dr hab. inż. Przemysław Dymarski, prof. PW
05520 Konstancin - Jeziorna
ul. Bielawska 31

3 grudnia 2020

Recenzja
rozprawy doktorskiej mgr inż. Jarosława Wojtunia
pt. "Efektywne metody skrytej synchronizacji akustycznych kanałów steganograficznych"
sporządzona dla Rady Dyscypliny Naukowej "Informatyka Techniczna i Telekomunikacja" WAT

1. Cel badań i teza rozprawy

Steganografia oraz znakowanie wodne dźwięku i obrazu stały się przedmiotem intensywnych badań naukowych w ostatnich dwóch dekadach. Stało się tak dzięki upowszechnieniu się Internetu i dzięki postępowi w dziedzinie cyfrowego przetwarzania sygnałów. W zagadnieniu steganografii zwraca się uwagę na ukrycie przekazywanej informacji w sygnale nośnika oraz na jej odporność na przetwarzanie tego sygnału (kompresja, filtracja).

Do przesłania ukrytej informacji, zakodowaniej w strumieniu binarnym, niezbędna jest synchronizacja elementowa i synchronizacja ramki. Sygnał synchronizacji powinien być ukryty w sygnale nośnika, niewykrywalny prostymi metodami steganalizy, odporny na przetwarzanie sygnału nośnika. Podstawowym problemem jest tu przepróbkowanie sygnału, wynikające z dryftu częstotliwości próbkowania w urządzeniu odbiorczym. Badaniem metod skrytej synchronizacji poświęcona jest praca mgr inż. Jarosława Wojtunia. Autor zajmuje się ukrywaniem strumienia binarnego w sygnale mowy o paśmie 4 kHz. Sygnał mowy z ukrytą informacją steganograficzną i sygnałem synchronizacji jest następnie transmitowany z wykorzystaniem kanału VoIP (wykorzystano 3 różne kodeki) lub bezprzewodowo z wykorzystaniem radiostacji pracujących w trybie analogowym lub cyfrowym. Pozwala to na zbadanie odporności transmisji steganograficznej i algorytmów skrytej synchronizacji na róznego rodzaju zakłócenia, powstające w procesie kodowania sygnału mowy i jego transmisji.

Celem pracy jest zaproponowanie szeregu algorytmów skrytej synchronizacji i porównanie tych algorytmów pod kątem ich skuteczności, odporności na zakłócenia i wpływu na jakość sygnału mowy z przekazem steganograficznym. Cel uznaje się za osiągnięty w przypadku znalezienia algorytmu synchronizacji umożliwiającego prawidłowe działanie łącza steganograficznego w warunkach zakłóceń i nieobniżającego w znacznym stopniu jakości sygnału mowy. Cel ten jest jasny, chociaż jego sformułowanie w formie tez jest nieco mniej przejrzyste. Teza I (skryta synchronizacja zwiększa efektywność transmisji danych steganograficznych) może być zrozumiana jako oczywistość, gdyż bez synchronizacji transmisja steganograficzna byłaby niemożliwa. Należało raczej napisać, że proponowane w rozprawie algorytmy syn-

chronizacji zapewniają efektywną transmisję danych steganograficznych. Teza II nie budzi wątpliwości - niniejsza rozprawa zawiera jej dowód. Reasumując, cel pracy jest właściwie postawiony i określa zadania badawcze na poziomie pracy doktorskiej.

2. Charakter rozprawy - teoretyczny, doświadczalny, konstrukcyjny

Praca ma charakter teoretyczny i doświadczalny. Warstwa teoretyczna obejmuje propozycję czterech algorytmów skrytej synchronizacji strumienia danych steganograficznych:
MPC (Monotonic Phase Correction) - Jest to metoda oparta na transmisji szeregu tonów pilotowych,
DSS: (Direct Spread Spectrum) - Metoda ta wykorzystuje tony pilotowe i sygnał DSS oparty na ciągu Golda,
PID: (Pattern Insertion Detection) - Polega na wstawieniu znacznika synchronizującego, który jest wzorcem wytłumienia sygnału mowy,
MES: (Minimal Error Synchronization) - Jest to metoda autosynchronizacji, niewymagająca modyfikacji sygnału mowy z przekazem steganograficznym.
Warstwa doświadczalna polega na wykonaniu szeregu symulacji i testów skuteczności transmisji oraz wpływu sygnałów synchronizujących na jakość sygnału mowy. Należy tu podkreślić zastosowanie metod obiektywnych (PESQ) i subiektywnych (odsłuchy) badania jakości mowy.
Charakter doświadczalno - konstrukcyjny ma implementacja algorytmów skrytej synchronizacji wraz z kanałem steganograficznym w systemie VoIP i w sytemie transmisji bezprzewodowej. Pozwoliło to na praktyczną weryfikację wyników otrzymanych metodami symulacyjnymi.

