

Krzysztof Jaskólski¹⁾

**ZASTOSOWANIE ŁAŃCUCHÓW MARKOWA
W BADANIU POTENCJALNYCH MOŻLIWOŚCI
WYKORZYSTANIA INFORMACJI AUTOMATYCZNEGO
SYSTEMU IDENTYFIKACJI (AIS)
W APLIKACJACH ANTYKOLIZYJNYCH**

**USING MARKOV CHAINS TO INVESTIGATE
THE POTENTIAL POSSIBILITIES OF EMPLOYING DATA
FROM AN AUTOMATIC IDENTIFICATION SYSTEM
FOR COLLISION AVOIDANCE APPLICATIONS**

STRESZCZENIE Urządzenia radiolokacyjne, jak radar czy ARPA, mają wiele ograniczeń związanych z przekazywaniem ciągłej i wiarygodnej informacji. Wynikają one z braku możliwości wykrycia ech radarowych w niekorzystnych warunkach atmosferycznych i podczas manewrowania jednostek. Istotne jest, by informację radarową uzupełniać inną, której dokładność porównywalna jest do tej pochodzącej z systemów GNSS. Takim systemem wydaje się AIS. W artykule analizę dostępności informacji o pozycji geograficznej przedstawiono, stosując dyskretne w stanach i czasie łańcuchy Markowa. Do analiz wykorzystano dane zarejestrowane w latach 2006–2007 i 2010–2012 w laboratorium Instytutu Nawigacji i Hydrografii Morskiej.

Słowa kluczowe:

AIS, łańcuchy Markowa.

ABSTRACT Radio-location devices such as radar or ARPA display several limitations concerned with transferring continuous and reliable data. These limitations stem from the absence of the possibility to detect radar echo under adverse weather conditions during a maneuver executed by a vessel. It is important that the radar data be supplemented with data coming from a system working on continuous basis, with data whose accuracy is comparable to the data coming from GNSS. AIS seems to be such a system. This paper presents analysis of data availability related to geographical position using discrete in states and time the Markov chains. Data recorded in the Institute of Marine Navigation and Hydrography in 2006–2007 and 2010–2012 have been used to conduct analyses.

Keywords:

AIS, Markov chains.

DOI: 10.5604/0860889X/1097964

¹⁾ Akademia Marynarki Wojennej, Wydział Nawigacji i Uzbrojenia Okrętowego, 81-103 Gdynia, ul. J. Śmidowicza 69; e-mail: k.jaskolski@amw.gdynia.pl

WSTĘP

AIS (*Automatic Identification System*), czyli system automatycznej identyfikacji statków, wprowadzono dla poprawy bezpieczeństwa żeglugi oraz możliwości wymiany na szczeblu krajowym i międzynarodowym danych o statkach płynących do lub z portów, a także wymiany informacji o przewożonych statkami pasażerach i ładunkach niebezpiecznych lub zanieczyszczających środowisko. Głównym założeniem wprowadzenia AIS było zapewnienie szerszego spektrum dostępnej, ciągłej i wiarygodnej informacji nawigacyjnej. Wykorzystanie informacji transmitowanej za pośrednictwem systemu AIS w celu podniesienia bezpieczeństwa żeglugi stało się więc zjawiskiem powszechnym. W szczególności, poza przydatnością w obszarze nadzoru ruchu na akwenie, informacje AIS mogą stanowić istotne źródło danych w procesie podejmowania decyzji związanych z unikaniem kolizji. W takich sytuacjach należy mieć jednak pełne zaufanie do informacji przekazywanej tym kanałem. Zasadne jest więc pytanie, na ile i w jakim stopniu informacja pochodząca z systemu AIS jest dostępna i wiarygodna. Studia nad specyfikacją techniczną systemu ITU R.M.1371 pozwalają wstępnie zakładać, że błędy systemu klasyfikowane jako brak wiarygodności zależą od pracy sensorów

INTRODUCTION

AIS (*Automatic Identification System*) is a data exchange system which has been introduced to improve shipping safety and the possibility of exchanging data, at a country level and international level, related to ships heading to or from ports, as well as exchanging data relating to passengers and dangerous or environment-polluting cargo carried by ships. The main purpose of introducing AIS was to offer a wider spectrum of available, continuous and reliable navigational data. Using data transmitted through AIS in order to enhance shipping safety has become common. Apart from usefulness for traffic control in a marine region AIS data can be a very important source of data to be used for making decisions related to collision avoidance. However, some reservations are voiced with regard to the unconditional reliance on the data transferred through this channel. Thus, it is justifiable to ask the question how much and to what extent data coming from AIS is available and reliable. The studies on the technical specification of the ITU R.M.1371 system can allow the assumption that errors classified as lack of reliability will depend on the performance of sensors co-working with AIS and human errors made by operators, and those fitting the devices.

współpracujących z AIS oraz błędów ludzkich popełnianych przez operatorów i instalatorów urządzeń.

Obecnie problem wyznaczania współrzędnych pozycji dla potrzeb nawigacji morskiej rozważany w kategoriach błędu pomiaru wydaje się już nie istnieć. Jego realizacja jest jedynie funkcją zastosowanego rozwiązania technicznego. W tej sytuacji istotnego znaczenia nabierają równie ważne, choć często pomijane, eksploatacyjne charakterystyki systemów radionawigacyjnych, takie jak dostępność, niezawodność, ciągłość czy wiarygodność.

Traktowanie AIS jako systemu radionawigacyjnego jest jednak dyskusyjne. Jest to w istocie radiowy kanał transmisji informacji. Dlatego w artykule posłużono się pojęciem dostępności informacji. Miara ta będzie wyrażana metodami statystycznymi.

POJĘCIE DOSTĘPNOŚCI INFORMACJI AIS ORAZ WSPÓŁCZYNNIKA DOSTĘPNOŚCI INFORMACJI

Dla celów przedstawionych wyników badań zdefiniowano dwa pojęcia.

Dostępność informacji: stopień zgodności odebranych przez użytkowników informacji z treścią wiadomości charakterystycznych dla AIS, zgodnych ze specyfikacją ITU R.M.1371, w dowolnym momencie, w strefie działania systemu.

At present the problem of fixing position co-ordinates for marine navigation purposes considered in categories of measurement error seems to have been solved. Its realization is only a function of the technical solution employed. In this situation, equally important but often neglected operation characteristics of radio-navigation systems, such as availability, continuity and reliability are of increasing importance.

