<-\beta, \beta>$, z którego losowana jest liczba dodawana do wyliczonej wartości jednego z prążków, w celu uniknięcia wprowadzania stałej różnicy wartości pomiędzy prążkami,
- współczynnik θ_{max} określający maksymalną wartość dzielnika możliwą do zastosowania podczas zmniejszania wartości prążka,
- długość informacji,
- macierz permutacji wykorzystywaną do mieszania ukrywanych danych oraz ponownego przywrócenia ich właściwej kolejności w procesie ekstrakcji,
- klucze: szyfrujący oraz deszyfrujący, służące do wstępnego zaszyfrowania dołączanej informacji dodatkowym algorytmem kryptograficznym.

3.4. Metoda odczytu dołączonej informacji

Wyodrębnienie informacji ze stegokontenera wymaga przeprowadzenia w odwrotnej kolejności operacji użytych do ukrycia informacji, tak jak to zostało zaprezentowane na rysunku 3.2. Oczywiście nie uda się tego dokonać bez znajomości klucza steganograficznego. Pierwszym etapem będzie podział stegokontenera na bloki. Wystarczy wyodrębnić fragmenty przenoszące informację. Bloki łączące można ominąć jako, że nie przenoszą informacji. Proces podziału na bloki jest identyczny do przeprowadzanego podczas ukrywania. Każdy z wyodrębnionych bloków

poddawany jest transformacie Fouriera a następnie obliczany jest wektor modułów jego widma (Y_r), z którego odczytywany jest ukryty bit. Odbywa się to według schematu:

- znalezienie prążka p_{max} i określenie odpowiadającej mu częstotliwości $f_{max,}$
- wybór prążków widma przenoszących informację p_1 i p_2 (identycznie jak w procesie dołączania),
- odczyt wartości dołączonego bitu poprzez porównanie wartości wybranych prążków w_1 i w_2, schemat NS tego procesu został zaprezentowany na rysunku 3.17.



Rysunek 3.17. Określenie wartości odczytywanego bitu [opracowanie własne]

W wyniku odczytu możliwe jest otrzymanie trzech różnych wartości bitu: 0, 1, -1. Wartość -1 oznacza, że odczytana różnica wartości R' nie znajduje się w przedziale odpowiadającym binarnemu zeru lub jedynce. Wprowadzenie dodatkowej wartości było konieczne, ze względu na to, że nie zawsze odczytana różnica wartości prążków widma mieści się w założonych przedziałach. Ponadto odbiorca może zgodnie z własnymi preferencjami ustalić szerokość przedziałów odpowiadających wartościom 0 i 1.

Wartości odczytanych bitów umieszczane są w tablicy, która następnie poddawana jest permutacji w celu uzyskania właściwej kolejności bitów przekazywanej informacji. Jeśli informacja przed dołączeniem była dodatkowo poddana szyfrowaniu, niezbędne jest jeszcze deszyfrowanie otrzymanych danych.

3.5. Implementacja opracowanego algorytmu steganograficznego bazującego na transformacie Fouriera

Implementacja algorytmu przedstawionego w rozdziale 3 została wykonana w środowisku obliczeniowym *Matlab 7.9 (R2009B)*. Bazę do opracowania i przetestowania algorytmu stanowiły funkcje umożliwiające przetwarzanie sygnałów dźwiękowych. Na ich podstawie opracowane zostały autorskie funkcje i m-skrypty umożliwiające praktyczną implementację i testowanie opracowanej metody. Najważniejsze z nich przedstawia tabela 3.4.

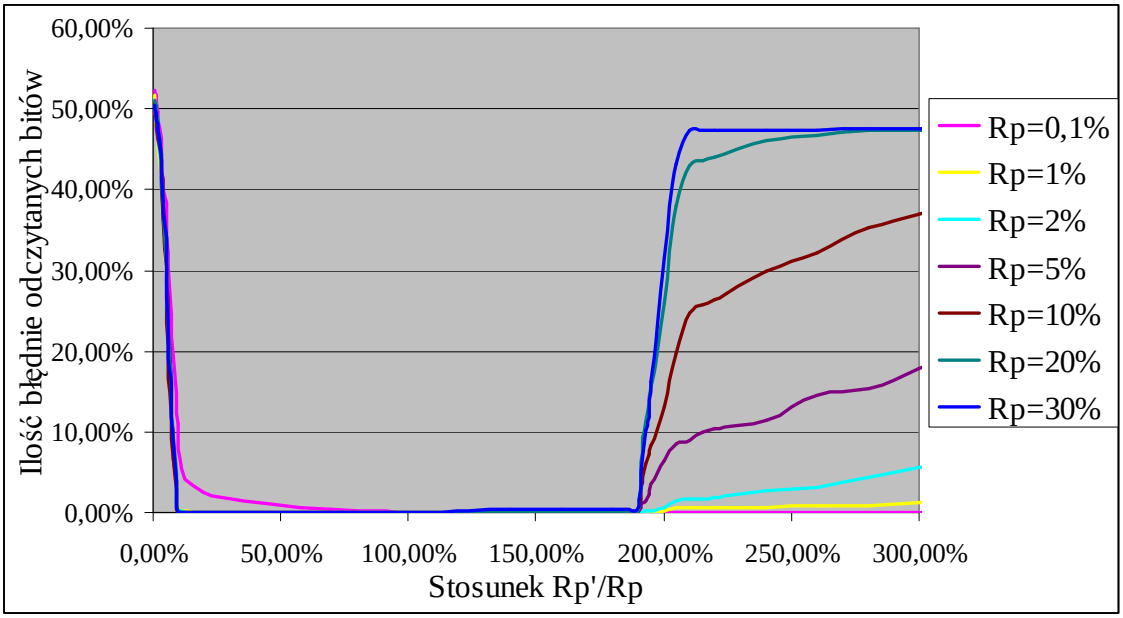
Tabela 3.4. Funkcje autorskie wykorzystywane w implementacji i weryfikacji opracowanego algorytmu [opracowanie własne]

Nazwa pliku	Opis działania i parametry
Fourier_ukryj	Funkcja ukrywająca bit informacji we fragmencie sygnału kontenera. Parametry funkcji: *s*- ciąg wartości próbek sygnału, w którym ma zostać ukryty bit informacji (kontener) *f_sampling* – częstotliwość próbkowania sygnału kontenera *bit* – wartość ukrywanego bitu *Rp* – moc dołączania (R_p) *beta* – wartość parametru β, określająca przedział, z którego losowana jest liczba dodawana do wyliczonej wartości prążka widma *Fdif1* – minimalna odległość wykorzystywanych prążków widma od największego (F_{dif1}) *f_dol* – najniższa częstotliwość na której mogą występować prążki widma przenoszące informację *obetnij* – wartość parametru *a* z równania 3.16, wpływająca na wysokość położenia funkcji ograniczającej wykorzystywane wartości prążków widma *Fdif2* – maksymalna odległość wykorzystywanych prążków widma od największego (F_{dif2}) *b*- wartość parametru *b* z równania 3.16, wpływająca na kształt funkcji ograniczającej wykorzystywane wartości prążków widma

	gamma – wartość γ z równania 3.16, wpływająca na wysokość położenia funkcji ograniczającej docelowe wartości prążków widma *ord* – tablica określająca kolejność wykorzystania prążków widma
Fourier_odczyt	Funkcja odczytująca bit ukrytej informacji z sygnału przekazanego jako parametr Parametry funkcji: *margines* – dopuszczalna rozbieżność pomiędzy oczekiwaną a odczytaną różnicą wartości prążków widma, która jest traktowana jako prawidłowo odczytana wartość *s, f_sampling, Rp, Fdif1, f_dol, obetnij, Fdif2, b, gamma, ord* opis parametrów taki sam jak w funkcji *Fourier_ukryj*
test_ukryj	Funkcja ukrywająca ciąg bitów informacji w sygnale przekazanym jako parametr. Aby ukryć bit informacji we fragmencie sygnału wywołuje funkcję *Fourier_ukryj*. Parametry funkcji: *wielk_fr* – wielkość fragmentu (bloku) przenoszącego informację *sklej* – wielkość bloku sklejania *nazwa* – nazwa pliku w jakim zostanie zapisany stegokontener zawierający ukrytą informację *info* – informacja do ukrycia, podana w postaci ciągu bitów *s, f_sampling, Rp, beta, Fdif1, f_dol, obetnij, Fdif2, b, gamma, ord* opis parametrów taki sam jak w poprzednich funkcjach
test_odczyt	Funkcja odczytująca dołączone dane ze stegokontenera Parametry funkcji: *wielk_fr, sklej, s, f_sampling, Rp, Fdif1, f_dol, obetnij, Fdif2, b, gamma, ord* - opis parametrów taki sam jak w poprzednich funkcjach *inf_ukr* – oryginalnie ukryta informacja (przekazywana w celu porównania z informacją odczytaną by określić ilość błędów odczytu)
zniekształcenia	Funkcja określająca poziom zniekształceń sygnału wprowadzany podczas modyfikacji sygnału Parametry funkcji: *s_org* – sygnał oryginalny *nazwa_s_mod* – ścieżka do pliku zawierającego sygnał przekształcony

3.6. Określenie wrażliwości na zmianę różnicy wartości pomiędzy prążkami widma sygnału

Niniejszy rozdział ma na celu określenie wpływu rozbieżności pomiędzy wartością parametru R_p użytego do ukrywania informacji a wartością R_p' użytego podczas odczytu. W tym celu przeprowadzono eksperyment, w którym modyfikowano wartość R_p', aż do wystąpienia błędów odczytu. Wartość stosunku R_p'/R_p została na wykresie przedstawiona na osi odciętych. Wyniki tego eksperymentu zaprezentowane zostały na rysunku 3.18.



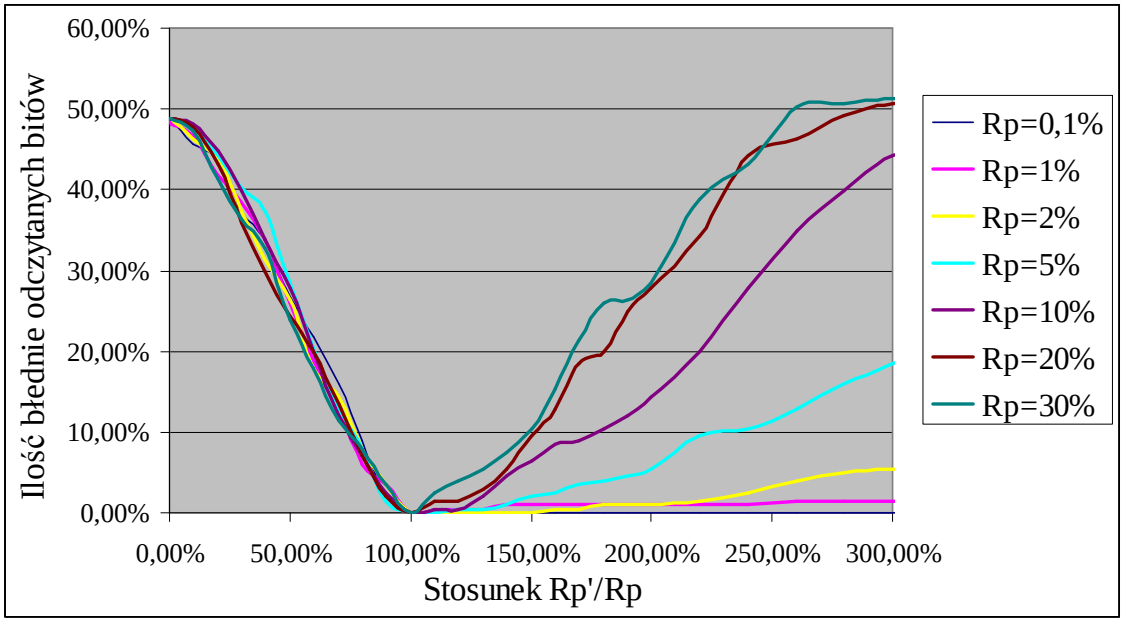
Rysunek 3.18. Ilość błędnie odczytanych bitów w zależności od stosunku R_p'/R_p dla parametru β=0.05 [opracowanie własne]

Postać wyników zaprezentowanych na rysunku 3.18 wynika z wartości parametrów użytych podczas ukrywania informacji. Szybkie narastanie ilości błędów dla R_p'/R_p<10% wynika z faktu dodawania do obliczonej wartości prążków widma wartości losowej z przedziału $<-5\%\cdot R_p\cdot W_{max}, 5\%\cdot R_p\cdot W_{max}>$ oraz odczytywania jako binarne zero wartości z przedziału $<0, 0.5\cdot R_p)$. Po dziesięciokrotnym zmniejszeniu wartości R_p', w stosunku do użytej podczas dołączania R_p, przedział dodawanych wartości losowych jest identyczny z przedziałem wartości odczytywanych jako binarne zero. Dalsze zmniejszanie tego przedziału powoduje, że coraz więcej bitów "zero" odczytywanych jest jako jedynki. Skutkuje to wzrostem ilości błędnie odczytanych bitów.

Górna granica stosunku R_p'/R_p=190% gwarantującego poprawny odczyt wynika z faktu, że w tym punkcie przedział wartości odczytywanych jako binarne "zero" przyjmuje wartości z przedziału <0, $0.95 \cdot R_p \cdot W_{max}$). Powyżej tej granicy kodowane są już wartości "jeden". Dalsze zwiększanie wartości R_p'/R_p powoduje nałożenie się przedziałów odpowiadających zakodowanej jedynce i odczytywanemu zeru, co generuje błędy. Daje się zauważyć, że tempo wzrostu ilości błędów odczytu jest zależne od użytej do dołączania wartości R_p. Dzieje się tak ze względu na brak ograniczenia maksymalnej wartości R wykorzystywanej podczas kodowania binarnej "jedynki". Jeśli w oryginale różnica wartości wykorzystywanych prążków widma była większa od wyliczonej to nie była modyfikowana, jako że spełniała założeń algorytmu.

Im większa jest zastosowana podczas dołączania wartość R_p tym szybciej podczas zwiększania stosunku R_p'/R_p są osiągane wyższe wartości R co powoduje szybsze osiąganie większych wartości skutkujące zwiększaniem stopy błędów.

W przypadku, gdy najbardziej istotne jest bezpieczeństwo steganograficzne, możliwe jest zmniejszenie przedziału R_p' w jakim prawidłowo odczytywana jest ukryta informacja lub zwiększenie β, kosztem zmniejszenia odporności na wprowadzane modyfikacje sygnału. Aby uzyskać najmniejszy zakres wartości R_p', w którym następuje prawidłowy odczyt dołączonych danych, należy podczas ukrywania przyjąć wartość β=0.5. Poziomy błędów odczytu uzyskanych dla $\beta=0.5$ zaprezentowane zostały na rysunku 3.19.



Rysunek 3.19. Ilość błędnie odczytanych bitów w zależności od stosunku R_p'/R_p dla parametru $\beta=0.5$ [opracowanie własne]

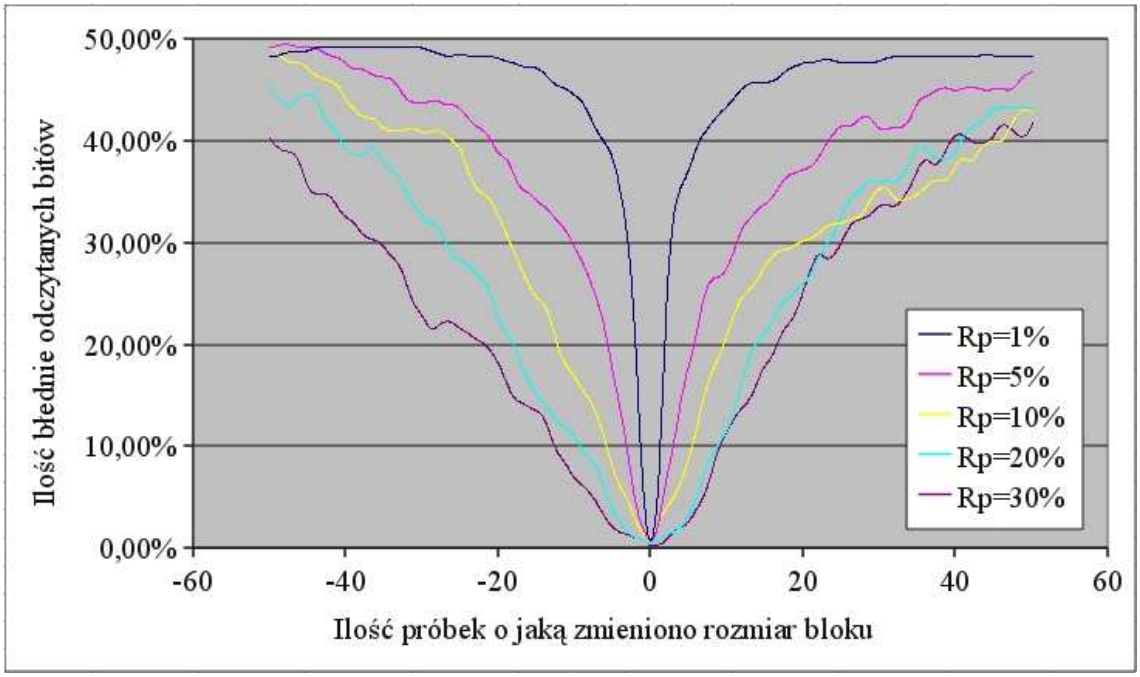
Na podstawie rysunku 3.19 można zauważyć, że prawidłowe określenie wartości R_p' jest niezbędne do bezbłędnego odczytania ukrytych danych. Parametr ten znacząco wpływa na bezpieczeństwo steganograficzne systemu.

