

Mirosław ŁAZORYSZCZAK

POLITECHNIKA SZCZECIŃSKA, WYDZIAŁ INFORMATYKI

Ukrywanie informacji w środowisku dźwiękowym z wykorzystaniem modyfikacji zawartości muzycznej

Dr inż. Mirosław ŁAZORYSZCZAK

Ukończył studia na Wydziale Elektrycznym Politechniki Szczecińskiej, obronił pracę doktorską w 2005 r. Jest adiunktem w Instytucie Architektury Komputerów i Telekomunikacji. Zainteresowania naukowe koncentrują się przede wszystkim wokół zagadnień związanych z umieszczaniem znaków wodnych w środowisku dźwiękowym, ale także dotyczą cyfrowych systemów i układów rekonfigurowalnych.



e-mail: mlazoryszczak@wi.ps.pl

Streszczenie

W artykule przedstawiono propozycję wprowadzania znaków wodnych w środowisku dźwiękowym, które wykorzystuje analizę kontekstu muzycznego. Idea kontekstowego znaku wodnego w tym przypadku polega na wprowadzaniu informacji do oryginalnego przebiegu dźwiękowego w sposób z jednej strony niesłyszalny, a z drugiej strony modyfikujący podstawowe nośniki „treści” muzycznej. W artykule opisano sposoby modyfikacji wybranych elementów składowych muzyki.

Słowa kluczowe: znak wodny, analiza treści**Data hiding in the audio environment using analysis of the content of music****Abstract**

In this paper the kind of the digital watermark embedding into an audio environment is presented. The watermark is based on analysis of music content. The main idea of the context digital watermark consists on data embedding into the source audio signal. From the one hand the watermarked signal should be inaudible, from the other hand it should change the main parts of music content. The modifications of the time duration of the sound, pitch and volume of sound are described below.

Keywords: digital watermark, content analysis**1. Wstęp**

Zastosowania cyfrowych znaków wodnych (CZW, ang. *digital watermark*) mogą być wielorakie. Jednym z nich jest ochrona szeroko pojętej własności intelektualnej. Przypadkiem szczególnym jest ochrona praw autorskich w zakresie nagrań muzycznych. Wagę problemu podkreśla skala zjawiska polegającego na nielegalnym obrocie nagraniami muzycznymi na różnych nośnikach. Wg IFPI [1] co trzeci nośnik w postaci płyty kompaktowej ma nielegalne pochodzenie.

Innym przykładem zastosowania znaków wodnych może być np. automatyczny monitoring stacji nadawczych zarówno radiowych jak i telewizyjnych pod kątem rzetelności emisji. Bardziej wyrafinowanym zastosowaniem jest np. sterowanie za pomocą informacji ukrytej w postaci CZW w sygnale nośnym.

W przypadku ochrony nagrań muzycznych zazwyczaj zadania znaków wodnych mogą polegać na [2, 3]:

- identyfikacji właściciela praw autorskich,
- potwierdzeniu prawa własności do nagrania,
- obserwacji drogi handlowej nośnika,
- potwierdzenia zawartości.

Istnieje szereg metod wprowadzania znaków wodnych do sygnałów reprezentujących w ogólności dane multimedialne w sposób niezależny od zawartości tych sygnałów. W niniejszym artykule zaproponowano takie wprowadzanie znaków wodnych do plików dźwiękowych, które ściśle zależy od zawartości muzycznej. Wykorzystanie zawartości muzycznej pozwala na osiągnięcie

zwiększonej odporności sygnału z wbudowanym znakiem wodnym na przekształcenia zamierzone i niezamierzone, określane wspólnym pojęciem ataków. Do szerzej stosowanych obecnie sposobów przekształceń technologicznych należy kompresja stratna dźwięku. Ponieważ jednak zadaniem algorytmów kompresji stratnej jest maksymalne wykorzystanie wszelkich możliwości redukcji danych, które są nadmiarowe z punktu widzenia ludzkiego układu słuchowego, to stosowanie metod kontekstowego znakowania wodnego może być alternatywą dla metod bezkontekstowych.

2. Nośniki treści muzycznej

Wbudowanie znaków wodnych w sposób kontekstowy wymaga określenia nośników treści muzycznej. Za punkt wyjścia przyjęto, iż nośnikami treści muzycznej są elementy utworu muzycznego wyróżniane na gruncie teorii muzyki takie jak wysokość dźwięku i rytm (traktowane łącznie tworzą melodię), głośność, tempo, barwa, faktura (zarówno w sensie liczby głosów jak i obsady wykonawczej) itp.

