UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH
W BYDGOSZCZY

ZESZYTY NAUKOWE
SCIENTIFIC JOURNAL
262

TELEKOMUNIKACJA I ELEKTRONIKA

TELECOMMUNICATIONS AND ELECTRONICS

17

WYDZIAŁ TELEKOMUNIKACJI,
INFORMATYKI I ELEKTROTECHNIKI

BYDGOSZCZ – 2013

UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH
W BYDGOSZCZY

ZESZYTY NAUKOWE
SCIENTIFIC JOURNAL
262

TELEKOMUNIKACJA I ELEKTRONIKA

TELECOMMUNICATIONS AND ELECTRONICS

17

BYDGOSZCZ – 2013

OPRACOWANIE TECHNICZNE
mgr inż. Daniel Morzyński

Praca powstała przy wsparciu projektu
„Realizacja II etapu Regionalnego Centrum Innowacyjności"
współfinansowanego ze środków Europejskiego Funduszu Rozwoju Regionalnego
w ramach Regionalnego Programu Operacyjnego
Województwa Kujawsko-Pomorskiego na lata 2007-2013

ISSN 1899-0088

Wydawnictwa Uczelniane Uniwersytetu Technologiczno-Przyrodniczego
ul. Ks. A. Kordeckiego 20, 85-225 Bydgoszcz, tel. 52 3749482, 3749426
e-mail: wydawucz@utp.edu.pl http://www.wu.utp.edu.pl/

Wyd. I. Nakład 60 egz. Ark. aut. 3,8. Ark. druk. 4,6.
Zakład Małej Poligrafii UTP Bydgoszcz, ul. Ks. A. Kordeckiego 20

Contents

UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY
ZESZYTY NAUKOWE NR 262
ELEKTRONIKA I TELEKOMUNIKACJA 17 (2013), 5-16

TOTAL SIGNAL DEGRADATION DUE TO RAIN PRECIPITATION IN THE TROPOSPHERE IN THE AREA OF KIELCE CITY

Jacek Łukasz Wilk

Division of Telecommunications, Department of Telecommunications, Photonics and Nanomaterials, Faculty of Electrical Engineering, Automatics and Computer Science
al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland
j.wilk@tu.kielce.pl

Summary. The article discusses the different effects contributing to the overall loss of signal propagation due to rain along Earth-space paths and analyses atmospheric attenuation. It deals with propagation concerns for satellite communications systems in the troposphere. The presence of rainfall has a considerable influence on the quality of microwaves links, especially due to absorption and reflection by raindrops, as rain acts as a rather strong source of noise. The discussion includes signal attenuation due to precipitation, an increase in noise due to precipitation and total signal degradation due to precipitation. In practice, statistical description is the only satisfactory way to show the results of modelling of microwave links in rainy weather in the area of Kielce city. The theoretical results were compared with real-time data (e.g. rain attenuation as a function of the rain rate). The data were used to determine the impact of rain intensity, polarization and radio waves frequency on the received satellite signal quality and finally to provide information about signal attenuation due to rain and suggests ways to improve the design of communication systems. The research results presented in the article can be useful to satellite communications engineers as they can improve the design and performance of satellite links and minimize the interruption or lack of communication between the terminal and the satellite to deliver the satellite services at the highest level.

Keywords: atmospheric attenuation, signal attenuation due to precipitation, an increase in noise due to precipitation, total signal degradation due to precipitation, the application of polynomial regressions, radio wave propagation in troposphere.

1. INTRODUCTION

After free-space path loss, one of the most important effect contributing to the overall loss of signal propagation along Earth-space paths is rain attenuation, which is attributable to absorption and scattering by raindrops. In the article, atmospheric losses are treated separately from the losses caused by adverse weather conditions, which are also atmospheric losses. In practice, rain attenuation is not predictable with very good precision and it is directly related to the properties of rainfall, frequencies and polarization of radio waves. The article covers attenuation of satellite links due to rain intensity, frequency and polarization of radio waves in the region of Kielce city in Poland in accordance with series of the ITU-R Recommendations in up-to-date. The long-term 1-min average rain-rate characteristics in this area is estimated to be 34,4 mm/h. All of data were used to estimate the impact of the rainfall intensity, polarization and radio wave frequency on the quality of received microwave satellite signals in Poland, especially in Kielce city [3]. In this article, the chosen experimental data are presented.

2. RESULT AND ANALYSIS

As you can see below (Fig. 1, Fig. 2), increase in the frequency bring about changes at the increase in the signal attenuation L_d, noise increase due to precipitation W_{sz} and total signal degradation *DND*, downlink degradation (sum of these components) for downlink for each type of signal polarization (horizontal and vertical, respectively) with a step of 1 GHz.

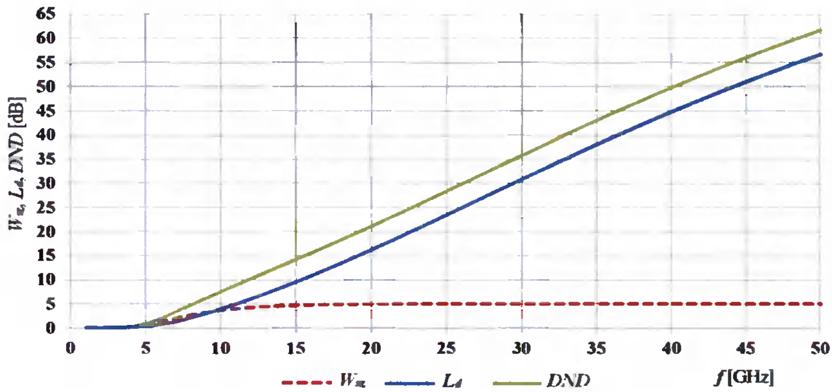

Fig. 1 . W_{sz}, L_d, *DND* [dB] vs frequency [GHz] (pol. H) [3]

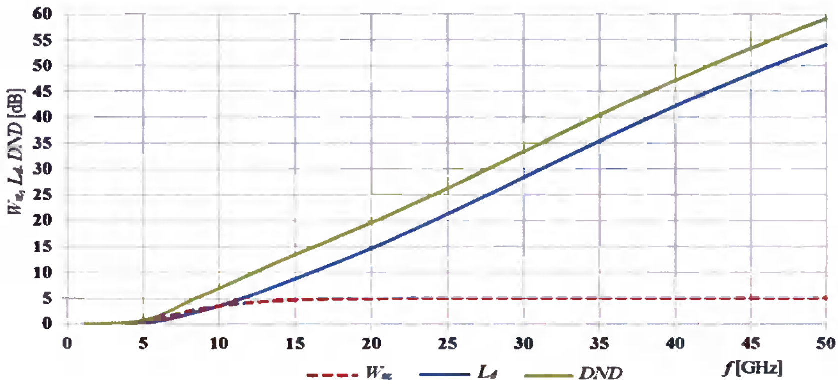

Fig. 2. W_{sz}, L_d, *DND* [dB] vs frequency [GHz] (pol. V) [3]

Although for over long distances, the atmosphere can be responsible for the polarization of a radio wave to fluctuate, the difference between horizontal and vertical becomes less significant. The results indicate (on the basis of the comparative results of the analysis of measurements for vertical and horizontal polarization) that the signal attenuation (with the same frequency and rainfall intensity) of horizontally polarized radio waves is greater than the signal attenuation of vertically polarized radio waves. Increase in the microwave signal frequency causes the increase in the difference in signal attenuation L_d between horizontally and vertically polarized radio waves. Generally, with the increase in the frequency, we can observe that the difference does not exceed 3 dB (max was approximately 2,67 dB in the frequency range from 39 GHz to 45 GHz and about 0-0,04 dB for the radio frequency exceeding 3 GHz). The procedures for the theory of correlation and regression allow – if it is necessary – to determine the analytical form of the regression function by estimating the parameters of this function for a given samples.

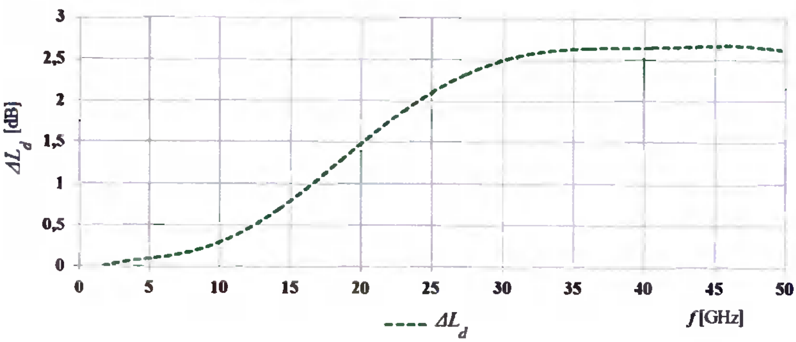

Fig. 3. Increase in L_d [dB] vs frequency [GHz] (pol. H-V) [3]

Figure 4 shows the relationship between the signal attenuation L_d [dB], increase in L_d [dB] for horizontally polarized radio waves related to vertically polarized radio waves, versus frequency to 35 GHz with a step of 10 MHz (0,01 GHz).

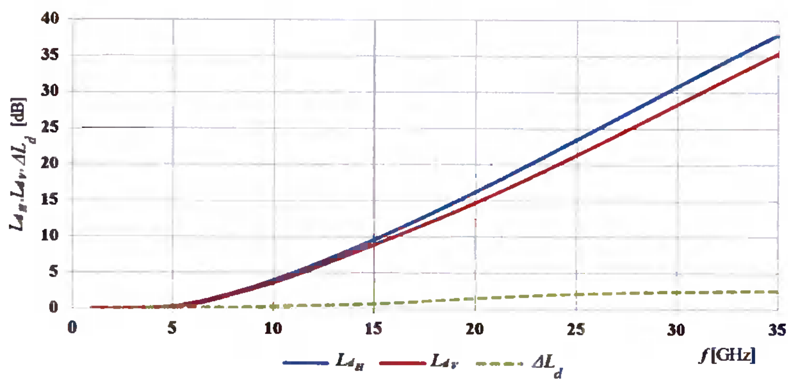

Fig. 4. L_d, increase in L_d [dB] vs frequency [GHz] (pol. H-V)

This article presents the application of polynomial regressions to provide information about signal attenuation due to rain. Each datum was processed using a bilinear interpolation scheme. The values of the signal attenuation L_d versus frequency for a selected rainfall statistics can be, as previously mentioned, analytically described by the principal component of regression model, which is the regression function. As a result of matrix operations, the predicted values of signal attenuation L_d were obtained versus frequency for both polarizations. The results are shown in Figure 5 (pol. H) and Figure 6 (pol. V).

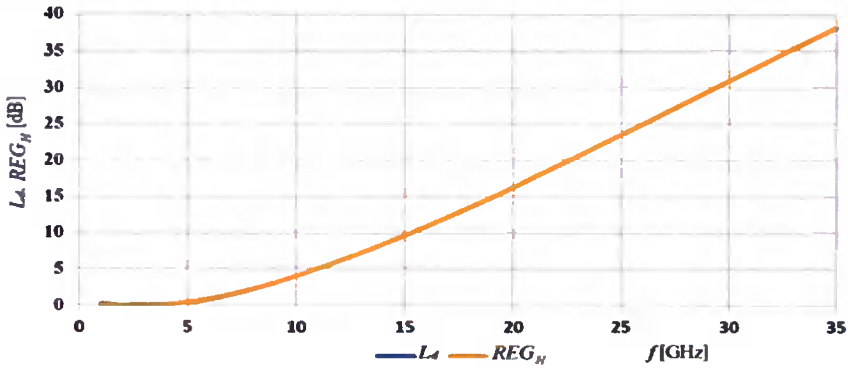

Fig. 5. REG_H [dB] vs frequency [GHz] (pol. H)

The proposed formulas of attenuation due to rain and the rain attenuation statistics in the area of Kielce city for horizontally polarized radio waves in the range of frequency of 1-35 GHz is the following (Fig. 5) [3]:

$$REG_H = -7{,}612 \cdot 10^{-9} f^6 - 4{,}933 \cdot 10^{-8} f^5 + 6{,}594 \cdot 10^{-5} f^4 - 4{,}287 \cdot 10^{-3} f^3 + 0{,}131 f^2 - 0{,}629 f^1 + 0{,}731$$

The maximum absolute value of attenuation calculated in agreement with the formula REG_H and the values obtained as the results of attenuation of radio waves in the area of Kielce city equals 0,22871 dB (for 1 GHz), whereas the minimum absolute value of attenuation is equal to from 0,00024 dB to 0,00079 dB (e.g. for 9,5 GHz, 16-16,2 GHz, 29,6 GHz, 30,1 GHz).

Similarly, the analogically determined relationship for vertically polarized radio waves versus frequency, whose takes into account the rain attenuation statistics in the area of Kielce city and changes in frequency with the step of 0,01 GHz, takes the formulas (Fig. 6) [3]:

$$REG_V = -4{,}967 \cdot 10^{-8} f^6 + 3{,}753 \cdot 10^{-6} f^5 - 5{,}066 \cdot 10^{-5} f^4 - 2{,}864 \cdot 10^{-3} f^3 + 0{,}122 f^2 - 0{,}637 f^1 + 0{,}779$$

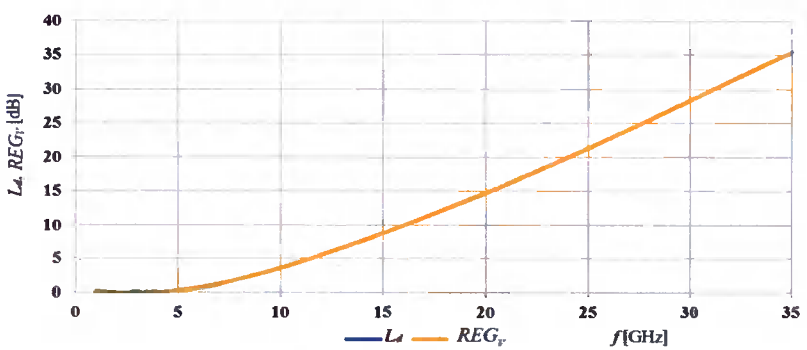

Fig. 6. REG_V [dB] vs frequency [GHz] (pol. V)

The maximum absolute value of attenuation calculated in agreement with the formula REG_V and the values obtained as the results of attenuation of radio waves in the area of Kielce city equals 0,26188 dB (for 1 GHz), but in most cases (89,44% of the absolute errors of approximation) it does not exceed the value of 0,1 dB. Because of the achieved results (Fig. 5, Fig. 6) whose correspond to the attenuation curves for horizontally polarized radio waves (Fig. 1) and vertically polarized radio waves (Fig. 2), it can be assumed that they provide a good estimate of actual values of L_d in the area of Kielce city (Fig. 7, comparatively).

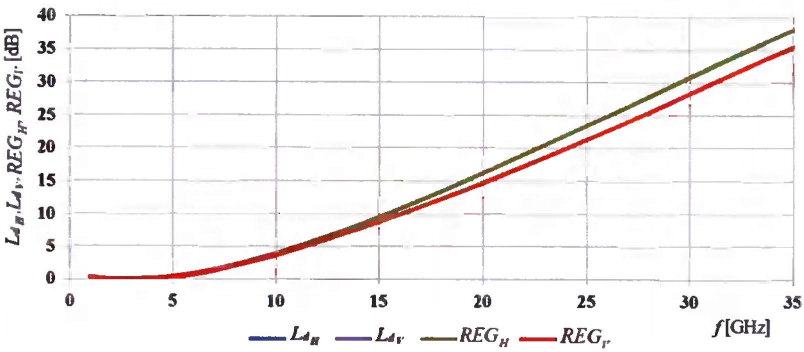

Fig. 7. L_{dH}, L_{dV}, REG_H, REG_V [dB] vs frequency [GHz] [3]

These estimates can be used to calculate the link budget analysis in the design of telecommunication systems by systems engineers as they can improve the design and performance of satellite links in adverse weather conditions and reduce the risk of interruption or lack of communication between the terminal and the satellite to deliver the satellite services at the highest level.

In practice, the signal attenuation L_d depends on the rainfall intensity as well as frequency and polarization. Because of climate changes whose affect both the mean and variability of climatic variables, the author in calculations take into account the limit values of attenuation L_d, which were predicted in Poland within 40 years (31 mm/h in Gorzow Wielkopolski and Gubin, 40,1 mm/h in Nowy Targ), additionally. This model uses bilinear interpolation to obtain an improved evaluation for a selected location from the grid neighbours and is based on many years of records collected by the European Space Agency (ESA). The Kielce University of Technology was the member of the international research project COST Action IC0802 – *Propagation Tools and Data for Integrated Telecommunication, Navigation and Earth Observation Systems*, whose main aim is to analyse the impact of weather conditions on the quality of satellite transmission [1,8]. So, the part of results of measurements which were realized in the Kielce University of Technology, as a Polish member of COST Action IC0802, is presented in this article.

The values of the signal attenuation L_d, which were obtained in this way, depending on the frequency and polarization of radio waves, in extreme cases, can illustrate the possible impact of the critical values of rainfall intensity on the extreme level of signal attenuation. Figure 8 and Figure 9 shows the diagram of the signal attenuation L_d, noise increase due to precipitation W_{sz} and total signal degradation *DND* (sum of these components) for downlink for each type of polarization (horizontally polarized radio waves and vertically polarized radio waves, respectively) in the case of the minimum rainfall intensity ($R_{0,01}$ = 31 mm/h), which was recorded in Poland within 40 years in Gorzow Wielkopolski and Gubin, versus frequency with a step of 1 GHz. Figure 11 and Figure 12 shows the diagram which was obtained for the horizontally and vertically polarized radio waves, respectively, in the case of the maximum rainfall intensity ($R_{0,01}$ = 40,1 mm/h), which was recorded in Poland in Nowy Targ within 40 years, versus frequency with a step of 1 GHz, too.