3.7. Wyznaczenie wpływu zmiany długości bloku danych na poprawność odczytu

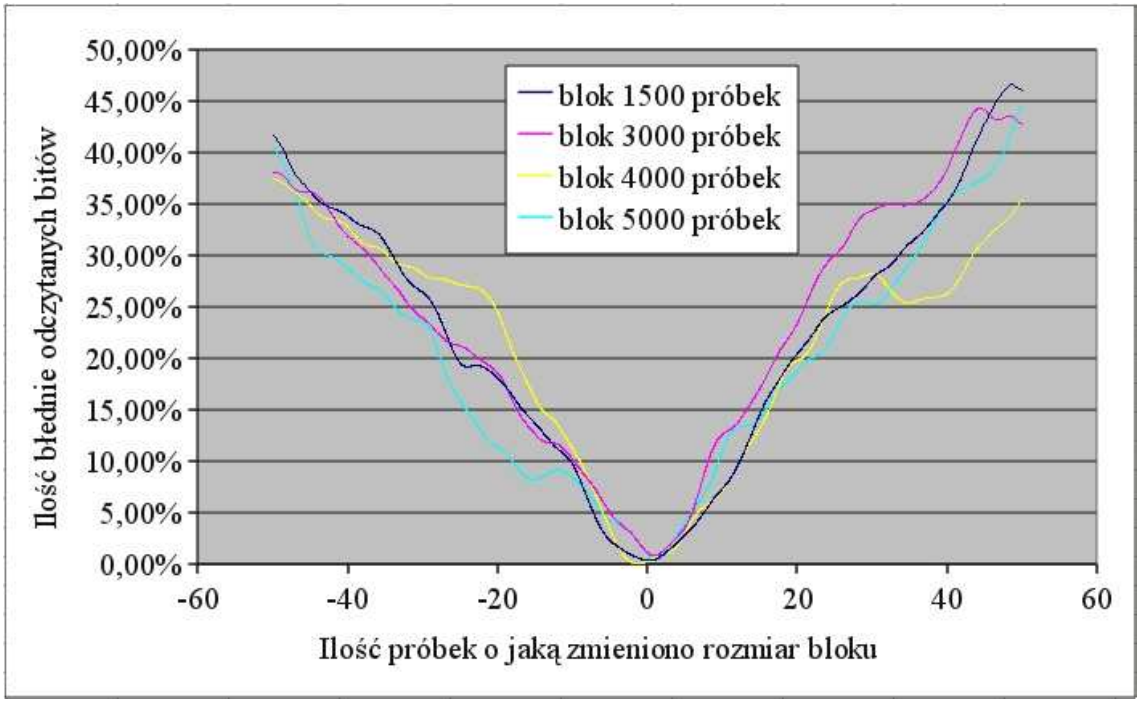
Bardzo istotnym parametrem wpływającym na bezpieczeństwo steganograficzne systemu jest znajomość rozmiaru bloku. Bez znajomości tego parametru osoba postronna nie powinna mieć możliwości poprawnego odczytania dołączonych danych. Wykonano więc badania mające na celu określenie wpływu zmiany rozmiaru bloku na liczbę błędnie odczytanych bitów. Eksperyment polegał na ukryciu informacji przy użyciu bloku o określonym rozmiarze, a następnie podejmowane były próby odczytu przy zastosowaniu bloku o innym rozmiarze. Przeprowadzono dwie serie badań, które miały określić czy uzyskane wyniki są zależne od siły dołączania R_p lub od rozmiaru bloku jaki był wykorzystany podczas ukrywania. Rysunek 3.20 prezentuje rezultaty badań, które dodatkowo uwzględniają wpływ siły dołączania R_p. Do ukrywania informacji wykorzystano tu bloki o rozmiarze 2000 próbek. Na rysunku 3.21 przedstawiono wyniki eksperymentu otrzymane przy zastosowaniu do ukrywania bloków o różnym rozmiarze. Dla wszystkich rozmiarów bloków zastosowano tą samą siłę dołączania R_p=20%. Na osi odciętych przedstawiono liczbę próbek o jaką zmodyfikowano długość bloku podczas odczytu. Ujemne wartości oznaczają blok o mniejszej długości, dodatnie blok wydłużony o zadaną liczbę próbek.

Podczas analizy rysunku 3.20 wyraźnie uwidacznia się wpływ siły dołączania na nieprecyzyjne dobranie rozmiaru bloku podczas odczytu. Najbardziej wrażliwe są stegokontenery zawierające dane dołączone z niewielką siłą. Dzieje się tak ponieważ zmieniając rozmiar bloku wpływamy również na rozdzielczość transformaty Fouriera. Powoduje to nieco inne odwzorowanie częstotliwości przez prążki a co za tym idzie zmianę ich wartości. Im większa zmiana rozdzielczości transformaty tym większe wystąpią różnice wartości pomiędzy prążkami. Jak widać na rysunku 3.20 dla mocy dołączania Rp=1% zmiana rozmiaru bloku o kilka próbek niszczy ukrytą informację.

Wpływ rozmiaru bloku użytego podczas ukrywania informacji nie ma istotnego wpływu na liczbę błędnie odczytanych bitów przy braku precyzji w doborze rozmiaru bloku.



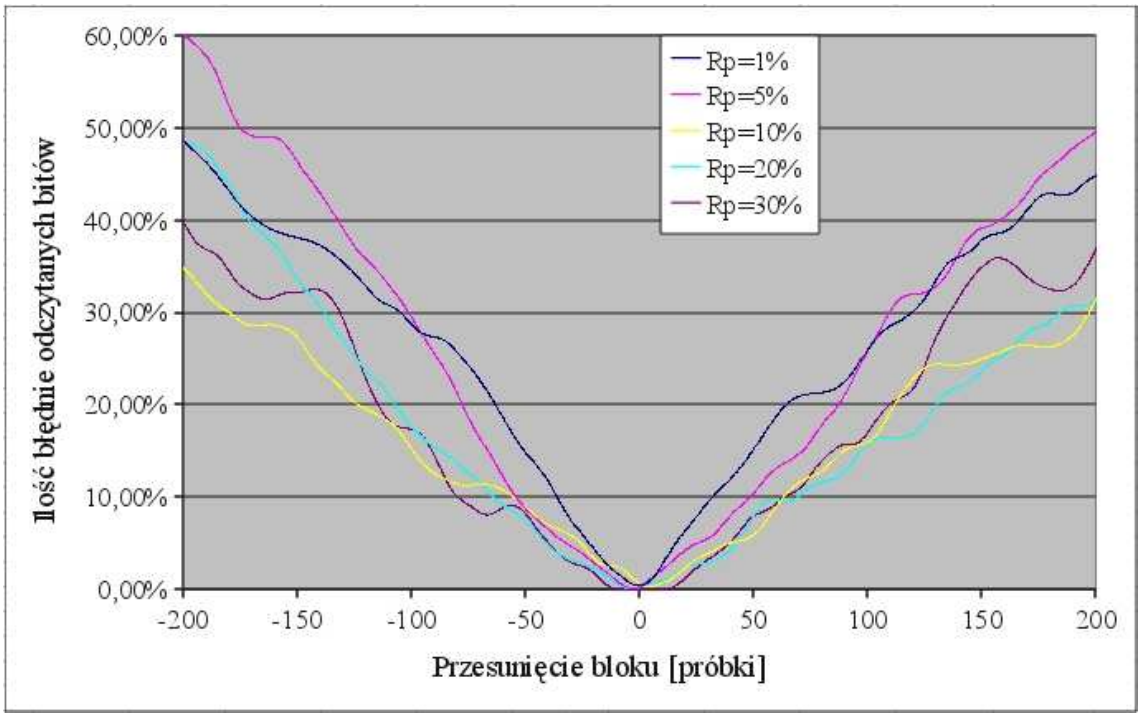
Rysunek 3.20. Ilość błędnie odczytanych bitów w funkcji zmiany rozmiaru bloku podczas odczytu z uwzględnieniem wpływu siły dołączania [opracowanie własne]



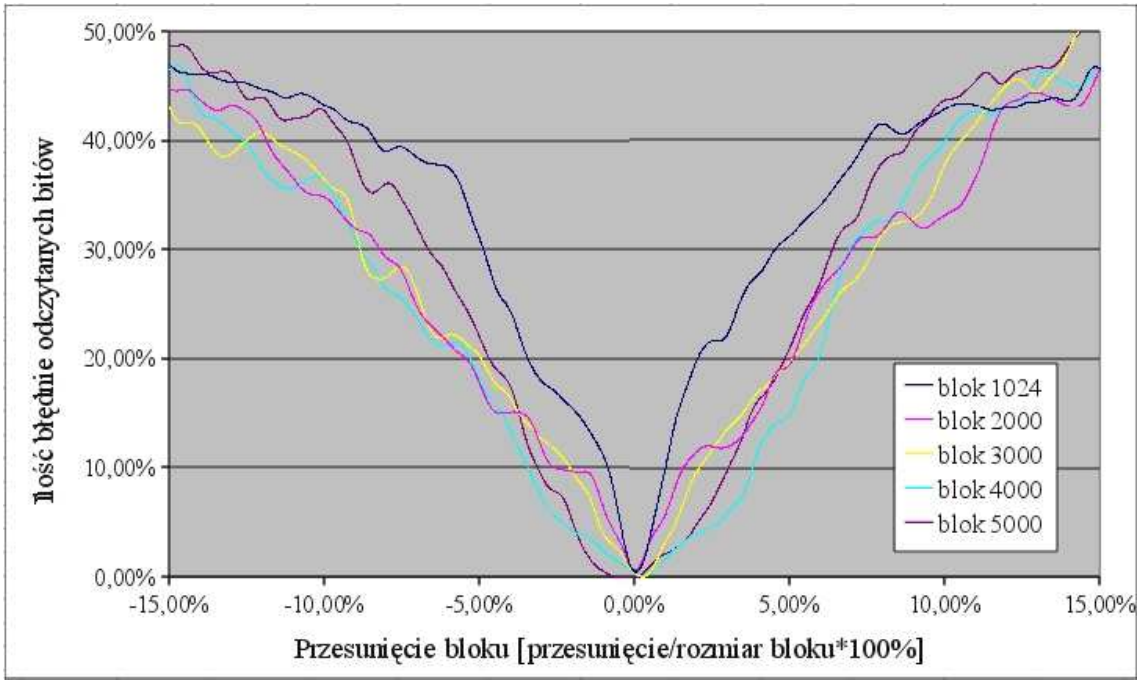
Rysunek 3.21. Ilość błędnie odczytanych bitów w funkcji zmiany rozmiaru bloku podczas odczytu z uwzględnieniem wpływu zastosowanego podczas ukrywania rozmiaru bloku [opracowanie własne]

3.8. Określenie odporności na przesunięcia bloku danych

Kolejnym parametrem niezbędnym do prawidłowego odczytania ukrytej informacji jest określenie położenia bloku przenoszącego dane. Pobranie próbek z niewłaściwego miejsca w sygnale powinno uniemożliwiać otrzymanie dołączonej wiadomości. W celu zbadania wpływu niewłaściwego określenia położenia bloku na poprawność odczytu przeprowadzony został eksperyment. Do kontenera dołączone zostały dane, a następnie podejmowano próby odczytu przesuwając położenie bloku o określoną liczbę próbek. Wyniki zostały zaprezentowane na rysunkach 3.22 oraz 3.23. Liczby na osi odciętych określają przesunięcie (w próbkach) bloku użytego podczas odczytu względem bloku, w którym ukryto dodatkową informację. Ujemne liczby oznaczają przesunięcie w stronę lewą, dodatnie w prawą.



Rysunek 3.22. Ilość błędnie odczytanych bitów w funkcji zmiany położenia bloku z uwzględnieniem wpływu rozmiaru bloku zastosowanego podczas ukrywania [opracowanie własne]



Rysunek 3.23. Ilość błędnie odczytanych bitów w funkcji zmiany położenia bloku z uwzględnieniem wpływu zastosowanej siły dołączania R_p [opracowanie własne]

Jak widać na rysunkach 3.22 i 3.23 błędne określenie położenia bloku wpływa na poprawność odczytu. Zmniejszanie rozmiaru bloku oraz siły dołączania skutkuje zwiększaniem ilości błędów podczas przesuwania bloku. Szybkość narastania ilości błędów jest uzależniona od siły dołączania R_p. Duży wpływ ma również rozmiar bloku. Ze względu na zależność proporcjonalną do rozmiaru bloku, na osi odciętych rysunku 3.23 zdecydowano się wyrazić przesunięcie (P) w procentach, obliczane według wzoru:

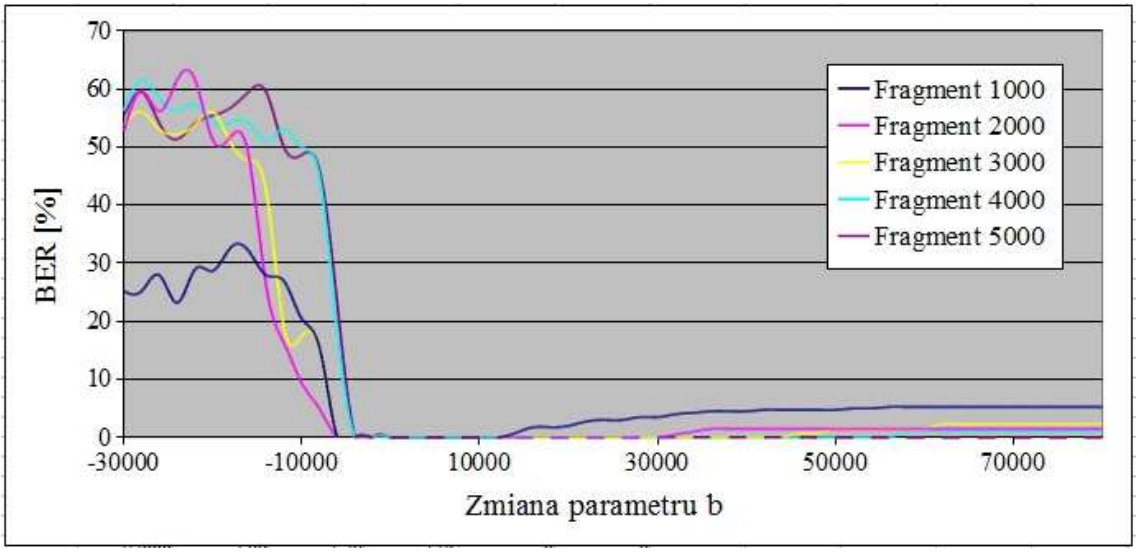
$$P = \frac{n}{N} \cdot 100\% \qquad (3.17),$$

gdzie: n – liczba próbek o jaką przesunięto położenie bloku, N – rozmiar bloku.

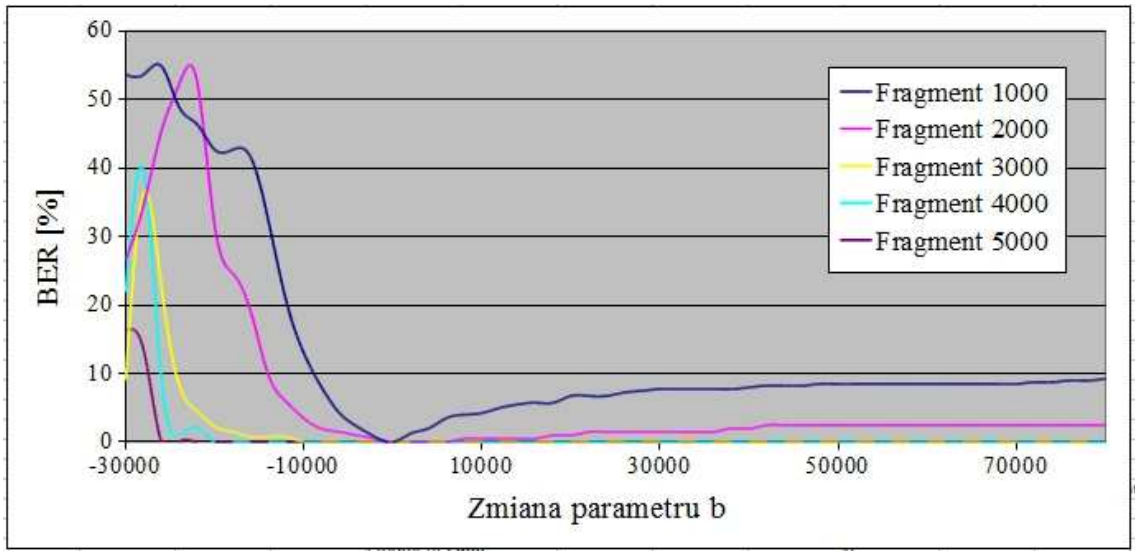
3.9. Wyznaczenie wpływu zmiany kształtu krzywej definiującej dopuszczalne wartości prążków na błędy odczytu ukrytej informacji

Funkcja ograniczająca wartości prążków widma wpływa na ich wybór. Zmiana kształtu tej krzywej może spowodować wybranie zupełnie innych prążków. W niniejszym rozdziale podjęta została próba określenia wpływu współczynników a i b z równania 3.12 określającego kształt omawianej krzywej na liczbę błędnie odczytanych bitów. Procedura testowa polegała na ukryciu danych z zadanymi parametrami a=0.6 i b=30000, a następnie na próbie odczytywania dołączonej informacji przy użyciu a i b zmodyfikowanych o pewną wartość. Wykonana została

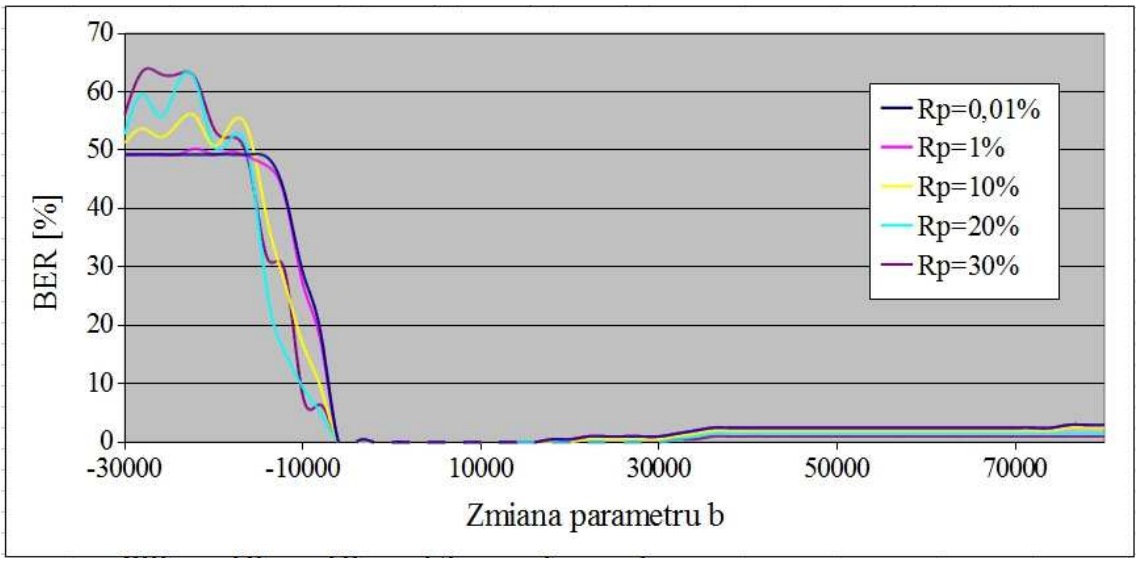
seria eksperymentów, która miała określić liczbę błędnie odczytanych bitów w funkcji zmiany parametrów *a* oraz *b* w trakcie odczytu. Podczas badań uwzględniony został również wpływ rozmiaru zastosowanego bloku oraz siły dołączania na badane zjawisko. Każdy eksperyment został wykonany w dwóch wersjach. W pierwszej wersji prążki do ukrycia informacji były wybierane, tak by znajdowały się jak najbliżej f_{max}. W drugiej, spośród znalezionych prążków widma wybierano najbardziej oddalone od f_{max}. Wyniki uzyskane podczas modyfikacji parametru *b* w momencie odczytu przedstawiono na rysunkach 3.24-3.27. Na osi odciętych przedstawiono wartości, jakie dodane zostały do parametru *b*.