3. Sposób przeprowadzenia analizy źródeł (w tym literatury światowej i stanu zagadnień w przemyśle). Sposób sformułowania wniosków z analizy

Bibliografia zgromadzona przez Autora liczy 104 pozycje - połowa z tej liczby to publikacje z lat 2010-2020. Świadczy to o dobrym rozeznaniu autora w dziedzinie steganografii i jego przygotowaniu do realizacji zadań badawczych w tym zakresie. Należy też zwrócić uwagę na dorobek publikacyjny autora, na który składa się 11 publikacji (plus 2 w recenzji).

Autor przedstawia przegląd źródeł w rozdz.2. Wyróżnia trzy grupy metod steganografii, w zależności od miejsca osadzenia i detekcji skrytej informacji (Rys.2.1 - 2.3). Wg tych kryteriów uporządkowano materiał w p. 2.3, a w p. 2.4 uszeregowano publikacje pod względem zasady działania opisywanych algotytmów. W przeglądzie metod steganograficznych uwzględniono najważniejsze tendencje w tej dziedzinie. Drobnym niedopatrzeniem jest pominięcie metody QIM (Quantization Index Modulation). Autor cytuje publikację [55], zawierającą ten skrót w tytule, ale czyni to w innym kontekście. Są to jednak szczegóły, generalnie sposób

przeprowadzenia analizy źródeł uznaję za prawidłowy.

4. Rozwiązanie postawionego zadania - właściwość przyjętych metod i założeń

Jako algorytm referencyjny transmisji steganograficznej autor zastosował modyfikację widma sygnału akustycznego w podpasmach częstotliwości. Algorytm ten, opisany w [49], [51] i [68], został przez niego zmodyfikowany, w szczególności zastosował on próg maskowania o regulowanym poziomie. Algorytm toleruje odstrojenie częstotliwości próbkowania o kilka Hz i przesunięcie ramki o 3 ms, co stanowi duże wyzwanie dla konstruktora algorytmu synchronizacji.

W p. 4.1 autor proponuje algorytm synchronizacji MPC, oparty na wstawieniu tonów pilotowych, generowanych w układzie OFDM. Sygnał OFDM, zawierający 14 harmonicznych, wstawia się na częstotliwościach poniżej 500 Hz i powyżej 3000 Hz. 6 tonów zapewnia korekcję odstrojenia częstotliwości próbkowania a 8 tonów zapewnia synchronizację elementową (szczelina czasowa wynosi 48 ms). Korekcja częstotliwości próbkowania, oparta na pomiarze odchylenia kąta fazowego dla kolejnych szczelin czasowych, wskazuje jednoznacznie wartość odstrojenia (Rys.4.5). Z kolei metoda synchronizacji elementowej, oparta na porównaniu skumulowanej fazy z fazą zmierzoną, daje wyniki mniej jednoznaczne (Rys. 4.8 - 4.10). Ewentualny błąd pozostaje jednak w dopuszczalnym zakresie 3 ms. Nie jest też wyjaśnione w tekście rozprawy, dlaczego tony wykorzystywane do korekcji częstotliwości próbkowania nie mogą być jednocześnie wykorzystane do synchronizacji elementowej.

W p.4.2 Autor połączył algorytm korekcji częstotliwości próbkowania, zastosowany w metodzie MPC, z algorytmem synchronizacji elementowej, opartym na sygnale o rozproszonym widmie (DSS). Sygnał z rozproszonym widmem, składający sie z ponad 6000 symboli o czasie trwania 1 ms, generowanych z wykorzystaniem ciągu Golda, zapewnia synchronizację elementową z dokładnością wystarczającą do niezawodnego działania referencyjnego systemu steganograficznego. Korelacyjna metoda detekcji sygnału DSS wskazuje jednoznacznie pozycję symboli i ramek wykorzystywanych w transmisji danych steganograficznych (Rys.4.17).

Inną metodę synchronizacji symboli i 16-bitowych ramek opisano w p. 4.3. Polega ona na wstawieniu znacznika synchronizującego przed ciągiem symboli wykorzystywanych w transmisji steganograficznej (metoda PID). W obrębie znacznika sygnał mowy ulega wytłumieniu, co nie musi prowadzić do znacznego obniżenia jakości mowy ze względu na stosowany w odbiorniku algorytm PLC (Packet Loss Concealment). Krótka ramka nie zmusza do korekcji częstotliwości próbkowania, co zmniejsza złożonośc obliczeniową algorytmu synchronizacji.