However, treating AIS as a radio-navigation system is disputable. It is actually a radio channel for data transmission. That is why the notion of data availability is used in the paper. This measure will be expressed with statistical methods.

THE NOTIONS OF AIS DATA AVAILABILITY AND OF DATA AVAILABILITY RATIO

For the purposes of investigating the results presented two concepts are defined.

Availability of data: degree of conformity of data received by users with the content of a message characteristic of AIS, compatible with the ITU R.M.1371 specification, at any moment, in the zone where the system is operating.

In this case AIS data availability refers to the conformity of data, related to geographical position, received by

W tym przypadku dostępność informacji AIS odnosi się do zgodności odebranej przez użytkownika informacji dotyczącej pozycji geograficznej z treścią wiadomości charakterystycznej dla depeszy nr 1 AIS, zgodnie z treścią specyfikacji technicznej systemu AIS. Wartości poprawnych i niepoprawnych danych dotyczących pozycji przedstawiono w tabeli 1.

Tak przedstawione podejście badawcze ułatwia opracowanie modelu badań dostępności informacji w oparciu o sposoby *postprocessingowe*.

Tabela 1. Zestawienie poprawnych i niepoprawnych wartości pozycji geograficznej depeszy nr 1 AIS

Table 1. Correct and incorrect values relating to geographical position of No 1 AIS message

Pozycja geograficzna Geographical position	Wartość poprawna Correct value	Wartość niepoprawna Incorrect value
LONGITUDE	±180	181
LATITUDE	±90	91

Źródło / Source: Draft Revision of Recommendation ITU-R.M.1371. Technical characteristics for a universal shipborne automatic identification system using time division multiple access in VHF maritime mobile band, Radiocommunication study Groups, Interenational Telecommunication Union, 2010.

Współczynnik dostępności: graniczna wartość wielkości zdefiniowanej jako dostępność informacji nawigacyjnej przesłanej w postaci binarnej użytkownikom systemu.

REJESTRACJA SYGNAŁÓW, OBSŁUGA SYSTEMU BAZODANOWEGO ZEBRANYCH INFORMACJI

W celu przeprowadzenia analiz statystycznych zarejestrowanych komunikatów systemu AIS zastosowano

a user with the content of a message characteristic of AIS type No 1, compatible with the technical specification for AIS.

The investigation approach presented this way substantially facilitates developing an availability investigation model based on postprocessing methods.

Availability ratio: the boundary value of a quantity defined as availability of navigational data, transmitted in binary form to a system's user.

RECORDING SIGNALS, OPERATING DATABASE OF DATA COLLECTED

The postprocessing method was used in order to conduct statistical analyses of AIS messages recorded. It consists in processing data after it is recorded. To record data required

metodę *postprocessingu* polegającą na późniejszym opracowaniu zebranych danych. Rejestracja danych wymagała przygotowania stanowiska badawczego składającego się z urządzenia odbiorczego AIS, transpondera, anteny, układu zasilania i rejestratora danych. Stanowisko rejestrujące sygnały AIS zostało przygotowane w Instytucie Nawigacji i Hydrografii Morskiej AMW. Na nośniku danych rejestratora sygnałów rejestrowano informacje z mnemonicem AIVDM. Do analizy danych wykorzystano wiadomość z numerem 1. Zarejestrowane dane pochodzą z lat: 2006–2007 i 2010–2012. Dane zarejestrowane zostały w plikach tekstowych.

making an investigation station consisting of an AIS receiver, AIS transponder, AIS antenna, power supply system, and data recorder. The station for recording AIS signals was prepared in the Institute of Navigation and Hydrography, at the Gdynia Naval Academy. The data was recorded on the data carrier of a signal recorder with mnemonic AIVDM. No 1 message was used to analyze the data. The data recorded date was for the years: 2006–2007 and 2010–2012. They were recorded in text files.



Rys. 1. Stanowisko rejestracji sygnałów wraz z urządzeniem odbiorczym AIS SAAB R4
Fig. 1. Station for recording signals along with AIS SAAB R4 receiver

Praca z zarejestrowanym sygnałem w postaci tekstowej w oparciu o metodę *postprocessingu* wymaga przygotowania odpowiedniego oprogramowania.

To work with a signal recorded in text file form using the postprocessing method requires appropriate preparation of software.



Rys. 2. Obsługa zapisanych danych AIS z wykorzystaniem oprogramowania IBSExpert typu Client do zarządzania bazą danych

Fig. 2. Use of AIS data recorded with software IBSExpert, type Client to manage database

Budowa narzędzia na potrzeby artykułu była zadaniem złożonym. AIS w transmisji wykorzystuje pakiety 6-bitowe. Jest to związane z wykorzystaniem ograniczonej liczby znaków w systemie i zmniejszeniem objętości transmitowanych paczek danych. Tak zarejestrowane dane muszą zostać zamienione na postać dziesiętną lub tekstową. Standard IEC 61162-1 [4] dostarcza informacji na temat sposobu kodowania binarnych pakietów danych. Dokumentami pomocniczymi w dekodowaniu danych binarnych AIS są ITU-R.M 1371 [1] i IEC 61993-2 ed.1 [5].

WSTĘPNE ZAŁOŻENIA DO BADAŃ DOSTĘPNOŚCI INFORMACJI SYSTEMU AIS

Jeśli proces przekazywania informacji skutkujący dostarczeniem użytkownikowi kompletnej informacji w dowolnym momencie jest procesem niezawodnym, to proces przekazywania informacji możemy uznać za funkcjonowanie kanału informacyjnego i badać go metodami znanymi z teorii niezawodności, tak jak w [2, 6, 10, 11]. W związku z tym do badania dostępności zaproponowano model oparty na teorii markowskich procesów eksploatacji obiektów technicznych.

Łańcuch Markowa charakteryzuje się tym, że stan procesu w chwili $n + 1$ zależy wyłącznie od stanu w chwili n ,

To develop a tool for the purpose of this paper was a complex task. AIS uses 6-bit packages for transmission. This is connected by the use of a limited number of signs in the system and decrease in the volumes of the packages transmitted. Data recorded this way has to be transformed into a decimal or a text form. Standard IEC 61162-1 [4] provides data on how to code binar data packages. Documents which help decode AIS binar data are ITU-R.M 1371 [1], IEC 61993-2 ed.1 [5].