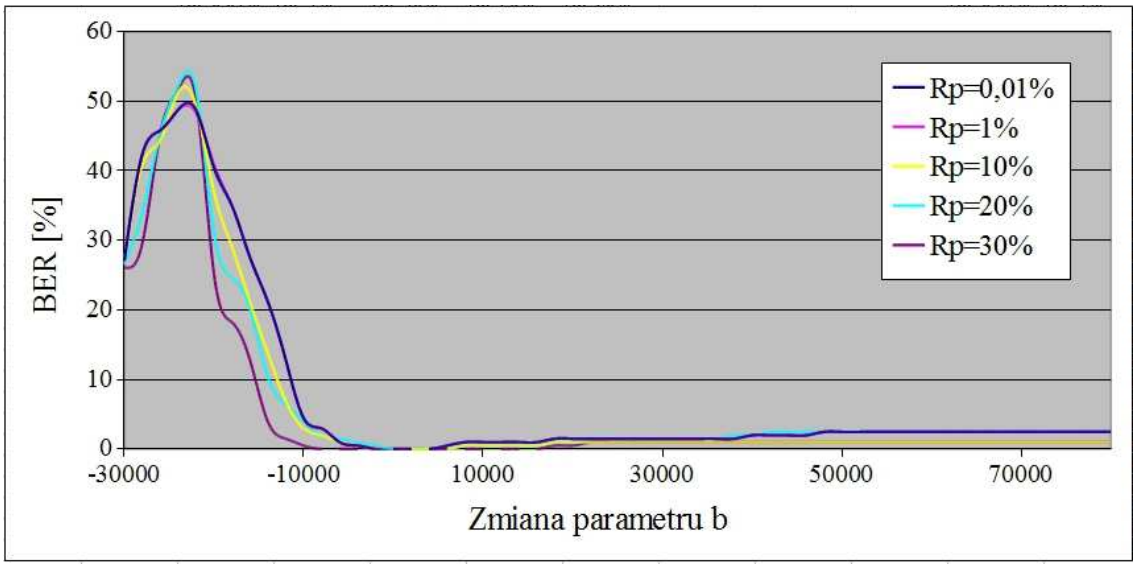
Rysunek 3.24. Wpływ zmiany parametru *b* na ilość błędnie odczytanych bitów, przy wyborze prążków najbardziej oddalonych od maskera [opracowanie własne]



Rysunek 3.25. Wpływ zmiany parametru *b* na ilość błędnie odczytanych bitów, dla prążków położonych najbliżej maskera [opracowanie własne]



Rysunek 3.26. Wpływ zmiany parametru *b* na ilość błędnie odczytanych bitów, przy wyborze prążków najbardziej oddalonych od maskera [opracowanie własne]



Rysunek 3.27. Wpływ zmiany parametru *b* na ilość błędnie odczytanych bitów, dla prążków położonych najbliżej maskera [opracowanie własne]

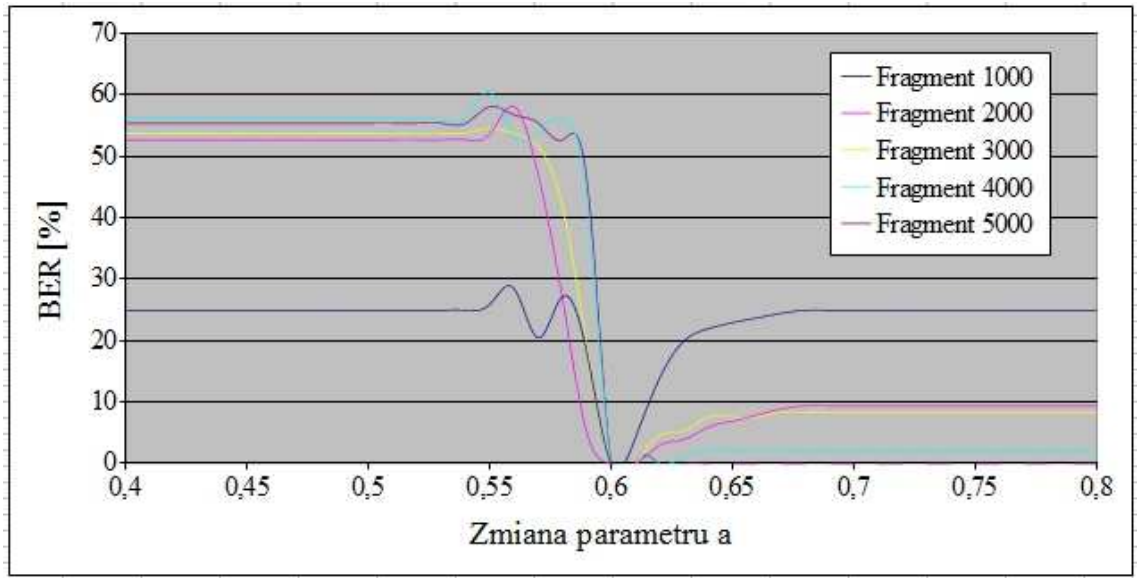
Porównując rysunki 3.24 z 3.25 i 3.26 z 3.27 można zauważyć, że większa podatność na zmiany parametru *b* wstępuje wówczas, gdy preferujemy wybór prążków położonych dalej od maskera. Wydaje się to dosyć oczywiste ze względu na to, że zmiana tego parametru powoduje coraz większe zmiany wartości badanej krzywej w miarę oddalania się od maskera. Ponadto im dalej od maskera tym mniejsze są wartości jakie mogą przyjmować prążki po modyfikacji. Dlatego przy dużych wartościach R_p będą coraz częściej przyjmowały maksymalne dopuszczalne wartości,

co będzie skutkowało gwałtownym narastaniem ilości błędów w miarę zwiększania wartości *b*, ze względu na przekroczenie przez nie wartości dopuszczalnej i uznanie ich za prążki niewykorzystywane. Dowód na to możemy zaobserwować na rysunku 3.27, gdzie widać szybsze narastanie BER dla większych R_p.

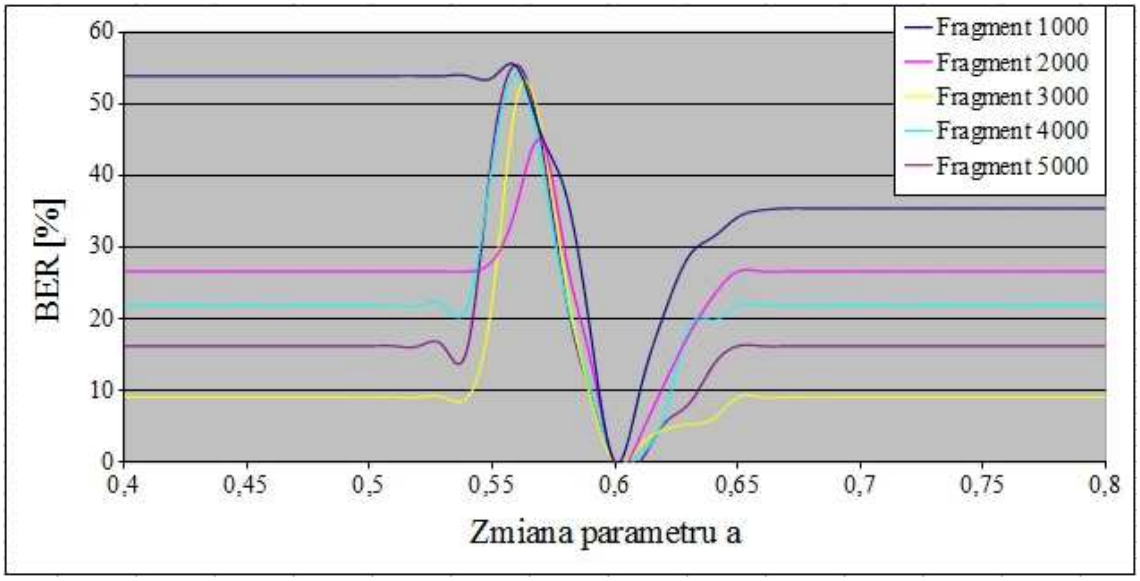
Analiza rysunku 3.25 wykazuje zależność BER od wielkości zastosowanego bloku. Im mniejszy blok tym większa podatność na zmianę *b*. Zaistniała sytuacja spowodowana jest zależnością rozdzielczości transformaty Fouriera od rozmiaru bloku. Mniejszy blok pozwala na uzyskanie mniejszej rozdzielczości, a co za tym idzie skutkuje większymi odległościami pomiędzy prążkami. A jako, że w miarę zwiększania odległości wzrasta wpływ zmiany *b* na zmianę wartości funkcji, to przekłada się to na szybkość narastania BER.

Analiza rysunków 3.26 i 3.27 pozwala zauważyć, że większą podatność na zmiany wykazują stegokontenery, w których użyto mniejszą siłę dołączania.

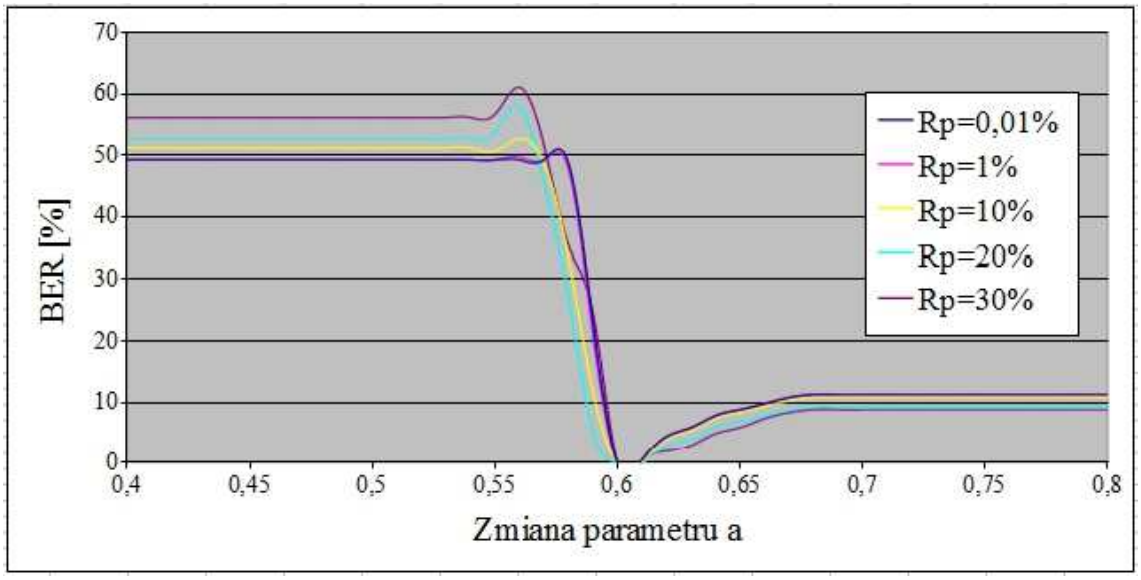
Wyniki uzyskane podczas modyfikacji parametru *a* przedstawiono na rysunkach 3.28-3.31. Na osi odciętych przedstawiono wartości, o jakie zmodyfikowano wartość parametru *a*.



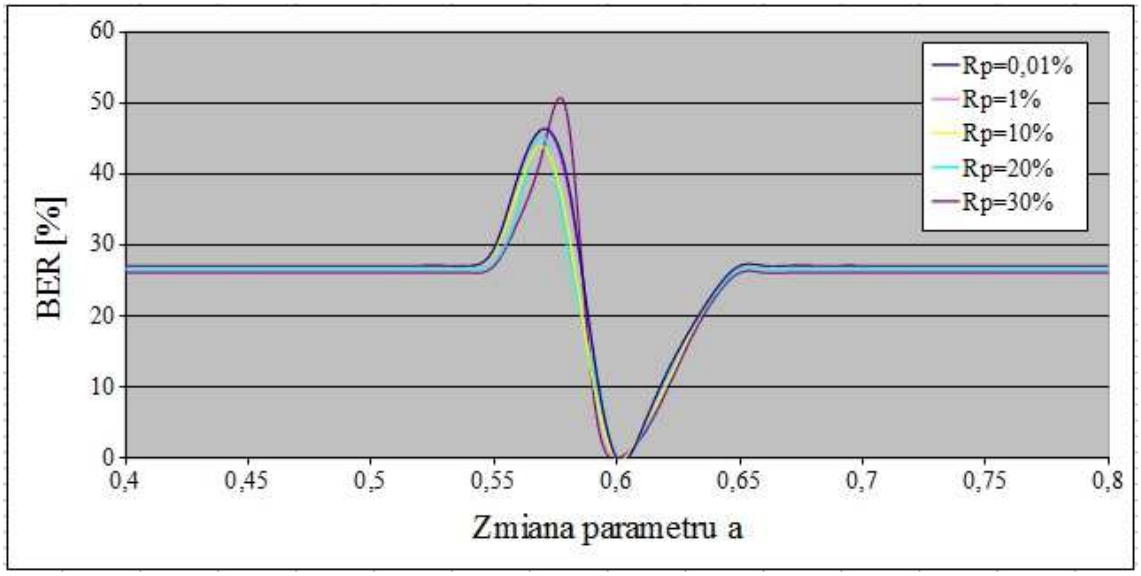
Rysunek 3.28. Wpływ zmiany parametru *a* na ilość błędnie odczytanych bitów, przy wyborze prążków najbardziej oddalonych od maskera [opracowanie własne]



Rysunek 3.29. Wpływ zmiany parametru *a* na ilość błędnie odczytanych bitów, dla prążków położonych najbliżej maskera [opracowanie własne]



Rysunek 3.30 Wpływ zmiany parametru *a* na ilość błędnie odczytanych bitów, przy wyborze prążków najbardziej oddalonych od maskera [opracowanie własne]



Rysunek 3.31. Wpływ zmiany parametru *a* na ilość błędnie odczytanych bitów, dla prążków położonych najbliżej maskera [opracowanie własne]

Analiza rysunków 3.28 – 3.31 pozwala zauważyć, że największą wrażliwość na zmianę parametru *a* wykazują stegokontenery, w których preferowano ukrywanie w prążkach oddalonych od maskera. BER wyraźnie narasta w momencie zmniejszania wartości *a*. Powodem tego jest przyjmowanie przez funkcję ograniczającą niewielkich wartości na krańcach przedziału częstotliwości przeznaczonych do dołączania danych. Podczas zmniejszania wartości współczynnika *a* niejednokrotnie wartość funkcji może tam być ujemna, co skutkuje zupełnym wyłączeniem tego obszaru z użycia i powoduje wybranie do odczytu zupełnie innych prążków.

Wyniki przedstawione w niniejszym rozdziale wykazują, że znajomość parametrów *a* i *b* jest niezbędna do prawidłowego odczytania dołączonej wiadomości, ze względu na ich znaczący wpływ na liczbę błędnie odczytanych bitów.

4. Ocena jakości opracowanego modelu systemu steganograficznego

W poprzednim rozdziale zaprezentowana została koncepcja budowy nowo opracowanej metody steganograficznej wykorzystującej dźwięk jako kontener do ukrycia informacji. Przeprowadzone zostały również badania mające na celu określenie dopuszczalnej zmienności parametrów zawartych w kluczu steganograficznym. Określony został również wpływ tych zmian na poziom wprowadzanych zniekształceń oraz ich słyszalności. Niniejszy rozdział ma na celu przedstawienie wyników badań określających jakość opracowanej metody steganograficznej i jego implementacji.

W niniejszym rozdziale opracowany algorytm będzie oznaczany symbolem MF (jest to skrót od słów Maskowanie oraz Fourier). Jakość badanej metody określona została w wyniku porównania z innymi popularnymi aplikacjami steganograficznymi używającymi dźwięku jako nośnika ukrytej informacji.

Wykonane testy porównawcze obejmowały takie aspekty jak:

1. pojemność steganograficzna,
2. przezroczystość dołączanej informacji,
3. poziom wprowadzanych zniekształceń,
4. odporność dołączonych danych na przekształcenia dźwięku oraz kompresję,
5. odporność na filtrowanie.

Ze względu na często występującą zależność wyników od rodzaju dźwięku użytego jako nośnika ukrytej informacji, ocena jakości badanej metody przeprowadzona została przy wykorzystaniu zestawu nagrań. Dobrano do niego dźwięki wytwarzane w różny sposób prezentujące różne grupy. W skład zestawu weszły:

- mowa – nagranie głosu jednej osoby, reprezentujące dźwięki naturalnej mowy ludzkiej,
- pianino – utwór wykonywany tylko przy użyciu pianina, charakterystyczny dla instrumentów strunowych,
- trąbka – jedno-instrumentowe wykonanie utworu muzycznego, zawierające dźwięki charakterystyczne dla instrumentów dętych,
- muzyka – popularne nagranie z gatunku pop, zawierające kompozycję dźwięków wielu różnych instrumentów oraz wokal.