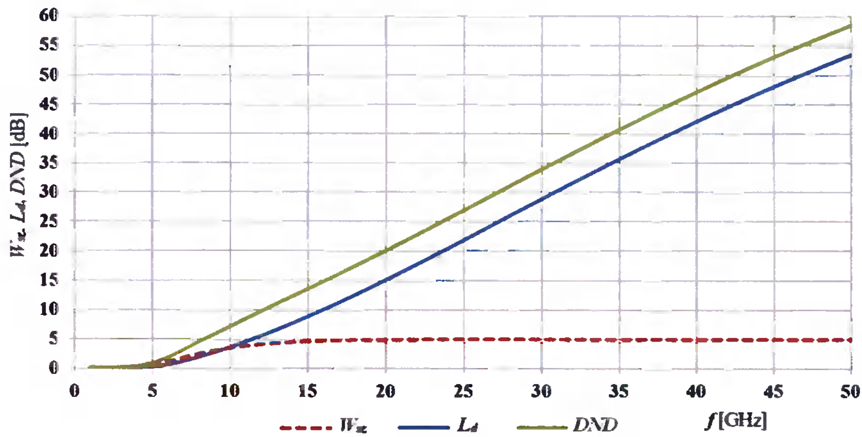

Fig. 8 . W_{sz}, L_d, DND [dB] vs frequency [GHz] (pol. H)

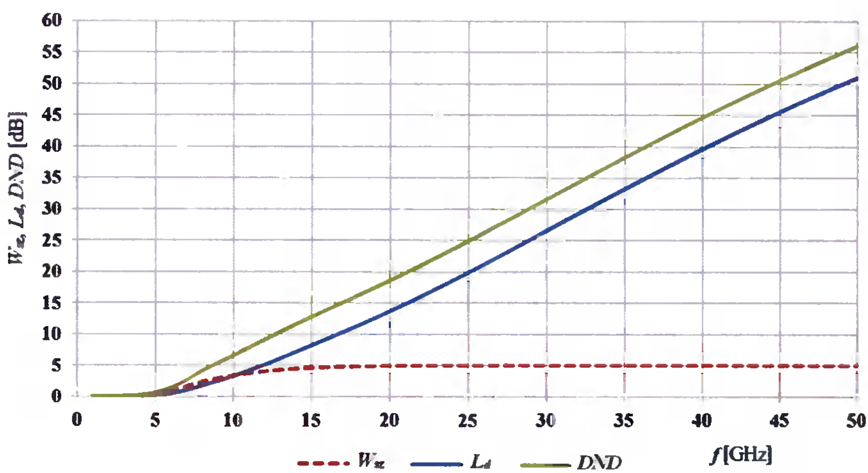

Fig. 9 . W_{sz}, L_d, DND [dB] vs frequency [GHz] (pol. V)

The results shown in Figure 10, directly and Figure 8, Figure 9, indirectly indicate that the maximum of increase in L_d [dB] between horizontally and vertically polarized radio waves (pol. H-V) versus frequency [GHz] in the case of Gorzow Wielkopolski i Gubin equals about 2,48 dB (above 40 GHz). Increase in the microwave signal frequency causes the increase in the difference between signal attenuation (Fig. 10). Generally, with the increase in the radio wave frequency we can observed that the difference exceeds 2 dB (above 26 GHz). Moreover, increase in L_d does not exceed 0,5 dB (below 13 GHz) and 1 dB (below 17 GHz).

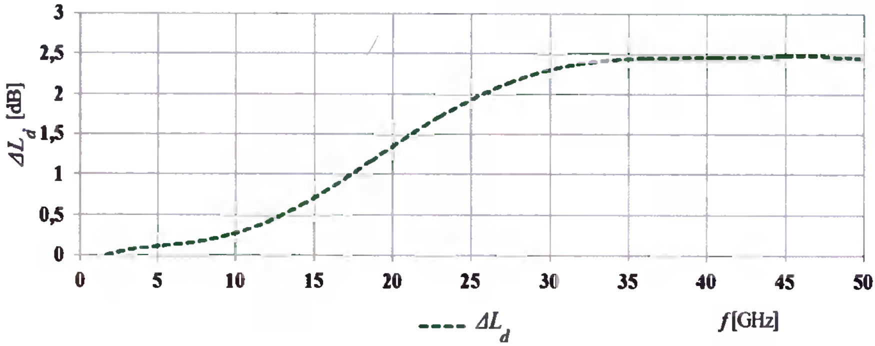

Fig. 10. Increase in L_d [dB] vs frequency [GHz] (pol. H-V)

The results obtained on the basis of the rainfall statistics from Nowy Targ seems to be the interested date on this background. The results shown in Figure 11 and Figure 12 indicate that the maximum of increase in L_d [dB] between horizontally and vertically polarized radio waves (pol. H-V) versus frequency [GHz] in the case of Nowy Targ is equal to 2,97 dB (for 41 GHz). With the increase in the frequency we can observe that the difference exceeds 2 dB (above 22 GHz) and 2,9 dB (above 35 GHz). Increase in L_d does not exceed 0,5 dB (below 10 GHz) and 1 dB (below 16 GHz) [see Fig. 13].

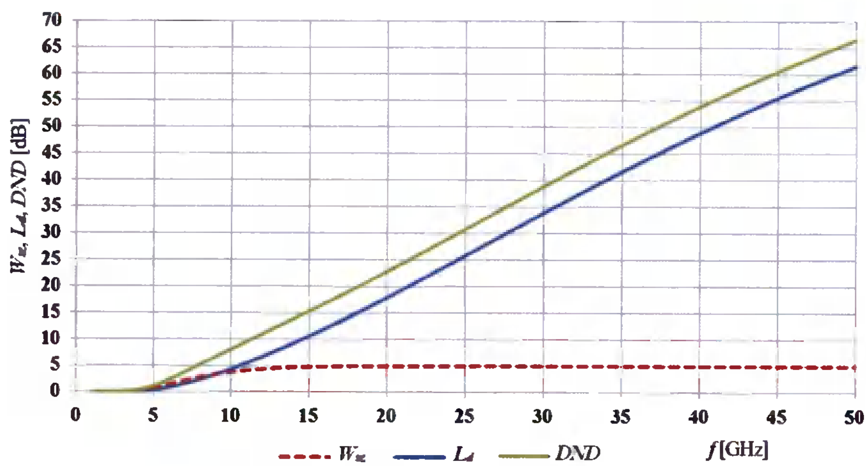

Fig. 11. W_{sz}, L_d, *DND* [dB] vs frequency [GHz] (pol. H)

By the results of the analysis (Fig. 14), we can present and compare the impact of extreme rainfall intensity in Poland on the receive microwave satellite signals in the selected area. On this basis, increase and decrease in L_d was determined, which refers to the rainfall statistics in the area of Kielce city (Fig. 15).

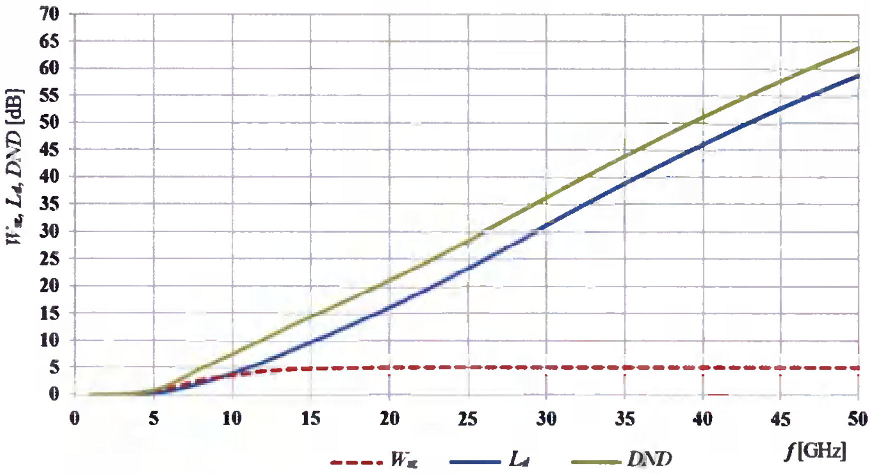

Fig. 12. W_{sz}, L_d, DND [dB] vs frequency [GHz] (pol. V)

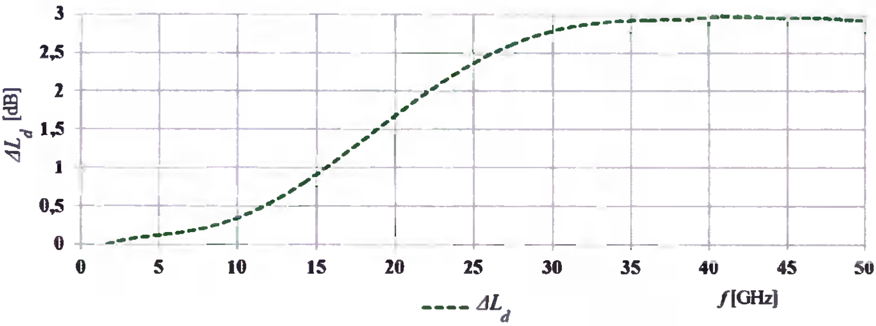

Fig. 13. Increase in L_d [dB] vs frequency [GHz] (pol. H-V)

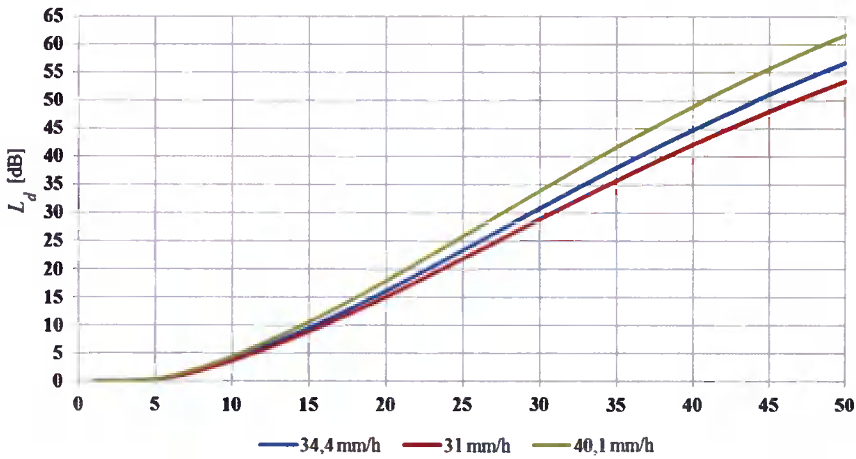

Fig. 14. L_d [dB] vs frequency [GHz] (pol. H)

Increase in L_d for frequencies below 10 GHz does not exceed the value of 0,5 dB, for frequencies below 14 GHz it does not exceed the value of 1 dB, for frequencies below 21 GHz – 2 dB, for frequencies below 29 GHz – 3 dB and for frequencies below 37 GHz – 4 dB (in the case of rainfall statistics of Nowy Targ). Starting from frequency which is equal to 38 GHz, the attenuation increases with frequency and it equals at least 4 dB.

On the other hand (in the case of rainfall statistics of Gorzow Wielkopolski i Gubin), decrease in L_d at 3,4 mm/h, which refers to rainfall statistics in the area of Kielce city, for frequency equals 50 GHz is responsible for the decrease in signal attenuation at 3,2 dB, below 12 GHz – 0,5 dB, below 18 GHz – 1 dB, below 30 GHz – 2 dB, below 45 GHz – 3 dB.

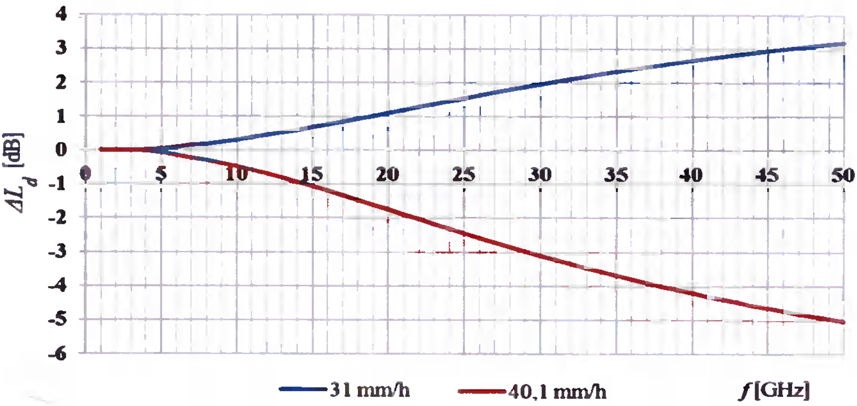

Fig. 15. Increase in L_d, decrease in L_d [dB] vs frequency [GHz] (pol. H)

Below in Figure 16 and Figure 17 you can see the analogically determined statistics for vertically polarized radio waves (versus frequency).

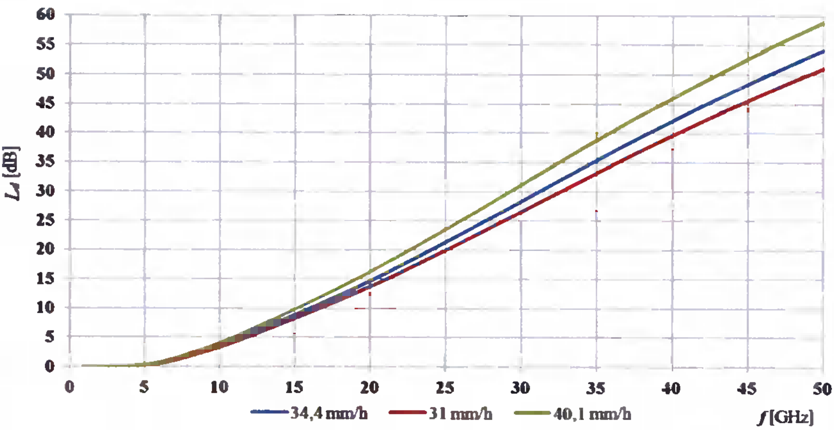

Fig. 16. L_d [dB] vs frequency [GHz] (pol. V)

If we summarize, we can see that the signal attenuation L_d for vertically polarized radio waves, as previously mentioned, reaches a lower value in comparison with horizontally polarized radio waves, so the influence of rainfall on their propagation becomes less significant. Decrease in precipitation of 3,4 mm/h (up to the level of Gorzow Wielkopolski and Gubin) is responsible for decrease in the attenuation of the vertically polarized radio waves in comparison with data obtained from the actual rainfall statistics. It fluctuates around 3 dB at 50 GHz. For microwaves with frequency up to 13 GHz – noted a difference less than 0,5 dB, for microwaves with frequency up to 20 GHz – it did not exceed 1 dB, for frequency up to 32 GHz – the difference is less than 2 dB and for frequency up to 49 GHz – it did not exceed the value of 3 dB.

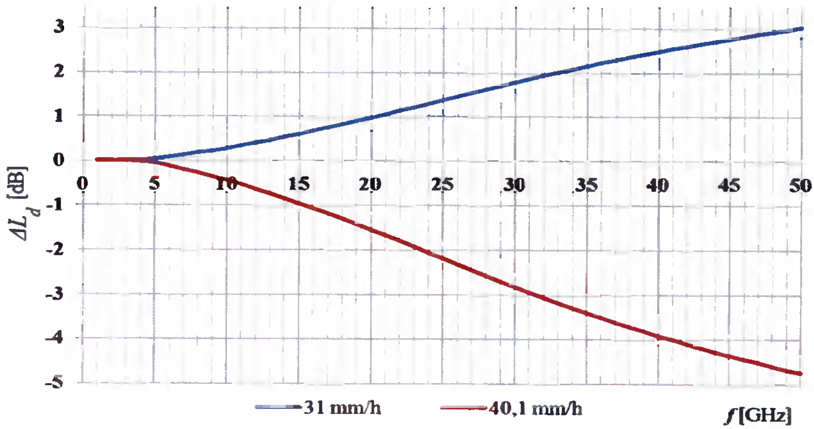

Fig. 17. Increase in L_d, decrease in L_d [dB] vs frequency [GHz] (pol. V)

Moreover, increase in the rainfall in the area of Kielce at 5,7 mm/h for vertically polarized radio waves (up to the level of Nowy Targ), in similar to the horizontally polarized radio waves, gives rise to increase in the signal attenuation relate to the rainfall statistics obtained in the area of Kielce city and Gorzow Wielkopolski and Gubin, too. In the frequency range from 1 GHz to 50 GHz attenuation does not exceed 5 dB (below 50 GHz). For the microwaves with frequency up to 10 GHz, the attenuation is less than 0,5 dB, up to 15 GHz – is less than 1 dB, below 23 GHz – is less than 2 dB , up to 29 GHz – it does not exceed 3 dB and below 41 GHz – is less than 4 dB. Above 41 GHz, attenuation was at least of 4 dB. The results seem to suggest that below 10 GHz (as Figure 15 and Figure 17 show) the impact of radio wave polarization on the microwave satellite signal reception is not meaningful.

3. CONCLUSION

The troposphere as the region of the atmosphere is very important for the mechanism of radio waves propagation [2,9]. In this context, this layer intervenes in meteorological phenomena exert the influence on radio waves especially due to rain. The troposphere, unlike the ionosphere, remains largely unpredictable. However, rain attenuation estimates can be made which allow links to be implemented and tested, obey certain laws of nature. Obviously dry seasons with low precipitation would not suffer greatly from this problem. Due to the wavelengths which are usually at least extremely short, a small-scale detailed description of this layer is quite often required. These losses (signal attenuation, increase in noise and total signal degradation due to precipitation), which increase with frequency was presented in Figure 1-17 in Poland (especially in Kielce city). Moreover, close by this city from 1974 at Psary-Kąty, was a large satellite ground station, operated by TP SA, with up to seven large parabolic antennas. Measurements in this region seems to be a good indication for system planners doing signal attenuation. The area of Kielce city is the representative region especially due to the central location in Poland, environment and morphology of terrain. Because of not spherically shape of hydrometeors, they can result in changes in the polarization, which is the another individual problem. The results of the analysis may be important when high reliability of microwave link is needed. If attenuation exceeds the link budget of a selected link, then (for example) an alternative technique is required to overcome the attenuation problem in microwaves link engineering. This is especially useful for satellite communications to have practical applications which has become more prominent as the communication community moves into higher spectrum of the available bandwidth, which in a future applications become available due to the optimization of satellite equipment [3,4,5,6,7]. The role is also to ensure the rational, efficient and economical use of the radio-frequency spectrum by many types of the radiocommunication services and carry out studies without limit of frequencies on the basis of the series of ITU-R Recommendations.

BIBLIOGRAPHY

[1] Memorandum of Understanding for the implementation of a European Concerted Research Action designated as COST Action IC0802 'Propagation tools and data for integrated Telecommunication, Navigation and Earth Observation systems', http://w3.cost.eu/fileadmin/domain_files/ICT/Action_IC0802/mou/IC0802-e.pdf

[2] Wilk J.Ł., 2010. Naturalne źródła szumów w transmisji satelitarnej. [W:] Rola Informatyki w Naukach Ekonomicznych i Społecznych. Innowacje i implikacje interdyscyplinarne, t. II, pod red. T. Grabiński. Wyd. Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach.

[3] Wilk J.Ł., 2011. The impact of radiowave polarization, frequency and rain intensity on the satellite signal reception in the area of Kielce city. [W:] TRANSCOM 2013, 10th European Conference of Young Research and Scientific Workers, Proceedings, Section 3 – Information And Communication Technologies. EDIS-Žilina University publisher, Žilina.

[4] Wilk J.Ł., 2010. The influence of the antenna aperture on the quality of a satellite signal. [w:] XII International PhD Workshop OWD 2010, pod red. G. Kłapyty. PTETiS, vol. 28, Wisła.