Należy jednak zauważyć, iż niektóre komponenty muzyczne mają charakter pierwszoplanowy w stosunku do pozostałych, jak np. melodia (wysokość + czas trwania dźwięku) czy głośność (komponenty pierwszoplanowe) i barwa dźwięku lub tempo (komponenty drugoplanowe).

W sensie fizycznym wysokości dźwięku odpowiada częstotliwości składowej podstawowej, rytmowi zaś odpowiada czas trwania dźwięku, a ściślej wzajemne relacje pomiędzy czasem trwania poszczególnych dźwięków, natomiast głośność czyli dynamika jest związana bezpośrednio z amplitudą sygnału i jej zmianą w czasie.

Ze względu na wpływ poszczególnych komponentów na treść muzyczną wybrano jako elementy pierwszoplanowe wysokość dźwięku, czas trwania dźwięku oraz jego głośność. W celu określenia możliwości ukrycia informacji należy wyznaczyć dopuszczalne granice modyfikacji wybranych komponentów, przy czym dopuszczalność może oznaczać różne aspekty. Najczęściej jest to niesłyszalność, ale może to także oznaczać akceptowalność w ramach normalnej praktyki wykonawczej. Np. w przypadku wysokości dźwięku dopuszczalność jest ograniczona niesłyszalnością, ponieważ każda słyszalna zmiana wysokości dźwięku nie może być akceptowalna w praktyce wykonawczej, natomiast w przypadku dynamiki muzycznej zmiana głośności dźwięku może być słyszalna, a mimo to akceptowalna przez słuchacza i wykonawcę.

Tab. 1. Wybrane parametry dźwięku i ich dopuszczalne zmiany w procesie wprowadzania znaków wodnych

Tab. 1. Selected parameters of the sound and their acceptable modifications

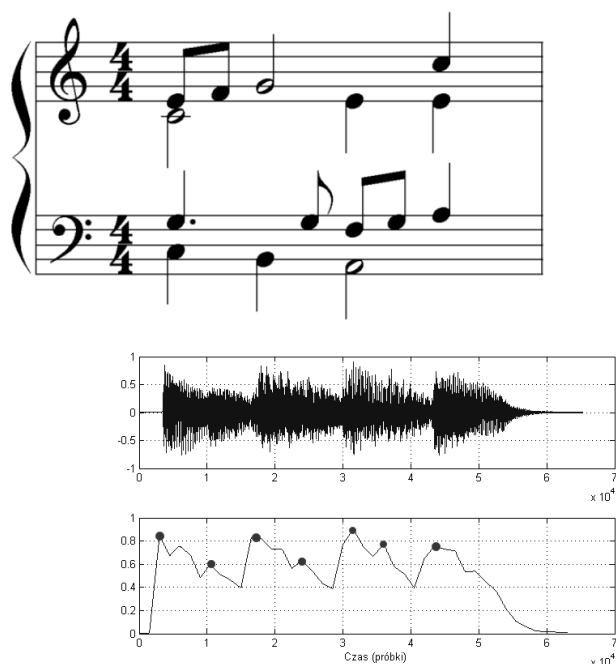
Element	Dopuszczalna zmiana
Wysokość	± 5 centów
Czas trwania	±3% długości modyfikowanego dźwięku
Dynamika	0,4dB

Dopuszczalne zmiany wybranych parametrów dźwięku wyznaczone pod kątem znakowania wodnego przedstawiono w tab. 1. Powyższe parametry wynikają bezpośrednio ze zdolności percepcyjnych aparatu słuchowego człowieka [4, 5]. Zostały one także zweryfikowane w zakresie wysokości oraz czasu trwania dźwięku dla dźwięków muzycznych [6]. Parametry w tabeli 1 dotyczą najgorszego przypadku z punktu widzenia ukrywania informacji

zawierającej znak wodny. Oznacza to, że badano wpływ modyfikacji jednego tylko rodzaju oraz wybrano wartości graniczne dla osób charakteryzujących się jednocześnie ponadprzeciętnym słuchem muzycznym i profesjonalną wiedzą muzyczną, a ponadto rozważano jedynie pojedyncze, izolowane dźwięki. Spełnienie w praktyce powyższych warunków jest w zasadzie niemożliwe, dlatego też w rzeczywistych przykładach muzycznych możliwa jest korekta dopuszczalnych granic zmienności podstawowych komponentów muzycznych.