Metoda MES, opisana w p.4.4, polega na autosynchronizacji i nie wymaga wstawiania sygnału synchronizacyjnego. Autosynchronizacja polega na wielokrotnym uruchomianiu detektora symboli steganograficznych, z przesunięciem o kilka próbek sygnału nośnika. W stanie synchronizacji amplitudy wykrytych symboli powinny być największe. Synchronizację ramki zapewniono kodując ciąg bitów steganograficznych kodem nadmiarowym (zastosowano kod BCH). W stanie synchronizacji powinna nastąpić detekcja bitów steganograficznych i korekcja błędów, o czym świadczy nieujemna wartość syndromu. Proponowana metoda zapewnia synchronizację elementową i synchronizację ramki. Nie jest jasne, w jaki sposób rozwiązano

problem korekcji częstotliwości próbkowania, która wydaje się być konieczna przy długiej ramce liczącej 127 bitów.

Reasumując, każda z czterech proponowanych metod jest prawidłowym rozwiązaniem problemu skrytej synchronizacji, choć pewne szczegóły wymagają wyjaśnienia.

Dowód tez rozprawy wymaga wykazania, że sygnał synchronizacji nie obniża w znacznym stopniu jakości mowy, a jednocześnie zapewnia niezawodną detekcję ukrytego strumienia binarnego. Badania porównawcze czterech metod skrytej synchronizacji przedstawiono w rozdziale 5.

Badania algorytmu MPC wykazały zmniejszenie MOS (pomiar metodą PESQ) o 0.07 po dodaniu sygnału synchronizującego, co oznacza znikome pogorszenie jakości sygnału mowy. Największe znaczenie mają jednak przeprowadzone badania odsłuchowe (pomiar współczynnika SDG). Wykazały one zauważalne, lecz niedokuczliwe pogorszenie jakości sygnału mowy. Badanie skuteczności synchronizacji wykazalo, że transmisja trwająca 6 s daje się synchronizować z błędem nieprzekraczającym dopuszczalnej wartości 3 ms. Badania jakości metodą obiektywną (MOS) i subiektywną (SDG) stanowią dowod tez rozprawy w odniesieniu do metody MPC. Badania SNR (Rys.5.4 - 5.6) nie wnoszą, moim zdaniem, nowych argumentów, gdyż sygnały steganograficzne wraz z synchronizacją są wstawiane metodą "psychoakustyczną" z uwzględnieniem maskowania, jakość mowy nie powinna być zatem oceniana z wykorzystaniem SNR, bez uwzględniania psychoakustyki.

Badania algorytmu DSS wykazały większą skuteczność synchronizacji (identyfikacja położenia symboli z dokładnością co do próbki), jednak kosztem obniżenia jakości mowy. To obniżenie jakości zostało określone w badaniach odsłuchowych jako niedokuczliwe.

Metoda PID wykazała się dobrą skutecznością synchronizacji (błąd lokalizacji symboli mniejszy niż 10 próbek), przy tym spadek jakości mowy, mierzony metodą subiektywną (SDG) okazał sie znikomy (zależny jednak od przetwarzanej frazy mowy). Należy jednak zauważyć, że nie uwzględniono wpływu kanału transmisyjnego, np. szum mógłby zakłócić identyfikację znaczników synchronizacyjnych.

Metoda autosynchronizacji w wykorzystaniem kodu BCH (algorytm MES) nie wykorzystuje dodatkowego sygnału synchronizacji, jednak sygnał transmisji steganograficznej powinien mieć wystarczającą moc, aby zapewnić synchronizację. Współczynnik SDG w odsłuchowych badaniach jakościowych wynosił w tych warunkach od -0.5 do -0.8, co oznacza niedokuczliwe obniżenie jakości mowy.

Ostateczną weryfikację poprawności zaproponowanych algorytmów skrytej synchronizacji stanowi ich implementacja wraz z referencyjnym algorytmem steganograficznym w systemie VoIP i w sytemie transmisji bezprzewodowej. Pozwoliło to na praktyczną weryfikację wyników otrzymanych metodami symulacyjnymi. W transmisji bezprzewodowej wykorzystano radiostacje UKF pracujące w trybie analogowym lub cyfrowym (kodek CVSD16). Uzyskano pozytywne wyniki testu dla wszystkich metod z wyjątkiem PID, gdyż w obecności szumu nie udawało się zlokalizować znacznika. W kanałach VoIP wyniki zależały od kodeka i rodzaniu sieci. Dla kodeka PCMA w każdych warunkach usyskiwano poprawną transmisję steganograficzną. Kodeki Speex i G.729 nie współpracowały z algorytmem PID. Pozostałe

metody skrytej synchronizacji zapewniały prawidłowy przebieg transmisji steganograficznej, co świadczy o osiągnięciu celu rozprawy.