[5] Wilk J.Ł., 2011. The influence of the antenna parameters on a digital satellite signal reception, CEEPUS CII-CZ-0404-01-1011-M-43894, Brno.

[6] Wilk J.Ł., 2011. The influence of the dish efficiency on the quality of a satellite signal, [W:] TRANSCOM 2011, 9th European Conference of Young Research and Scientific Workers, Proceedings, Section 3 – Information And Communication Technologies, EDIS-Žilina University publisher, Žilina.

[7] Wilk J.Ł., 2011. The measurment processing of satellite signal, CEEPUS CII-CZ-0404-01-1011-M-43894, Brno.

[8] Wilk J.Ł., 2012. Współpraca naukowa w ramach Projektu Europejskiego COST IC0802. [W:] Wschód i Zachód w wymiarze globalnym. Doświadczenia z przeszłości a perspektywy na przyszłość, pod red. M. Miłek, G. Wilk-Jakubowski, R.S. Brzoza. Wyd. Stowarzyszenia Współpracy Polska-Wschód. Oddział Świętokrzyski, Kielce.
[9] Wilk J.Ł., 2010. Wybrane zagadnienia dotyczące szumów w komunikacji satelitarnej. [W:] Rola Informatyki w Naukach Ekonomicznych i Społecznych. Innowacje i implikacje interdyscyplinarne, t. II, pod red. T. Grabiński. Wyd. Wyższej Szkoły Handlowej im. B. Markowskiego w Kielcach.

CAŁKOWITA DEGRADACJA SYGNAŁU WKUTEK WYSTĄPIENIA OPADÓW DESZCZU W OBSZARZE MIASTA KIELCE

Streszczenie

W artykule omówiono różne efekty przyczyniające się do całkowitej degradacji sygnału, wskutek wystąpienia opadów deszczu na drodze propagacji fali radiowej oraz przeanalizowano tłumienie atmosferyczne. Skupiono się na analizie propagacji fal radiowych w systemach łączności satelitarnej w troposferze ziemskiej. W praktyce, obecność opadów znacząco wpływa na jakość sygnałów mikrofalowych, szczególnie ze względu na zjawisko absorpcji i refrakcji występujące na kroplach deszczu, stanowiących silne źródło szumów. Analizy dotyczą tłumienia sygnału wskutek wystąpienia opadów deszczu, towarzyszącego ich wystąpieniu wzrostu szumów systemowych i całkowitej degradacji sygnału. W ujęciu praktycznym, tylko poprzez statystyczny opis możliwe staje się pokazanie wyników modelowania mikrofalowych łączy satelitarnych w warunkach opadów deszczu w obszarze miasta Kielce. Otrzymane wyniki porównano z danymi rzeczywistymi, dotyczącymi np. tłumienia fal radiowych w zależności od intensywności opadów deszczu. Dane te posłużyły do określenia wpływu intensywności opadów, polaryzacji i częstotliwości fal radiowych na jakość odbieranego sygnału satelitarnego oraz poprzez przekazanie informacji – do optymalnego projektowania systemów telekomunikacyjnych. Wyniki badań są przydatne dla inżynierów łączności satelitarnej, ponieważ mogą przyczynić się optymalizacji i poprawy wydajności łączy satelitarnych oraz zminimalizowania ryzyka utraty sygnału lub przerw w ciągłości świadczenia usług, co umożliwia zaoferowanie usług satelitarnych na najwyższym poziomie.

Słowa kluczowe: tłumienie atmosferyczne, tłumienie sygnału wskutek wystąpienia opadów deszczu, wzrost szumów systemowych wskutek wystąpienia opadów deszczu, całkowita degradacja sygnału wskutek wystąpienia opadów deszczu, zastosowanie regresji wielomianowej, propagacja fal radiowych w troposferze ziemskiej.

UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY
ZESZYTY NAUKOWE NR 262
ELEKTRONIKA I TELEKOMUNIKACJA 17 (2013), 17-26

RADON-ZERNIKE MOMENTS METHOD APPLIED IN POSTAL APPLICATION

Mirosław Miciak, Roman Wiatr, Tomasz Andrysiak, Mirosław Maszewski

Institute of Telecommunications and Computer Science
Faculty of Telecommunications, Computer Science and Electrical Engineering
University of Technology and Life Sciences (UTP)
ul. Kaliskiego 7, 85-789 Bydgoszcz, Poland
[miciak, rowiat, andrys, mmasz]@utp.edu.pl

Summary: In this paper a new method of a handwritten characters recognition is introduced. The proposed algorithm is applied to classification of post mails on the basis of postal code information. In connection with this work the research was conducted with numeric characters used in real post code of mail pieces. Moreover, the article contains image processing, for instance, filtration of Radon transformation of the character. The main objective of this article is to use the Radon transform parameter space to obtain a set of moment features on basis of which postal code will be recognized.

Keywords: Character recognition, Radon transform, Zernike moments.

1. INTRODUCTION

The today's systems of automatic sorting of the post (Fig. 1) mainly use the OCR (Optical Character Recognition) mechanisms. In the present, recognizing of addresses (particularly written by hand) is still insufficient and needs to be improved.

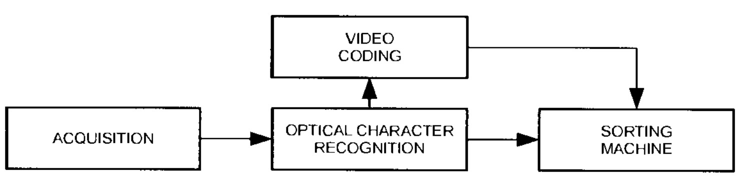

Fig. 1. The automatic sorting system – mail flow

The typical system of post mails sorting consists of the image acquisition unit, video coding unit and OCR unit. The image acquisition unit sends the mail piece image for interpretation. If the OCR unit is able to provide the sort of information required (this technology has 50% effectiveness for all mails [7]), it sends this data to the sorting system, otherwise the image of the mail pieces is sent to the video coding unit where the operator writes down the information about mail pieces. The main problem is that operators of the video coding unit have lower throughput and induce higher costs [7]. Therefore, the OCR module is improving, particularly in the field of recognition of the characters [16]. Although these satisfactory results were received for printed writing, the handwriting is still difficult to recognize. Taking into consideration the fact that manually described mail pieces make over 30% of the whole mainstream it is important to improve the possibility of segment recognizing the handwriting [20]. This paper presents the proposal of a system for recognition of handwritten characters for reading post code from mail pieces.

2. PROPOSED SYSTEM OVERVIEW

The process of character recognition (Fig. 2) can be divided into stages: preprocessing, Radon transform calculating, Zernike moment calculation, feature vector building, and character recognition stage. The first step of the image processing is preprocessing stage, where the colourful image mainly represented by 3 coefficients (red, green and blue) from the acquisition unit must be converted to the image with grey scale. The next step of processing of the image of postal code character is Radon transform. As a result, we obtain the parametric space where subsequent operations will be carried out. The main attention is drawn to the possibility of calculating the Zernike moments as features of the post mail code characters on the basis of the Radon parameter space. Application of Radon transform allows skipping several stages of preprocessing occurring in other publications, such as: binarization, filtration and normalization [16,19]. An additional advantage of this approach is that we do not need to realise the inverse Radon transform to calculate the image moments of the post mail code characters.

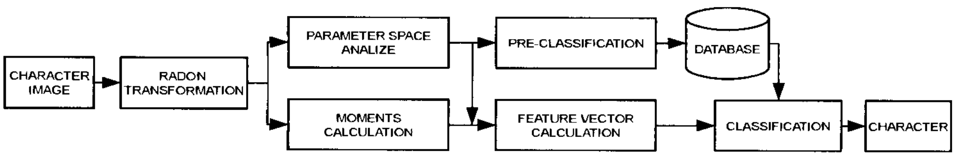

Fig. 2. The proposed method of character recognition

3. RADON TRANSFORMATION

In recent years the Radon transform has received much attention. This transform is able to transform two dimensional images with lines into a domain of possible line parameters, where each line in the image will give a peak positioned at the corresponding line parameters. This have lead to many line detection applications within image processing and computer vision. The Radon transformation is a fundamental tool which is used in various applications such as radar imaging, geophysical imaging, nondestructive testing and medical imaging [14,19]. The Radon transform computes projections of an image matrix along specified directions. A projection of a two-dimensional function $f(x, y)$ is a set of line integrals. The Radon function computes the line integrals from multiple sources along parallel paths, or beams, in a certain direction. The beams are spaced 1 pixel unit apart. To represent an image, the Radon function takes multiple, parallel-beam projections of the image from different angles by rotating the source around the centre of the image [2].

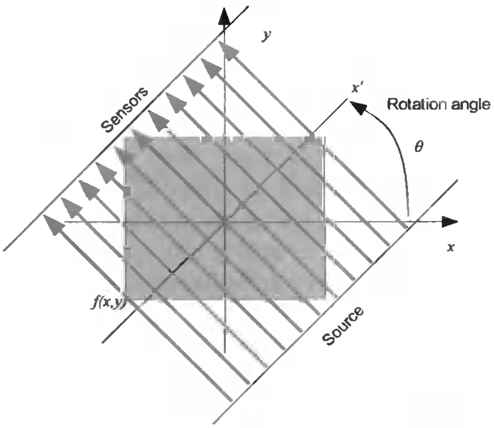

Fig. 3. Single projection at a specified rotation angle

Figure 3 shows a single projection at a specified rotation angle. The Radon transform is the projection of the image intensity along a radial line oriented at a specific angle. The radial coordinates are the values along the $x'-axis$, which is oriented at θ degrees counter clockwise from the $x'-axis$. The origin of both axes is the center pixel of the image.

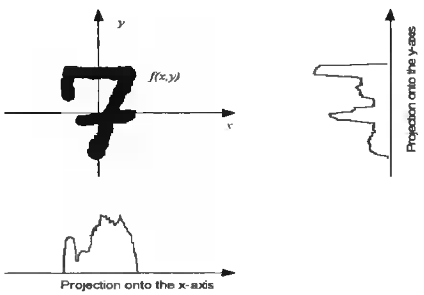

Fig. 4. Horizontal and vertical projections of a simple function

For example, the line integral of $f(x,y)$ in the vertical direction is the projection of $f(x,y)$ onto the $x'-axis$; the line integral in the horizontal direction is the projection of $f(x,y)$ onto the $y'-axis$. Figure 4 shows horizontal and vertical projections for a simple two-dimensional function.

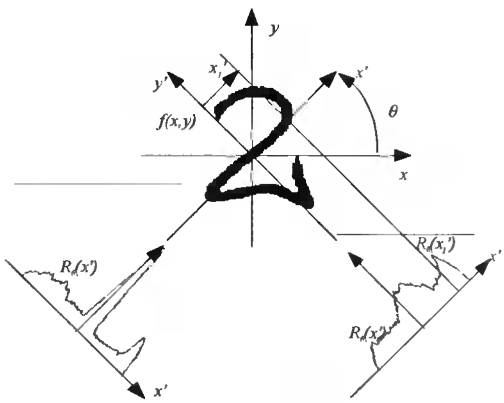

Fig. 5. Geometry of the Radon transform.

Projections can be computed along any angle θ by use of general equation of the Radon transformation [3,10]:

$$R_{\Theta}(x') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\delta(x\cos\Theta + y\sin\Theta - x')dxdy \quad (1)$$

where $\delta(\cdot)$ is the delta function with value not equal zero only for argument equal 0, and:

$$x' = x\cos\Theta + y\sin\Theta \quad (2)$$

where x' is the perpendicular distance of the beam from the origin, and θ is the angle of incidence of the beams. The Fig. 5 illustrates the geometry of the Radon transformation. The Fig. 6 shows sample of accumulator data of Radon transformation.

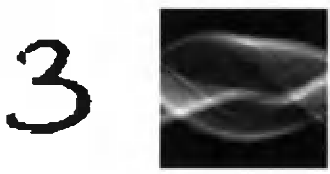

Fig. 6. Sample of accumulator data of Radon transformation.

The very strong property of the Radon transform is the ability to extract lines (curves in general) from very noise images. Radon transform has some interesting properties relating to the application of affine transformations. We can compute the Radon transform of any translated, rotated or scaled image, knowing the Radon transform of the original image and the parameters of the affine transformation applied to it. This is a very interesting property for symbol representation because it permits to distinguish between transformed objects, but we can also know if two objects are related by an affine transformation by analyzing their Radon transforms [15]. It is also possible to generalize the Radon transform in order to detect parameterized curves with non-linear behavior [6, 11,14].

4. ANALYSIS OF THE RADON PARAMETER SPACE

In the next stage of processing, Radon parameter space (accumulator) is analyzed to receive the vector of features of the character. To obtain one of the first part of vector's parameter we used the local peaks of the accumulator. Therefore, the initial accumulator data reduction is fulfilled by matching of local maximum operation. As a result of this stage, the amount of local peaks is limited as to the value of the threshold, which minimizes the intra-class variance and defined as a weighted sum of variances of the two classes, calculated basis on histogram for all accumulator cells. A second restraint is the distance between each of peak, which is determined on basis double value of mean thickness of the characters (keeping the unitary of scale transformation). Thus, on this stage we can notice that a series of characters are divided into some class because of the peaks amount. The Fig. 7 depicts a Radon transformation accumulators local peaks and their localization for sample characters.

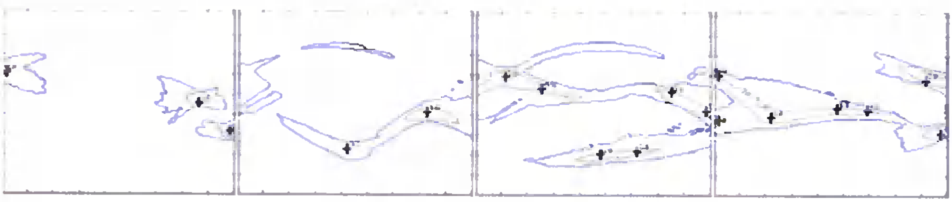

Fig. 7. The accumulator peaks and their localization for few digits

5. MOMENT CALCULATION STAGE

The group of moments calculated on the basis of radial functions are parameters determined on the basis of the Zernike polynomials. For the function of two real variables (such as an image) the polynomials can be written in the general form:

$$Z_{nm} = \frac{n+1}{\pi}\int_0^{2\pi}\int_0^1 V_{nm}^*(r,\theta) f(r,\theta) dr d\theta \tag{3}$$

where V is the complex Zernike polynomial.

Figure 8 shows Zernike polynomials of various orders *mn*.

We considered the use of polynomial $P_{mn}(s)$ with coefficients A_{nmk} [8]:

$$P_{nm}(s) = \sum_{k=0}^{n} A_{nmk} s^k \tag{4}$$

by substituting the general form of radial polynomials we get:

$$\int_0^1 \int_0^{2\pi} P_{mn}(s) e^{0im\phi} g(r,\theta) ds d\phi = \sum_{k=0}^{n} A_{nmk} H_{km} \tag{5}$$

where $g(r,\theta)$ is a parametric space of Radon transformation.

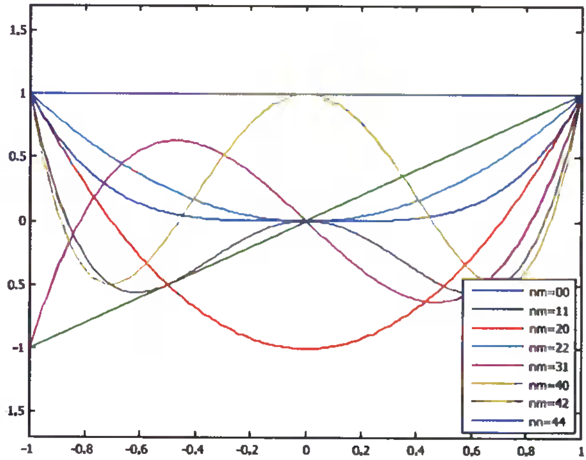

Fig. 8. Zernike polynomials of various orders *mn*.

We can determine the coefficients A_{nmk} in the form:

$$A_{nmk} = (-1)^{(n-k)/2} \frac{n+1}{2\pi^2} \frac{2}{k!} \frac{((n+k)/2)!}{((n-k)/2)!} \tag{6}$$

applying the [4]:

$$P_n(s) = \frac{n+1}{2\pi^2} \sum_{k=0}^{n}{}'(-1)^{\frac{n-k}{2}} \frac{\frac{n+k}{2}!}{k!\frac{n-k}{2}!}(2s)^k \tag{7}$$

we prove that the sum:

$$\sum_{k=0}^{n}{}'(-1)^{\frac{n-k}{2}} \frac{\frac{n+k}{2}!}{k!\frac{n-k}{2}!}(2s)^k \tag{8}$$

is the Chebyshew polynomial second kind order of n denote as $U(s)$ [8]. This will determine the complex Zernike moments which can be represented as:

$$Zc_{nm} = \frac{m+1}{2\pi} \int_0^{2\pi} \int_0^1 U_n(s) e^{-im\phi} g(r,\phi) ds d\phi \tag{9}$$

where

$$|m| \le n, \frac{n-|m|}{2} \in N. \tag{10}$$

Defined in this way, moment values allow us to formulate the feature vector of the postal code character consisting of modules of the Chebyshew polynomials Zc_{mn}. For each character feature vector consist of 8 values, defined as:

$$FV_{Mcz} = \{Mcz_{00}, Mcz_{11} Mcz_{20} Mcz_{22} Mcz_{31} Mcz_{33} Mcz_{40} Mcz_{42}\}. \tag{11}$$

In order to observe changes in the value of the feature vector of character postal code was rotated. The Fig. 9 shows change in the feature vector of the rotated character image (presented on Fig. 10).

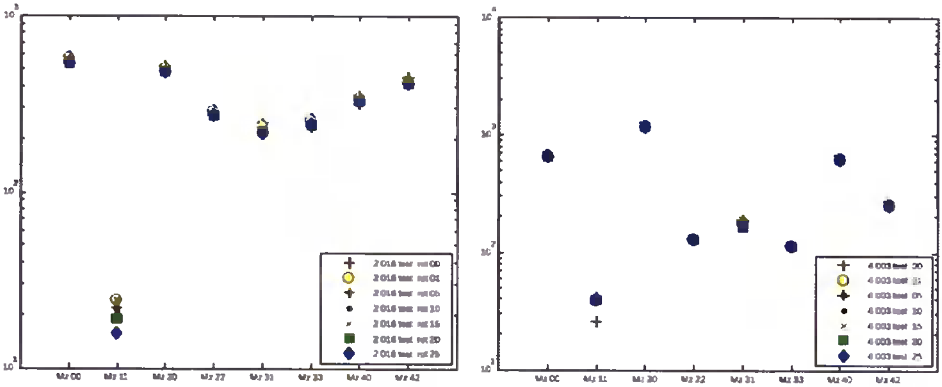

Fig. 9. The changes of FV depending from rotation of character image.