3. Wprowadzanie znaków wodnych poprzez modyfikacje wybranych komponentów muzycznych

Jednym z postulatów dotyczących wprowadzania kontekstowego znaku wodnego w środowisku muzycznym jest taki sposób modyfikacji wybranych komponentów, który nie wywiera znaczącego wpływu na pozostałe komponenty. Zarówno w przypadku modyfikacji dynamiki dźwięku, czasu trwania oraz wysokości dźwięku konieczne jest wyznaczenie elementarnych fragmentów muzycznych odpowiadających najkrótszym dźwiękom tworzącym analizowane współbrzmienie (rys. 1).



Rys. 1. Wyznaczanie czasów trwania współbrzmień głównych
Fig. 1. Determining the time duration of the main sound combinations

Uproszczony algorytm wyznaczania początków współbrzmień oparty na rozwiązaniach opisanych w pracach [7, 8] jest następujący:

- Krok 1: podział sygnału dźwiękowego na pasma częstotliwości odpowiadające oktawom,
- Krok 2: filtracja za pomocą bloku filtrów pasmowoprzepustowych o częstotliwościach odpowiadających poszczególnym pasmom,
- Krok 3: wyznaczenie chwil czasu odpowiadających maksimum lokalnym dla przefiltrowanych przebiegów
- Krok 4: określenie chwil czasu rozpoczynających kolejne współbrzmienie na podstawie „początków lokalnych” wyznaczonych w poprzednim kroku.

W ramach fragmentu muzycznego, w którym wyznaczono przedziały czasowe odpowiadające elementarnym współbrzmieniom możliwe jest następnie dokonanie modyfikacji dynamiki oraz czasu trwania i wysokości dźwięków dla całych współbrzmień.

Modyfikacja wysokości dźwięku (M1)

Modyfikacja wysokości dźwięku zgodnie z powyższymi założeniami nie może oddziaływać na inne parametry dźwięku. Zatem należy zmienić wysokość bez zmiany czasu jego trwania. Realizację tego zadania umożliwia np. wokoder fazy [9] lub algorytm WSOLA [10], co wymaga wyznaczenia wartości parametrów przeskoku analizy (H_a) oraz syntezy (H_s).

Z tabeli 1 wynika, iż dopuszczalna zmiana wysokości (w górę) dźwięku wynosi $w = \pm 5$. Oznacza to, iż dla zwiększenia częstotliwości mamy:

$$\frac{\Delta f}{f_0} \leq 2^{\frac{w}{1200}}, \text{ dla } 0 < w < 5, \quad (1)$$

gdzie:

Δf – rozstrojenie,
 f_0 – częstotliwość podstawowa,

co prowadzi do

$$\frac{H_s}{H_a} \leq 2^{\frac{w}{1200}}. \quad (2)$$

Przy założeniu, że H_a oraz H_s są kolejnymi liczbami naturalnymi, mamy:

$$H_s = H_a + 1. \quad (3)$$

Otrzymujemy zatem:

$$H_a = 347 \text{ i } H_s = 346, \quad (4)$$

co odpowiada zwiększeniu częstotliwości dźwięków.

Analogicznie w przypadku zmniejszenia częstotliwości mamy:

$$H_a = 346 \text{ i } H_s = 347, \quad (5)$$

Wbudowanie bitów znaku wodnego w przypadku modyfikacji wysokości dźwięku odbywa się w sposób następujący: „0” oznacza zwiększenie wysokości określonego, wyznaczonego uprzednio fragmentu dźwiękowego, „1” oznacza zmniejszenie wysokości dźwięku dla danego fragmentu, przy czym zmiana częstotliwości dźwięku nie powoduje zmiany jego długości.

Pojedynczy bit znaku wodnego może zatem zostać umieszczony w jednym przedziale czasowym wyznaczającym ramy współbrzmienia.