Taki dobór sygnałów testowych pozwolił na przeanalizowanie wyników uzyskiwanych dla różnych metod w znacznie się różniących pod względem charakteru utworach, będących reprezentantami bardzo popularnych i szeroko rozpowszechnionych grup nagrań.

Wszystkie badania przedstawione w tym rozdziale wykonane zostały przy użyciu do ukrywania informacji dźwięku zapisanego w formacie "wav", próbkowanego z częstotliwością 44100Hz o próbce zapisanej na 16 bitach.

4.1. Aplikacje porównawcze

Przegląd dostępnych aplikacji i wstępna analiza ich działania pozwoliły wybrać grupę programów do porównań. Podczas wstępnej selekcji odrzucone zostały programy, które nie dołączały ukrywanej informacji bezpośrednio do danych oraz te, które działały błędnie. Do prac badawczych wybrano trzy aplikacje oraz implementację algorytmu wykorzystującego do ukrywania informacji modyfikację prążków widma uzyskanego za pomocą transformaty Fouriera. Aplikacjami tymi są:

1. **h4pgp** – program ukrywa dane przy wykorzystaniu najmniej znaczących bitów. Przed dołączeniem informacja jest szyfrowana algorytmem PGP. Ma to na celu zwiększenie bezpieczeństwa chronionych danych oraz zwiększenie losowości dołączanego ciągu. Ponadto uniemożliwia wykrycie dołączonej informacji przez prostą analizę ciągu bitów odczytanych z najmniej znaczących pozycji zapisu[86].
2. **mp3stego** – aplikacja ukrywa dane podczas kodowania dźwięku do formatu mp3. Wprowadza zmiany do algorytmu kodującego zależne od dołączanej sekwencji bitów. Informacja przed ukryciem jest dodatkowo kompresowana i szyfrowana w celu zmniejszenia objętości i zwiększenia losowości dołączanego ciągu bitów[87].
3. **s-tools** – aplikacja ukrywająca dane techniką LSB. Nie wykorzystuje jednak wszystkich próbek, lecz losowo wybiera miejsca, w których dołączana jest informacja[91]. Dane przed dołączeniem poddawane są szyfrowaniu algorytmem wybranym spośród: IDEA, DES, tripple DES, MDC.

Algorytm oparty o transformatę Fouriera (oznaczany skrótem TF) zaprogramowany został w celu porównania zaprojektowanej metody z dotychczas istniejącym rozwiązaniem. Jego zasada działania opiera się na zmianie wartości dwóch prążków widma p1 i p2 tak, by prążek p1 był większy o *R* od prążka p2 jeśli ukrywana

jest binarna jedynka. Ukrywanie binarnego zera polega na doprowadzeniu do sytuacji w której prążek p1 jest mniejszy od p2 o wartość *R*. Położenie prążków określane jest poprzez podanie odpowiadających im częstotliwości. We wszystkich fragmentach wykorzystywane są te same prążki. Jednakowa dla wszystkich fragmentów jest również różnica wartości *R*. Dobrano ją tak, by odpowiadała średniej wartości różnic wprowadzonych w sygnale przetworzonym metodą MF. Ze względu na znaczne różnice zależne od użytego pasma częstotliwości metodą TF przygotowano dwa różne fragmenty:

- pierwszy oznaczony symbolem TF_{aud}, w którym do ukrycia wykorzystano częstotliwości słyszalne 308Hz oraz 330Hz,
- drugi oznaczony symbolem TF_{inaud}, w którym do ukrycia wykorzystano częstotliwości słyszalne 20,021kHz oraz 20,066kHz,
- Jeśli w prezentowanym eksperymencie wyniki uzyskane dla dwóch przedstawionych wersji metody TF będą się różniły wówczas zostaną zaprezentowane oddzielnie. W przeciwnym wypadku uzyskane wyniki zostaną umieszczone tylko raz i zostaną oznaczone symbolem TF.
- W przypadku metody MF do porównań wykorzystywane są stegokontenery, w których wykorzystano bloki o rozmiarze 2000 próbek oddzielone od siebie blokami łączącymi długości 100 próbek. Siłę dołączania R_p oraz odstęp pomiędzy używanymi i nieużywanymi wartościami prążków przyjęto na poziomie $15\% \cdot W_{max}$. Jeśli w danym eksperymencie nie zostaną podane inne parametry to będzie to oznaczało, że zastosowane zostały przedstawione powyżej wartości parametrów podczas tworzenia stegokontenera.

4.2. Porównanie pojemności steganograficznej

Ze względu na odmienny charakter działania, każda z badanych metod dołącza informację w inny sposób. Największą pojemność steganograficzną oferuje program h4pgp. Uzyskuje ją poprzez wykorzystanie czterech najmniej znaczących bitów w każdej 16 bitowej próbce do zapisania dodatkowych danych metodą LSB. Podczas ukrywania wykorzystywane są wszystkie próbki, dlatego więc pojemność steganograficzna metody równa jest jednej czwartej rozmiaru nośnika.

Program s-tools pozwala na ukrycie mniejszej ilości danych pomimo tego, że również ukrywa informację metodą LSB. Spowodowane jest to wykorzystaniem

tylko niektórych próbek. Algorytm analizuje sygnał a następnie wybiera do ukrycia próbki, które według zaprogramowanych zasad najlepiej się do tego nadają. Liczba wykorzystywanych próbek zależna jest od charakteru wykorzystywanego kontenera. Przeprowadzone eksperymenty wykazały, że jest ona kilkukrotnie mniejsza od pojemności oferowanej przez metodę h4pgp. Jednak ciągle są to bardzo duże pojemności ze względu na operowanie na pojedynczych próbkach.

Metody, które nie mają możliwości analizy pojedynczych próbek i wprowadzania zmian niezależnie w każdej z nich nie są w stanie osiągnąć dużych pojemności steganograficznych. Uwidacznia się to na przykładzie programu mp3stego, dołączającego dane podczas kompresji mp3.

Metody MF oraz TF również nie są w stanie osiągnąć dużych pojemności steganograficznych ze względu na rozpraszanie bitu informacji w całym obszarze przetwarzanego sygnału zamiast wykorzystania do tego celu jednej próbki. Ponadto, ze względu na możliwość wystąpienia słyszalnych zakłóceń przy stosowaniu zbyt małych bloków konieczne jest stosowanie bloków o dużym rozmiarze. W tabeli 4.1 Przedstawiono uzyskane pojemności steganograficzne porównywanych metod dla zapisu stereofonicznego. W metodach TF oraz MF zastosowano do ukrycia danych bloki o rozmiarze tysiąca próbek. Różnice w pojemności oferowanej przez te metody wynikają z zastosowania w metodzie MF dodatkowych bloków łączących.

Tabela 4.1. Porównanie pojemności steganograficznej [opracowanie własne]

	MF	h4pgp	mp3stego	TF	s-tools
muzyka	84 bit/s	176400 bit/s	153 bit/s	88 bit/s	45200 bit/s
trąbka	84 bit/s	176400 bit/s	153 bit/s	88 bit/s	44093 bit/s
mowa	84 bit/s	176400 bit/s	153 bit/s	88 bit/s	45144 bit/s
pianino	84 bit/s	176400 bit/s	153 bit/s	88 bit/s	45112 bit/s

4.3. Ocena przezroczystości ukrytej informacji

Najważniejszym aspektem w steganografii jest dołączenie dodatkowych danych w taki sposób by osoby postronne nie były w stanie zidentyfikować ich istnienia. Aby było to możliwe muszą zostać spełnione dwa warunki: wartości statystyczne

nośnika nie mogą w znaczącym stopniu ulec zmianie oraz nie mogą się pojawić zakłócenia, które człowiek będzie w stanie usłyszeć.

Do eksperymentu przygotowano sygnały zawierające jednakową ilość danych dołączonych za pomocą prezentowanych metod. W przypadku, gdy pojemność steganograficzna metody była większa, część nośnika nie była wykorzystywana.

W celu porównania zniekształceń wprowadzanych przez metodę MF z zniekształceniami generowanymi przez pozostałe metody porównawcze wybrano następujące miary jakości dźwięku:

- błędu średniokwadratowego (MSE),
- znormalizowanego błędu średniokwadratowego (NMSE),
- odległości sygnału od szumu (SNR),
- szczytowej wartości odległości sygnału od szumu (PSNR),
- normy LP (LP),
- maksymalnej różnicy pomiędzy sygnałem oryginalnym i zmodyfikowanym (MD),
- średniej bezwzględnej różnicy pomiędzy sygnałami (AD),
- znormalizowanej średniej bezwzględnej różnicy pomiędzy sygnałami (NAD),
- przezroczystości ukrytych danych (AF).

Ze względu na zależność uzyskanych wartości w metodzie MF od siły dołączania *Rp* w tabeli 4.2 zaprezentowane zostały poziomy zniekształceń wprowadzane przy użyciu różnych sił dołączania dodatkowej informacji. W tabeli 4.3 przedstawione zostały wyniki wszystkich porównywanych metod dla różnych rodzajów dźwięku.

Tabela 4.2. Zniekształcenia wprowadzane przez metodę MF [opracowanie własne]

Rp	MSE	NMSE	SNR [dB]	PSNR [dB]	LP	MD	AD	NAD	AF
1%	7E-6	1,3E-3	29,0	99,9	7 E-6	0,06	1E-3	0,02	1
10%	9E-6	1,7E-3	27,6	98,5	9 E-6	0,06	1,6E-3	0,03	1
15%	1,4E-5	2,7E-3	25,7	96,6	1,4E-5	0,06	2E-3	0,04	1
20%	2,3E-5	4,3E-3	23,6	94,5	2,3E-5	0,06	2,8E-3	0,06	1
30%	5,6E-5	1E-3	19,8	90,7	5,6E-5	0,06	4,3E-3	0,09	0,99
40%	1E-4	2E-3	17,0	87,9	1E-4	0,08	5,8E-3	0,12	0,98

Tabela 4.3. Zniekształcenia wprowadzane przez metody porównawcze [opracowanie własne]

metoda	sygnał	MSE	NMSE	SNR [dB]	PSNR [dB]	LP	MD	AD	NAD	AF
MF	muzyka	2E-4	4E-3	24,1	85,2	2E-4	0,19	0,008	0,05	1
	trąbka	1,9E-5	4E-3	23,9	95,3	1,9E-5	0,05	0,002	0,04	1
	piano	1,4E-5	2,7E-3	25,7	96,5	14E-6	0,06	0,002	0,04	1
	mowa	1,4E-4	5E-3	22,8	86,6	1,4E-4	0,14	0,006	0,05	0,99
h4pgp	muzyka	2,5E-13	5E-12	113	174	2,5E-13	3E-5	8,4E-9	5E-8	1
	trąbka	2,0E-13	4,4E-11	103	175	2,1E-13	3E-5	6,8E-9	1,3E-7	1
	piano	2,3E-13	4,4E-11	103	174	2,3E-13	3E-5	7,5E-9	1,5E-7	1
	mowa	2E-13	8E-12	110,7	174,7	2E-13	3E-5	7E-9	6E-8	1
s-tools	muzyka	4,6E-13	9,1E-12	110	171	4,6E-13	3E-5	1,5E-8	9,3E-8	1
	trąbka	4,3E-13	9,2E-11	100	171	4,3E-13	3E-5	1,4E-8	2,7E-7	1
	piano	6,8E-13	1,3E-10	99	170	6,8E-13	3E-5	2,2E-8	4,4E-7	1
	mowa	6,6E-13	2,5E-11	106	170	6,6E-13	3E-5	2,2E-8	2E-7	1
TF_{aud}	muzyka	2,8E-4	0,0047	15,3	75,4	2,3E-4	0,021	0,018	0,12	0,97
	trąbka	2,4E-4	0,05	22,9	94,3	2,4E-4	0,04	0,004	0,07	0,99
	piano	4E-4	0,077	11,1	82,1	4E-4	0,17	0,009	0,18	0,92
	mowa	2,8E-4	0,01	19,7	83,7	2,8E-4	0,17	0,01	0,1	0,99

TF_{inaud}	muzyka	3E-4	0,0059	22,3	83,3	3E-4	0,025	0,015	0,095	0,99
	trąbka	3,4E-5	0,007	21,4	92,9	3,4E-5	0,008	0,005	0,1	0,99
	piano	4,5E-5	0,0087	20,6	91,6	4,5E-5	0,01	0,0061	0,12	0,99
	mowa	1,67E-4	0,0063	22	85,9	1,7E-4	0,02	0,01	0,1	0,99
mp3stego	muzyka	4,7E-3	0,09	10,35	71,4	0,0047	0,88	0,045	0,27	0,9
	trąbka	1,8E-4	0,046	13,3	85,5	1,8E-4	0,71	0,008	0,17	0,95
	piano	3,1E-4	0,06	12,2	83,2	3,1E-4	0,21	0,01	0,22	0,94
	mowa	16E-4	6E-2	12,23	76,1	16E-4	0,7	0,02	0,19	0,94

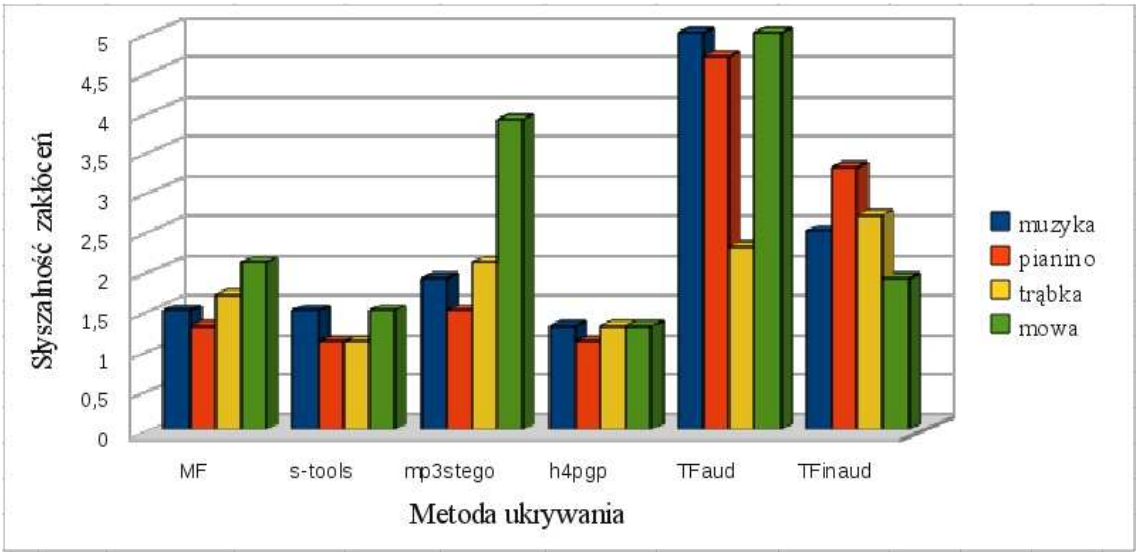
Analizując wyniki przedstawione w tabeli 4.3 można łatwo zauważyć, że najmniejsze zniekształcenia generują metody h4pgp oraz s-tools ukrywające informację w najmniej znaczących bitach próbek sygnału. Dziwnym może się wydać nieco wyższy poziom zniekształceń generowanych przez program s-tools pomimo mniejszej pojemności. Wynik taki spowodowany jest nieco innym rozmieszczeniem informacji w nośniku. Metoda h4pgp wykorzystuje najmniej znaczące bity każdej z próbek. Algorytm s-tools wybiera natomiast miejsca, w których zmiany będą najlepiej maskowane i wykorzystuje większą liczbę bitów w modyfikowanych próbkach. Tak więc zmiana wprowadzona w sygnale będzie miała o wiele większy wpływ na wartości przedstawionych miar zakłóceń i pomimo mniejszej ilości ukrytych bitów statystyki będą nieco gorsze od uzyskanych za pomocą metody h4pgp.

Najgorzej w przedstawionym porównaniu wypadła metoda mp3stego. Należy jednak pamiętać, że na poziom zniekształceń ma również wpływ sama kompresja do formatu mp3. Nie ma możliwości określenia jaki jest poziom zniekształceń wprowadzanych przez samą metodę.

Metoda MF generuje większe zakłócenia niż metody oparte na zasadzie działania algorytmu LSB, lecz jest pod tym względem lepsza od mp3stego oraz wcześniejszych metod bazujących na transformacie Fouriera. Podczas porównania z metodami TF widać, że poziom ten spełnia tylko metoda działająca w zakresie pasma niesłyszalnego. Jednak nawet ona nie jest w stanie osiągnąć równie dobrych wyników w zakresie drugiej ważnej miary jaką jest przezroczystość znaku wodnego (AF).

Przedstawione miary wykazują, że poziom generowanych zniekształceń jest dostatecznie niski by wprowadzone zmiany nie były zauważane. Potwierdzają to przeprowadzone testy odsłuchowe których wyniki zaprezentowane zostały w rozdziale 3.2.2.

Dodatkowo aby porównać słyszalność zakłóceń wprowadzanych przez metodę MF z zakłóceniami wprowadzanymi przez pozostałe metody przeprowadzono test porównawczy. Zaprezentowano słuchaczom nagrania przetworzone za pomocą każdej z porównywanych metod i poproszono o ocenę słyszalności ich zakłóceń. Zastosowano metodologię przeprowadzania testu i oceniania opisaną w rozdziale 3.2.1. Uzyskane wyniki zaprezentowane zostały na rysunku 4.1.



Rysunek 4.1. Ocena słyszalności zakłóceń wprowadzanych przez porównywane metody [opracowanie własne]

Jak widać na rysunku 4.1 metoda MF została dobrze oceniona w teście odsłuchowym pod kątem wprowadzanych zakłóceń. Jej wyniki są niewiele gorsze od metod s-tools i h4pgp bazujących na algorytmie LSB znanym z wprowadzania bardzo małego poziomu zakłóceń. Pozostałe badane metody wprowadzają wyższy poziom zakłóceń. Dziwne mogą się wydać słabe wyniki uzyskane przez metodę TF_{inaud} działającą w paśmie niesłyszalnym, która z definicji nie wprowadza słyszalnych zmian w sygnale. Jednak szczegółowa analiza materiału testowego wykazała, że występują w nim charakterystyczne zakłócenia powstające na łączeniach bloków. One to były przyczyną gorszej oceny uzyskanej przez metodę.