Additionally, beside the number of peaks LP from the Radon parameter space analysis stage, vector contains code of known character ZN as a Unicode [17]. After all, the feature vector consists of values $FV = \{ZN, LP, FV_{Mcz}\}$ for each character from training set.

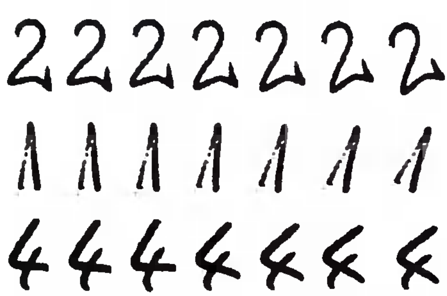

Fig. 10. Sample of rotated digits form postal codes.

6. MODULE OF THE PRE-CLASSIFICATION

The aim of the preliminary classification is to reduce the number of possible candidates for an unknown character to a subset of the total character set. For this purpose, the selected domain is categorized into six groups with number of local maximum. It is worth mentioning that, after applying the preliminary classification, the number of wrongly classified characters was decreased. The analysis of the initial classification groups does not allow indicating the clear membership rules classes of character but rather may show their geometrical features. The preliminary classification is based on the amount of local maximum calculation in the Radon parameter space analysis stage.

7. CHARACTER CLASSIFICATION STAGE

The classification in the recognition module compared features from the pattern to model features sets obtained during the learning process. Based on the feature vector *FV* recognition, the classification attempts to identify the character based on the calculation of Euclidean distance between the features of the character and of the character models [1,18]. The distance function is given by:

$$D(PR,CR)=\sum_{j=1}^{N}\left[FV_B(j)-FV_R(j)\right]^2 \tag{12}$$

where:

- PR – is the predefined character,
- CR – is the character to be recognized,
- FV_B – is the feature vector of the predefined character,
- FV_R – is the feature vector of the character to be recognized,
- N – is the number of features.

The minimum distance D between unknown character feature and predefined class of the characters is the criterion choice [1].

8. EXPERIMENTS RESULTS

For evaluation experiments, we extracted some digit data from various paper documents from different sources e.g. mail pieces post code, bank cheque, etc. In total, the training datasets contain the digit patterns of above 200 writers. Our database include 1440 different digits patterns for training set and 600 digits for testing set. Comparing the results for handwritten character with other researches is a difficult task because there are differences in experimental methodology, experimental settings and handwriting database. In [12] authors presented a handwritten character recognition system with modified quadratic discriminant function, for which they recorded recognition rate of above 98%. Work [9] employed Hidden Markov Models for digits recognition. They obtained a recognition rate of 87%. The method described in [1] using Normalized Fourier Descriptors for character recognition obtained a recognition rate above 96%. Bellili using the MLP-SVM achieves a recognition rate of 98% for real mail zip code digits recognition task [5]. For comparison, proposed in this article method (Radon-Zernike) reaches recognition rate above 94% (see Tab. 1). Learning to recognition rate is the ratio of the number of learning characters to the number of testing characters. The results were compared with the methods Zernike-Image and Radon Profil [13].

Table 1. Character recognition rate for various testing set size.

	Learning to testing ratio				
	30	25	20	15	10
Method	Character recognition rate [%]				
Radon-Zernike	91,3	92.1	93,5	94,1	94,6
Zernike-Image	90,5	91,7	92,7	94,2	94,5
Radon-Profil	89,3	90,7	91,2	93,7	93,9

The Table 2 below shows the moment calculation time for different reference methods.

Table 2. The moment calculation time for different methods.

Image size (nxn)	Zernike moments	Radon-Zernike
64	0,231	0,535
96	0,524	0,783
128	0,936	1,044
160	1,461	1,316
192	2,110	1,573
224	2,874	1,841
256	3,770	2,112

9. CONCLUSIONS AND FUTURE WORK

Selecting the features for character recognition can be problematic. Moreover, the fact that the mail pieces have different sizes, shapes, layouts, etc., makes this process more complicated. The paper describes often used character image processing such as Radon transformation calculating. The main objective of this article is to use the Radon transform parameter space to obtain a set of moment features on basis of which postal code will be recognized. The presented method is less computationally complex for larger images (Tab. 2). For the future work, presented method will be upgraded to remaining all alphanumeric signs and special signs often placed on regular post mails.

BIBLIOGRAPHY

[1] Aissaoui A., 1999. Normalised Fourier Coefficients for Cursive Arabic Script recognition, Universite Mohamed, Morocco.

[2] Asano A., 2002. Radon transformation and projection theorem, Topic 5, Lecture notes of subject Pattern information processing, http://kuva.mis.hiroshima-u.ac.jp/~asano/Kougi/02a/PIP/

[3] Averbuch A., Coifman R.R., 2001. Fast Slant Stack: A notion of Radon Transform for Data in a Cartesian Grid which is Rapidly Computible, SIAM J. Scientific Computing.

[4] Poularikas A.D., 1990. Chebyshew Polynomials, The Handbook of Formulas and Tables for Signal Processing, CRC Press LLC.

[5] Bellili A.; Giloux M. (2003). An MLP-SVM combination architecture for handwritten digit recognition, International Journal on Document Analysis and Recognition, Springer-Verlag.

[6] Bracewell R.N., 1995. Two-Dimensional Imaging. Englewood Cliffs. Prentice Hall, 505-537.

[7] Forella G., 2000. Word perfect. Postal Technology. UKIP Media & Events Ltd UK.

[8] Mukundan R., Ramakrishnan K.R., 1998. Moment Functions in Image Analysis: Theory and Applications. World Scientific.

[9] Kaufmann G., Bunke H., 2000. Automated Reading of Cheque Amounts. Pattern Analysis & Applications, Springer-Verlag.
[10] Kupce E., Freeman R., 2004. The Radon Transform: A New Scheme for Fast Multidimensional NMR. Concepts in Magnetic Resonance, Wiley Periodicals 22, 4-11.
[11] Lim J.S., 1990. Two-Dimensional Signal and Image Processing. Englewood Cliffs, Prentice Hall, 42-45.
[12] Liu C., Sako H., 2002. Performance evaluation of pattern classifiers for handwritten character recognition. IJDAR, Springer-Verlag.
[13] Miciak M., 2013. Inteligentny system Rozpoznawania i Klasyfikacji Przesyłek pocztowych. Ph.D. Thesis, University of Technology and Life Sciences.
[14] Peter T., 1996. The Radon Transform - Theory and Implementation. PhD thesis, Dept. of Mathematical Modelling Section for DSP of Technical University of Denmark.
[15] Ramos O., Valveny T. E., 2003. Radon Transform for Lineal Symbol Representation. The 7th International Conference on Document Analysis and Recognition.
[16] Srihari S., Govindaraju V., Bouchaffra D., 1999. Recognition of Strings Using Nonstationary Markovian Models: An Application in ZIP Code Recognition. IEEE CSC on CV& PR 2, USA.
[17] The UTF-8 encoding form, 2003. The UTF-8 encoding scheme. UCS Transformation Format 8, defined in Annex D of ISO/IEC 10646:2003.
[18] Turk M.A., Pentland A.P., 1991. Face recognition using eigenfaces. Computer Soc. Conference on Computer Vision and Pattern Recognition, 586-591.
[19] Venturas S., Flaounas I., 2005. Study of Radon Transformation and Application of its Inverse to NMR. Algorithms in Molecular Biology.
[20] Volgunin A., Filatov A., 1997. Handwritten ZIP Code Recognition. 4th International Conference on Document Analysis and Recognition, USA.

ZASTOSOWANIE MOMENTÓW RADONA-ZERNIKE DO APLIKACJI POCZTOWYCH

Streszczenie

W artykule przedstawiono nowe rozwiązanie zadania rozpoznawania znaków pisanych ręcznie dla zastosowań pocztowych. Zaproponowano algorytm klasyfikacji przesyłek pocztowych działający na podstawie informacji zawartej w zapisie kodu pocztowego. Główny nacisk położono na wykorzystanie transformaty Radona i momentów Zernike do uzyskania zbioru cech, na podstawie, których rozpoznawano kod pocztowy. Otrzymane wyniki eksperymentów pozwoliły wykazać skuteczność proponowanej metody.

Słowa kluczowe: rozpoznawanie znaków, transformata Radona, momenty Zernike

UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY
ZESZYTY NAUKOWE NR 262
ELEKTRONIKA I TELEKOMUNIKACJA 17 (2013), 27-36

APPROACH TO WORKING FEATURES DICTIONARY CONSTRUCTION BASED ON DEFINED PRIORITIES OF THE PRIMARY FEATURES

Vasyl M. Zayats, Galyna Ya. Shokyra

National University Lviv Polytechnic, Department Ecology and Information Systems
130 Heneral Chuprynka str., 79057, Lviv, Ukraine
1st academic building, 405 rooms, e-mail: zvm01@rambler.ru

Summary. The simulation model for features dictionary based on the authors developed approach to priority primary features in the process of objects and phenomena recognition process is proposed. The main directions of applied applications results are determined.

Keywords: primary symptoms priority, working dictionary, simulation model

1. DESCRIPTION OF THE MODEL

The paper presents the simulation model and approach to the construction and optimization of working signs dictionary in the process of recognition systems development based on defined signs priority. It enables to optimize working features dictionary by adding features with the highest priority. The initial model for recognition process is the determined features of object recognition, which have specific numeric values or lie in a certain range of values (determinate model) and/or probabilistic features, which are the probability of an object assignment to a class and the conditional probability of sign's relation to the object on condition that it is included in a certain class (probabilistic model). This approach enables under unfavorable terms of the inevitable financial allocations to ensure development of objects or phenomena recognition and identification system with a given accuracy, incomplete a priori information and economy of technical means.

2. CREATION OF PRIORITY FEATURES FOR RECOGNITION OBJECTS

The rational selection of parameters for recognition objects description greatly simplifies the choice of the recognition tools and increases the probability of correct recognition [1 - 10, 11].. It should be indicated that in the process of the automatic pattern recognition (objects written in features language) it is impossible to create a device that provides absolutely accurate object recognition. Model always has some measuring inaccuracy. The main task is to ensure its minimization.

There is always some probability of an error. It should be ensured that the identification device gives the most possible and correct classification on the selected parameters with known determinate or probabilistic properties and specified allowable error probability.

The design and technical implementation of the recognition device, especially the part that serves for perception (measurement) of image description image parameters, depend on the physical nature of the recognition object [1, 2]. The same object can be characterized not only by its geometrical properties (height, width), but also mechanical. electrical, chemical, thermal and other properties. Thus an object can be evaluated from different perspectives. The combination of all features creates the image of the object. Thus, if a certain number of objects have the same features they can be combined into a general class and treated as a single image, nevertheless there are also other features of the same objects that are quite different and such objects should be assigned to different classes. For example, if one defines the image of natural oscillation frequency the naval ship and pendulum belong to the same class of objects but if their geometric size and shape are taken into consideration they should be treated as different classes.

Therefore, at the first stage of the recognition we should determine features priority for images distinction, and accordingly choose physical or any other parameters for objects description.

Methods for selecting features for a variety of application uses are presented in [3-5]. The analysis of selected features is carried out in [6, 7].

For example, in the works dedicated to disease diagnostics [8] as disease features were used, except its symptoms, laboratory analysis data and calculated probabilities of certain features presence in the given disease. It served as a source statistical material for their recognition. As it can be seen from the experiments, the probability of correct disease recognition with the use of formal algorithm in electronic computers is significantly higher than the recognition of the same disease by a medical specialist.

Recognition theory methods are also used for solving the most difficult tasks in minerals classification on the base of geophysical measurements: for example, the process the aquifers and oil reservoirs in boreholes recognition with a very small probability of errors occurrence is described in [9].

The collected statistical material and the use of optimal recognition algorithms method (mean distance for determinate features and strategy for Bayesian probabilistic features) and priority primary features [2, 11, 12] enables to make correct decisions with higher probability.

Thus, the choice of recognition parameters for certain physical quantities is defined by:

a) parameters invariance to various transformations;
b) simplicity of a selected device for parameters measurement and a device that converts the parameters in code helpful in a decision making device;
c) parameters sensitivity (hindrance immunity) to countermeasures;
d) complexity of systems teaching and self-teaching.

While solving the specific recognition problem the most important step is the optimal choice of the most informative parameters set of the recognition image, thus it is determination of features priority. This will determine the better design of the device and broader prospective of its use [10, 11].

Difficult recognition system – heterogeneous information is utilized physically. For example, systems of medical diagnostics (analyses, cardiograms, temperature, pressure).

recognition systems for geological physical, chemical researches. If in difficult recognition system – heterogeneous information is utilized physically. For example, systems of medical diagnostics (analyses, cardiograms, temperature, pressure). Recognition system for geological secret service are physical and chemical properties. If in quality of principle of classification to choose the method of receipt of a posteriori information, the difficult systems are divided by an onelevel and multilevel.

The onelevel difficult systems are a posteriori information about the signs of objects which are recognized determined the direct measuring directly on the basis of treatment of results of experiments.

The multilevel difficult systems are a posteriori information about signs, determined on the basis of the indirect measuring, as a result of functioning of auxiliary recognition system.

If in quality of principle of classification to choose a priori initial information content, recognition system is divided by the systems without studies, which study with self-training. For the multilevel systems such distributing is not synonymous.

The systems are without teaching. Initial information sufficiently in order that on select principle of classification to divide all of plural of objects into classes Ω_i, $i = 1,2,...,m$, to define the dictionary of signs of K_j, $j = 1,2,...,N$, and on the basis of the direct working of weekend of information to describe every class of objects in language of signs.

Systems with teaching. Initial information sufficiently in order that for select principle of classification to divide all of plural of objects into classes Ω_i, $i = 1,2,...,m$, to define the dictionary of signs of K_j, $j = 1,2,...,N$, but not enough, to describe classes in language of signs.

The systems are with self-training. A priori initial information sufficiently for determination of dictionary of signs of $X_1, X_2,..., X_N$. But not enough for conducting of classification, in place of pointing what class an object belongs to, the set of rules is given, in accordance with what RS makes classification. A priori information absents must be produced in the process of studies or self-training.

Regardless of on what principle the system of recognition is built it is important to set priority of the chosen signs, as authenticity of recognition depends on quality of decision of this task. At establishment of priority of signs it follows to go out from such considering:

- informative descriptions of sign are about the object of recognition.
- expenses are on determination of quantitative estimations of sign,
- influence of concrete sign is on authenticity of recognition.
- the use of method of progressive approximations is to establishment of priority.

Task of choice of the most informing subsystem of signs from some initial system – the task of objects recognition theory is important, because:

- reduction of number of signs diminishes different losses, including the losses of useful information,
- reduction of number of signs leads, as a rule, to the improvement of quality of decision [6],
- there is a limit on the number of signs which a deciding rule can be founded on at the fixed sample size. Yes, for quadratic functions and selection of $N = 100$, to take rule governed guilty to contain 10 signs or more.

Let $X = \{x_1, x_2, ..., x_n\}$ it is the virgin system of signs.

It is needed to define the most informing subsystem from m signs $(m < n)$ from the point of view some criterion of F.

Will consider the subsystem of $V = \{x_{t_1}, x_{t_2}, ..., x_{t_m}\}$, $1 \le t_1 \le t_2 \le ... \le t_m \le n$.

Will enter n-dimension vector of $U = \{u_1, u_2, ..., u_n\}$. Then subsystem V it is possible to put in accordance of $U = \{u_1, u_2, ..., u_j, ..., u_n\}$ so, that $u_j = 1$, if $x_j \in V$ and $u_j = 0$, if $x_j \notin V$.

A task of choice of the most informing subsystem of signs is an extreme task on a single hypercube with possible by part number of tops of C_n^m.

$F(u)$ – multiextreme analytical expression of $F(u)$ is unknown.

Examples of task of $F(u)$ are in ORT for the most widespread classes of deciding rules. For the construction of deciding rule a selection is needed (educational sequence). Let this will be the table of

$$V_p = \{X_{ij}^p\},\ p = 1,2,...,k,\ j = 1,2,...,n,\ i = 1,2,...,N.$$

$$\sum_{p=1}^{k} V_p = N,$$

where:

- p – a number of appearance,
- k – an amount of appearances,
- V a sample size,
- N – a number of signs.

$\forall$ subsystem with m signs $\{x_{t_1}, ..., x_{t_m}\}$. With use a table V a some system is selected

$$V_p = \{X_{it}^p\},\ t = t_1, t_2, ..., t_m,\ p = 1,2,...,k.$$

where V_p is a total number of errors of recognition on an educational selection at the use of some fixed class of deciding rules $F(u) = \alpha(u)$.

3. APPROACES TO WORKING DICTIONARY CONSTRACTION IN CONDITIONS OF LIMITED RESOURCES

In the process of features working dictionary construction the following restrictions are met:

1) dictionary includes only features with a priori information necessary for description of the classes,
2) some of the features should not be included into the a priori dictionary because of their low informatively [11],
3) some of the features, which are usually the most informative, can not be identified due to the lack of appropriate measuring devices, and limit of the resources allocated to the recognition process.

Therefore the a priori features dictionary can generally be used only as a basis for building the real dictionary used in the recognition system, which includes only the features, which, on the one hand, are the informative and, on the other hand, can be defined with the help of existing or specially created measurement equipment.

Thus, the task of features working dictionary construction in general is limited to determination of the hardware number needed for features with the highest priority with the help of allocated resources. This priority can be established based on methods and approaches proposed in [12, 13].

It is possibility to increase efficiency of the system of recognition due to the increase of distance between objects from different classes. Then at implementation of some terms on middle square variation of objects into a class $S(\Omega_k)$ a task is taken to finding of maximal value for all μ from a minimum to functional $R(\Omega_k,\Omega_r)$, that it is needed to find max_ min $R(\Omega_k,\Omega_r)$ at limitations μ $k,r=1,...,n$:

$$S(\Omega_k)\le S_k^*,\quad \sum_{j=1}^{N} C_j\mu_j \le C^* .$$

Let $L=C_n^2=n(n-1)/2$ is an amount of possible pair from n classes, the square of distance between the pair of classes is evened

$$R^2(\Omega_k,\Omega_r)=\sum_{i=1}^{N}\mu_i\left\{1/m_k/m_r\sum_{j=1}^{m_k}\sum_{l=1}^{m_r}(x_{kj}{}'-x_{lp}{}')^2\right\},$$

where $k,r=1,...,n$.