INITIAL ASSUMPTIONS RELATED TO AIS DATA AVAILABILITY

If a process of data transfer, resulting in providing a user with complete data at any moment, is a reliable process it can be viewed as a functioning of an information channel and can be investigated with methods included in the theory of reliability as in [2, 6, 10, 11]. In connection with the above a model based on the Theory of the Markov Processes related to Operating Technical Objects was proposed.

The characteristic of Markov Chain is that the process state at the moment $n + 1$ depends exclusively on the moment state n and does not depend on states at previous moments [3]. Based on this assumption a stochastic probability matrix of transitions between availability states of AIS for incoming data with disregard to the

a nie zależy od stanów w chwilach wcześniejszych [3]. Przy takim założeniu wyznaczono macierz stochastyczną prawdopodobieństwa przejść między stanami dostępności systemu AIS dla napływającej informacji bez uwzględnienia źródła informacji — kryterium „wiersze” i kryterium „statki”, co pozwala wydzielić źródło niedostępnej informacji. Macierz stochastyczna prezentuje intensywność przejść między stanami. Analizy przeprowadzono w odniesieniu do całości informacji zaimportowanej do bazy danych, przyjmując jako kryterium poszczególne wiersze zapisane w tabelicy (kryterium wiersze), a także kryterium statki, co oznacza rozpoznawanie zapisów do numeru MMSI (*Maritime Mobile Service Identity*) identyfikującego jednostkę pływającą. Należy zwrócić uwagę, że statki nadają swoje depesze z różną częstotliwością, co oznacza, że zależnie od przyjętego kryterium można oczekiwać różnych wyników analiz statystycznych.

Kontynuując, przyjęto, że kryterium „wiersze” odnosić się będzie do liczby wierszy z mnemonikiem AIVDM z dostępną informacją dotyczącą pozycji geograficznej, natomiast kryterium „statki” — do liczby statków odpowiadających za stan tej informacji.

source of data — criterion ‘lines’ and criterion ‘ships’ were determined, which makes it possible to identify the source of unavailable data. Stochastic matrix presents intensity of transitions between states.

The analyses were conducted in relation to the whole data imported to the database taking as a criterion the particular lines written on the board (criterion lines) as well as the criterion ships, which means identifying recordings with an MMSI number (Maritime Mobile Service Identity) showing the identity of a vessel afloat. It should be noted that, for various reasons, ships send their messages using various frequencies, which means that depending on the criterion adopted, various results of the statistical analysis can be expected.

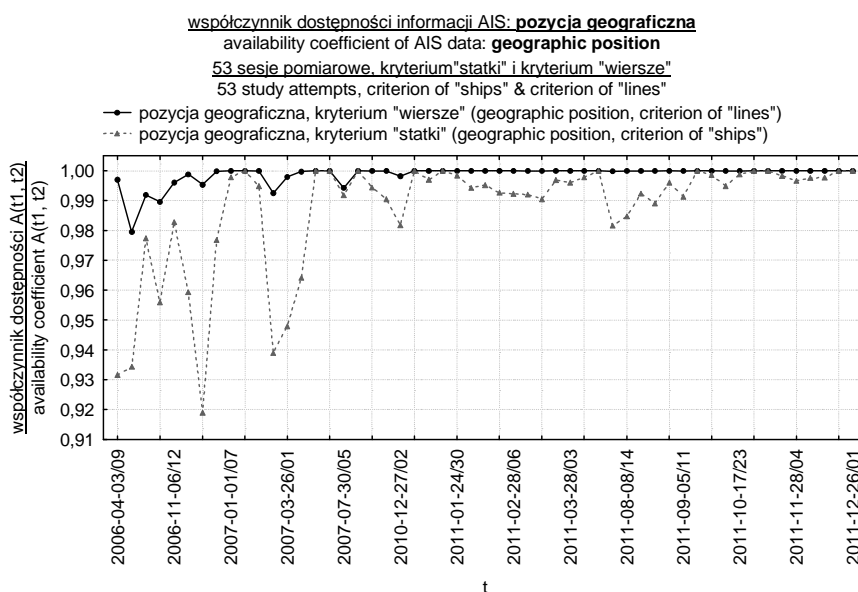
Furthermore, a decision was made that the criterion ‘lines’ would be referred to as the number of lines with mnemonic AIVDM with available data related to geographical position, whereas criterion ‘ships’ would refer to the number of ships responsible for the state of this data.

STATYSTYCZNE ZESTAWIENIE WYNIKÓW BADAŃ DOSTĘPNOŚCI INFORMACJI DOTYCZĄCEJ POZYCJI GEOGRAFICZNEJ

Badanie informacji dotyczącej pozycji geograficznej przeprowadzono na podstawie komunikatów zarejestrowanych w plikach tekstowych zawierających komunikaty odebrane za pośrednictwem AIS o statkach znajdujących się w rejonie Zatoki Gdańskiej w przedziale od 04.04.2006 r. do 08.01.2012 r. (wybrane 54 tygodnie — 378 dni). Wstępne wyniki badań w postaci wykresów współczynnika dostępności informacji o pozycji geograficznej przedstawiono poniżej.

STATISTICAL PRESENTATION OF INVESTIGATION RESULTS OF DATA AVAILABILITY AS RELATED TO GEOGRAPHICAL POSITION

The investigation of the data relating to geographical position was based on messages recorded in text files containing messages received through AIS about vessels being in the area of the Bay of Gdańsk from 4 April, 2006 to 8 January, 2012 (selected 54 weeks — 378 days). The initial investigation results in the form of a diagram of data availability ratio related to geographical position is presented below.



Rys. 3. Wykres zmian współczynnika dostępności informacji dotyczącej pozycji geograficznej: linia ciągła — kryterium „wiersze”, przerywana — kryterium „statki”

Fig. 3. Diagram of changes in data availability ratio related to geographical position: line continuous — criterion 'lines', dotted — criterion 'ships'

Na podstawie analizy statystycznej, w oparciu o 24-godzinny rozkład wyników badań za każde siedem dni (tydzień) przedstawiono wyniki dostępności informacji o pozycji geograficznej, ich wartości minimalne, maksymalne i odchylenia standardowe.

A statistical analysis, based on 24-hour distribution of investigation results for each 7 days (week) was conducted to present the results of data availability related to geographical position, their minimum and maximum values, and standard deviations.