Modyfikacja czasu trwania dźwięków (M2) [11, 12]

Zmiana czasu trwania dźwięku wykorzystuje mechanizm analogiczny do zmiany częstotliwości dźwięku. Podobnie zatem jak w pierwszej metodzie wykorzystano jeden z algorytmów umożliwiający zmianę czasu trwania dźwięku bez zmiany jego wysokości. Biorąc pod uwagę ograniczenia z tabeli 1 otrzymujemy dla wydłużenia dźwięku:

$$\frac{t_2 - t_1}{t_1} \leq 0,03 \text{ dla } t_2 > t_1, \quad (6)$$

co daje w wyniku

$$H_a = 34 \text{ i } H_s = 35, \quad (7)$$

natomiast dla skrócenia dźwięku mamy:

$$\frac{t_1 - t_2}{t_1} \leq 0,03 \text{ dla } t_1 > t_2, \quad (8)$$

i w rezultacie:

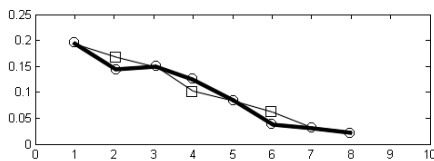
$$H_a = 34 \text{ i } H_s = 33, \quad (9)$$

Nieco inaczej wygląda samo kodowanie bitów znaku wodnego przy zastosowaniu modyfikacji czasu trwania dźwięków. Kodowanie binarnego „0” polega na skróceniu czasu trwania dźwięku (współbrzmienia) w dopuszczalnych dla danego dźwięku granicach oraz jednoczesnym wydłużeniu dźwięku po nim następującego. Kodowanie binarnej „1” polega na wydłużeniu czasu trwania dźwięku w dopuszczalnych granicach oraz jednoczesnym skróceniu dźwięku (współbrzmienia) po nim następującego. Zatem zarówno kodowanie „0” jak i „1” wymaga zachowania łącznego czasu trwania obu dźwięków oraz ewentualnych przerw między nimi. Ma to na celu eliminację wpływu zmian długości pojedynczych dźwięków na ogólną strukturę rytmiczną utworu.

Pojedynczy bit znaku wodnego w tej metodzie wymaga zaangażowania dwóch sąsiednich przedziałów czasowych oznaczających współbrzmienia.

Modyfikacja głośności dźwięków (M3) [13]

Modyfikacja głośności dźwięku wymaga nieco innego podejścia. Rozwiązanie zaproponowane w [6] polega na zwiększeniu dla binarnej „1” znaku wodnego głośności dźwięku w środku przedziału czasowego ograniczającego dźwięk (współbrzmienie), bądź jego zmniejszenie dla binarnego „0” (rys. 1). Dodatkowo na końcach przedziałów wyznaczających poszczególne dźwięki zachowano oryginalną dynamikę, tak aby wyeliminować potencjalne, nagłe zmiany dynamiki. Na rys. 2 linia cienka oznacza oryginalną obwiednię dźwięku, linia gruba zaś obwiednię zmodyfikowaną.



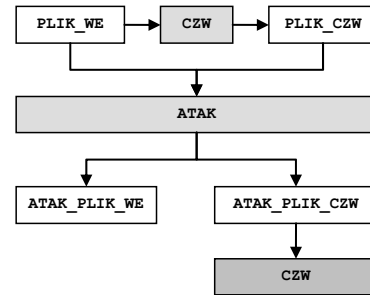
Rys. 2. Wbudowanie bitów znaku wodnego za pomocą modyfikacji głośności dźwięków

Fig. 2. Watermark embedding using sound volume modification

W przypadku umieszczania znaków wodnych z wykorzystaniem zmiany głośności dźwięku do zakodowania pojedynczego bitu wykorzystano pojedynczy dźwięk, choć możliwy jest także podział przedziału czasowego wyznaczającego poszczególne dźwięki na mniejsze podprzedziały i zwiększenie tym samym liczby bitów przypadających na jeden dźwięk.

4. Odporność znaków wodnych na wybrane przekształcenia

W wybrane przykłady muzyczne wbudowano znaki wodne za pomocą zaproponowanych powyżej sposobów (rys. 3). Następnie otrzymane w ten sposób wynikowe pliki dźwiękowe poddano wybranym atakom wymienionym w tabeli 2. Ma to na celu określenie odporności poszczególnych sposobów wbudowania znaków wodnych na popularne przekształcenia sygnałów dźwiękowych. W opisie uzyskanych wyników w dalszej części artykułu będą wykorzystywane jedynie symbole poszczególnych ataków (A1, A2, ..., A5) oraz symbole poszczególnych, opisanych kolejno metod (M1, M2, M3).