Then, entering informing of i sign for the s pair of classes

$$\sigma_{si}=1/m_k/m_r\sum_{j=1}^{m_k}\sum_{l=1}^{m_r}(x_{kj}{}'-x_{lp}{}')^2,\ s=1,...,L,\ i=1,...,N,$$

where the square of distance for a s pair is evened $R_s^2\sum_{j=1}^{N}\mu_j\sigma_{sj}$, $s=1,...,L$, will get a next task [2, 10]:

$$\max_{\mu_j=0,1}\ \min_{1\le s\le L} R_s^2=\max_{\mu\in B}\ \min_{1\le s\le l}(\mu,\sigma_s)$$

$$(C,\mu)=\sum_{l=1}^{N}\mu_l C_l < C^*,$$

where $\sigma_s=(\sigma_{s1},\sigma_{s2},...,\sigma_{sl}),\mu=(\mu_1,\mu_2,...,\mu_N)$.

In investigation of continuity of function $\min_{s=1,\dots,L}(\mu,\sigma_s)$, and also to simple maxmin have the method of report of maxmin, that just the following assertion:

$$\max\ \min(\mu,\sigma_s)=\lim\ \max\Psi(\mu,\sigma),$$

where $D = L_1 \times [0, \max \sum_{j=1}^{N} \sigma_{sj}]$ is cartesian product N-dimension a cube and segment $[0, \max \sum_{j=1}^{N} \sigma_{sj}]$. It is possible to take in quality a function $\Psi_{s\,\overline{1,l}}(\mu, v)$, for example, function

$$\Psi(\mu, v) = v - k \sum_{s=1}^{l} \left[(\mu, \sigma_s) + kG(\mu) - v - \left| (\mu, \sigma_s) + kG(\mu) - v \right| \right]^2,$$

where

$$G(\mu) = \sum_{j=1}^{N} (\mu_j^2 - \mu_j) - \left[C^* - (\mu, C) - \left| C^* - (\mu, C) \right| \right]^2,$$

and μ examined on the N-dimension single cube of $L_1 = \{\mu : 0 \le \mu_j \le 1; j = 1, \dots, N\}$.
It is modification to show that $\forall \varepsilon > 0 \exists K^*$ such, that $\forall k > K^*$ k just inequality

$$\max_{\mu \in B} \min_{s=\overline{1,L}} (\mu, \sigma_s) - \min_s (\mu_{\max}(k), \sigma_s) < \varepsilon .$$

Will consider a case, when except for a priori information on the location of objects in N-dimension of space and values of signs of objects the known and a priori probability of appearance of objects from the classes Ω_k, Ω_r of $k, r = 1, \dots, n$, that determine $P(\Omega_k)$ and $P(\Omega_r)$.

Then, assuming independence of these events, a priori probability of appearance of objects of both classes

$$P_s = P(\Omega_k) * P(\Omega_r), \quad s = 1, \dots L, \quad and \ L = C_n^2 .$$

Now a task can be formulated as follows: to find such set of working dictionary of signs, which will provide the maximal value of the expected value

$$M[R_s^2] = \sum_{s=1}^{l} R_s^2 * P_s$$

at limits on resources

$$(C, \mu) = \sum_{l=1}^{N} \mu_l C_l < C^* .$$

On condition that efficiency of the system of recognition can be attained only due to the increase of relation of distance between classes to middle square variation of objects into classes, the functional Φ can look like [2, 11]

$$\Phi = R^2(\Omega_k, \Omega_r) / S(\Omega_k) / S(\Omega_r).$$

In some cases the task of finding of working dictionary of signs succeeds to be untied due to more compact location of objects of every class, that by minimization of maximal variation of objects into a class. In this case it is possible to choose functional in a kind

$$\Phi = S(\Omega_k), \quad k = 1,...,n \ .$$

signs by the sequential removal of signs. The basis algorithm deterministic priority signs recognition is the process of establishing recognition error in sequential withdrawal symptoms.

Let's consider the process of recognition. Suppose that we have an object with the following set of features x = { $x_1, x_2,...x_n$ }, working dictionary that contains classes and class objects with descriptions of all the features of these classes ($\Omega_i, i = \overline{1,m}, \omega_{ij}, i = \overline{1,m}, j = \overline{1,k}, a_i, i = \overline{1,b}$).Let's take a recognition process using a distance function. This simple classification method is a very effective tool for solving such problems in which classes are characterized by degree changes limited to certain limits. Suppose that each class can be described by not the only but by multiple reference images, that is, any image that belongs to a class tends to group around one of the benchmarks $z_i^1, z_i^2, ..., z_i^{N_i}$, where N_i - number of reference images that define the i-th class. Function, which determines the distance between an arbitrary image x and class ω_i will look as:

$$D = \min_l \left\| x - z_i^l \right\|, \quad l = 1,2,..., N_i;$$

It means that D_i is the smallest of the distances from each image to class ω_i standard. Let's compute the value of distances D_i, $i = 1,2,...,M$, and classifying image belongs to the class ω_i if the condition is $D_i < D_j$ fair for all $j \neq i$. In case of equal distances decision is made randomly.

In case of equal distances decision is made randomly.

The next step is the establishment of the recognition error. Ideally, the error = 0, the characteristic value of the investigated object would match the average values with feature vector class:

$$k_s = \frac{x_{i,2} - x_{i,1}}{2}.$$

But it is quite rare, so the conditional interpretation that the average grade point is ideal and error estimation at these points = 0 is introduced. Instead deviation values of attributes of the object will be closer to 50%, this is due to the fact that the working dictionary attributes the difference between the features of different classes can be "1". Therefore:

$$S_s = \sum_i (k_{i,2} - k_s)^2;$$

Error estimating can be presented:

$$K = \frac{S_s}{50}.$$

The error recognition process is equal to:

$$P = 50 - \frac{S_{min}}{K},$$

where S_{min} the minimum value of properties belonging to a known class.

The next step is the removing if one of the signs performing the recognition process with maintaining all the results. This process is carried out until all the symptoms of the object are disguised.

The next step is the comparison of the results of alternate recognition without one feature and pre result taking into account all attributes. First the signs, the absence of which changed the outcome of the test object class membership ω_i, are selected. They will be more informative. The next step will be the selection of the signs that led to controversial situation. They will have the medium priority, because their absence causes a partial error (50%) of belonging to a certain class. Signs that remain will be understood as little information as their absence does not affect the result of recognition (the error is 0%).

Identify of the priorities is adding new features to the selected base. This method is rather cumbersome, because it would include $\frac{(n-1)!}{2}$ operations, while only preliminary *(n+1)* operations. To reduce the number of operations we should consider combining those features which give absence of 100 or 50% error. It will reduce the number of operations $\frac{(t-1)!}{2}$ where $t << n$ much. Algorithm of realization is similar to sorting algorithm in sequential shutdown of one trait.

Thus, we can identify the most informative features and clarify the working vocabulary of signs. The significant solving of this problem is extremely important, as it is affecting on the efficiency of the recognition system and is appropriate in decision management.

Based on the proposed algorithm, we can select the priority signs as the most informative features and clarify the working vocabulary of signs. The significant solving of this problem is extremely important, as it is affecting on the efficiency of the recognition system and is appropriate in decision management.

On the basis of these studies, the real technical limitations and simulation results of the developed system for determinate and probabilistic objects computer recognition was determined the approach to features working dictionary construction. The main outcomes of this approach are:

- recognition process will be more accurate if the objects belonging to each of the given classes in feature space will be compactly arranged, and the distance between the classes will be large enough,

- with the increase of classes number, features with less variation values should be incorporated into the dictionary in order to reduce the average risk,
- reducing the recognition efficiency by increasing the number of classes can be compensated by the increase of feature vector dimension, but it requires additional financial means,
- while solving the recognition efficiency problem by increasing the feature vector dimension much attention should be paid to the fact that the described increase is largely increasing number of measurement means, each of which provides definition of one or a group of features. It requires additional costs for building recognition systems and resources are often limited.

In conditions of the limited resources on recognition systems (limited financial resources; unavailable necessary measurement hardware; uninformative, non-priority features) only a compromise between the size of the classes alphabet and features dictionary volume provides a solution to the recognition problem and correct objects and phenomena identification in the most optimal way [3, 9. 11, 13].

4. CONCLUSIONS

Application of the proposed model for features working dictionary construction after their priority determination enables to optimize the procedure of the working dictionary formation, keeping only the most informative of the features. This allows effective implementation of the recognition procedure under incomplete information space and with already determined accuracy. This significantly increases the reliability of the recognition system.

BIBLIOGRAPHY

[1] Пересада В.П.. 1970. Автоматическое распознавание образов. / В.П. Пересада, Л.: Энергия, 92 с.

[2] Заяць В.М., 2003. Методи розпізнавання образів. / В.М. Заяць, Р.М. Камінський, Навчальний посібник. Львів: Львівська політехніка, 2004, 176 с.

[3] Аникина Г.О.,. Непараметрические алгоритмы распознавания образов в пространстве дискретных признаков. / Г.О. Аникина, Информатика и системы управления: межвуз. сб. науч. тр. Красноярск: ГУ НИИ ИПУ КГТУ, Вып. 9. с. 55-64.

[4] Барило К.В., 2011. Алгоритм оптимізації рівня селекції координат еталонних векторів класів розпізнавання. / К.В. Барило, Тези доповідей Третьої міжнародної науково-практичної конференції (Суми, 2-4 листопада 2011 року). Інтелектуальні системи в промисловості і освіті (ІСПО), Суми «Видавництво СумДУ», Том I, с. 12-13.

[5] Тарасов Е.М.. 2012. Выбор признаков и решающих правил при распознавании состояний рельсовых линий сортировочных горок. / Е.М. Тарасов, М.В. Трошина. Уфа: УГАТУ, Т. 16, № 3 (48), с. 194-197.

[6] Aeberhard S., 1994. Comparative analysis of statistical pattern recognition methods in high dimensional settings. / S. Aeberhard, D. Coomans, O. de Vel., Pattern Recognition Vol. 27, № 8, p. 1065-1077.

[7] Ball G.H., 1995. A novel Method of Data Analysis and Pattern Classification. / G.H. Ball, D.J. Hall. Menlo Park, CA: Stanford Research Institute.

[8] Кравченко А.И., 1996. О решении на ЭЦВМ задачи распознавания логичного образа, заданного дискретной информацией. / А.И. Кравченко, М.И. Кравченко, Известия ВУЗов. Электромеханика № 4.

[9] Вапник В.М., 1965. Системы обучения распознаванию образов при помощи обобщенных портретов. / В.М. Вапник, Известия АН СССР. Техническая кибернетика № 1.

[10] Горелик А.Л., 1965. Об одном методе решения задач классификации объектов или явлений / А.Л. Горелик, В.А. Скрипник, Техническая кибернетика № 1, с. 10-19.

[11] Заяць В.М., 2009. Доцільність встановлення пріоритету первинних ознак при побудові систем розпізнавання та ідентифікації об'єктів і процесів на основі детермінованих та імовірнісних методів. / В.М. Заяць, Радіоелектроніка. Інформатика. Управління № 1, с. 92-98.

[12] Заяць В.М., 2012. Алгоритм корекції пріоритету первинних ознак при побудові систем розпізнавання. / В.М. Заяць, Г.Я. Шокира, Збірник наукових праць. Інститут проблем моделювання в енергетиці. Вис. 63, Київ, с. 105--111.

[13] Zayats V, 2012. Algorithm for Priority Deterministic and Probabilistic Primary Signs in Pattern Recognition Problems / V. Zayats, G. Shokyra. Proc. of the XX International Ukraine-Polish Conference CADMD'2012, Lviv, UKRAINE, p. 131-133.

PODEJŚCIE DO BUDOWANIA CECH SŁOWNIKA NA PODSTAWIE OKREŚLONYCH PRIORYTETÓW GŁÓWNYCH ELEMENTÓW

Streszczenie

W artykule przedstawiono opracowane podejście pozwalające na tworzenie funkcji słownika dla modelu symulacyjnego na podstawie priorytetowych funkcji. Zaproponowane podejście może zostać wykorzystane w procesie rozpoznawania obiektów, zjawisk i procesów. W pracy przedstawiono główne kierunki zastosowań dla aplikacji i otrzymanych wyników.

Słowa kluczowe: wybór priorytetów, słownik funkcji roboczych, model symulacyjny

UNIWERSYTET TECHNOLOGICZNO-PRZYRODNICZY
IM. JANA I JĘDRZEJA ŚNIADECKICH W BYDGOSZCZY
ZESZYTY NAUKOWE NR 262
ELEKTRONIKA I TELEKOMUNIKACJA 17 (2013), 37-73

REFERENTIAL GRAPHS

Damian Ledziński, Beata Marciniak, Mścisław Śrutek

University of Technology and Life Sciences in Bydgoszcz
{damian.ledzinski}{beata.marciniak}{mscislaw.srutek}@utp.edu.pl

Summary: The authors of this paper have defined the notion of referential graphs which allow to model data and telecommunication networks in order to optimize them. Parameters of this type of graphs can be compared with parameters of modified chordal rings three and four degree. A description of the program developed for graphs of this type and its effects have been presented.

Keywords: graphs, chordal rings, data communication networks.

1. INTRODUCTION

The research subject of this paper is to present a method for optimization of network structures built on the basis of chordal rings (multi top networks) which connect nodes forming data communication distributed systems. The term 'distributed system' is understood as a data communication structure consisting of a certain number of identical, intelligent nodes who are responsible for providing the system users with specific services ensuring adequate quality, speed and reliability [1,2,3,4,5]. Acceptance of such a solution positively contributes to reduction of investment costs, facilitates operation and maintenance of data communication systems as well as multi-process systems whose operation involves the concept of parallel computing processes carried out by a set of processes connected by a network with adequate configuration. The right choice of the topology of connections between its components is of key importance for the design and analysis of distributed computer systems. It determines effectiveness of the whole system operation.

In case of distributed computing systems, the main purpose is to reach the maximum computing power, whereas for telecommunication system it is of key importance to reach [4]:

- Minimal Connection Costs expressed by a summary number of used links;
- Minimal Communication Delay it is measured by diameter length and average path length;
- High fault tolerance which is characterized by the number of independent paths between two nodes connectivity or the minimal number of nodes and edges after removal of which the network ceases to be coherent Node and Edge Connectivity; Regularity and Symmetry;
- Easy of Routing [6,7];
- Extensibility.

The main elements of distributed systems are: the type of transmission medium and configuration of the network connecting modules used for collecting, transmission and distribution of data [8].

The discussed networks can be modeled using undirected graphs, and more specifically, using Symmetric Digraphs [9]. A characteristic feature of undirected graph G with a set of vertices $V(G)$ and a set of edges $E(G)$, is that if edge $[v_i, v_j]$ belongs to set $E(G)$, then also edge $[v_j, v_i]$ belongs to the same set. Thus, the diagraph edge connecting nodes v_i and v_j can be replaced with two directed edges $[v_i, v_j]$ and $[v_j, v_i]$, which, in implementation of an optical network, corresponds to a connection of two random nodes by two optical fibers, one of which to be used for transmitting signals between v_i and v_j, whereas the other between nodes v_j and v_i.

It is obvious that the network described by complete graph has the best transmission properties. Such a network is the highly reliable (in case of node or link damage the most bypass paths are available), it has the shortest graph diameter (equal to 1), the lowest average connection path length (equal to 1) and the lowest connectivity delay.

The main fault of a complete graph involves a necessity of physical or logical connection of all the nodes with each other which is specified by the formula:

$$L = \frac{p(p-1)}{2} \tag{1}$$

where:

L denotes the number of links,
p the number of nodes.

For this reason, the most desirable is such a topology which will provide the nodes with connectivity in the most satisfying way, ensuring reliability, speed and quality of provided services with economically justified outlays and implementable technical solutions.

Conclusions drawn on the basis of the studied papers indicate that the parameters which have the largest influence on the network operation properties are the graph diameter and the average path length.

Definition 1

The diameter of graph $D(G)$ is the longest path among the minimal length paths connecting two arbitrary nodes of a graph:

$$D(G) = \max_{i,j} \{d(v_i, v_j)\} \tag{2}$$

where v_i, v_j denote the number of nodes, $d(v_i,v_j)$ length of the path (number of nodes) connecting nodes v_i, v_j.

Definition 1

The average path length in a graph is defined as follows:

$$d_{av} = \frac{1}{p(p-1)} \sum_{i=0}^{p-1} \sum_{j=0}^{p-1} d_{min}(v_i, v_j) \tag{3}$$

where d_{min} denotes the minimal number of edges between nodes v_i, v_j, whereas $i \neq j$, p – is the number of nodes forming a graph.

If a graph is symmetric, the average path length is expressed by the formula:

$$d_{av} = \frac{1}{(p-1)} \sum_{i=0}^{p-1} d_{min}(v_0, v_i) \text{ [10]}. \quad (4)$$

A ring is a base structure of most data communication systems. Commutation modules or computers can be the system nodes, whereas two-direction transmission channels are the edges. This kind of network is described by the Hamilton one cycle digraph [4]. Thus, there are always two straight routes connecting two random nodes of the network. In practice, the links between nodes are provided using a two optical fiber, in one of which the transmission is clockwise and in the other it is counter clockwise. Acceptance of such a solution increases reliability of data transmission as damage to one of the nodes or links does not involve a collapse of the entire system, although it does impair its operation. Another advantage of such a structure is standardization of nodes, small number of inter-node links and good extensibility, the disadvantage though is poor transmission parameters of this type of topology (diameter, average path length). Numerous publications contain instructions how to improve the above properties of networks built on the basis of rings proposing introduction of additional links, called chords. The structures developed in this way are referred to as chordal rings.

The notion "chordal ring" was introduced by Arden and Lee in their work [11], where they describe application of this type of topology for building a network of multi-computers. They have accepted the assumption that this structure will be described by a regular degree three graph.

Definition of degree three chordal rings is as follows:

Definition 2

A degree three chordal ring (Fig. 1) is a ring in which each odd number node v_i, $i \in \{1, 3, 5, \ldots, p-1\}$ is additionally connected with node $v_{(i+q) \bmod p}$, and each even number node v_j, $j \in \{0, 2, 4, \ldots, p-2\}$ is connected with node $v_{(j-q) \bmod p}$. Value p denotes the number of nodes forming a ring and needs to be of even number (which results from the fact that each pair of nodes is connected by a chord), whereas $q \leq p/2$ is the chord length and is an odd value equal to the multitude of the edge forming a ring. It has been accepted that this kind of structure is described by notation CHR3(p; q).

A more general definition is included in paper [12].