Tabela 2. Analiza statystyczna wyników badań za każdy tydzień w oparciu o 24-godzinne przedstawienie wyników kompletności informacji dotyczącej pozycji geograficznej

Table 2. Statistical analysis of the investigation results for each week, based on a 24-hour presentation of the results of the data complete related to geographical position

Pozycja geograficzna Geographical position	Kryterium „wiersze” Criterion ‘lines’				Kryterium „statki” Criterion ‘ships’			
	DATA	$A(t_1, t_2)$	δ	$A(t_1, t_2)_{max}$	$A(t_1, t_2)_{min}$	$A(t_1, t_2)$	δ	$A(t_1, t_2)_{max}$
2006-04-03/09	0,9969	0,0016	0,9993	0,9938	0,9316	0,0199	0,9710	0,9094
2006-04-17/23	0,9795	0,0029	0,9831	0,9752	0,9343	0,0075	0,9450	0,9258
2006-04-24/30	0,9918	0,0114	1	0,9837	0,9774	0,0225	1	0,9548
2006-11-06/12	0,9896	0,0115	0,9999	0,9741	0,9559	0,0239	0,9894	0,9351
2006-11-13/19	0,9960	0,0124	1	0,9741	0,9828	0,0241	1	0,9362
2006-11-20/26	0,9988	0,0017	1	0,9963	0,9592	0,0338	1,0000	0,9157
2006-11-27/03	0,9953	0,0040	0,9996	0,9905	0,9190	0,0140	0,9354	0,9010
2007-01-01/07	0,9998	0,0008	0,9996	0,9985	0,9767	0,0050	1	0,9893
2007-01-08/14	0,9999	0,0001	1	0	0,9978	0,0043	1	0,9893
2007-01-15/21	1	0	1	1	1	0	1	1
· · ·								
2011-12-12/18	0,9999	0,0001	1	0,9999	0,9976	0,0059	1	0,9880
2011-12-19/25	0,9999	0,0001	1	0,9999	0,9977	0,0045	1	0,9887
2011-12-26/01	1	0	1	1	1	0	1	1
2012-01-02/08	1	0	1	1	1	0	1	1
Wartość oczekiwana: Expected value:	0,9986	0,0011	0,9996	0,8649	0,9867	0,0074	0,9953	0,9765

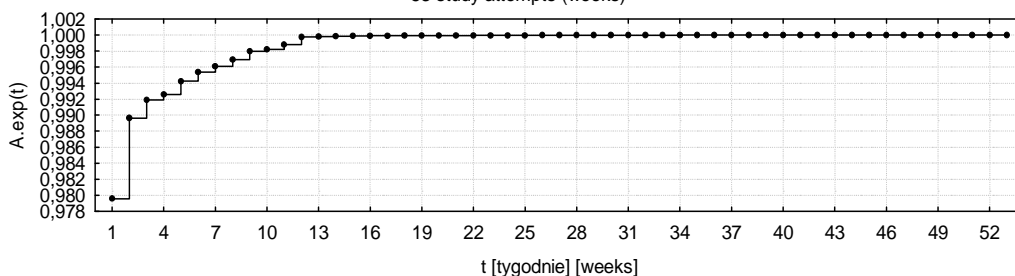
$A(t_1, t_2)$ — dostępność informacji w zdefiniowanym przedziale czasu, δ — odchylenie standardowe, $A(t_1, t_2)_{max}$ — wartość maksymalna dostępności w zdefiniowanym przedziale czasu, $A(t_1, t_2)_{min}$ — wartość minimalna dostępności w zdefiniowanym przedziale czasu.

$(A(t_1, t_2)$ — data availability in the defined time interval, δ — standard deviation, $A(t_1, t_2)_{max}$ — maximum availability value in the defined time interval, $A(t_1, t_2)_{min}$ — minimum availability value in the defined time interval.

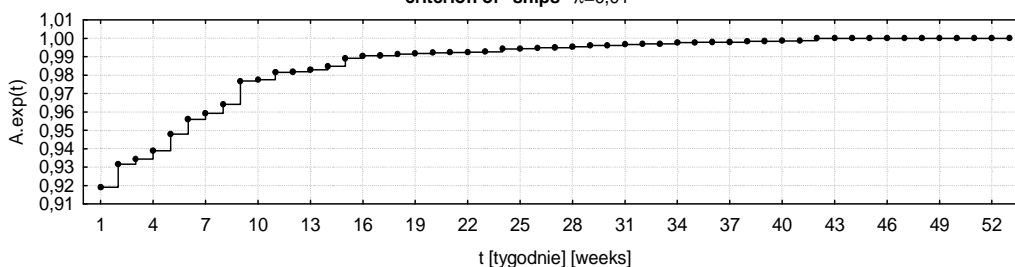
Funkcja rozkładu wykładniczego $A.exp(t)$, dostępność informacji AIS: pozycja geograficzna, kryterium "wiersze" $\lambda=0,001$
 Cumulative Distribution Function $A.exp(t)$, AIS Availability data: geographic position, criterion of "lines"

$\lambda=0,001$

53 sesje pomiarowe (tygodnie)
 53 study attempts (weeks)



kryterium "statki" $\lambda=0,01$
 criterion of "ships" $\lambda=0,01$



Rys. 4. Rozkład wykładniczy dostępności $A.exp(t)$ danych AIS: pozycja geograficzna, u góry kryterium „wiersze”, na dole kryterium „statki” (λ — intensywność uszkodzeń, $A.exp(t)$ — wykładnicza postać współczynnika dostępności informacji AIS)

Fig. 4. Exponential distribution of AIS data availability $A.exp(t)$: geographical position, top criterion 'lines', bottom criterion 'ships' (λ — damage intensity, $A.exp(t)$ — exponential form of AIS data availability ratio)

METODYKA BADANIA DOSTĘPNOŚCI INFORMACJI AIS

Zmieniający się w czasie stan dostępności informacji systemu jest wielkością, która może być opisana za pomocą procesów stochastycznych. W związku z powyższym opracowano metodykę badania dostępności informacji AIS, wykorzystując w tym celu łańcuchy Markowa.

METHODOLOGY USED TO INVESTIGATE AIS DATA AVAILABILITY

Data availability state changing in time is a quantity that can be described with stochastic processes. For this reason Markov Chains were used to develop the methodology to investigate AIS data availability.