4.4. Ocena odporności nowego systemu steganograficznego

Nie zawsze przekazanie informacji możliwe jest do zrealizowania za pomocą idealnego kanału komunikacyjnego. Często wprowadza on do stegokontenera pewien poziom zakłóceń. Ponadto z definicji stegosystemu wynika, że kanał komunikacyjny jest publiczny, więc przesyłane dane mogą podlegać różnym modyfikacjom, zarówno celowym jak i przypadkowym. Mogą być one związane zarówno z kompresją danych podczas przesyłu jak również z modyfikacjami jakim nośnik jest poddawany przez osoby trzecie. Niniejszy rozdział został więc poświęcony ocenie odporności danych do ukrytych za pomocą analizowanych metod.

4.4.1. Porównanie odporności na filtrowanie

Podczas przesyłania lub przetwarzania sygnału dźwiękowego powstają w nim zakłócenia. W celu ich usunięcia sygnał często poddawany jest filtrowaniu. Operacja ta usuwa z niego niepożądane składowe. Najczęściej polega ona na usunięciu niepożądanych pasm częstotliwości lub odjęciu sygnału szumu. Aby ocenić odporność dołączonych danych na filtrowanie przeprowadzono badania określające wpływ odfiltrowywania różnych pasm częstotliwości na liczbę błędnie odczytanych bitów informacji. Wyniki eksperymentu przedstawione zostały w tabeli 4.4.

Tabela 4.4. Odporność porównywanych metod na filtrowanie [opracowanie własne]

Metoda	Dźwięk	Ilość błędnie odczytanych bitów [%]									
Pasmo przepustowe:		20Hz - 22kHz	60Hz - 22kHz	100Hz - 22kHz	200Hz - 22kHz	1kHz - 22kHz	1Hz - 15kHz	1Hz - 10kHz	1Hz - 5kHz	1Hz - 2kHz	1Hz - 1kHz
MF	Pianino	0.7	2.8	2.8	8.6	42.1	1.4	1.4	2.1	3.6	10.0
	Trąbka	3.6	4.5	5.0	5.0	21.8	1.4	1.4	2.3	3.2	32.3
	Mowa	10.0	16.0	20.6	33.3	48.7	6.0	6.7	6.0	12.7	11.3
	Muzyka	9.3	8.3	11.8	15.3	49.3	0	0.5	1.0	2.0	3.9
TF_{inaud}	Pianino	0	0	0	0	0	0	4.3	46.6	47.2	47.8
	Trąbka	0	0	0	0	0	0.4	8.6	48.2	49.8	49.8
	Mowa	0	0	0	0	0	5.8	20.3	46.5	48.8	48.8
	muzyka	0	0	0	0	0	0	7.0	47.1	49.3	49.3
TF_{aud}	pianino	0	0	0	0	34.8	0	0	0	0	0
	trąbka	0	0	0	0	55.1	0	0	0	0	0
	mowa	0	0.5	2.3	6.4	25.0	0	0	0	0	0.6
	muzyka	0	0.8	2.2	3.1	44.1	0	0	0	0	4.0

Wyniki przedstawione w tabeli 4.4 dotyczą tylko metod MF i TF. Próba wykonania najmniejszej nawet modyfikacji w stegokontenerach wygenerowanych przez pozostałe z analizowanych metod powodowała, że generowały one błędy podczas próby odczytu lub działały w inny nieprawidłowy sposób. Żaden z tych programów nie zwrócił wyniku. Ze względu na niemożliwość odczytania informacji ze zmodyfikowanego stegokontenera należy uznać, że żaden z analizowanych programów bazujących na metodzie LSB nie wykazuje odporności na modyfikacje stegokontenera. Dlatego też w kolejnych eksperymentach, w których stegokontener będzie podlegał modyfikacjom, metody oparte na technice LSB zostaną pominięte jako nie wykazujące żadnej odporności.

Analiza wyników uzyskanych dla metod wykorzystujących DFT pozwala zauważyć, że dane dołączone metodą TF_{aud} są wrażliwe na odfiltrowywanie częstotliwości słyszalnych z zakresu, w którym ukryte zostały dane, natomiast odporne są na filtrowanie dolnoprzepustowe.

Metoda TF_{inaud} jest zupełnie niewrażliwa na usuwanie niskich częstotliwości. Jednak ze względu na wykorzystanie ultradźwięków do ukrycia informacji wykazuje ona wrażliwość na filtrowanie dolnoprzepustowe.

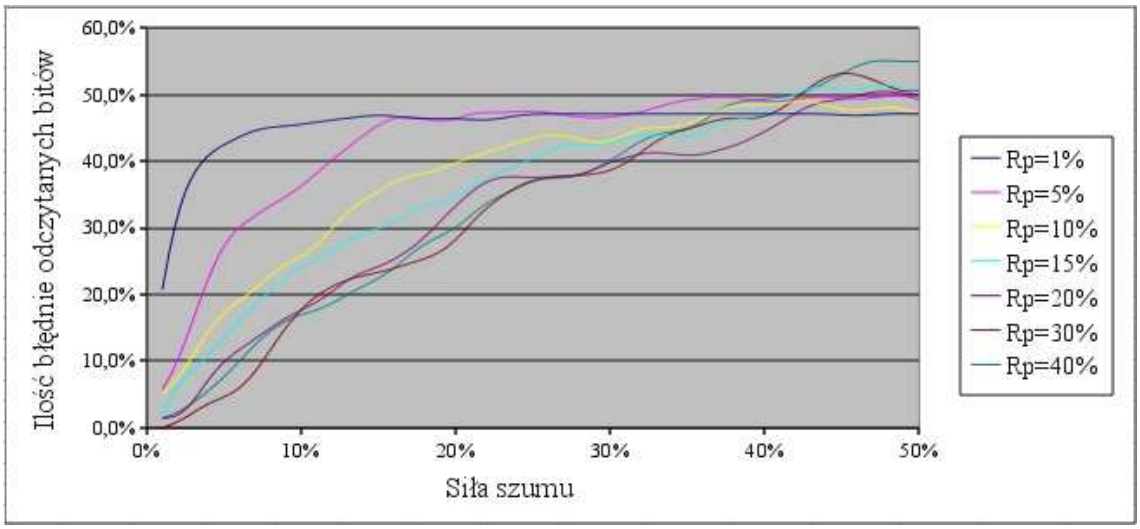
W przypadku metody MF nie da się jednoznacznie określić wrażliwości na odfiltrowywanie określonego pasma częstotliwości. Wrażliwość ta będzie inna dla każdego sygnału. Powodem takiego stanu rzeczy jest zupełnie inne rozmieszczenie danych w każdym sygnale zależne od jego parametrów. Analizując uzyskane wyniki można zauważyć, że wzrost ilości błędnie odczytanych bitów jest ściśle powiązany z odfiltrowywaniem większego zakresu częstotliwości słyszalnych. Jest to zgodne z zasadą działania metody, która ukrywa dane w pobliżu prążka mającego największą wartość. A jako że w każdym z sygnałów audio największy udział mają częstotliwości słyszalne to wraz z ich usuwaniem następuje również uszkadzanie ukrytych danych.

W metodzie MF częstotliwości użyte do ukrycia danych zależne są od miejsc występowania prążków o największej wartości w poszczególnych fragmentach sygnału. Ich rozkład będzie zupełnie inny w każdym sygnale. Stanowi to zaletę zaprojektowanej metody ze względu na niemożliwość określenia uniwersalnego pasma częstotliwości, którego odfiltrowanie niszczyło by ukryta informację. Ponadto rozproszenie danych

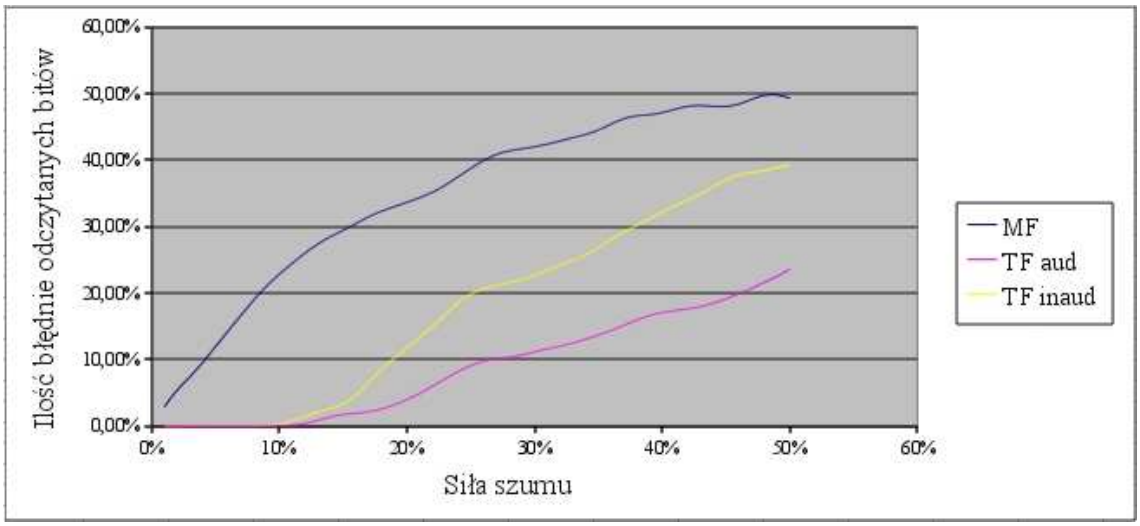
w szerokim paśmie częstotliwości słyszalnych uniemożliwia poważne uszkodzenie dołączonych danych bez zniszczenia sygnału kontenera.

4.4.2. Analiza wpływu szumu na uszkodzenia dołączonych danych

Istnieje wiele źródeł generujących szum, na który narażony jest sygnał. Wprowadzają one do sygnału zmiany, które niejednokrotnie przybierają formę słyszalnych zakłóceń. W niniejszym rozdziale przedstawiony został rezultat eksperymentu mającego na celu ocenę wpływu dołączenia do sygnału stegokontenera sygnału szumu białego na uszkodzenia ukrytych danych. Ocenie takiej poddano najpierw dane dołączone za pomocą metody MF, uwzględniając wpływ siły dołączania R_p. Uzyskane wyniki zaprezentowane zostały na rysunku 4.2. Następnie porównano uzyskane wyniki z metodami TF. Rezultaty prezentuje rysunek 4.3.



Rysunek 4.2 Odporność metody MF na dołączenie szumu do stegokontenera [opracowanie własne]



Rysunek 4.3. Odporność porównywanych metod na dołączanie szumu [opr.własne]

Analizujący rysunek 4.2 można zauważyć, że wraz ze wzrostem siły szumu rośnie liczba błędów odczytu. Ponadto uwidacznia się zależność pomiędzy odpornością na dołączanie szumu a zastosowaną siłą dołączania R_p. Większa wartość R_p skutkuje większymi różnicami wartości pomiędzy prążkami, Uszkodzenie dołączonej informacji wymaga więc zmian o większej mocy. Uzyskane wyniki są więc zgodne z teoretycznymi założeniami metody. Na rysunku 4.3 możemy zauważyć, że metody TF mają większą odporność od MF. Dzieje się tak ze względu na to, że w metodach TF wprowadzana różnica wartości pomiędzy prążkami przenoszącymi informację jest jednakowa dla wszystkich fragmentów. Uszkodzenie informacji następuje dopiero w momencie osiągnięcia pewnej siły zmian, co łatwo zauważyć na analizowanym wykresie. Dalszy wzrost siły dołączanego szumu powoduje zwiększanie się ilości błędów. Nieco niższa wrażliwość metody TF_{aud} wynika z tego, że w paśmie słyszalnym oryginalnie występują duże różnice wartości. Jeśli pozwalają one na poprawne odczytanie bitu, który jest ukrywany w danym fragmencie, wówczas wartości prążków pozostają niezmienione. W efekcie niektóre z fragmentów mają większą różnicę wartości pomiędzy prążkami przenoszącymi informację, co jednocześnie przekłada się na konieczność zastosowania większej siły zmian do uszkodzenia ukrytego bitu informacji. W paśmie niesłyszalnym oryginalnie występujące wartości prążków są bardzo małe, więc różnica wartości pomiędzy prążkami przenoszącymi informację będzie zbliżona do wprowadzanej przez algorytm, co skutkuje mniejszą odpornością.

Najmniejsza odporność metody MF wynika z zastosowania adaptacji mocy wprowadzanych zmian do mocy sygnału. We fragmentach sygnału o niewielkich

wartościach prążków, wprowadzane zmiany mają bardzo małe wartości wobec czego nawet zmiany o niewielkiej mocy są w stanie uszkodzić zawartą w nich informację. Dołączenie szumu o mocy 3% powoduje, że około 7% bitów ukrytej informacji jest odczytywanych błędnie. Szum o takiej mocy jest jednak wyraźnie słyszalny w dźwięku. Uszkodzenie informacji dołączonej metodą MF przez dołączanie szumu wymagałoby więc uszkodzenia sygnału i byłoby jednoznaczną wskazówką dla odbiorcy, że informacja została celowo uszkodzona.

4.4.3. Ocena odporności na przekształcenia dynamiczne

Przekształceniem dynamicznym nazywamy dowolne przekształcenie wpływające na zmianę amplitudy sygnału. Aby ocenić wpływ takiego rodzaju modyfikacji na liczbę błędnie odczytanych bitów, poddano przygotowane stegokontenery modyfikacjom poprzez wyciszanie sygnału, oraz normalizację połączoną z usuwaniem składowej stałej. Uzyskane wyniki zostały zaprezentowane w tabeli 4.5.

Tabela 4.5. Odporność porównywanych metod na przekształcenia dynamiczne (w tabeli przedstawiono wartości BER) [opracowanie własne]

metoda	wyciszanie				normalizacja			
	muzyka	mowa	pianino	trąbka	muzyka	mowa	pianino	trąbka
MF	4,4%	4,6%	2,1%	1,3%	4,4%	5,3%	2,1%	0%
TF_{aud}	56,6%	50%	58,3%	48,9%	47,7%	49,4%	60,8%	49,7%
TF_{inaud}	0%	0%	0%	0%	0%	0%	0%	0%

Wyniki przedstawione w tabeli 4.5 wykazują, że dane dołączane w paśmie niesłyszalnym nie są wrażliwe na przekształcenia dynamiczne. Pasmo słyszalne natomiast ulega znacznym modyfikacjom, które niszczą dane dołączone do niego za pomocą metod TF. Zupełnie inna technika ukrywania danych zastosowana w metodzie MF pozwala na uzyskanie wysokiego stopnia odporności na przekształcenia dynamiczne. Metoda MF oblicza wartości prążków na podstawie klucza oraz wartości największego z prążków występujących w danym fragmencie. Jeśli zastosowane fragmenty są niewielkie, to możemy zmiany wprowadzane w obrębie fragmentu traktować jako jednakowe dla wszystkich próbek. Proporcje pomiędzy poszczególnymi

próbkami pozostają zbliżone do oryginalnych i tylko w minimalnym stopniu wpływają na zmiany widma sygnału. Dzięki temu metoda MF jest w stanie dostosować się do zmian w dynamice sygnału, co daje w efekcie wysoką odporność na omawiane przekształcenia przy jednoczesnym wykorzystaniu pasma słyszalnego.

4.4.4. Analiza podatności na uszkodzenie przez operacje kształtujące efekt przestrzenny

Niejednokrotnie w celu uzyskania bogatszego brzmienia dźwięku do sygnału dodaje się sygnał echa. Jest on również wykorzystywany w znakowaniu wodnym utworów. Ze względu na dużą popularność operacja ta została wykorzystana do ocenienia podatności ukrytych danych na przekształcenia kształtujące efekt przestrzenny. Do stegokontenerów przygotowanych za pomocą porównywanych metod dodano sygnał echa o sile równej 10% sygnału oryginalnego, opóźniony względem oryginału o 2ms. Następnie podjęto próbę odczytania dołączonej informacji z tak przygotowanych nośników. Uzyskane wyniki przedstawione zostały w tabeli 4.6.

Tabela 4.6. Odporność porównywanych metod na operacje kształtujące efekt przestrzenny [opracowanie własne]

metoda	muzyka	mowa	pianino	trąbka
MF	7,9%	13,3%	4,3%	10,0%
TF_{aud}	60,2%	51,2%	55,3%	46,9%
TF_{inaud}	11,9%	0,0%	0,0%	0,0%

Najbardziej odporna na dołączanie echa okazała się metoda TF_{inaud}. Powodem tego był znikomy udział wysokich częstotliwości w analizowanych sygnałach, wobec czego echo, które wpływało na wartości prążków pochodziło z dołączonego sygnału. Pozwalało to utrzymać jego wartość na niskim poziomie nie mającym wpływu na poprawność odczytu danych.

Metoda TF_{aud} okazała się zupełnie nieodporna na dołączenie echa. Energia sygnału echa jest proporcjonalna do enrgii poprzedzającego je sygnału. W przypadku metod TF wartość zmian jest stała i niezależna od energii sygnału. Tak więc głośne fragmenty generowały tak duże zakłócenia, że zmieniały wartość odczytanego bitu.

Metoda MF wykazuje dosyć dużą odporność na operację dołączenia echa, pomimo tego, że ukrywa dane w zakresie pasma słyszalnego. Możliwe jest to ze względu na dostosowywanie siły dołączania do energii sygnału. Wobec czego wielkość zmian konieczna do uszkodzenia dołączonych danych również jest większa niż w przypadku metod TF.