Definition 3

A chordal ring is a special case of a Circulant Graph defined by pair *(p, Q)*, where p denotes the number of nodes and Q a set of chords, $Q \subseteq \{1, 2, \ldots, \lfloor p/2 \rfloor\}$. Each chord $q_i \in Q$ connects each pair of nodes forming a ring, whereas a chordal ring is described by notation $G(p; q_1, \ldots, q_i)$, and $q_1 = 1 < q_2 < \ldots < q_i$. According to the rule the degree of nodes is equal to $d(V) = 2i$, except for the case when the chord length is $p/2$ when p is even and the node degree is $2i - 1$.

The above given definition refers to graphs of chordal type in which each chord is an element forming the Hamilton cycle.

Chordal rings of this type can be found in a series of publications and on this basis the concepts of optimal and ideal graphs have been introduced [9,13,14].

Definition 4

An ideal chordal ring of $d(V)$ degree is a regular graph with pi nodes which is described by the formula:

$$p_i = 1 + \sum_{d=1}^{D(G)-1} |p_d| + |p_{D(G)}| \tag{5}$$

where p_d denotes the number of nodes belonging to the d-th layer (layer is a subset of nodes equally distant from a randomly chosen source node with d edges), $p_{D(G)}$ denotes the number of the remaining nodes which belong to the last layer and $D(G)$ is the diameter of the analyzed graph. For each n and $m < D(G)$ $p_n \cap p_m = \emptyset$. If subset $p_{D(G)}$ reaches the maximum value possible for the last layer, then such a graph is considered to be optimal.

Average path length d_{avi} for an ideal chordal ring is defined by the expression:

$$d_{avi} = \frac{\sum_{d=1}^{d(G)-1} d|p_d| + D(G)|p_{d(G)}|}{p_i - 1} \tag{6}$$

whereas in an optimal graph d_{avo} is equal to:

$$d_{avo} = \frac{\sum_{d=1}^{d(G)} d\, p_d}{p_o - 1} \tag{7}$$

where:

d – the layer number,
p_d number of nodes in the d-th layer
p_o – number of the optimal graph nodes.

Optimal and ideal nodes are characteristic of a selected type of graphs and their theoretically computed basic parameters will be used for evaluation of parameters of graphs obtained in real world systems.

According to the analyzed papers and the authors' own experiments, it has been found that it is possible to build modified structures with better properties than commonly known chordal rings on the basis of rings [15,16,17,18,19,20]. In order to find out whether the proposed structures are characterized by the best possible parameters it is necessary to define referential graphs whose parameter values will provide the lowest achievable limit. For this purpose the notion of Referential Graph has been introduced.

The minimal spanning tree comprising all p nodes of a regular graph was used as the base for construction of Referential Graphs structures and determination of their parameters (Fig. 1), and its parameters – radius and the average path length, computed from a randomly chosen source node, were accepted as values of the Referential Graph parameters – the diameter and the average path length.

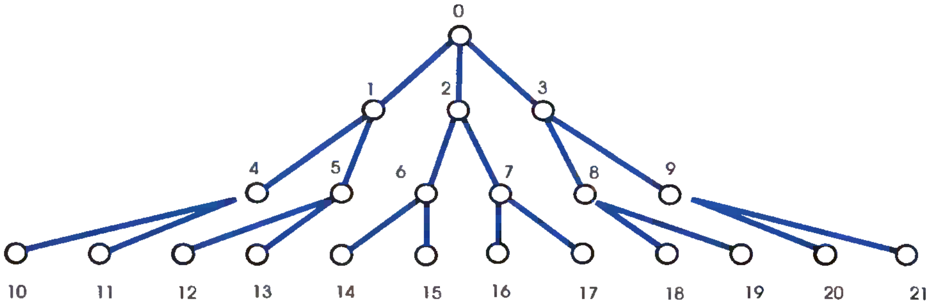

Fig. 1. A sample tree used for construction of an optimal Referential Graph third degree

The notion of Referential Graph has been defined on the basis of a carried out theoretical analysis, accounting for the degree of its nodes.

Definition 5

Referential graph of $d(V)$ degree is called a regular structure with the following properties:

– Node number p_{doGR} in an optimal graph d-th layer is defined by the formula:

$$p_{doGR} = d(V)\cdot(d(V)-1)^{d-1} \tag{8}$$

– In an ideal graph this number refers to all layers $d < D(G)$, whereas in layer $d = D(G)$ this number is :

$$p_{diGR} = p_i - d(V)\cdot(d(V)-1)^{d-2} \tag{9}$$

– Number of nodes $p_{D(G)GR}$ in an optimal graph in the function of its graph diameter:

$$p_{D(G)_{GR}} = \frac{d(V)\cdot(d(V)-1)^{D(G)}-2}{d(V)-2} \tag{10}$$

– The graph diameter in the function of the number of nodes forming a graph of this type is expressed by the formula:

$$D(G) = \log_{d(V)-1}\frac{p_{GR}\left(d(V)-2\right)+2}{d(V)} \tag{1}$$

– In an ideal graph the diameter is defined by the following expression:

$$D(G) = \left\lceil \log_{d(V)-1}\frac{p_{GR}\left(d(V)-2\right)+2}{d(V)} \right\rceil \tag{2}$$

– The mean value of path length d_{avGR} in the function of the optimal graph diameter is:

$$d_{avo} = \frac{d(V)\dfrac{(d(V)-1)^{D(G)}\cdot\left(\left(d(V)-2\right)\cdot D(G)\cdot-1\right)+1}{\left(d(V)-2\right)^2}}{\dfrac{d(V)\cdot(d(V)-1)^{D(G)}-2}{d(V)-2}-1} = \frac{(d(V)-1)^{D(G)}\cdot\left(\left(d(V)-2\right)\cdot D(G)\cdot-1\right)+1}{\left(d(V)-2\right)\cdot(d(V)-1)^{D(G)}-1} \tag{3}$$

The average path length of an ideal graph is defined by the expression:

$$d_{avr} = \frac{d(V)\frac{(d(V)-1)^{D(G)-1}\cdot((d(V)-2)\cdot(D(G)-1)\cdot-1)+1}{(d(V)-2)^2} + (p_{iGR} - p_{(D(G)-1)GR})\cdot D(G)}{p_{iGR}-1} \tag{4}$$

where p_{iGR} denotes the number of the deal graph nodes.

- A graph is symmetric, meaning that its parameters remain identical no matter which node they are computed from.

2. COMPARISON OF PARAMETERS OF MODIFIED CHORDAL RINGS AND REFERENTIAL GRAPHS

In this section third and fourth degree Referential Graphs will be presented including comparison of their parameters with parameters of the earlier studied best modified structures.

2.1. Third degree referential graphs

The basis for determination of the main parameters for graphs of each type is provided by a description of the distribution of nodes that occur in their successive layers. For degree three Referential Graphs, theoretical distributions of the maximal number of nodes have been determined, on the basis of a minimum spanning tree comprising all the nodes that form the structure which is shown in Table 1.

Table 1. Distribution of the maximal number of nodes occurring in successive layers of Referential Graphs

d	1	2	3	4	5	6	7	8	9	10
p_{dRF}	3	6	12	24	48	96	192	384	768	1536

On the basis of the above presented table, the summary number of nodes that occur in optimal graphs in the function of their diameter, has been calculated which is presented in table 2.

Table 1. Summary number of nodes occurring in an optimal Referential Graph third degree

$D(G)$	1	2	3	4	5	6	7	8	9	10
p_{dRF}	4	10	22	46	94	190	382	766	1534	3070

A third degree optimal Referential Graph has been defined on the basis of the obtained values presented in the tables.

Definition 6

A three degree optimal Referential Graph (CHR_{RG3}) is a regular structure with the following properties:

– Number of nodes p_{dGR3} in the d-th layer is defined by:

$$p_{d_{GR3}} = 3 \cdot 2^{(d-1)} \tag{5}$$

– Number of nodes $p_{D(G)GR3}$ in the graph diameter function is:

$$p_{D(G)_{GR}3} = 3 \cdot 2^{D(G)_{GR3}} - 2 \tag{6}$$

– The graph diameter in the function of the number of nodes forming its structure is expressed by:

$$D(G) = \log_2 \frac{p_{D(G)_{GR3}} + 2}{3} \tag{7}$$

– The mean path length d_{avGR3} in the function of its diameter is:

$$d_{avGR3} = 3\frac{\left(D(G)_{GR3} - 1\right) \cdot 2^{D(G)_{GR}3} + 1}{3 \cdot \left(2^{D(G)_{GR3}} - 1\right)}\,. \tag{8}$$

Charts in Fig. show a comparison between the parameters of modified structures CR3m and CR3n having better properties than the discussed in work [21] standard chordal rings, and the parameters of Referential Graphs.

Chordal rings CR3m i CR3n shown in Fig. 2 and 3 are defined in the following way:

Definition 7

Modified chordal ring CR3m is a ring whose each node with number v_i (i = 0 *mod* 4) is connected with node $v_{(i+q1) \bmod p}$, and each node v_j (j = 2 mod 4) is connected with node $v_{(i+q2) \bmod p}$; where p denotes the number of nodes in a ring that needs to be divisible by 4, whereas q_1 i q_2 are chord lengths which need to be even and satisfy condition $3 \leq q_i \leq p - 3$. Values p, q_1 and q_2 define the structure (graph) CR3m (p; q_1,q_2).

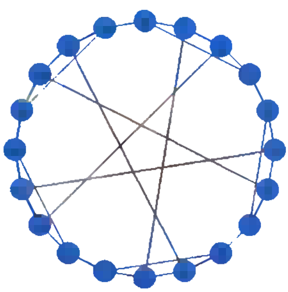

Fig. 2. An example of structure CR3m(20; 3,9)

Definition 8

Chordal ring CR3n is a ring where each node with number v_i (i = 0 mod 4) is connected with node $v_{(i+q1) \bmod p}$, node v_j (j = 1 mod 4) is connected with node $v_{(j-q1) \bmod p}$, v_k (k = 2 mod 4) with node $v_{(k+q2) \bmod p}$, a v_l (l = 3 mod 4) with node $v_{(l-q2) \bmod p}$. Values p denoting the number of nodes forming a ring is divisible by 4, q_1 and q_2 are chord lengths which fulfill condition $3 < q_i < p/2$. This graph is defined as CR3n (p; q_1,q_2).

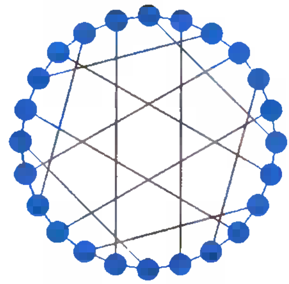

Fig. 3. An example of graph CR3n(24; 6,10)

In Figure 4 there are comparative charts for parameters of the modified graphs and Referential Graphs.

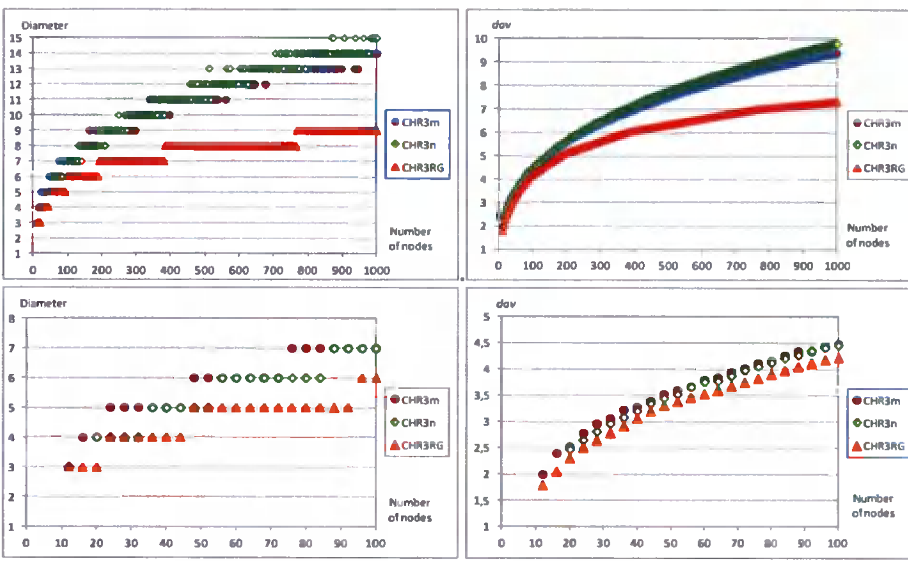

Fig. 4. Comparison of the studied structures with Referential Graphs in the function of the number of nodes

According to the charts, in case of a higher number of nodes, the parameters of modified structures vary considerably from the parameters of a Referential Graph, whereas if the number of nodes does not exceed 100, the differences are insignificant. Moreover, it has been found that there is no ideal graph possessing features of a Referential Graph among the studied structures CHR3m and CHR3n (optimal graphs are assumed not to be existing as according to the carried out studies , the number of nodes of optimal node GR3 is not divisible by 4).

Theoretically, there should exist an optimal Referential Graph CHR3 having 10 nodes as the computed sum of the determined nodes distribution equals 9 for the first two layers, however, if we use the below presented mosaic (fig. 5) it is easy to prove that there is no possibility to obtain a graph with parameters of a referential graph.

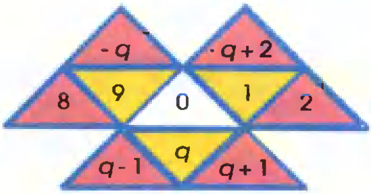

Fig. 5. Scheme of a mosaic

In the mosaic row where there is a zero number node, there occur nodes with numbers 1, 2, 9 and 8 which is determined according to the construction rule for chordal rings.

In order to provide a graph with properties of a Referential Graph, the remaining nodes of the second layer need to have numbers: 3, 4, 5, 6 and 7. The length of chord q cannot assume values 3 and 7 as in the lower layer there would appear nodes with numbers, respectively 2 and 8. If the chord length was equal to 4, 5 or 6 then the nodes with numbers, respectively: 5, 4 and 6 would be doubled in the second layer. Hence, it can be said that there is no CHR3 graph with properties of a optimal Referential Graph third degree.

The only found ideal Referential Graphs belonging to the so far discussed degree three structures are standard chordal rings built of 6, 8 and 14 nodes whose chord lengths are, respectively: 3, 4 and 5 (Fig. 6).

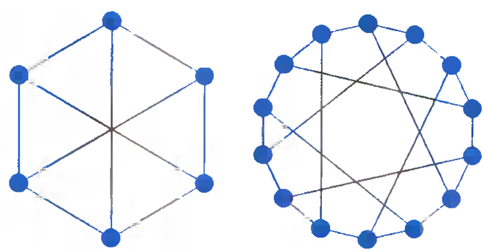

Fig. 6. Referential Graphs CHR3

The second group of the analyzed graphs of third degree, form the structures having a double ring. The definition of this type of graphs is as follows:

Definition 9

Structure NdR (fig. 7), defined as NdR(2p), is called a graph formed by two rings each consisting of p nodes:

- External ring where each node o_k is connected with two adjacent nodes $o_{k-1(\text{mod } p)}$ i $o_{k+1(\text{mod } p)}$,
- Internal ring where each node i_{k+p} is connected with two adjacent nodes $i_{k+p-1\ (\text{mod } 2p)}$ and $i_{k+p+1\ (\text{mod } 2p)}$,
- Each node of internal ring i_{k+p} is connected with a corresponding node of the external ring o_k.,
- Each node is a degree three node.

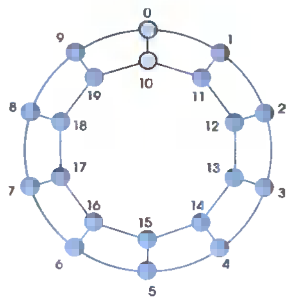

Fig. 7. An example of structure NdR(20)

Structure NdR($2p$) can be considered as 'not flexible' – its parameters are invariable and they depend merely on the number of nodes. Its modification has been proposed to remove this disadvantage. Graphs NdRm, described in papers [22,23,24] having the best properties of all the analyzed graphs, are a good example of these structures modification.

Definition 10

Two rings, each containing p nodes, form the described NdRm($2p;q$) structure:

- External ring where each node o_k is connected with two adjacent nodes $o_{k\text{-}1(\text{mod } p)}$ i $o_{k+1(\text{mod } p)}$;
- Internal ring where each node $i_{kwo\ adjacent\ nodes\ +p}$ is connected with two adjacent nodes $i_{k+p-q\ (\text{mod } 2p)}$ i $i_{k+p+q\ (\text{mod } 2p)}$ through a chord of q length being a multitude of the external ring length;
- Each node of the internal ring i_{k+p} is connected with its corresponding node of external node *ok.*;
- Degree of all the nodes equals 3.

NdRm structures can be divided into two classes:

- The first, in which chords of both rings form the Hamilton cycles. In this case the number of nodes and the chord length of inner need to be relatively prime (Fig. 8A).
- The second, in which the external ring chord forms the Hamilton cycle with property $j = i \oplus 1\ mod\ p$, whereas the internal ring consists of a certain number of disjoint cycles of the same length, smaller than p (Fig. 8B).

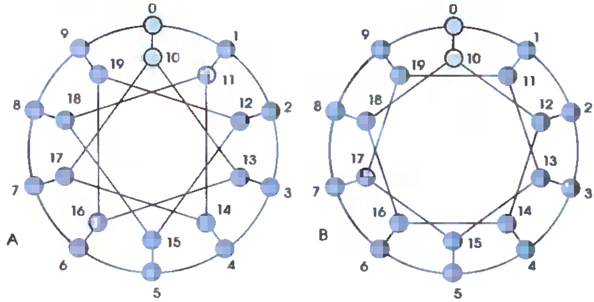

Fig. 8. Examples of two kinds of structures –NdRm(20;3) and NdRm(20;2)

Definition 11

NdRc (fig. 9) is a structure described as NdRc($2p$; q_1,q_2), formed by two rings:

- External ring consisting of p nodes (p must be divisible by 2), in which each node o_k is connected with two adjacent nodes $o_{k-1(\mathrm{mod}\ p)}$ i $o_{k+1(\mathrm{mod}\ p)}$;
- Internal ring also consisting of p nodes. Each even number node i_{2k+p} is connected with two adjacent nodes $i_{2k+p-q1\ (\mathrm{mod}\ 2p)}$ and $i_{2k+p+q1\ (\mathrm{mod}\ 2p)}$, and each odd number node i_{2k+1+p} is connected with two nodes $i_{2k+1+p+q2\ (\mathrm{mod}\ 2p)}$ and $i_{2k+1+p-q2\ (\mathrm{mod}\ 2p)}$;
- *Each* node of internal ring i_{k+p} is connected with its corresponding node o_k of the external ring;
- *Parameters* q_1 and q_2 denoting chord lengths must be of even number;
- All *nodes* are of three degree.