Założenia:

Proces stochastyczny oznaczono symbolem:

$$\{X(t): t \in T\}. \quad (1)$$

Specjalnym przypadkiem procesu stochastycznego jest ciąg zdarzeń losowych $X_n: n \in (0, 1, 2, 3, \dots, n)$ nazywany łańcuchem losowym. Wartości zmiennych losowych $X_n: n = 0, 1, 2, 3, \dots, 53$ reprezentują stan dostępności informacji AIS. Charakter zmian stanów można przyjmować jako łańcuch Markowa w zbiorze stanów:

$$S = \{S_1, S_2\}, \quad (2)$$

gdzie:

S_1 — (stan 1) oznacza, że informacja jest niedostępna; system jest w stanie niezdatności / (state 1) means that the data is unavailable; the system is in the state of uselessness;

S_2 — (stan 2) oznacza, że informacja jest dostępna; system jest w stanie zdatności / (state 2) means that the data is available; the system is in the state of usefulness.

Momenty przebywania systemu w stanie awarii są momentami odnowy struktury nawigacyjnej.

Zgodnie z założeniami łańcuch Markowa (Markov chain) charakteryzuje się tym, że stan procesu w chwili $n + 1$ zależy wyłącznie od stanu w chwili n , a nie zależy od stanów w chwilach wcześniejszych [3].

Assumptions:

The stochastic process was designated with the symbol:

A special case of the stochastic process is a chain of random occurrences $X_n: n \in (0, 1, 2, 3, \dots, n)$ referred to as random chain. Values of random variables $X_n: n = 0, 1, 2, 3, \dots, 53$ represent the AIS data availability state. The character of state changes can be regarded as a Markov Chain in the set of states:

where:

Moments when the system is in a state of failure are moments of the reconditioning of navigational structure.

According to the assumptions the characteristic of Markov Chain is that the process state at the moment $n + 1$ depends exclusively on the moment state n , and it does not depend on the state in previous moments [3].

Łańcuch Markowa można zdefiniować, jeżeli rozkład początkowy

Markov chain can be defined, if the initial distribution

$$P(X_0 = i) = p_i, i \in S \quad (3)$$

oraz macierz prawdopodobieństw przejść

and matrix of transition probabilities

$$P = [p_{ij}: i, j \in S]; \quad (4)$$

$$p_{ij} = P(X_{n+1} = j | X_n = i), n = 0, 1, 2, 3, \dots, 53 \quad (5)$$

jest znana.

is known.

Wówczas macierz stochastyczna wyznacza prawdopodobieństwo przejścia między stanami dostępności informacji AIS. Macierz stochastyczna prezentuje intensywność przejść między stanami.

Then, the stochastic matrix determines transition probabilities between AIS data availability states. The matrix presents the transition intensity between the states.

Kontynuując, p_{ij} oznacza prawdopodobieństwo przejścia ze stanu $i \in S$ do stanu $j \in S$ w chwili $n + 1$.

Further, p_{ij} designates the probability of transition from the state $i \in S$ to the state $j \in S$ at the moment $n + 1$.

W naszym przypadku macierz przyjmuje postać:

In our case the matrix takes the form:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad (6)$$

gdzie:

where:

p_{ij} — prawdopodobieństwo przejścia ze stanu i do stanu j / probability of transition from the i to the state j .

Ponadto do badań przyjęto rozkład początkowy $p(0)=[0, 1]$. Oznacza to, że system znajduje się w stanie pracy S_2 .

In addition, for the purposes of the investigation the initial distribution $p(0)=[0, 1]$ was adopted. This means that the system is in the work state S_2 .

Macierz prawdopodobieństw przejść wyznaczamy na podstawie procesu intensywności przejść między stanami.

The process of transition intensity is used to determine the matrix of transition probabilities between states.

Prawdopodobieństwa graniczne wyznaczono z zależności:

The boundary probability was determined from the dependence:

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \lim_{n \rightarrow \infty} P(X_{n+1} = j | X_n = i) = \lim_{n \rightarrow \infty} P(X_{n+1} = j) = \pi_j. \quad (7)$$

Aby osiągnąć zamierzony cel, należy rozwiązać układ równań liniowych:

In order to reach the intended goal it is necessary to solve the set of linear equations:

$$\sum_{i \in S} \pi_i p_{ij} = \pi_j, \quad j \in S \quad (8)$$

oraz

and

$$\sum_{i \in S} \pi_i = 1, \quad (9)$$

gdzie:

where:

π — prawdopodobieństwo graniczne / boundary probability.

Prawdopodobieństwa π_1, π_2 wyznaczają rozkład stacjonarny jednorodnego łańcucha Markowa $\pi = [\pi_1, \pi_2]$ z macierzą prawdopodobieństw przejść:

Probabilities π_1, π_2 determine stationary distribution of homogenous Markov Chain $\pi = [\pi_1, \pi_2]$ with matrix of transition probabilities:

$$P = p_{ij}; \quad i, j \in S. \quad (10)$$

Prawdopodobieństwo graniczne obliczamy, rozwiązując układ równań na podstawie iloczynu macierzy:

The boundary probability is calculated through solving a set of equations based on the matrix product:

$$[\pi_1 \quad \pi_2] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = [\pi_1, \quad \pi_2]. \quad (11)$$

W związku z powyższym otrzymamy układ równań liniowych:

$$\begin{cases} p_{11}\pi_1 + p_{21}\pi_2 = \pi_1 \\ p_{12}\pi_1 + p_{22}\pi_2 = \pi_2. \\ \pi_1 + \pi_2 = 1 \end{cases} \quad (12)$$

Po rozwiązaniu układu równań otrzymamy prawdopodobieństwa graniczne π_1, π_2 .

ZASTOSOWANIE METODYKI BADAŃ DOSTĘPNOŚCI INFORMACJI AIS

Niech wartości zmiennych losowych $\{X_n: n = 0,1,2,3, \dots, 53\}$ będą reprezentowały stan dostępności informacji o pozycji geograficznej statku. Charakter zmian stanów określono w zbiorze $S = \{S_1, S_2\}$, gdzie:

S_1 — oznacza, że szerokość geograficzna przyjmuje wartość 91 lub długość geograficzna przyjmuje wartość 181 oraz graniczna wartość dostępności badanej zmiennej danej populacji przyjmuje wartość (0; 0,95); informacja jest w stanie niezdatności;

S_2 — oznacza, że szerokość geograficzna przyjmuje wartości z zakresu (-90; 90) oraz długość geograficzna przyjmuje wartość z zakresu (-180; 180), a graniczna wartość kompletności informacji badanej zmiennej danej populacji przyjmuje wartość (0,95; 1); informacja jest w stanie zdatności.