Rys. 3. Schemat blokowy systemu do badania odporności znaku wodnego
Fig. 3. Block diagram of the system for testing the immunity of the watermark

Tab. 2. Parametry ataków testowych

Tab. 2. Parameters of the test attacks

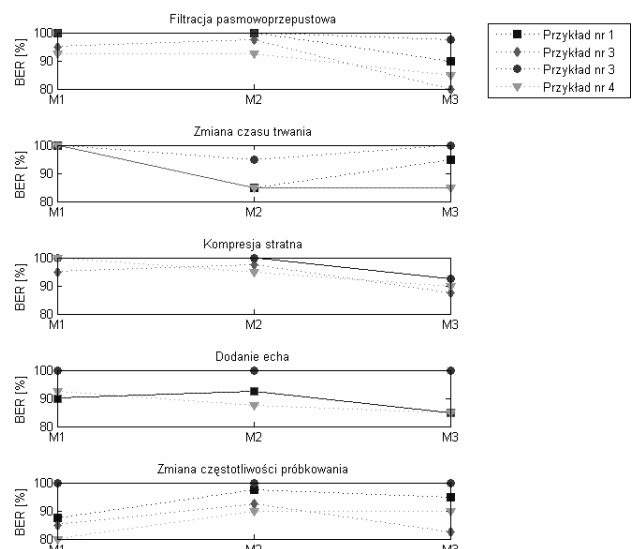
Symbol	Rodzaj ataku	Szczegóły
A1	Filtracja pasmowo-przepustowa	pasmo: 100Hz – 5 kHz
A2	Zmiana czasu trwania	± 10% (skracanie i wydłużanie)
A3	Kompresja stratna	Pakiet Ogg Vorbis, kompresja 128 kbit/s
A4	Dodanie echa	Opóźnienie 150 ms, współczynnik sprzężenia 0,5, czas opadania 1s
A5	Zmiana częstotliwości próbkowania	44,1 kHz → 22,1 kHz

Odporność przykładowych plików dźwiękowych zawierających znak wodny na wybrane ataki charakteryzuje bitowy współczynnik błędów *BER* (ang. *bit error rate*):

$$BER = \frac{B_o}{B_z} \cdot 100\%, \quad (10)$$

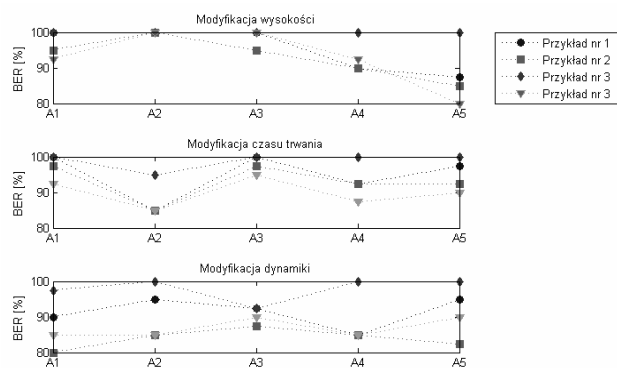
gdzie:

B_o – liczba bitów odczytanych z sygnału zawierającego znak wodny po poddaniu go działaniu przekształcenia
 B_z – liczba bitów znaku wodnego umieszczonych w sygnale źródłowym



Rys. 4. Bitowy współczynnik błędów (*BER*) dla wybranych ataków w funkcji różnych metod wbudowania CZW

Fig. 4. Bit error rate (*BER*) for the selected attacks as function of different methods of the watermark embedding



Rys. 5. Bitowy współczynnik błęd (BER) dla różnych metod wbudowania CZW w funkcji wybranych ataków

Fig. 5. Bit error rate (BER) for the selected methods of the watermark embedding as function of different attacks

Na rys. 4 pokazano wartości uzyskanego bitowego współczynnika błędów dla najgorszego przypadku spośród wszystkich wybranych przykładów przy różnych metodach wprowadzania znaku wodnego.

Rys. 5 przedstawia ten sam współczynnik, ale w funkcji poszczególnych ataków, dla różnych przykładów muzycznych (najgorszy przypadek) oraz różnych metod wbudowania znaku wodnego.