4.4.5. Odporność na zmianę formatu i kompresję

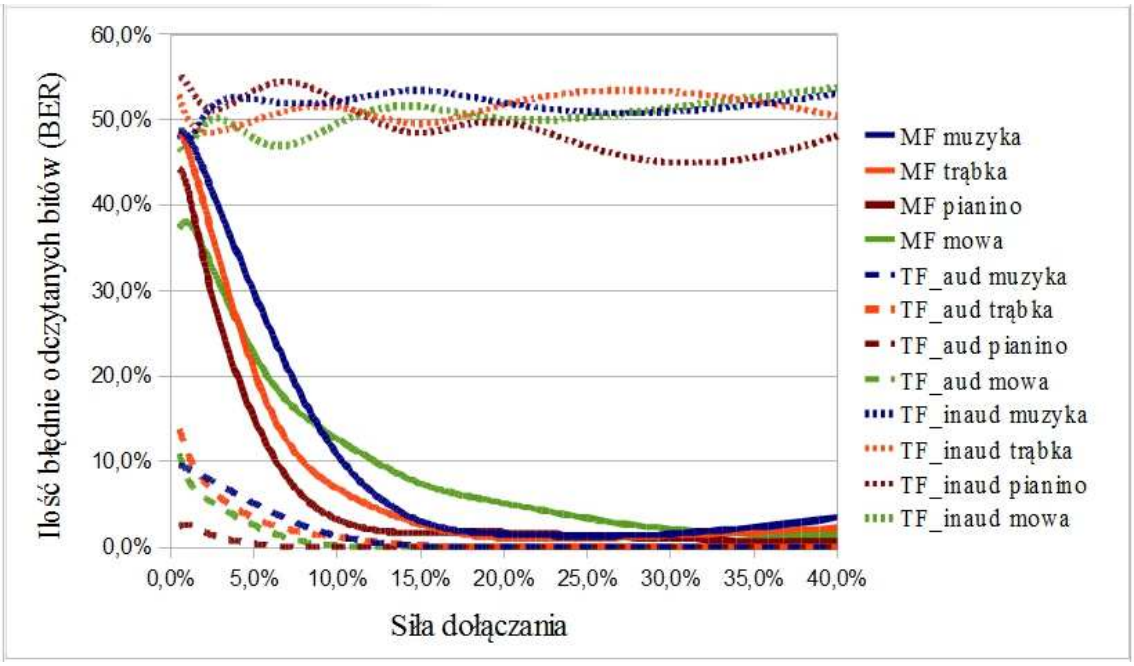
W przypadku, gdy istnieje możliwość zmiany formatu nośnika w czasie przekazywania danych od nadawcy do odbiorcy, konieczne jest zapewnienie odporności na przekształcenia, jakim może zostać poddany nośnik. Wiele metod steganograficznych nie zapewnia odporności lub zapewnia bardzo mały stopień odporności.

Aby zweryfikować stopień odporności prezentowanej metody na kompresję przeprowadzono badania mające na celu określenie ilości błędnie odczytanych bitów (BER) po skompresowaniu nośnika do różnych formatów zapisu dźwięku. Testy polegały na skonwertowaniu przygotowanego wcześniej stegokontenera do określonego formatu, powrotną konwersję do formatu "wav", po której odczytywano ukryte dane i określano stopę błędów. Aby określić, czy występuje zależność pomiędzy odpornością a rodzajem dźwięku, użyto do testów zestawu nagrań zawierającego cztery fragmenty utworów muzycznych o różnym charakterze.

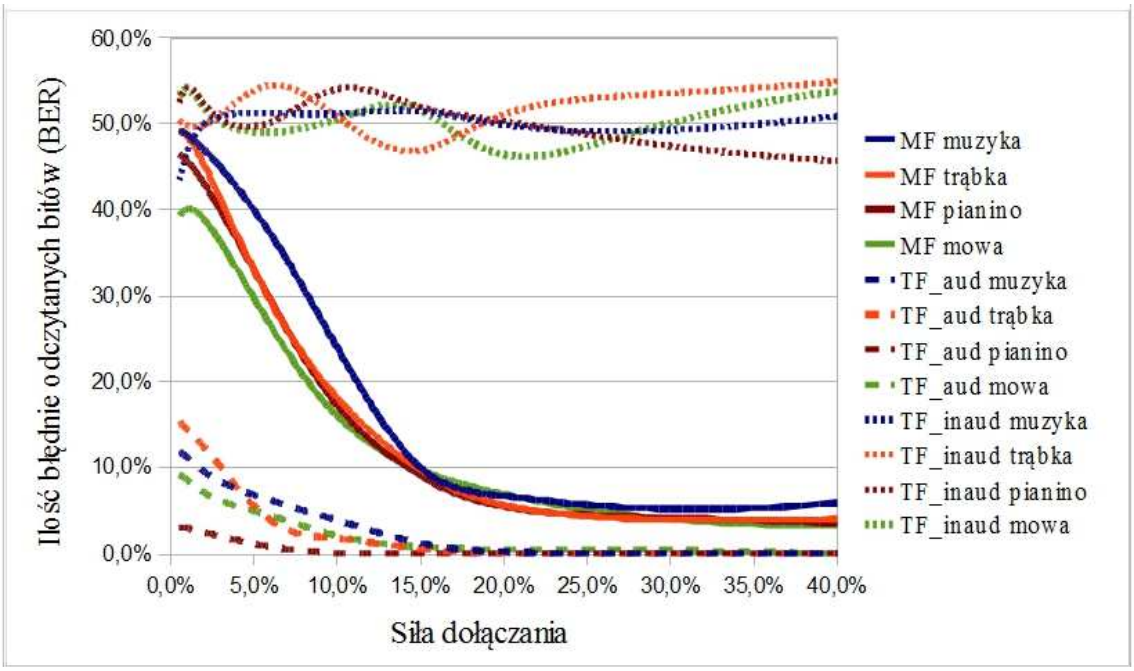
Analogiczne badanie przeprowadzono na tych samych nośnikach przy wykorzystaniu metod TF_{inaud} i TF_{aud}. Aby możliwe było porównanie odporności tych metod z odpornością metody MF zastosowano tu różne siły dołączania. Dobierane były one tak, by były równe średnim wartościom zmian wprowadzanych przez metodę MF. Należy jednak pamiętać, że siła dołączania jest parametrem określającym jedynie wartość wprowadzonych zmian i nie przekłada się na stopień wprowadzanych zniekształceń, ich słyszalność i łatwość wykrycia dołączonej informacji.

Na zamieszczonych poniżej wykresach umieszczone zostały wykresy odporności wszystkich porównywanych metod. Aby ułatwić analizę wykresów zastosowano konwencję oznaczeń, w której każdemu badanemu fragmentowi dźwięku przyporządkowano jeden określony kolor, natomiast każdej z metod przyporządkowano inny rodzaj linii: metodzie MF linię ciągłą, metodzie TF_{aud} przerywaną a TF_{inaud} wykropkowaną.

Określenie odporności na kompresję do formatu "aac" wykonano dla przepływności bitowych formatu "aac" wynoszących po skompresowaniu 128kbit/s i 64kbit/s. Uzyskane wyniki przedstawione zostały na rysunkach 4.4 oraz 4.5.



Rysunek 4.4. Odporność na kompresję do formatu "aac" o przepływności 128kbit/s [opracowanie własne]

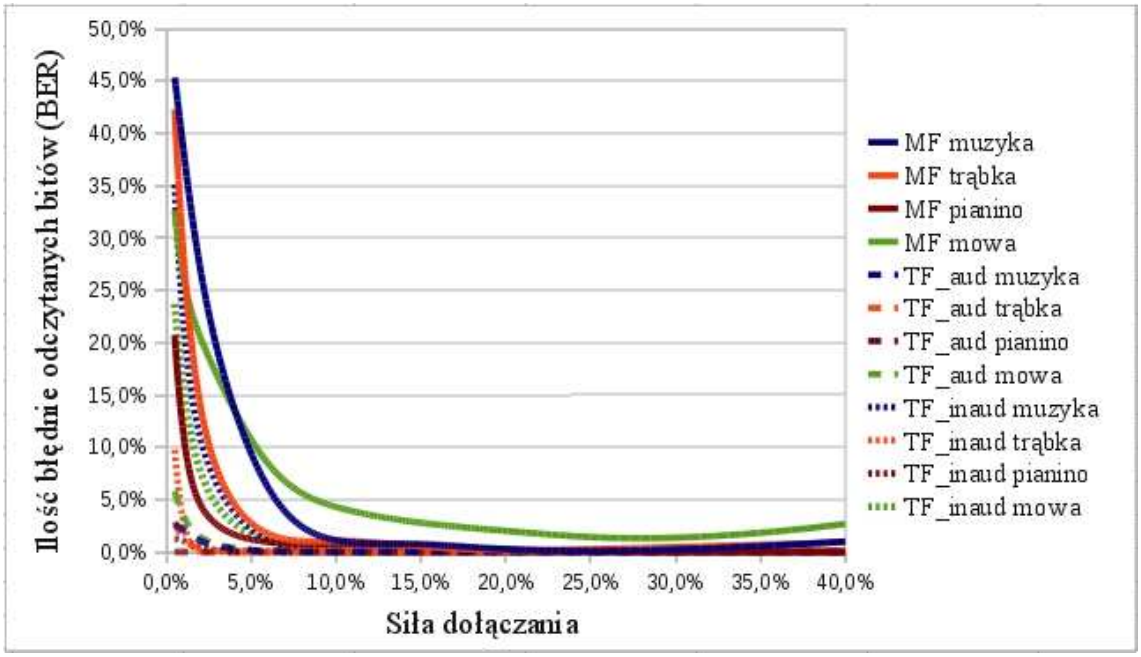


Rysunek 4.5. Odporność na kompresję do formatu "aac" o przepływności 64kbit/s [opracowanie własne]

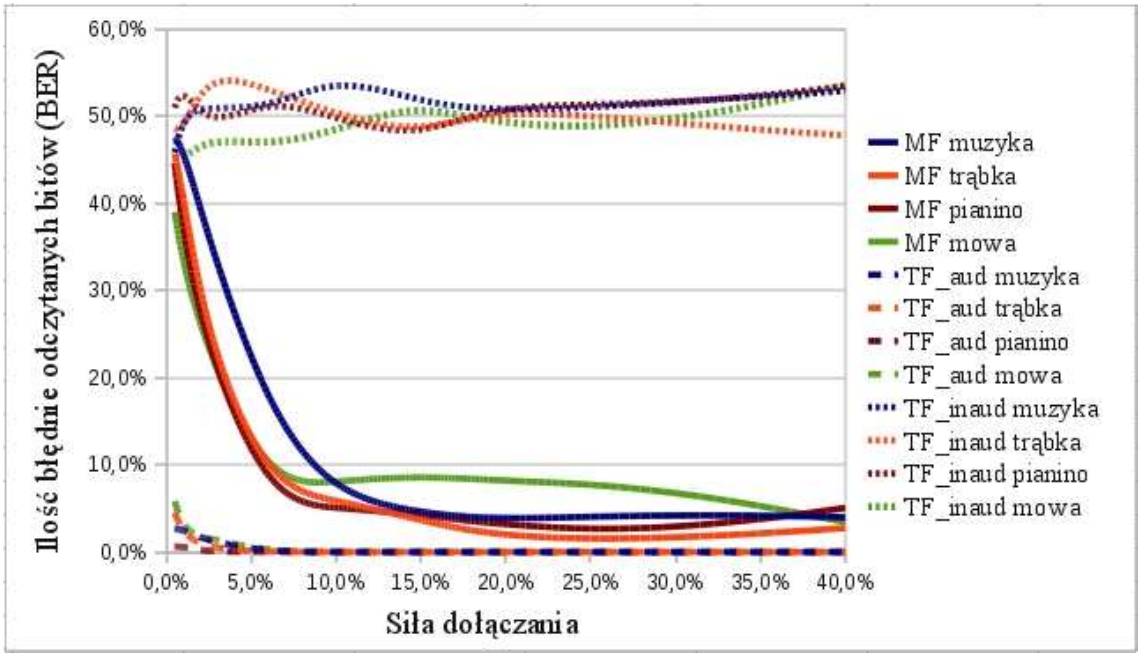
Jak widać na rysunkach 4.4 oraz 4.5 odporność na kompresję do formatu "aac" wzrasta wraz ze wzrostem siły dołączania R_p. Zjawisko to wynika z faktu, że kompresja wprowadza określony poziom zmian. Im większe zmiany wprowadzone zostały w nośniku, tym mniejsze prawdopodobieństwo, że zmiany wykonane podczas kompresji będą na tyle silne, by uszkodzić dołączoną informację. W miarę wzrostu siły dołączania zmniejsza się liczba błędów odczytu, aż do osiągnięcia pewnej ustalonej wartości. Łatwo zauważyć, że silniejsza kompresja dająca w wyniku sygnał o mniejszej przepływności bitowej powoduje większe uszkodzenie dołączonych danych. Jest to spowodowane utratą większej ilości informacji o nośniku będącej efektem silniejszej kompresji stratnej.

Porównanie z innymi metodami wykazuje, że metoda TF_{aud} posiada większy stopień odporności na kompresję "aac", co uwidacznia się zwłaszcza w przypadku silniejszej kompresji. Wyniki uzyskane dla metody TF_{inaud} pokazują jej brak odporności na ten format kompresji.

Określenie odporności na kompresję "mp3" również przeprowadzono dla dwóch przepływności bitowych 128kbi/s i 64kbit/s. Otrzymane wyniki przedstawione zostały na rysunkach 4.6 i 4.7.



Rysunek 4.6. Odporność na kompresję do formatu "mp3" o przepływności 128kbit/s [opracowanie własne]

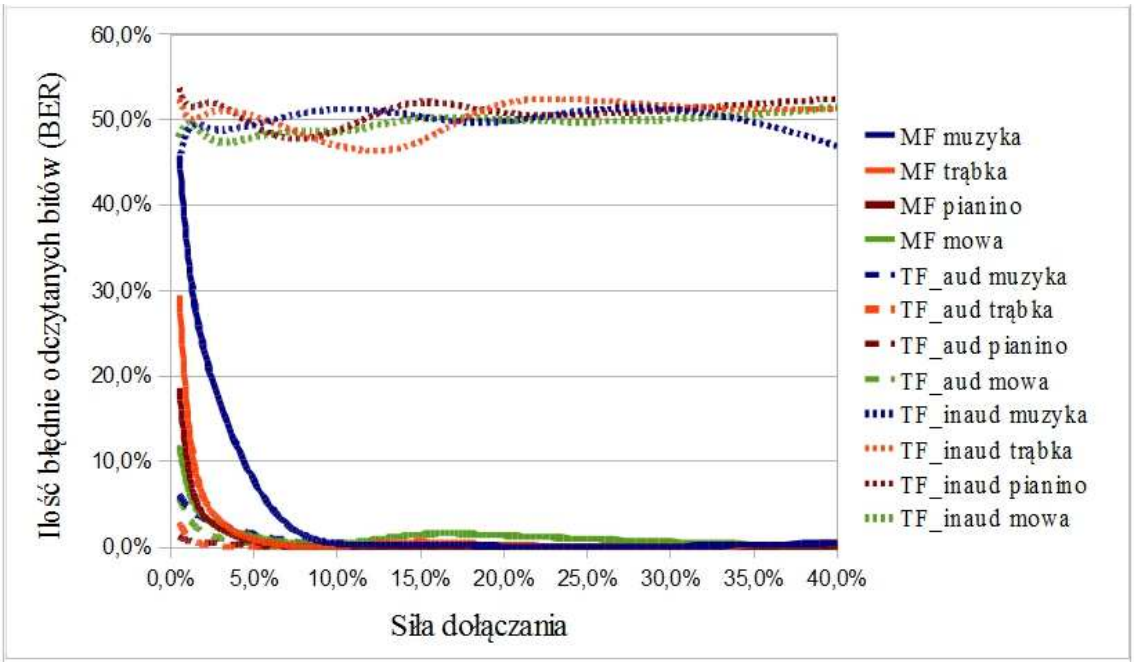


Rysunek 4.7. Odporność na kompresję do formatu "mp3" o przepływności 64kbit/s [opracowanie własne]

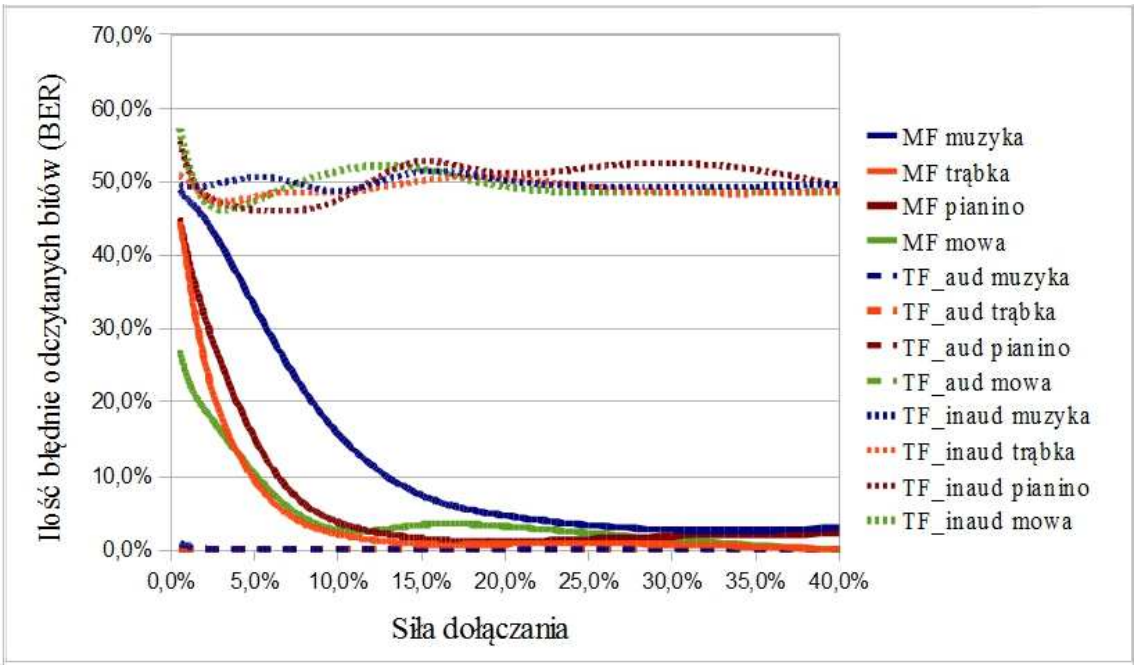
Badany algorytm wykazuje podobny stopień odporności na kompresję mp3 jak na kompresję do formatu "aac". Daje się zauważyć, że nieco mniejsza odporność dotyczy sygnałów mowy. W tym przypadku również zwiększenie odporności następuje w miarę wzrostu siły dołączania.