5. Oryginalność rozprawy - samodzielny dorobek autora - pozycja rozprawy w stosunku do stanu wiedzy i poziomu techniki prezentowanego w literaturze światowej

Steganografia zaczęła się burzliwie rozwijać w ostatnich latach ubiegłego stulecia. Od tego czasu pojawiło się wiele prac, jednak dziedzina ta jest wciąż w fazie rozwoju. Praca mgr inż. Jarosława Wojtunia wypełnia tu pewną lukę, gdyż niewiele jest prac poświęconych wyłącznie problemowi skrytej synchronizacji. Autorzy prac proponujący algorytmy steganografii często nawet pomijają problem synchronizacji, zakładając, że jest ona zapewniona.

Proponowane przez mgr inż. Jarosława Wojtunia metody synchronizacji zawierają elementy nowości. W metodzie MPC jest nim algorytm synchronizacji elementowej, opaty na pomiarze skumulowanej fazy, w metodzie DSS łączne wykorzystanie sygnałów DSS i OFDM. Wykorzystanie znaczników symulujących utracone pakiety, w celu aktywacji algorytmu PLC po stronie odbiorczej, jest też pomysłem oryginalnym. Idea autosynchronizacji pojawia się w publikacjach, jednak algorytmu wykorzystującego kod BCH z kontrolą syndromu nie napotkałem w literaturze.

Praca stanowi też osiągnięcie inżynierskie, udana implementacja skrytej synchronizacji w praktycznych systemach steganograficznej łączności przewodowej i bezprzewodowej jest tego dowodem.

6. Poprawność przedstawienia uzyskanych wyników - zwięzłość, jasność, umiejętność przekonywania, poprawność redakcyjna

Ogólna struktura rozprawy jest przejrzysta - ma w zamyśle prowadzić do dowodu tez rozprawy. Opis proponowanych metod skrytej synchronizacji (rozdział 4) jest poprawny, choć pojawiają się skróty w wywodzie, jak np. nazbyt zwięzły opis działania skanera kąta fazowego. Z kolei w rozdziale 5 (badania porównawcze metod synchronizacji) występuje pewien nadmiar informacji, np. wykresy wartości SNR mogłyby być pominięte, bo najważniejsze są tu wartości MOS uzyskane metodą PESQ i wyniki badań odsłuchowych. Praca jest poprawna pod względem językowym, usterki redakcyjne są nieliczne, np. w bibliografii pojawia się praca C.Nedeljko [9] i N.Cvejica [29,30], a chodzi o tę samą osobę.

7. Słabe strony rozprawy, jej główne wady

Praca nie ma wad, które mogłyby w znacznym stopniu obniżyć jej wartość. Powyżej przedstawiłem pewne zastrzeżenia, jak trochę niezręcznie sformułowana Teza I i nadmiernie skrócony opis metody MPC. Występuje niejednolita notacja we wzorach 4.2 i 4.4. W tekście (str.18) pojawia się kontrowersyjne stwierdzenie "Dla metod obietywnych jako próg znacznego pogorszenia jakości przyjęto obniżenie średniej o oceny o 10% a dla metod subiektywnych jako ten próg przyjęto dwukrotne obniżenie średniej oceny". Na str.90 jako "niesłyszalne" znie-

kształcenia kwalifikuje się wartości SDG ponad (-1), a przecież chodzi o subiektywną reakcję słuchaczy: jeśli nie słyszą zniekształceń to wystawiają SDG=0. Wątpliwości też budzą zerowe wartości prawdopodobieństwa błędu pierwszego rodzaju w Tab.4.3.

Są to jednak drobne niedociągnięcia nieobniżające wartości rozprawy.

8. Przydatność rozprawy dla nauk technicznych, przemysłu, obronności kraju itp.

Ze względu na rozwój Internetu, steganografia i znakowanie wodne rozwijają się stale, wciąż pojawiają się nowe algorytmy wstawiania tajnej informacji i jej detekcji, a także steganalizy. Każda nowa propozycja jest w tej sytuacji cenna. Proponowane przez autora algorytmy skrytej synchronizacji mogą być wykorzystane w systemach steganograficznych i znakowania wodnego sygnałów akustycznych. Implementacja w działających w czasie rzeczywistym łączach steganografii bezprzewodowej i przewodowej wykazała ich przydatność. Niektóre zastosowania wymagałyby zbadania odporności sygnałów skrytej synchronizacji na steganalizę.

9. Zaliczenie rozprawy do jednej z kategorii

Moim zdaniem pracę można zakwalifikować jako spełniającą wymagania.

P.Dymarski