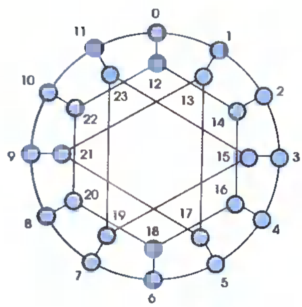

Fig. 9. Example of structure NdRc(24; 2,4)

In figure 10 there is a comparison of diameters and the average path length of the best structures, built on the basis of NdRc graphs, with Referential Graphs. The presented charts prove that parameters NdRc are closer to the parameters of Referential Graphs, though their values are far from being ideal.

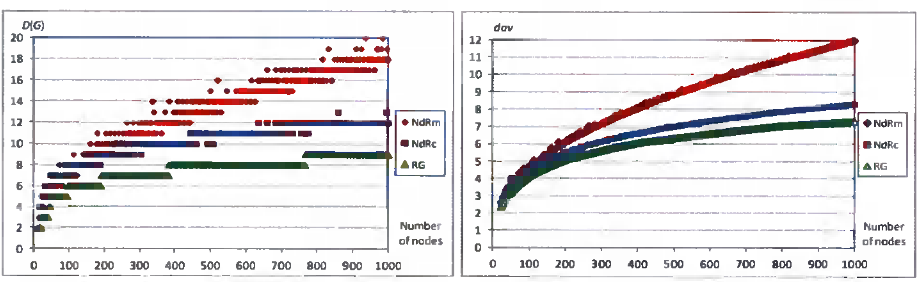

Fig. 10. Comparison of parameters of NdRm, NdRc real structures and Referential Graphs in the function of the node number.

The charts also account for parameters of NdRm structures. It results from the fact that although these parameters significantly differ from values characteristic of Referential Graphs, they form the only optimal graph NdRm (5;2) (Petersen graph) that has been identified, and two ideal graphs NdRm(7;2) and NdRm(13; 5) (Fig. 11).

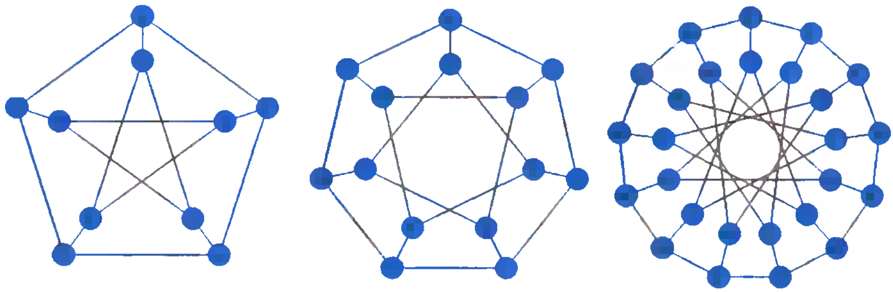

Fig. 11. Structures NdRm forming Referential Graphs

A comparison of parameters of the analyzed structures CHR3, NdR with Referential Graph has been shown in Figure 12 as a summary of this section of the paper.

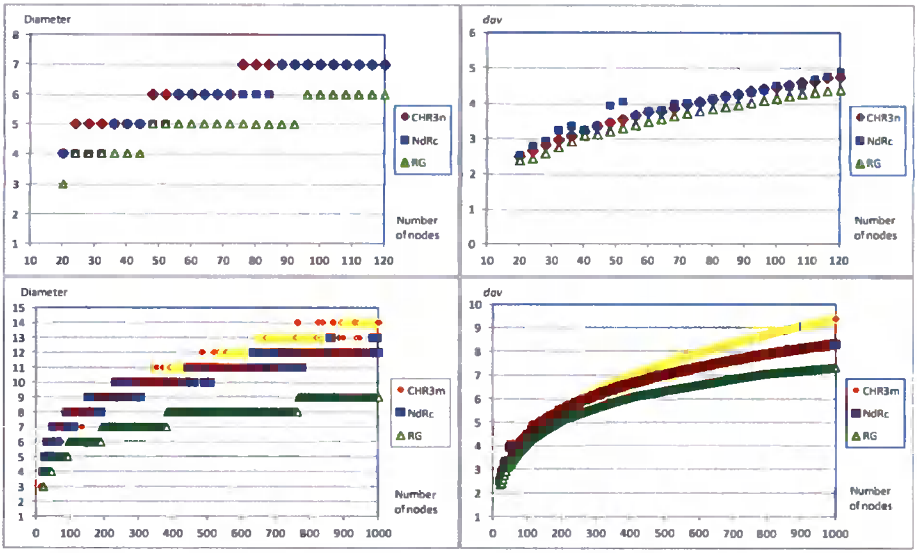

Fig.12. Comparison of real CHR3m, NdRc structures with Referential Graphs in the function of the node number

The charts show that parameters of the NdR type structure are more closer to the parameters of a Referential Graph though if the number of nodes connecting comparable modified structures does not exceed one hundred they do differ insignificantly.

This conclusion confirms that, despite being more expensive, the use of the NdR type structures for building a network increases its reliability.

2.1. Fourth degree referential graphs

In table 3 there is a distribution of nodes occurring in the successive layers of an optimal degree four Referential Graphs.

Table 2. Distribution of the maximal number of nodes occurring in the successive layers of a referential graph

d	1	2	3	4	5	6	7	8
p_{dRF}	4	12	36	108	324	972	2916	8748

The total number of nodes in the function of the graph diameter, computed using the above given number of nodes that occur in the successive layers, has been presented in table 4.

Table 3. The summary number of nodes in the function of the graph diameter, computed using the above given number of nodes that occur in the successive layers of four degree Referential Graph

$D(G)$	1	2	3	4	5	6	7	8
p_{dRF}	5	17	53	161	485	1457	4373	13121

Below there is the definition of a four degree Reference Graph.

Definition 12

Referential Graph is a graph whose parameters are equal to parameters of a minimum spanning tree defined in the following way:

– The number of nodes in the d-th layer of an optimal Referential Graph is defined by the formula:

$$p_{dRG} = 4 \cdot 3^{(d-1)} \tag{9}$$

– The summary number of nodes $p_{D(G)RG}$ in the function of diameter is described by the formula:

$$p_{D(G)RG} = \frac{4 \cdot 3^{D(G)RG} - 2}{2} = 2 \cdot 3^{D(G)RG} - 1 \tag{10}$$

– The graph diameter in the function of the number of nodes can be calculated with the use of the formula:

$$D(G) = \left\lceil \log_3 \frac{p_{D(G)RG} + 1}{2} \right\rceil \tag{11}$$

– The average path length in optimal Referential Graph d_{avRG} in the function of diameter is given by the formula:

$$d_{avRG} = \frac{\left(2D(G)-1\right) \cdot 3^{D(G)_{RG}} + 1}{2 \cdot \left(3^{D(G)_{RG}} - 1\right)} \tag{12}$$

Like in case of third degree graphs, definitions of chordal rings which have been compared with Referential Graphs will be presented [26].

The definition of a standard chordal ring sounds as follows:

Definition 13

The described by CHR4($p;Q$) chordal ring fourth degree, is a graph consisting of p nodes. Each node v_i is connected with four nodes: $v_{i-1(mod\ p)}$, $v_{i+1(mod\ p)}$ forming a ring and additionally with two adjacent nodes $v_{i-q(mod\ p)}$ and $v_{i+q(mod\ p)}$, where q denotes the length of the additional link (chord) corresponding to the number of nodes on the ring when the numbers corresponding to values p and q must be relatively prime. Figure 13 shows a graph of this type.

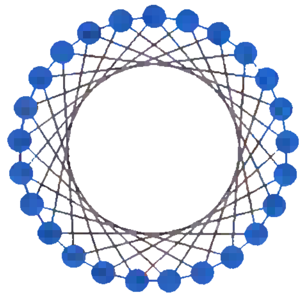

Fig.13. An example of a chordal ring CHR4(25; 7)

This base structure has been modified in order to find structures with better base parameters.

It has been found that of all the analyzed, modified, degree four chordal rings [26] the best parameters are characteristic of graphs called CHR4b and CHR4d which have been defined in the following way:

Definition 14

Chordal ring CHR4b is a structure defined as CHR4b(p; q_1,q_2), consisting of p nodes, whereas p, if it is not a full graph with five nodes, has to be of an even number. Each node v_i is connected with nodes $v_{i-1(\mathrm{mod}\ p)}$ and $v_{i+1(\mathrm{mod}\ p)}$. Moreover, each even number node v_{2i} is connected with nodes $v_{2i-q1(\mathrm{mod}\ p)}$ i $v_{2i+q1(\mathrm{mod}\ p)}$, and an odd number node $v_{2i+1(\mathrm{mod}\ p)}$ with nodes $v_{2i-q2(\mathrm{mod}\ p)}$ i $v_{2i+q2(\mathrm{mod}\ p)}$. Values of parameters q_1 and q_2 are of even numbers, whereas $p/2$ and $q_1/2$, $q_2/2$ are relatively prime, and denote chord lengths.

An example of the discussed structure is shown in Figure 14.

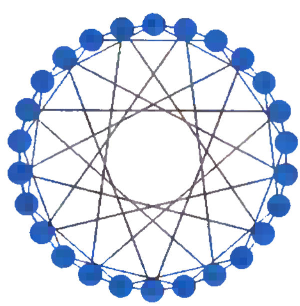

Fig.14. Graph CHR4b(26; 4,10)

Definition 15

Structure CHR4d (example Fig. 15) is a chordal ring defined as CHR4d(p; q_1,q_2,q_3), consisting of p nodes, whereas p must be divisible by 4. Each node v_i is connected with nodes $v_{i-1(\mathrm{mod}\ p)}$ and $v_{i+1(\mathrm{mod}\ p)}$. Additionally, even node v_{2i}, if $2i = 0\ (\mathrm{mod}\ 4)$, is connected with nodes $v_{2i+q1(\mathrm{mod}\ p)}$ and $v_{2i+q2(\mathrm{mod}\ p)}$, if $2i = 2\ (mod\ 4)$, it is connected with nodes $v_{2i+q1(\mathrm{mod}\ p)}$ and $v_{2i-q2(\mathrm{mod}\ p)}$, whereas odd node $v_{2i+1(mod\ p)}$ if $2i+1 = 1\ (\mathrm{mod}\ 4)$, is connected with nodes $v_{2i+1-q1(mod\ p)}$ and $v_{2i+q3(mod\ p)}$, and if $2i+1 = 3\ (\mathrm{mod}\ 4)$ it is connected with nodes $v_{2i+1-q1(mod\ p)}$ and $v_{2i+1-q3(mod\ p)}$. The value of parameter q_1 is even, whereas parameters q_2 and q_3 are even, whereas they must satisfy condition: $q_2 - q_3 = 0\ (\mathrm{mod}\ 4)$. Parameters q_1, q_2 and q_3 define the chords lengths.

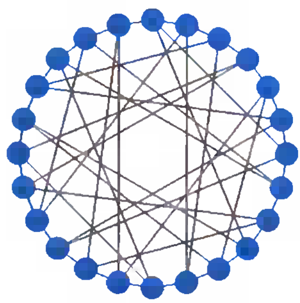

Fig. 15. An exemplary structure CHR4d(24; 9,6,10)

In Figure 16 there are results of the comparison of these structures parameters with the parameters of Referential Graphs.

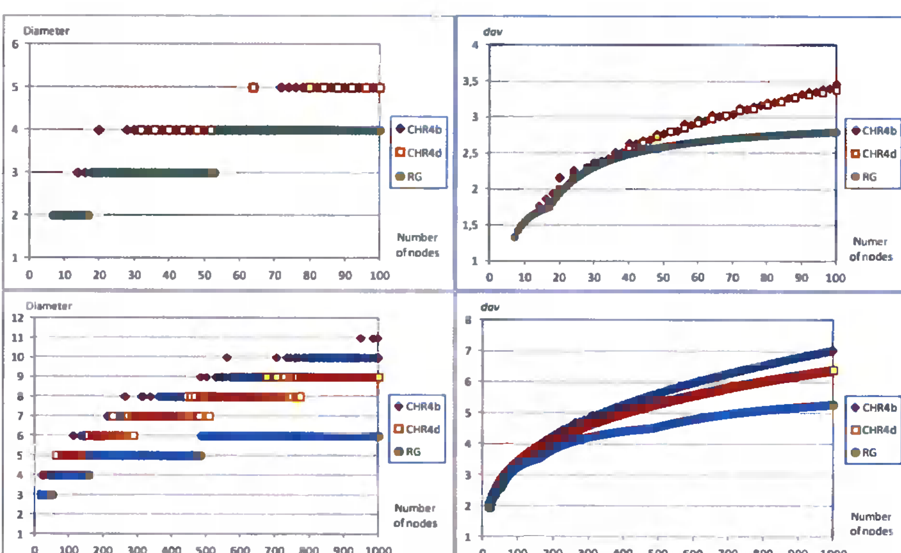

Fig. 16. Comparison of structures CHR4b i CHR4d from degree four Referential Graphs

The presented results prove that the parameters of Referential Graphs, if the number of nodes exceeds 40, differ significantly from the best proposed structures of type CHR4.

However, on the basis of the carried out tests it was found that standard chordal rings CHR4(7; 3), CHR4(8; 3), CHR4(9; 4), CHR4(11; 4), CHR4(13; 5) and graphs CHR4b(6; 2,2) CHR4b(10; 2,4), CHR4b(22; 4,8), CHR4b(26; 4,10) are ideal Referential Graphs.

These structures are shown in figures 17 and 18.

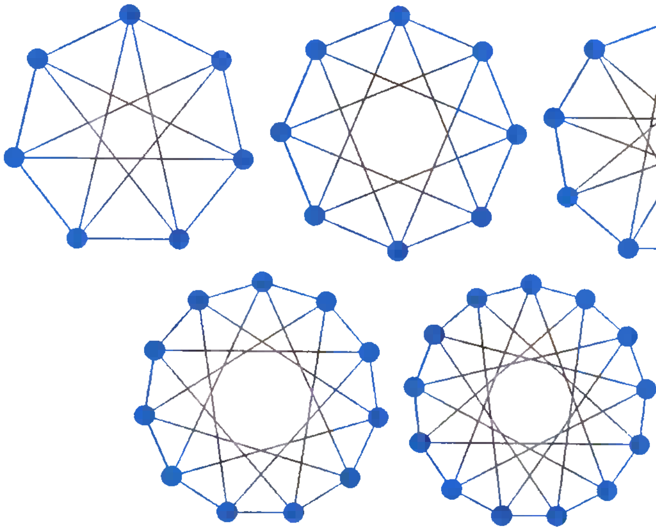

Fig. 17. Chordal rings of type CHR4

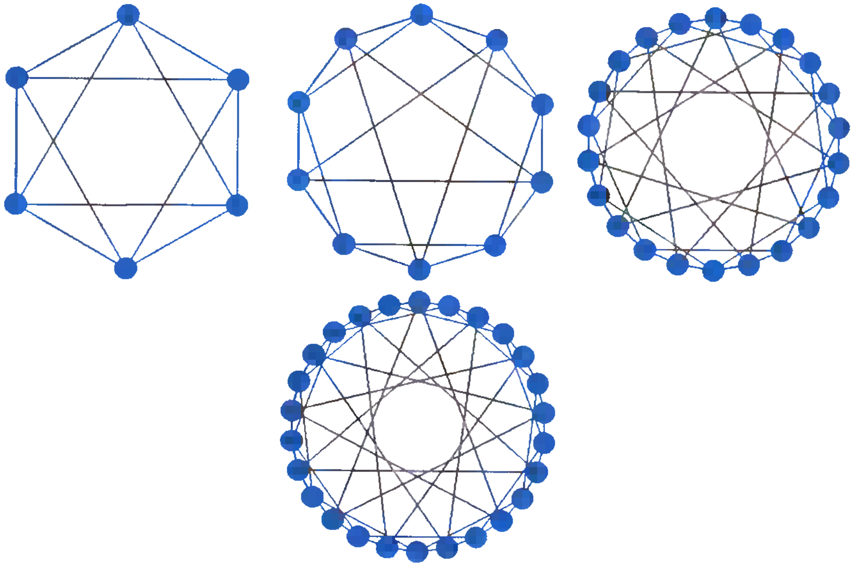

Fig. 18. Chordal Rings of CHR4b type

On the basis of the carried out analysis it can be said that:

Since Referential Graphs have been found to exist in reality it is advisable to continue searching for such structures which use rings as the base of their construction

3. SEARCH OF REFERENTIAL GRAPHS

The authors of this paper aim at finding other graphs with the characteristics of Referential Graphs (RG) parameters. If on the basis of the conducted research it has been found that RG graphs exist in the real world, the developed program enabled to survey all the structures possible to obtain for the assumed number of nodes forming these graphs.

To begin with, a new mode of operation will be presented, that is the tool to be used for finding the considered structures, will be described.

3.1. The method used for finding referential graphs

A tool which enables to survey all possible ring structures in order to find out whether their basic parameters reach values characteristic of Referential Graphs is the program referred to as GraphFinder4.

Graph Finder4 offers a wide range of possibilities to examine graphs. For this reason, before activating the program, it is necessary to specify features of the examined structures:

- Node Number,
- Node Degree,
- Minimal cycle, function is introduced in order to limit the number of the examined connections,
- Minimal number of parts on which decomposition of the problem will follow enabling performance of computing on numerous processors.

Introduction of the function „Min. Cycle size" is the effect of the following observation. According to the definition of Referential Graph the tree constructed from an arbitrary node being a root of this graph will have an identical form. The condition that a random node of the *k-th* layer, where $k < D(G)$ can be connected with only one node of the *l-th* layer, but there cannot exist an edge connecting it with a node belonging to the *l-th* layer, where $l < k - 1$ must be satisfied. However, this node can possess a connection with a node belonging to the same layer or a layer with number $k + 1$. Thus, the minimal length of the cycle (measured by the number of nodes) from a node considered to be the root, which can appear in a degree three graph is:

$$l_{min} = 2D(G) - 1 = 2\log_{d(V)-1} \frac{p_{GR}\left(d(V) - 2\right) + 2}{d(V)} - 1 \tag{13}$$

Hence, it can be concluded that the number of examined structures can be limited thanks to elimination of shorter length cycles, which reduces the examination time.

To illustrate the study, an example of a degree three referential graph has been presented in Figure 19. Node 0 was assumed to be a root of the tree. The l minimal cycle length is 5 and the length of the exemplary one covers nodes: 0, 6, 7, 8, 17, and all the remaining cycles formed with the use of chords, are longer (e.g. 0, 1, 10, 9, 8, 17) or equal to the shortest one (e.g. 1, 10, 9, 3,2).