Therefore a set of linear equations will be obtained:

On solving the set of equations the boundary probability π_1, π_2 will be obtained.

USE OF THE METHODOLOGY TO INVESTIGATE AIS DATA AVAILABILITY

Let the values of random variables $\{X_n: n = 0,1,2,3, \dots, 53\}$ represent the state of data information availability related to the ship's geographical position. The character of transition changes was determined in the set $S = \{S_1, S_2\}$, where:

S_1 — means that the geographical latitude takes value 91 or the geographical longitude takes value 181 and the boundary availability value of a population variable investigated takes value (0; 0,95); this means that the data is in a state of uselessness.

S_2 — means that the geographical latitude takes values (-90; 90) and the geographical longitude takes values (-180; 180), and the boundary value of data completeness of a population variable investigated takes the value (0,95; 1). This means that the data is in the state of usefulness.

Do badań przyjęto rozkład początkowy $p(0) = [1 \ 0]$. Oznacza to, że system znajduje się w stanie awarii S_1 .

Macierz prawdopodobieństw przejść wyznaczamy na podstawie procesu intensywności przejść między stanami, toteż zgodnie z zależnościami (5) i (6) macierz prawdopodobieństw przejść w łańcuchu Markowa przyjmuje postać:

Kryterium „wiersze” Criterion ‘lines’	Kryterium „statki” Criterion ‘chips’
$P = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$	$P = \begin{bmatrix} 0,5 & 0,5 \\ 0,0426 & 0,9574 \end{bmatrix}$

Szczegółowy rozkład przejść systemu do kolejnych stanów dostępności informacji przedstawiają tabela 3. i rysunek 5. W efekcie kwantowania przydzielono każdemu wynikowi dostępności badanej zmiennej wartość binarną odpowiadającą numerowi stanu systemu.

For the purpose of the investigation the initial distribution $p(0) = [1 \ 0]$ was adopted. This means that the system is in the state of failure S_1 .

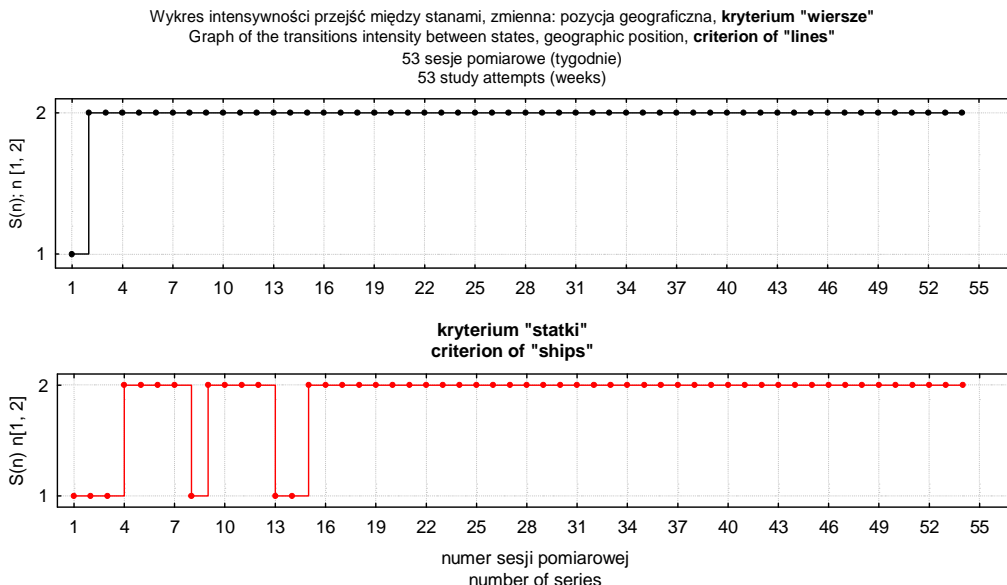
The intensity of transition between the states is used to determine the matrix of transition probability. Therefore following the dependence (5) and (6) the matrix of transition probability in Markov Chain took the form:

Table 3 and figure 5 present the detailed distribution of transitions into the particular states of system data availability. As a result of quantization each variable availability investigated was assigned a binary value corresponding to the number of the state.

Tabela 3. Rozkład przejść systemu do poszczególnych stanów dostępności informacji o pozycji geograficznej statku

Table 3. Distribution of system transitions into particular states of data availability related to ship’s geographical position

Kryterium „wiersze” Criterion ‘lines’				Kryterium „statki” Criterion ‘ships’			
P_{ij}	Intensywność przejść między stanami Intensity of transition between states	P_{ij}	Intensywność przejść między stanami Intensity of transition between states	P_{ij}	Intensywność przejść między stanami Intensity of transition between states	P_{ij}	Intensywność przejść między stanami Intensity of transition between states
p11	0	p21	0	p11	3	p21	2
p12	1	p22	52	p12	3	p22	45



Rys. 5. Wykres intensywności przejść między poszczególnymi stanami dostępności informacji dotyczącej pozycji: u góry kryterium „wiersze”, na dole kryterium „statki”

Fig. 5. Diagram of intensity of transitions between particular states of data availability related to geographical position: top criterion 'lines', bottom criterion 'ships'

Po rozwiązaniu układu równań (12), otrzymano prawdopodobieństwa graniczne π_1, π_2 .

On solving the set of equations (12) boundary probabilities π_1, π_2 were obtained.