Analiza wyników uzyskanych dla pozostałych metod pokazuje, że metoda TF w porównaniu do MF wykazuje nieco lepszą odporność na kompresję dającą w wyniku sygnał o przepływności bitowej 128kbit/s. W przypadku silniejszej kompresji, do przepływności 64kbit/s informacja dołączana w paśmie niesłyszalnym jest usuwana co skutkuje brakiem odporności metody TF działającej w paśmie niesłyszalnym.

Kolejny test dotyczył odporności na kompresję do formatu "wma". Tak jak poprzednie wykonano go dla dwóch poziomów jakości nagrań. Wyniki przedstawione zostały na rysunkach 4.8 oraz 4.9.



Rysunek 4.8. Odporność na kompresję do formatu "wma" o przepływności 128kbit/s [opracowanie własne]

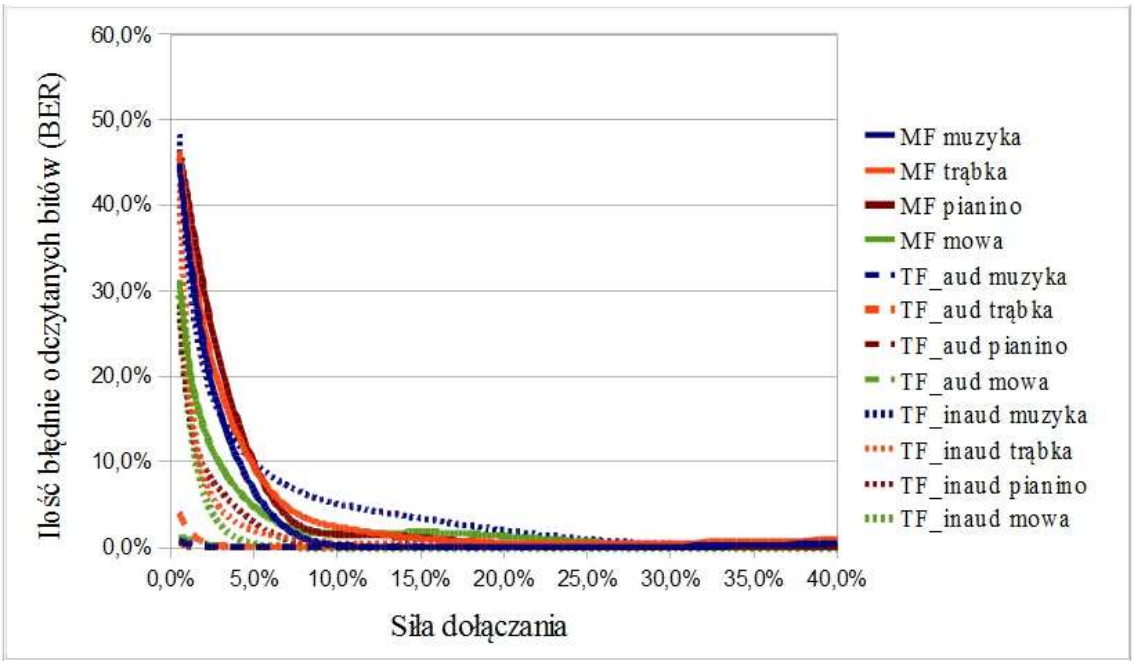


Rysunek 4.9. Odporność na kompresję do formatu "wma" o przepływności 64kbit/s [opracowanie własne]

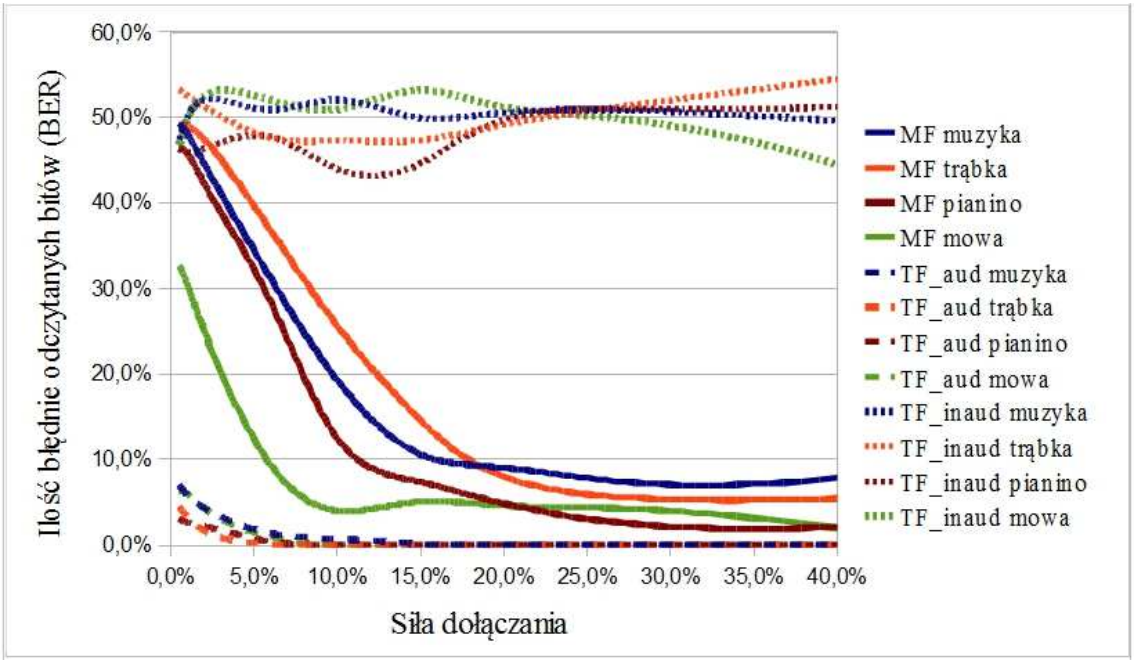
Analiza rysunków 4.8 i 4.9 pokazuje że w przypadku kompresji "wma" występują takie same zależności jak w przypadku analizowanych wcześniej formatów kompresji. Nieco gorsze wyniki dla metody MF uzyskane zostały dla sygnału muzyki.

Porównanie z innymi metodami pokazuje lepszą odporność metody TF_{aud} oraz brak odporności metody TF_{inaud}.

Kolejny test dotyczył określenia odporności na kompresję "ogg". Wyniki badań wykonanych dla dwóch stopni kompresji zaprezentowane zostały na rysunkach 4.10 i 4.11.



Rysunek 4.10. Odporność na kompresję do formatu "ogg" o przepływności 128kbit/s [opracowanie własne]



Rysunek 4.11. Odporność na kompresję do formatu "ogg" o przepływności 64kbit/s [opracowanie własne]

Analiza rysunków 4.10 i 4.11 pozwala zauważyć, że w przypadku kompresji sygnału do formatu "ogg" o przepływności 128kbit/s wszystkie badane metody osiągnęły porównywalne wyniki, z niewielką przewagą metody TF_{aud}. Zwiększenie stopnia kompresji ujawnia wyraźne dysproporcje pomiędzy nimi. Najlepsze wyniki uzyskuje metoda TF_{aud} natomiast metoda TF_{inaud} okazuje się być zupełnie nieodporna na kompresję "ogg" o przepływności 64kbit/s. W przypadku metody MF wyższy stopień kompresji spowodował wystąpienie większej ilości błędów, który dla wyższych sił dołączania R_p nie przekracza 10%.

Przeprowadzone badania wykazały, że metoda MF wykazuje dobrą odporność na badane rodzaje kompresji. Ponadto zależność odporności od siły dołączania i przepływności jest zbliżona dla wszystkich formatów. Przewaga proponowanej metody wyraźnie zarysowuje się w przypadku kompresji mp3, która okazała się destrukcyjna dla danych dołączonych za pomocą pozostałych metod. Metody bazujące na algorytmie LSB oraz program mp3stego nie zostały uwzględnione w badaniach ze względu na brak odporności na jakiekolwiek przekształcenia nośnika, co jednoznacznie świadczy o braku odporności na kompresję.

4.5. Porównanie metody autorskiej z innymi bazującymi na transformacie Fouriera

Aby ocenić skuteczność wykorzystania zalet transformaty Fouriera w opracowanym algorytmie, porównano go z innymi technikami steganograficznymi korzystającymi z tego samego przekształcenia, dla których znaleziono w literaturze opis uzyskanych rezultatów.

W [4] autorzy prezentują metodę znakowania wodnego utworów muzycznych. Pozwala ona na dołączanie ciągu bitów, zapewniając pojemność steganograficzną rzędu 1 bita na 1.2s nagrania próbkowanego z częstotliwością 44100Hz. W opracowaniu [4] autorzy prezentują uzyskane wyniki odporności na kompresję mp3 o przepływności bitowej 128kbit/s, filtrowanie pasmowo-przepustowe, tłumiące sygnał o -9dB dla częstotliwości mniejszych od 441Hz i większych od 4410Hz, zmianę częstotliwości próbkowania z 44100Hz na 22050Hz oraz zmianę rozdzielczości próbki z 16 na 8 bitów. Wyniki przedstawione przez autorów [4] zaprezentowane zostały w ostatniej kolumnie tabeli 3.7.

Aby porównać wyniki uzyskiwane przez metodę MF wykonano operacje opisane w [4] i określono BER. Jako że w badaniach użyto bloków o rozmiarze dwóch tysięcy próbek oraz bloków łączących o rozmiarze stu próbek, to w czasie 1,2s udało się ukryć ponad dwadzieścia pięć bitów danych podczas gdy metoda z [4] pozwalała na ukrycie w tym samym czasie tylko jednego bitu. W celu porównania odporności tych metod przy zbliżonej pojemności steganograficznej ukryto metodą MF powtórzony 25 krotnie ciąg bitów. Po odczytaniu danych obliczono średnią wartości powtórzonych bitów. Jeśli była ona mniejsza od 0,5 traktowano ją jako poprawnie odczytane binarne "zero". W pozostałych przypadkach uzyskaną wartość traktowano jako binarną "jedynkę". Dla tego ciągu bitów obliczono BER. Uzyskane wyniki zaprezentowane zostały w tabeli 4.7

Tabela 4.7 Porównanie odporności metody MF i prezentowanej w [4] [opracowanie własne na podstawie 4].

Przekształcenie	MF BER [%]	MF (25x powtórzenie) BER [%]	Arnold [4] BER [%]
mp3	4	**0**	**1,22**
filtrowanie	10,3	**0**	**1,47**
Zmiana częstotliwości próbkowania	0	**0**	**1,29**
Zmiana rozdzielczości bitowej próbki	0	**0**	**1,43**

Przeprowadzone badania wskazują, że metoda MF charakteryzuje się wyższą odpornością od metody znakowania wodnego prezentowanej w [4] przy wykorzystaniu takiej samej pojemności steganograficznej.

4.6. Praktyczna weryfikacja poziomu bezpieczeństwa oferowanego przez opracowaną metodę

Stegoanaliza kontenera stworzonego za pomocą nieznanej metody jest niezmiernie trudna do przeprowadzenia. Najczęściej sprowadza się do analizy cech statystycznych nośnika. Jeśli znacząco odbiegają one od normy to wówczas podejrzewa się, że dźwięk został zmodyfikowany i poddaje się go szczegółowej analizie w celu

stwierdzenia czy została do niego dołączona informacja oraz ewentualnego jej odczytania lub uszkodzenia.

Ze względu na trudność analitycznej oceny odporności metod steganograficznych na stegoanalizę postanowiono przeprowadzić praktyczną weryfikację odporności opracowanej metody. Wykonane zostało to poprzez umieszczenie w Internecie nośnika z ukrytymi danymi, zawierającymi adres e-mail. Do nośnika dołączony został opis zawierający informację, że wewnątrz została zawarta ukryta informacja.

Dodatkowo wiadomość o umieszczeniu tego nośnika została rozpropagowana na popularnych forach hakerskich takich jak hack.pl, www.binrev.com czy hak5.org. Umieszczono tam informację o lokalizacji pliku i jego ukrytej zawartości oraz zaproszono wszystkich zainteresowanych do prób złamania zabezpieczenia i powiadomienia autora o pomyślnym rezultacie. W celu weryfikacji czy metoda została naprawdę złamana poproszono uczestników testu o przesłanie informacji na adres email, który został ukryty wewnątrz udostępnionego nośnika.

Do momentu złożenia rozprawy, czyli przez ponad dwa miesiące od daty umieszczenia nośnika w Internecie informacja ta dotarła do kilkuset osób, jednak nikomu nie udało się odczytać ukrytej wiadomości. Tak długi okres ochrony świadczy o dużej sile i dobrej jakości ochrony oferowanej przez zaprojektowaną metodę steganograficzną.

4.7. Podsumowanie badań porównawczych

Badania porównawcze metody steganograficznej opartej o transformatę Fouriera i maskowanie oraz istniejącej wersji algorytmu bazującego na tym przekształceniu i programów mp3stego, h4pgp i s-tools wykazały, że opracowany algorytm dorównuje aplikacjom porównawczym a w niektórych przypadkach osiąga lepsze od nich wyniki. Pojemność steganograficzna metody MF jest co prawda najmniejsza, ale dzięki temu możliwe było uzyskanie wysokiego poziomu odporności na przekształcenia stegokontenera.

Badana metoda osiągnęła przewagę odporności nad metodami s-tools, h4pgp i mp3stego we wszystkich badanych rodzajach przekształceń. Porównanie z metodą TF w zakresie operacji kształtujących efekt przestrzenny, zmieniających dynamikę, czy dołączających szum wykazało, że najlepsze wyniki osiąga metoda TF ukrywająca

w paśmie niesłyszalnym. Jednak odbywa się to kosztem braku odporności na kompresję. Metoda TF_{aud} wykazuje uzyskała lepsze wyniki odporności na dołączanie szumu i filtrowanie, w przypadku operacji kształtujących efekt przestrzenny i dynamikę sygnału ustępowała metodzie MF.

Analiza odporności na kompresję wykazała, że z opracowaną metodą może się mierzyć tylko metoda TF_{aud}, która uzyskała we wszystkich badanych przypadkach lepsze wyniki.

Przewaga metody TF_{aud} uwidaczniająca się podczas analizy numerycznej nie przenosi się jednak na wyniki uzyskane w testach odsłuchowych. Metoda ta wprowadza wysoki poziom zakłóceń, praktycznie czyniący ją nieprzydatną w steganografii sygnałów dźwiękowych.

Przeprowadzone badania wykazują, że opracowana metoda umożliwia utworzenie stegokontenera, bez wprowadzania zakłóceń słyszalnych dla przeciętnego odbiorcy przy w sposób zapewniający wysoką odporność na różne przekształcenia dźwięku oraz kompresję. Poziom błędów wprowadzanych podczas różnych przekształceń jest na tyle niski, że zastosowanie kodów korekcji błędów powinno umożliwić bezbłędny odczyt ukrytych danych.

5. Wnioski

1. Przeprowadzona została analiza porównawcza teoretycznych metod cyfrowego przetwarzania sygnałów pod kątem oceny ich przydatności w steganografii sygnałów dźwiękowych. Posłużyła ona za podstawę do wybrania transformaty Fouriera jako przekształcenia bazowego nowej metody steganograficznej, która oferowała by wystarczającą pojemność steganograficzną do realizacji anonimowej komunikacji przy jednoczesnym zachowaniu przezroczystości i wysokiego poziomu odporności na uszkodzenie ukrytej informacji.
2. Zaproponowano ***nowe autorskie podejście do problemu steganografii komputerowej polegające na połączeniu zjawiska maskowania częstotliwościowego sygnału dźwiękowego z dyskretną transformatą Fouriera, które pozwoliło uzyskać skuteczną i efektywną metodę oraz algorytm ukrywania informacji w kontenerach dźwiękowych.***
3. Zastosowane podejście polega na odnalezieniu w widmie prążka zdolnego zamaskować zmiany wprowadzone w innych prążkach o zbliżonej częstotliwości (maskera) a następnie na określeniu skutecznie maskowanych prążków widma częstotliwościowego.

   Dodatkową zaletą tego rozwiązania jest uzależnienie położenia wykorzystywanych prążków widma częstotliwościowego sygnału od położenia maskera, skutkujące niezależnym wyborem prążków w każdym z fragmentów sygnału. Powoduje to rozproszenie dołączanej informacji w szerokim spektrum częstotliwości, znacząco utrudniając celowe uszkodzenie ukrytych danych (np. przez zastosowanie filtrowania) – zwłaszcza, że ***modyfikowane częstotliwości prawie zawsze znajdują się w paśmie słyszalnym.***
4. Zastosowanie zaproponowanej metody pozwala uzyskać stegokontener wykazujący wysoką odporność na ataki poprzez filtrowanie. Ze względu na dołączenie informacji do pasma słyszalnego oraz rozproszenie jej w szerokim zakresie częstotliwości tego pasma, uszkodzenie ukrytych danych niemożliwe jest bez wprowadzenia słyszalnych uszkodzeń sygnału. Ponadto usunięcie pewnych zakresów częstotliwości sygnału stanowi wskazówkę, że został przeprowadzony atak na stegokontener.

5. Wielkość wprowadzanych zmian wartości prążków widma uzależniona jest od wartości największego prążka w widmie. Pozwala to na uzyskanie optymalnego rozwiązania zarówno pod kątem odporności, jak i adaptacji wartości wprowadzanych zmian do amplitudy sygnału w danym fragmencie tego sygnału oraz na wykorzystanie wszystkich fragmentów do ukrycia niezależnie od ich głośności. Ponadto rozwiązanie to pozwala uzyskać wysoką odporność na przekształcenia dynamiczne sygnału, gdyż fragmenty przetwarzane są niezależnie a dodatkowa informacja ukryta jest w proporcjach pomiędzy prążkami widma. Zmiany w krótkim przedziale czasu są praktycznie jednakowe dla całego fragmentu więc nie modyfikują wykorzystywanych proporcji wartości prążków widma.
6. Wykonana praktyczna ocena zaproponowanej metody wykazuje duży wpływ niepoprawnego dobrania współczynników metody podczas odczytu na liczbę błędnie odczytanych bitów. Świadczy to o prawidłowym dobraniu składników klucza steganograficznego i konieczności ich znajomości do prawidłowego odczytania ukrytych danych.
7. Kolejnym nowym rezultatem pracy jest opracowanie metody konstrukcji bloków łączących stosowanych do scalania sygnału stegokontenera, co umożliwiło wyeliminowanie słyszalnych zakłóceń powstających na łączeniach bloków sygnału przenoszących ukryte dane.
8. Przeprowadzone rozważania teoretyczne oraz badania doświadczalne pozwalają na stwierdzenie, że cel pracy został osiągnięty.
9. Wykonane badania wykazały, że zmodyfikowane metody elektrotechniki teoretycznej a konkretnie cyfrowego przetwarzania sygnałów stanowią doskonałą podstawę do rozwoju teoretycznych i praktycznych metod steganografii komputerowej.