5. Podsumowanie

Przedstawione wyniki dla najgorszego przypadku (największy współczynnik błęd) ilustrują stosunkowo dobrą odporność proponowanych metod na ataki polegające na kompresji sygnału. Ataki polegające na kompansji w dziedzinie czasu wywierają największy wpływ na pliki zawierające znak wodny wbudowany za pomocą modyfikacji czasu trwania poszczególnych dźwięków, choć wydaje się, iż w przyjętym rozwiązaniu jest to wynikiem konstrukcji klucza, który posłużył w procesie dekodowania. W przypadku ataków polegających na zmianie czasu trwania całego fragmentu muzycznego możliwe jest uwzględnienie tego faktu w konstrukcji zarówno kodera jak i dekodera. Należy jednak przyznać, że nieco bardziej zaawansowane ataki polegające np. na przypadkowej zmianie długości poszczególnych segmentów muzycznych mogą spowodować degradację wbudowanego znaku wodnego.

6. Literatura

- [1] IFPI, The recording industry 2005 commercial piracy report, <http://www.ifpi.org/site-content/library/piracy2005.pdf>.
- [2] M. Arnold, M. Schmucker, S. D. Wolthusen: Techniques and Applications of Digital Watermarking and Content Protection. Artech House, Boston-London 2003.
- [3] I. J. Cox, M. L. Miller, J. A. Bloom: Digital watermarking. Morgan Kaufmann Publishers, 2002.
- [4] B. C. J. Moore: Wprowadzenie do psychologii słyszenia, PWN, Warszawa-Poznań, 1999.
- [5] E. Ozimek: Dźwięk i jego percepcja. Aspekty fizyczne i psychoakustyczne, PWN, Warszawa-Poznań 2002.
- [6] J. Korostil, M. Łazoryszczak: Digital Watermarking of an audio environment based on melody retrieval of musical composition. Proceedings of the ACS-CISIM 2004 Conference: Computer Information Systems and Applications, Białystok 2004.
- [7] T. Tolonen, M. Karjalainen: A computationally efficient multipitch analysis model. IEEE Transactions on Speech and Audio Processing, Vol. 8, No. 6, 2000.
- [8] A. Klapuri: Sound Onset Detection by Applying Psychoacoustic Knowledge. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 1999.
- [9] M. R. Portnoff: Implementation of the Digital Phase Vocoder Using the Fast Fourier Transform. IEEE Transactions on Acoustic, Speech and Signal Processing, Vol. ASSP-24, No. 3, 1976.
- [10] W. Verhelst, M. Roelands: An Overlap-Add Technique Based on Waveform Similarity (WSOLA) for High Quality Time-Scale Modification of Speech. In Proc. International Conference on Acoustics, Speech and Signal Processing, Minneapolis, 1993.
- [11] M. Łazoryszczak: Wbudowanie cyfrowych znaków wodnych w wybrane komponenty utworu muzycznego na przykładzie modyfikacji rytmu. Roczniki Informatyki Stosowanej Wydziału Informatyki Politechniki Szczecińskiej, Nr 8, Szczecin 2005.
- [12] M. Łazoryszczak: Digital watermarking of an audio environment based on musical rhythm analysis. Image Analysis, Computer Graphics, Security Systems and Artificial Intelligence Applications, Białystok 2005.
- [13] J. Korostil, M. Łazoryszczak: Wykorzystanie dynamiki muzycznej do umieszczania cyfrowych znaków wodnych w środowisku dźwiękowym. Roczniki Informatyki Stosowanej Wydziału Informatyki Politechniki Szczecińskiej, Nr 7, Szczecin 2004.

Artykuł recenzowany

IMPREZY NAUKOWO-TECHNICZNE

KONGRESY ☆ KONFERENCJE ☆ SYMPOZJA ☆ TARGI ☆ WYSTAWY

NAUKA

04-07 grudnia 2006 - Zakopane, POLSKA
TRANSCOMP 2006 - KOMPUTEROWE SYSTEMY WSPOMAGANIA NAUKI, PRZEMYSŁU I TRANSPORTU
 Organizator: Politechnika Radomska,
 Instytut Automatyki i Telematyki Transportu
 ul. Malczewskiego 29, 26-600 Radom
 tel. (048) 361-77-05, fax. (048) 361-77-42
 e-mail: transcomp@pr.radom.net

NAUKA

14-16 marca 2007 - Warszawa, POLSKA
AUTOMATION 2007 - Konferencja Naukowo-Techniczna AUTOMATYZACJA - NOWOŚCI I PERSPEKTYWY
 Organizator: Przemysłowy Instytut Automatyki i Pomiarów PIAP
 Al. Jerozolimskie 202, 02-486 Warszawa
 tel. (022) 8740-205, fax. (022) 8740-220
 e-mail: konferencja@piap.pl
 www.piap.pl