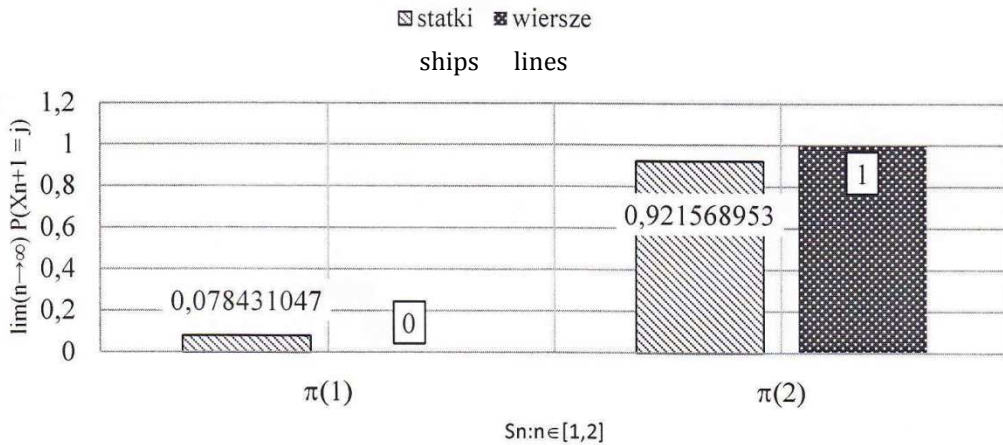
Kryterium „wiersze” Criterion 'lines'	Kryterium „statki” Criterion 'ships'
$\pi_1 \cong 0$	$\pi_1 \cong \frac{42553}{542553}$
$\pi_2 \cong 1$	$\pi_2 \cong \frac{500000}{542553}$

Dla kryterium „wiersze” prawdopodobieństwo przebywania systemu w stanie pracy S_2 wynosi 1, a w stanie awarii S_1 jest równe 0. Dla kryterium „statki” prawdopodobieństwo przebywania systemu w stanie pracy S_2 wynosi 0,9216, a w stanie awarii

Thus, for the criterion 'lines' the probability for the system to be in the work state S_2 is 1, and in the failure state S_1 the probability is 0. For the criterion 'ships' the probability for the system to be in the state work S_2 is 0,9216, and in the failure state S_1

S_1 jest równe 0,0784. Różnice w wynikach współczynników dostępności pochodzących z poszczególnych serii pomiarowych mieszczą się w przedziale (0; 0,0763).

the probability is 0,0784. The differences in availability ratios coming from the particular measurement series fall into the range (0; 0,0763).



Rys. 6. Rozkład prawdopodobieństw granicznych kompletności informacji dotyczącej pozycji geograficznej statku: kryterium „statki” i kryterium „wiersze”

Fig. 6. Distribution of boundary probabilities of data completeness related to ship’s geographical position: criterion ‘ships’ and criterion ‘lines’

Jeżeli rozkład początkowy $p(0)$ przyjmuje postać

Further, if the initial distribution $p(0)$ takes the form

$$p(0) = [0 \ 1],$$

to macierz prawdopodobieństw przejścia po n krokach przyjmuje postać

the matrix of transition probability after n steps takes the form

$$P(0, n) = P^n, \tag{13}$$

natomiast rozkład prawdopodobieństwa stanów procesu w chwili t przyjmuje postać

Whereas the distribution of process state probabilities at the moment t takes the form

$$P(n) = p(0) \cdot P(0, n). \tag{14}$$

Wyniki estymacji macierzy prawdopodobieństw po n krokach oraz rozkładu prawdopodobieństwa stanów procesu w chwili t przedstawiono w tabeli 4.

The estimation results of matrix probabilities following n steps and distribution of process state probability at the moment t is presented in table 4.

Tabela 4. Macierze prawdopodobieństw przejścia oraz rozkład prawdopodobieństw stanów informacji dotyczącej pozycji geograficznej

Table 4. Matrixes of transition probabilities and distribution of data states related to geographical position

n	Kryterium „wiersze” Criterion ‘lines’		Kryterium „statki” Criterion ‘ships’	
	$P(0, n) = P^n$	$P(n) = p(0) \cdot P(0, n)$	$P(0, n) = P^n$	$P(n) = p(0) \cdot P(0, n)$
2	$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$	[0 1]	$\begin{bmatrix} 0,2713 & 0,7287 \\ 0,0620 & 0,9380 \end{bmatrix}$	[0,0620 0,9380]
3			$\begin{bmatrix} 0,1664 & 0,8333 \\ 0,0709 & 0,9291 \end{bmatrix}$	[0,0709 0,9291]
4			$\begin{bmatrix} 0,1188 & 0,8812 \\ 0,0750 & 0,9250 \end{bmatrix}$	[0,0750 0,9250]
5			$\begin{bmatrix} 0,0969 & 0,9031 \\ 0,0769 & 0,9231 \end{bmatrix}$	[0,0709 0,9291]
6			$\begin{bmatrix} 0,0869 & 0,9131 \\ 0,0777 & 0,9223 \end{bmatrix}$	[0,0777 0,9223]
7			$\begin{bmatrix} 0,0823 & 0,9177 \\ 0,0781 & 0,9219 \end{bmatrix}$	[0,0781 0,9219]
8			$\begin{bmatrix} 0,0802 & 0,9198 \\ 0,0783 & 0,9217 \end{bmatrix}$	[0,0783 0,9217]
9			$\begin{bmatrix} 0,0792 & 0,9208 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]
10			$\begin{bmatrix} 0,0788 & 0,9212 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]
11			$\begin{bmatrix} 0,0786 & 0,9214 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]
12			$\begin{bmatrix} 0,0785 & 0,9215 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]
13			$\begin{bmatrix} 0,0785 & 0,9215 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]
14			$\begin{bmatrix} 0,0784 & 0,9216 \\ 0,0784 & 0,9216 \end{bmatrix}$	[0,0784 0,9216]

Dokonując analizy predykcji stanu informacji systemu, stwierdzono, że rozkład stabilizuje się już po drugim

On conducting the analysis of the system’s data state it was found that the distribution stabilized as early as after

kroku dla kryterium „wiersze” i po czternastym kroku dla kryterium „statki”.

Należy zauważyć, że wraz ze wzrostem czasu (liczby kroków) elementy macierzy $P(0, n)$ znajdujące się w tych samych kolumnach coraz mniej się od siebie różnią. Gdy $n \rightarrow \infty$, różnica ta dąży do 0. Ze wzoru (14) wynika, że w miarę wzrostu n rozkład $p(n)$ stabilizuje się. Na podstawie danych zawartych w tabeli 4. zaobserwowano stabilizację rozkładu dla $n = 14$. Właściwości te są istotnymi właściwościami pewnej klasy łańcuchów Markowa.