Teoretyczne podstawy metody przedstawionej w niniejszej rozprawie doktorskiej zostały zaprezentowane w autorskich publikacjach [49,50] znajdujących się na liście filadelfijskiej.

Za własne osiągnięcia autor uważa:

1. opracowanie nowej metody steganograficznej bazującej na połączeniu zjawiska maskowania częstotliwościowego z transformatą Fouriera, oraz opracowanie sposobu wyboru prążków widma sygnału do ukrycia informacji,

2. opracowanie metody adaptacji wartości wprowadzanych zmian do amplitudy sygnału, przy jednoczesnym zachowaniu wysokiego poziomu odporności i przezroczystości,
3. opracowanie sposobu łączenia zmodyfikowanych bloków stegokontenera,
4. określenie składników klucza steganograficznego, oraz zbadanie dopuszczalnej zmienności jego składników oraz wpływu ich błędnego określenia na poprawność odczytu ukrytych danych.

Kierunkami dalszych badań będą:

- prace nad zwiększaniem pojemności steganograficznej metody poprzez ukrywanie większej ilości bitów za pomocą dwóch prążków,
- adaptacja opracowanego algorytmu do wykorzystania w sygnałach ciągłych,
- zwiększanie odporności przez wykorzystanie większej ilości prążków do ukrycia jednego bitu, oraz matematycznych podstaw oceny odporności ukrytej informacji na uszkodzenie.

Literatura

1 Agaian S.S., Akopian D., Caglayan O., D'Souza S.A.: *"Lossless adaptive digital audio steganography"*, Proc. IEEE Int. Conf. Signals, Systems and Computers, pp. 903–906, 2005

2 Al-Haj A., Mohammad A.: *"Digital Audio Watermarking Based on the Discrete Wavelets Transform and Singular Value Decomposition"*, European Journal of Scientific Research Vol.39 No.1 (2010), pp. 6-21, 2010

3 Anderson R.: *"Stretching the Limits of Steganography"*, Lecture Notes of Computer Science pp. 39–48, Springer 1996

4 Arnold M.: *"Audio watermarking: Features, applications and algorithms"*, IEEE Int. Conf. Multimedia and Expo 2000, vol. 2, pp. 1013-1016 2000

5 Arttameeyanant P., Kumhom, P., Chamnongthai K.: *"Audio watermarking for Internet"*, IEEE International Conference on Information Technology: Coding and Computing, pp.237-242, 2001

6 Bassia P., Pitas I.: *"Robust audio watermarking in the time domain"*, Rhodes 1998

7 Bauer F. L.: *"Sekrety Kryptografii"*, Helion 2002

8 Bender W., Gruhl D., Morimoto N., Lu N.: *"Techniques for data hiding"*, IBM system Journal (5) 1996

9 Boney L., Tewfik A. H., Hamdy K. N.: *"Digital watermarks for audio signals"*, Proceedings of the 1996 International Conference on Multimedia Computing and Systems pp. 473-477, 1996

10 Cachin C.: *"An Information- Theoretic Model for Steganography"*, Information and Computation , 192(1) pp. 41-56, 2004

11 Chang L.: *"Issues in Information Hidding Transform Techniques"*, NRL/MR/5540-02-8621, Center for High Assurance Computer Systems (CHACS), Naval Research, 2002

12 Chen B., Wornell G. W.: *"Quantization index modulation: A class of provably dood methods for digital watermarking and information embedding of multimedia"*, Journal of VLSI Signal Processing Systems Volume 27 , Issue 1/2 (February 2001) pp. 7-33, 2001

13 Craver S.,Yeo B., Yeung M.: *"Technical trials and legal tributations"*, Communications of the ACM Volume 41 , Issue 7 (July 1998) pp. 45-54, 1998

14 Cvejic N.: *"Algorithms for audio watermarking and steganography"*, Oulu University Press 2004

15 Cvejic N., Seppanen T.: *"A wavelet domain LSB insertion algorithm for high capacity audio steganography"*, Proc. IEEE Digital Signal Processing Workshop, pp. 53–55, 2002

16 Cvejic N., Seppanen T.: "*Increasing robustness of LSB audio steganography using a novel embedding method*", Proc. IEEE Int. Conf. Info. Tech. Coding and Computing, Vol. 2, pp. 533–537, 2004

17 Cvejic N., Seppanen T.: *“Increasing the capacity of LSB-based audio steganography”*, IEEE Workshop on Multimedia Signal Processing, pp. 336–338, 2002

18 Czerwinski S., Fromm R.: *"Digital music distribution and audio watermarking"*, http://citeseerx.ist.psu.edu, Berkeley 1999

19 Cox I.,Kilian J.,Leighton T.,Shamon T.: *"A secure, robust watermark for multimedia"*, Information hiding, Lecture Notes of Computer Science(1147) pp. 185-206, Springer-Verlag, 1996

20 Czyżewski A., *"Dźwięk cyfrowy"*, EXIT, Warszawa 2001

21 Delforouzi A., Pooyan M.: *"Adaptive Digital Audio Steganography Based on Integer Wavelet Transform"*, Circuits Syst Signal Process Vol. 27 pp. 247–259, 2008

22 Dugelay J. L., Roche S.: *"A survey of current watermarking techniques"*, Information hiding: Techniques for steganography and digital watermarking pp. 121-148, 2000

23 Dymarski P.: *"Filtracja sygnałów dźwiękowych jako metoda znakowania wodnego i steganografii"*, Krajowe Sympozjum Telekomunikacji 2006 s.12, 2006

24 Dymarski P. , Pobłocki A., Baras C., Moreau N.: *"Algorytmy znakowania wodnego sygnałów dźwiękowych"*, Krajowe Sympozjum Telekomunikacji 2003 s. 26-34, 2003

25 Dymarski P.: *"Znakowanie wodne sygnałów dźwiękowych w pętli zamkniętej"*, XX Krajowe Sympozjum Telekomunikacji, 2004., Tom A. s. 221-231.

26 Gruhl, D., Lu, A.: *"Echo Hiding"*, Information Hiding Workshop pp. 295-315, 1996

27 Garay A.: *"Measuring and evaluating digital watermarks in audio files"*, master thesis 2002

28 Garbarczuk W., Kopniak P.: *"Steganologia: współczesne metody ochrony informacji"*, Pomiary, Automatyka, Kontrola 2005r. s. 21-26, 2005

29 Garbarczuk W., Kozieł G.: *"Review of sound-based steganographic techniques"*, ИССКУСТВЕННЫЙ ИНТЕЛЛЕКТ 4’2008 pp. 4-14, 2008

30 Garbarczuk V., Kozieł G.: *"The method of correlation between channels of audio record for steganographic purposes"*, Искуственный Интеллект 3’2006 s. 702-708, 2006

31 Garbarczuk W., Świć A.: *"Podstawy ochrony informacji"*, Politechnika Lubelska 2005

32 Gordy J., Burton L.: *"Performance evaluation of digital audio watermarking algorithms"*, Lecture Notes in Computer Science Volume 3261/2005, pp. 430-440, 2004

33 "*Herodotus: The Histories"*, J. M. Dont & Sons, Ltd, London, England, 1992.

34 Hsieh, C. T. Sou, P. Y.: *"Blind Cepstrum Domain Audio Watermarking Based on Time Energy Features"*, 14th IEEE International Conference on Digital Signal Processing, Vol. 2, pp. 705-708, 2002.

35 Huang J., Wang Y., Shi Y. Q., *"A Blind Audio Watermarking Algorithm with Self-synchronization"*, IEEE International Symposium on Circuits and Systems, Vol. 3, pp. 627-630, 2002.

36 Husrev T. Sencar, Ramkumar M., Akansu N.: *"Data Hiding Fundamentals and Applications"*, Elsevier 2004

37 Ikeda, M., Takeda, K., Itakura, F.*: "Audio Data Hiding by Use of Band-limited Random Sequences"*, ISASSP, Vol. 4, pp. 2315-2318, 1999.

38 Johnston J., Brandenburg K.: *"Wideband coding perceptual considerations for speech and music"*, Advances in Speech Signal Processing pp. 109-148, 1992

39 Johnson N. F., Katzenbeisser S. C.: *"A survey of steganographic techniques"*, Information hiding: Techniques for steganography and digital watermarking pp. 43-78, 2000

40 Jorasz U.: *"Selektywność układu słuchowego"*, Wydawnictwo UAM, 1999

41 Katzenbeisser S., Petitcolas F. A. P.: *"Information hiding techniques for steganography and digital watermarking"*, Artech House 2000

42 Kim S., Kwon H., Bae K.: *"Modification of polar echo kernel for performance improvement of audio watermarking"*, Lecture notes in computer science: international workshop on digital watermarking N^{o}2, Seoul , COREE, REPUBLIQUE DE (22/10/2003) 2004 , vol. 2939, pp. 456-466, 2004

43 Kirovski D.,Malvar H.: *"Robust covert communication over a public audio channel using spread spectrum"*, Lecture Notes In Computer Science; Vol. 2137 pp.354-368, 2001

44 Kopniak P.: *"Metody cyfrowego przetwarzania sygnałów na potrzeby steganologii komputerowej"*, rozprawa doktorska, Lublin 2007

45 Koukopoulos D., Stamatiou Y.: *"A watermarking scheme for MP3 audio files"*, International Journal of Signal Processing 2.3 pp. 206-214, 2005

46 Kozieł G.: *"Bezpieczeństwo szyfrowania"*, Systemy i sieci komputerowe s. 75-80, 2004

47 Kozieł G.: *"Cyfrowy zapis dźwięku w zadaniach steganograficznych"*, Bezpieczeństwo informacji s. 157-166, 2005

48 Kozieł G.: *"Cyfrowe znaki wodne w steganologii"*, Varia Informatica, Obliczenia i Technologie s. 229-236, 2005

49 Kozieł G.: *"Fourier transform and masking in sound steganography"*, Актуальні Проьлеми Економіки 12'2009, 2009

50 Kozieł G.: *"Method of concealing information in sound based on Fourier transform and masking"*, Polish Journal of Environmental Studies pp. 181-186, 2009

51 Kozieł G., Garbarczuk W.: *"Metody steganograficzne operujące w dziedzinie czasu odporne na ataki desynchronizujące"*, Varia Informatica s. 143-150, 2008

52 Kozieł G.: *"Steganografia w zabezpieczeniach nośników multimedialnych"*, Varia Informatica- technologie i bezpieczeństwo s. 193-200, 2006

53 Kozieł G.: *"Using rhythm shifts in steganography"*, Polish Journal of Environmental Studies. Vol. 17, No. 3B pp. 177-182, 2008

54 Kozieł G.: *"Using steganographical methods in multimedial carrier"*, Актуальні Проьлеми Економіки 2'2005, 2005

55 Kozieł G.: *"Zastosowanie Sieci Neuronowych w Steganografii Dźwięku"*, Informatyka stosowana, Planowanie s.89-94, 2007

56 Lee S. K., Ho, Y. S.: *"Digital Audio Watermarking in the Cepstrum Domain"*, IEEE Transactions on Consumer Electronics, Vol. 46, No. 3, pp. 744-750, 2000

57 Li W., Li X., Lu P.: *"A Novel Feature-based Robust Audio Watermarking for Copyright Protection"*, IEEE International Conference on Information Technology: Coding and Computing, pp. 554-558, 2003

58 Li W., Xue X.: *"Audio Watermarking Based on Music Content Analysis: Robust against Time Scale Modification"*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) pp. 289-300, 2004

59 Li W., Xue X., Lu P.: *"Localized audio watermarking technique robust against time-scale modifications"*, IEEE Transactions on Vol. 8, Issue 1 pp. 60 – 69, 2006

60 Liu K. C. , Tsai S. E., Yang S. M.: *"Development of an Effective Watermark Based on Energy Balance of DCT for Audio Signals"*, Proceedings of the 2005 Workshop on Consumer Electronics and Signal Processing, 2005

61 Mahbour R., Syed M.: *"Multimedia Technologies: concepts, methodologies, tools, and applications"*, London 2008

62 Mansour M., Tewfik A.: "*Audio Watermarking by Time-Scale Modification"*, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings 3, pp. 1353-1356, 2001

63 Nakayama A., Lu J.,Nakamura S., Sikano K.: *"Digital watermarks for audio signal based on psychoacoustic masking model"*, Electronics and Communications in Japan (Part III: Fundamental Electronic Science) Vol. 86 Issue 12 pp. 65 – 75, 2003

64 Petitcolas F. A. P., Ross J., Kuhn G.: *"Information Hiding—A Survey"*, Proceedings of the IEEE, special issue on protection of multimedia content 87(7) pp. 1062–1078, 1999

65 Petitcolas F., Steinbach M., Raynal F., Dittmann J., Fontaine C.: *"A public automated web-based evaluation service for watermarking schemes: StirMark Benchmark"*, Proc. SPIE Vol. 4314, pp. 575-584, Security and Watermarking of Multimedia Contents III, 2001

66 Petrovic R., Winograd J., Jemili K., Metois E.: *"Data hiding within audio signals"*, Facta Universitatis Series: Electronics and Energetics vol. 12 no. 2 pp. 103-122, 1999

67 Pfitzmann B.: *"Information Hiding Terminology"*, Proc. Information Hiding Workshop, LNCS, Springer-Verlag pp. 347 - 350, 1996

68 Pobłocki A.: *"Cyfrowe znakowanie wodne sygnałów dźwiękowych z wykorzystaniem echa"*, praca dyplomowa 2003

69 Ramkumar M., Akansu A.N., Alatan A.: *"On the choice of Transforms for Data Hiding in Compressed Video"*, Proceedings of the Acoustics, Speech, and Signal Processing pp. 3049-3052, 1999

70 Rivas E.: *"Fourier phase domain steganography: Phase bin encoding via interpolation"*, Mobile multimedia/image processing for military and security applications. Conference, Orlando FL , ETATS-UNIS (2007) vol. 6579, pp. 65790W.1-65790W.6, 2007

71 Santosa, R.A. Bao, P.: *"Audio-to-image wavelet transform based audio steganography"* , ELMAR 2005. 47th International Symposium pp. 209-212, 2005

72 Seok, J. Hong, Kim J.: *"A Novel Audio Watermarking Algorithm for Copyright Protection of Digital Audio"* , ETRI Journal, vol.24, no.3 pp.181-189, 2002

73 Shine K.P., Krishna Kumar S.: *"Extended Bipolar Echo Kernel for Audio Watermarking"*, Artcom 2009, International Conference on Advances in Recent Technologies in Communication and Computing, IEEE Computer Society 2009 pp. 487-489, 2009

74 Simmons G. J.: *"The prisoners' problem and the subliminal channel, Advances in Cryptology"*, Proceedings of Crypto 83, pp. 51-67, 1984

75 Tomaszewski M.: *"Analiza algorytmów steganograficznych"*, praca dyplomowa, Białystok 2006

76 Wang H., Wang S.: *"Cyber Warfare: Steganography vs. Steganoanalisys"*, Communications of the ACM, vol. 47 No. 10 2004 pp. 76-82, 2004

77 Wang X., Zhao H.: *"A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT"*, IEEE Transactions on Signal Processing, Vol. 54, Issue 12, 2006, pp. 4835-4840, 2006

78 Wayner P.: *"Disappearing cryptography Information hiding: steganography & watermarking"*, Morgan Kaufmann Publisher 2002

79 Xiang S., Huang J.: *"Robust Audio Watermarking Against the D/A and A/D Conversions"*, http://arxiv.org/abs/0707.0397, 2007

80 Xiang S., Huang J., Yang R.: *"Time-Scale Invariant Audio Watermarking Based on the Statistical Features in Time Domain"*, Artificial Intelligence and Lecture Notes in Bioinformatics 2007 pp. 93-108, 2007

81 Xiang S., Kim H., Huang J.: *"Audio watermarking robust against time-scale modification and MP3 compression"*, Signal Processing Volume 88 , Issue 10 pp. 2372-2387 , 2008

82 Yi-Wen L., Smith J. O.: *"Audio watermarking through deterministic plus stochastic signal decomposition"*, EURASIP Journal on Information Security Volume 2007 Article No.: 4, 2007

83 Zhi-jun W., Wei Y., Yi-xian Y.: *"ABS-based speech information hiding approach"*, Electronics Letters 2003 Vol. 39 No. 22

84 Zieliński T. P.: *"Przetwarzanie sygnałów cyfrowych"*, Wydawnictwa Komunikacji i Łączności, 2005

Źródła internetowe

85 A detailed look at steganographic techniques and their use in an Open-Systems Environment, http://www.sans.org/

86 Dokumentacja programu h4pgp, http://www.heinz-repp.onlinehome.de/index.html

87 http://www.petitcolas.net/fabien/steganography/mp3stego/

88 http://pl.wikipedia.org/wiki/Transformacja_Fouriera

89 http://pl.wikipedia.org/wiki/Bit_Error_Rate

90 Opracowanie Katedry Systemów Multimedialnych Politechniki Gdańskiej: "Perceptualne skale częstotliwości", http://sound.eti.pg.gda.pl/student/materialy.html

91 Rude T. , "Steganography", www.crazytrain.com/stegclassexam.doc, 2000

92 SDMI Call For Proposals, Phase II, 2000. http://www.sdmi.org/download/.

93 www.trl.ibm.com/projects/RightsManagement/datahiding/dhstep_e.htm.

Spis rysunków

Spis tabel