WNIOSKI

W artykule zaprezentowano wyniki badań dostępności informacji AIS, stosując dwa podejścia badawcze, tj. kryterium „wiersze” i kryterium „statki”. Wyznaczenie granicznych wartości dostępności informacji systemu AIS miało na celu określenie stopnia przydatności informacji AIS dotyczącej pozycji geograficznej w zastosowaniach antykolizyjnych. Wartości współczynników kompletności kształtujące się na poziomie 0,9986 dla kryterium wiersze oraz 0,9867 dla kryterium statki potwierdzają te możliwości. Stan zdatności dla kryterium „wiersze” został zaobserwowany w 52 przypadkach, a dla kryterium „statki” w 47 przypadkach na 53 przeprowadzone sesje pomiarowe, co również może

two steps for the criterion ‘lines’ and after the 14th step for the criterion ‘ships’.

It should be noted that with the rise in time (number of steps), the elements of matrix $P(0, n)$ which are in the same columns differ from each other less and less. When $n \rightarrow \infty$ the difference tends to 0. From the formula (14) follows that with the rise in n the distribution $p(n)$ stabilizes. On analyzing the data included in the table it was found out that the distribution stabilized for $n = 14$. These properties are significant properties of a certain class of Markov chains.

CONCLUSIONS

This paper presents the results of investigations into AIS data availability. The investigations were based on two investigation approaches, i.e. criterion ‘lines’ and criterion ‘ships’. Boundary values of AIS data availability were determined in order to determine the usefulness degree of AIS data related to geographical position with regard to collision avoidance applications. The values of completeness ratios 0.9986 for criterion lines and 0.9867 for criterion ships prove such application possibilities. The state of usefulness for the criterion lines was observed in 52 cases for the criterion lines and in 47 cases for the criterion ships in 53 measurement sessions, which can

świadczą o wysokim stopniu przydatności informacji dotyczącej pozycji statków pozyskanej za pośrednictwem urządzeń AIS w aplikacjach antykolizyjnych. Podczas przeprowadzonych badań zaobserwowano, że kryterium „wiersze” daje wyższe wyniki badań dostępności informacji o pozycji statku. Odnosząc się do liczby statków charakteryzujących się brakiem dostępności informacji o pozycji, łatwo zaobserwować pogorszenie wyników badań. Liczba statków transmitujących niekompletne komunikaty nie jest tożsama z liczbą niekompletnych komunikatów. Różnica ta wynika z liczebności populacji statków i wierszy spełniających warunek dostępności informacji, a także z interwału czasowego transmitowanych danych przez pokładowe urządzenia AIS, który przyjmuje wartości z zakresu od 2 do 180 sekund. Kryterium „statki” wydaje się być bardziej obiektywnym podejściem badawczym w stosunku do kryterium „wiersze”. W tym przypadku graniczna wartość dostępności informacji wyznaczona została na podstawie czasu przebywania informacji w stanie awarii i w stanie pracy AIS dla danego statku. W badaniach pominięto interwał czasowy nadawach wiadomości ze statków oraz charakterystyki dokładności informacji pochodzącej z sensorów GNSS.

point to the high level of data usefulness with regard to ships' positions fixed with AIS devices in collision avoidance applications. During the investigations it was found that the criterion 'lines' gives higher data availability results related to a ship's position. While referring to the number of ships characterized by the absence of position-related data availability it is easy to notice the deterioration of the results. The number of ships transmitting incomplete messages is not the same as the number of incomplete messages. The difference results from the number of ship population and lines satisfying the condition of data availability, as well as from the time interval related to data transmitted by shipboard AIS devices, which ranges from 2 to 180 seconds. Criterion 'ships' seems to be the most objective investigation approach in relation to criterion 'lines'. In this case determination of the boundary value of data availability was based on the time spent by the message in the state of failure and the state of AIS work. The time interval of messages transmitted from ships and the characteristics of data accuracy sensors were not included in the investigations.

BIBLIOGRAFIA / REFERENCES

- [1] *Draft Revision of Recommendation ITU-R.M.1371. Technical characteristics for a universal shipborne automatic identification system using time division multiple access in VHF maritime mobile band*, Radiocommunication study Groups, Interenational Telecommunication Union, 2010.
- [2] Felski A., Jakubowski B., *Niezawodność systemu nawigacyjnego w kontekście przetwarzania informacji*, "Zeszyty Naukowe AMW", 2004, No 1 (*Reliability of a navigational system in relation to data processing — available only in Polish*).
- [3] Grabski F., Jaźwiński J., *Funkcje o losowych argumentach w zagadnieniach niezawodności, bezpieczeństwa i logistyki*, WKiŁ, Warszawa 2009 (*Random argument functions in issues relating to reliability, security and logistics — available only in Polish*).
- [4] IEC 61162-1 ed.2: *Maritime navigation and radiocommunication equipment and systems — digital interfaces*, Part 1: *Single talker and multiple listeners*, International Electrotechnical Commission, 2000.
- [5] IEC 61993-2 ed. 1 *Maritime navigation and radiocommunication equipment and systems — automatic identification systems*, Part 2: *Class A shipborne equipment of the universal automatic identification system (AIS) — operational and performance requirements, methods of test and required test results*, International Electrotechnical Commission, 2001.
- [6] Jakubowski B., *Niezawodnościowe aspekty przetwarzania informacji w zintegrowanym systemie nawigacyjnym* (rozprawa doktorska / dissertation), AMW, Gdynia 2005 (*Reliability-based aspects of data processing in an integrated navigational system — available only in Polish*).
- [7] Jaskólski K., *Model wiarygodności i dostępności informacji pozyskiwanej za pośrednictwem systemu AIS* (rozprawa doktorska / dissertation), AMW, Gdynia 2013 (*Reliability and availability model of data acquired by means of AIS — available only in Polish*).
- [8] Pietrzykowski Z., Magaj J., *The problem of route determination in ship movement in a restricted area*, 'Annual of Navigation', 2012, No 19, Part 2, Gdynia 2012.
- [9] Pietrzykowski Z., *Fuzzy Control in Solving collision Situation at Sea. Computational Intelligence Methods and Applications*, EXIT, Warszawa 2008.
- [10] Specht C., *Availability, reliability and continuity of differential GPS transmission*, 'Annual of Navigation', 2003, No 5, Gdynia 2003.
- [11] Specht C., Nowak A., *Reliability characteristics of the position solution statistics in the DGPS system*, European Navigation Conference — GNSS 2005, Monachium 2005.