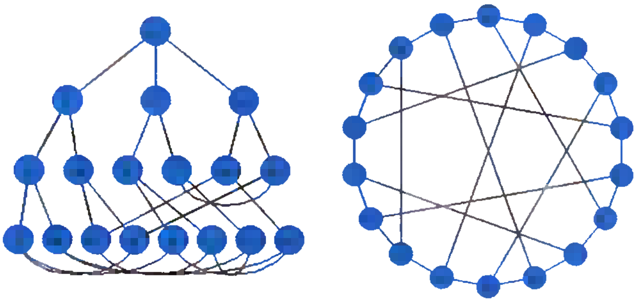

Fig. 19. An example of degree three Referential Graph

The operation principle of the program is given below. The described applications have been implemented in Java language, however, for better understanding, the program codes are presented in Phyton language.

- *Method of a graph denotation*

The program saves a graph in the form of a set of lists.
Initiation of a graph with no links looks as follows:

```
graph = [[-1 for x in range(degree)] for x in range(nodeNumber)]
```

where „degree" denotes degree of nodes , a „nodeNumber" – number of the graph nodes.

This notation can be interpreted, as a two dimensional matrix in which rows represent the successive nodes and columns connections with other nodes. Index nodes are from 0 to nodeNumber -1. Value "-1" means lack of connection.

- *Algorithm for determination of diameter and mean path length.*

The program determines a diameter and the average path length by means of a simplified Dijkstra algorithm. Such a simplification is possible thanks to acceptance of the assumption that the weight of the graph all edges are equal to 1.

Figure 20 shows the algorithm to be used for determination of the graph diameter and the average path length.

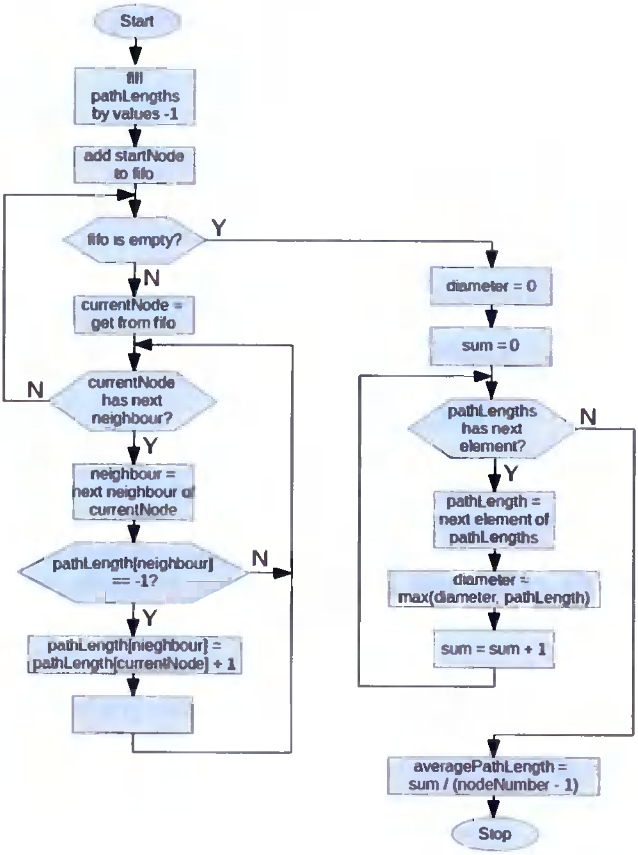

Fig. 20. Algorithm the diameter and average path length determination

Implementation of the function using the algorithm is of the following form.

```
# calculate diameter and average path length for given source node
def calculate(graph, sourceNode):
 # degree of node
 degree = len(graph[0])
 # create empty list of path lengths
 pathLengths = [-1 for x in range(degree)]
 pathLengths[sourceNode] = 0
 # create a queue of node to handle
 fifo = deque()
 fifo.append(sourceNode)
 # loop executing until the queue is empty
 while len(fifo) > 0:
 # getting and removing from queue a node to servicing
 currentNode = fifo.popleft()
 # service node neighbours
 for neighbour in graph[currentNode]:
 # checking if neighbour has been previously serviced
 if pathLengths[neighbour] == -1:
 # calculate the path length to the neighbour
```

```
pathLengths[neighbour] = pathLengths[currentNode] + 1
# adding a neighbour to the queue
fifo.append(neighbour)
# [illegible]
diameter = 0
# sum of path lengths
pathLenhtgSum = 0.0
for pathLength in pathLengths:
# searching of the longest path
diameter = max(diameter, pathLength)
# adding to the sum
pathLenhtgSum += pathLength
# calculate of average path length
averagePathLength = pathLenhtgSum / (len(graph) - 1)
# return results
return (diameter, averagePathLength).
```

This function computes the diameter and average path length from the point of view of a given source node, thus to determine these parameters values for a graph it is necessary to provide each graph with this function. According to the definition, the graph diameter will be the longest of the shortest paths obtained with the use of the function, whereas the average path length is the mean arithmetic from all the results returned by this function.

- *Method of finding referential graphs*

The purpose of this program is to search for referential graphs. This process can be divided into two stages. In the first stag, all possible structures of a minimum spanning tree are determined for a given number of nodes of the graph chosen degree, and in the other stage – each tree is developed.

Step I

First, all possible isomorphic simple trees are constructed from number 0 node, accepted to be their root. These trees are created only up to the last but one layer, (when optimal graphs are searched for – up to the last layer), and the other layers which belong to the last layer, remain unconnected.

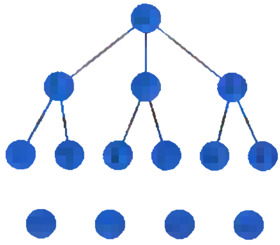

Fig. 21. A sample minimum spanning tree of the search for an ideal Referential Graph consisting of 14 nodes

In the next stage, all possible combinations of nodes of the last layer with the nodes of the last but one layer are created. Repetitions are omitted, that is, such connection topologies which after transformation would give the same effect.

In turn, a division of connections with the nodes belonging to the last layer is performed in order to build trees with 0 node root.

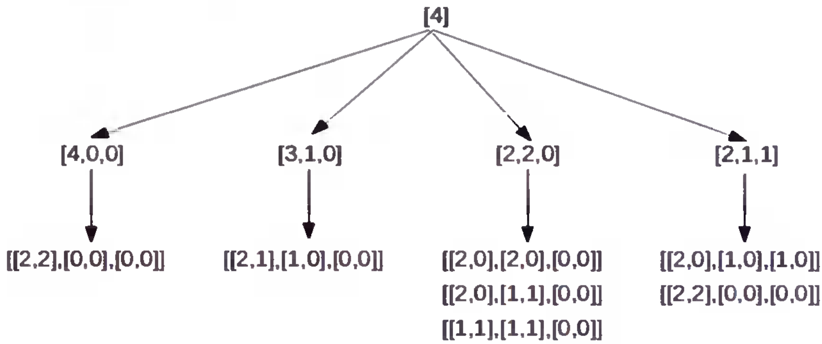

Fig. 22. Possible divisions of connections for the analyzed example

In the analyzed example four nodes are available. In Figure 22 there are presented possibilities of the second layer nodes to perform possible connections eliminating repetitions of identical connections and in Figure 23 there are trees developed for the above example which have the following form:

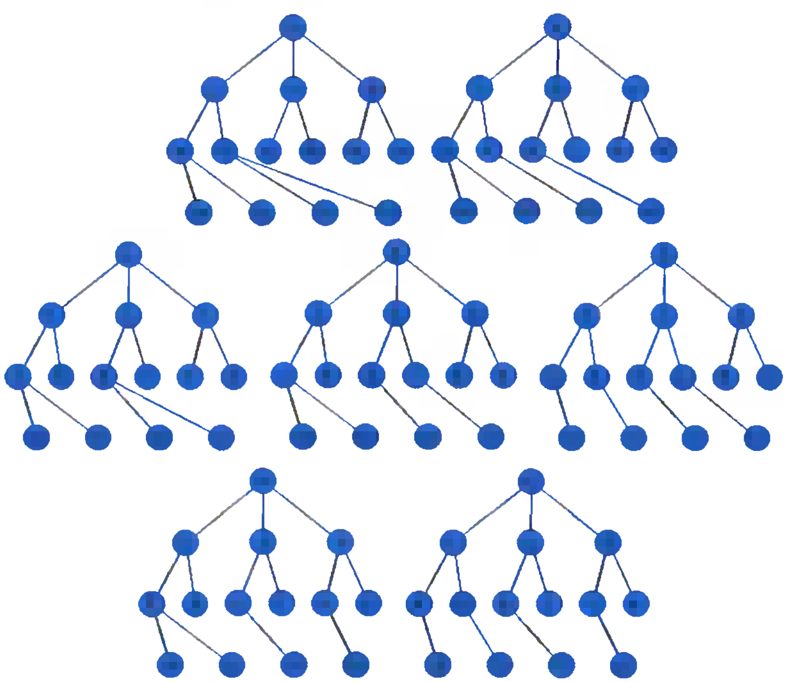

Fig. 23. Trees provided for the studied example

The obtained trees provide the basis for the next stage of Referential Graphs seeking.

Step II

The minimum spanning trees provided in the first step of Stage 1 are starting point of the algorithm. Additional connections from successively chosen nodes appear, every time the condition whether the graph radius and diameter are equal to parameters of the primary tree, is checked. I f the graph construction is finished (a set of connections is completed) and if it is a Referential Graph, the algorithm stops acting. If not – a new set of subgraphs, with one more connection, is created for the graph. This set is added to the list of graphs to be examined. The algorithm finishes its action after having examined all graphs from the list.

General algorithm for finding referential graphs looks as follows (Fig. 24):

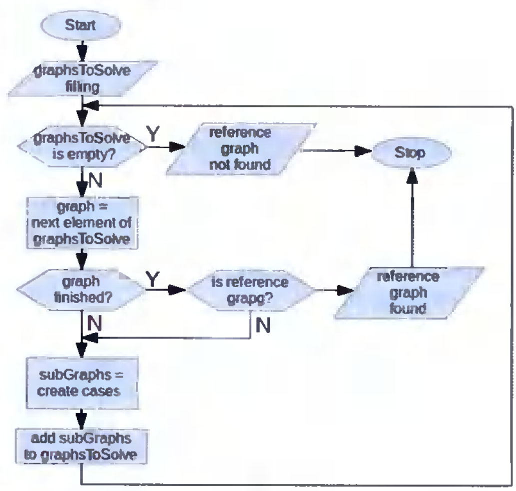

Fig. 24. General algorithm for finding Referential Graphs

Implementation of the function using recurrence for examination of graphs has the following form:

```
# decomposition function
def decompose(graphsToSolve):
 # loop on graphs to solve
 for graph in graphsToSolve:
 # checking if graph construction is finished, all nodes connected
 if isFinished(graph):
 # checking if graph is reference graph
 if isReference(graph):
 return graph
 # graph is not finished
 else:
 # sub graphs creating
 subGraphs = createSubGraphs(graph)
 # recursive call
 result = decompose(subGraphs)
 if result != None:
 return result
 # reference graph not found
 return None.
```

This function provides the possibility to decompose a problem in order to perform calculations in a concurrent way (e.g. on many cores of the processor or in a cloud computing).

In order to study all of the possible connections of a given graph, two lists of candidate nodes are made. The first one is called a list of source nodes and the second is called a list of target nodes. Next, graphs are formed for a combination of the first source node with all the target ones. The first successive node on the list of source nodes is the node nearest the start node which has no connection with z $d(V)$ neighbors. Target nodes are nodes from the last but one or the last layer from the start node.

The start node is determined by a separate algorithm. The distance between the source node and the target one cannot be shorter than the length of the shortest cycle that can occur in this graph, reduced by 1.

The algorithm providing the connections variants are shown in Figure 25.

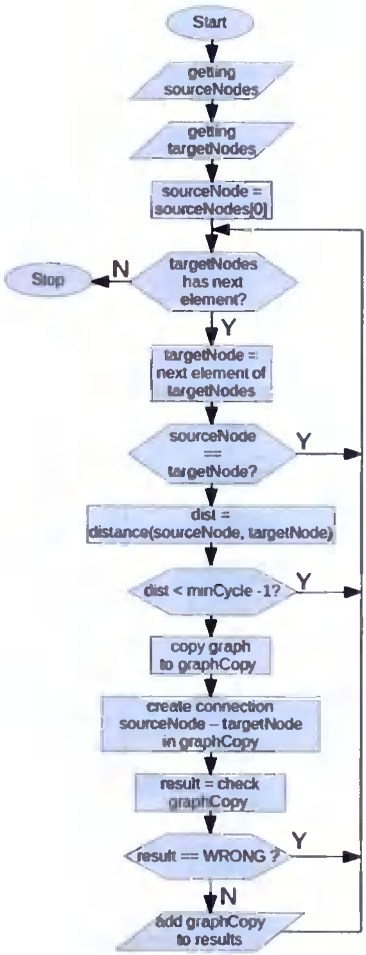

Fig. 25. Algorithm determining variants of connections

Implementation of the function searching for variants looks as follows:

```
def createSubGraphs(graph):
 startNode = getStartNode(graph)
 sourceNodes, targetNodes = findCandidates(graph, startNode)
 results = []
 if len(sourcenodes) == 0:
 return result
 sourceNode = sourceNodes[0]
 for targetNode in targetNodes:
 if sourceNode == targetNode:
 continue
 if distance(sourceNode, targetNode) < minCycle - 1:
 continue
 graphCopy = copy(graph)
 setConnection(graphCopy, sourceNode, targetNode)
 if checkGraph(graphCopy) != WRONG:
 results.append(graphCopy)
 return results.
```

The algorithm for determination of candidate nodes is shown in Figure 26.

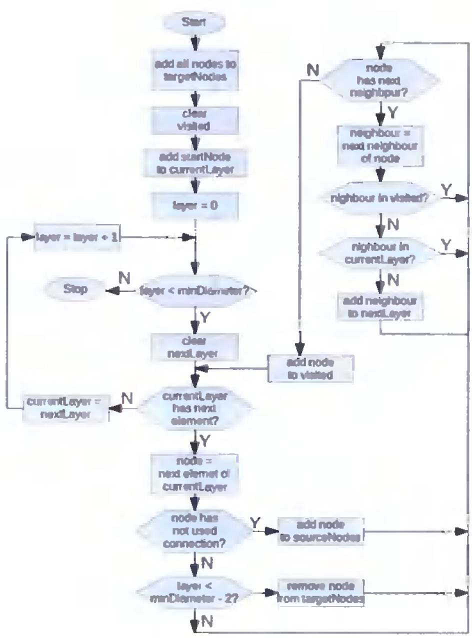

Fig. 26. The algorithm determining candidate source and target nodes

The function using shown above algorithm has following appearance:

```
# source and target nodes designate function
def findCandidates(graph, startNode):
 sourceNodes = []
 # setting a set of target nodes, which will be removed
 targetNodes = range(len(graph))
 # visited nodes
 visited = []
 # nodes in given layer
 currentLayer = [startNode]
 # loop on layers
 for layer in range(minDiameter):
 # nodes in next layer
 nextLayer = []
 # loop on nodes in layer
 for node in currentLayer:
 # checking if node has free connections
 if -1 in graph[node]:
 # adding to source nodes
 sourceNodes.append(node)
 # checking if this is last or penultimate layer
 if layer < minDiameter - 2:
 # removing the target nodes of the previous layer
 targetLayer.remove(node)
 # loop on neighbours to prepare next layer
 for neighbour in graph[node]:
 if neighbour != -1:
 # checking if neighbour has been visited
 if neighbour not in visited:
 # checking if node is not in the layer
 if neighbour not in currentLayer:
 # adding node to next layer
 nextLayer.append(neighbour)
 # marking node as visited
 visited.append(node)
 # changing layer
 currentLayer = nextLayer
 # result return
 return (sourceNodes, targetNodes.
```

The next element of the method for finding Referential Graphs is an algorithm classifying graphs. It examines the graph from the point of view of each node and provides one of the following answers:

- PERFECT – means that from the point of view of this node the distribution is the best of all possible,
- WRONG – means that the algorithm detected a situation excluding the graph as a referential one,
- UNKNOWN – lack of answer, graph can be further examined.

If after having examined all the nodes, the answer WRONG appears even once, there is no point in continuing the examinations. If the answer PERFCT is obtained for all the nodes it means that it is a Referential Graph.

The algorithm examining the graph for a single node looks as follows (Fig. 27):

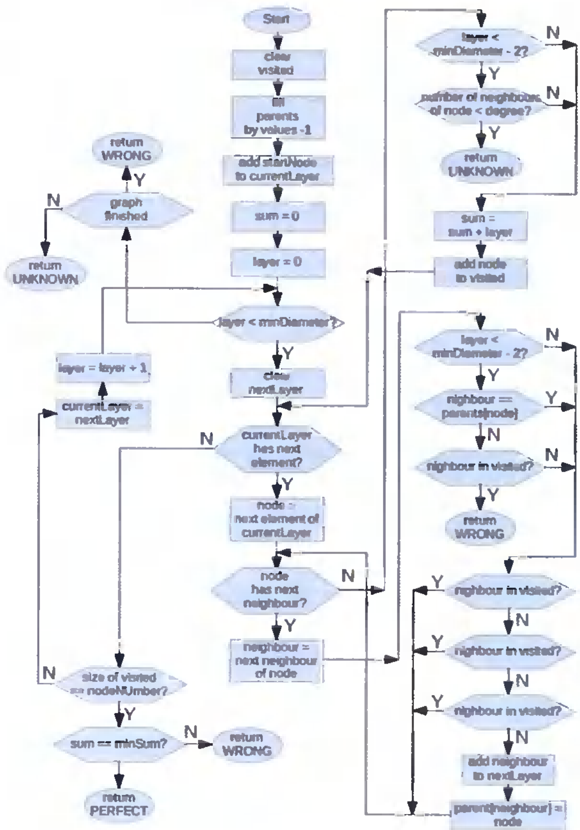

Fig. 27. Algorithm for graph examining from a single node

Implementation of this algorithm is as follows:

```
# constant definition
PERFECT = 0
WRONG = 1
UNKNOWN = 2
# graph examination function
def checkGraphFromNode(graph, startNode):
 visited = []
 # parents list
 parents = [[-1 for x in range(nodeNumber)]
 # current layer
 currentLayer = [startNode]
 # sum of path lengths
 pathLengthSum = 0
 # loop on layers
 for layer in range(minDiameter):
 # next layer nodes
 nextLayer = []
 # loop on nodes in layers
 for node in currentLayer:
```

```
# loop on neighbours
for neighbour in graph[node]:
if neighbour != -1:
# checking if this is last or penultimate layer
if layer < minDiameter - 2:
# checking if neighbour is not parent
if neighbour != parents[node]:
# checking if neighbour has been visited
if neighbour in visited:
# this combination is forbidden
return WRONG
# checking if neighbour hasn't been visited
if neighbour not in visited:
# checking if neighbour is not in current layer
if neighbour not in currentLayer:
# checking if neighbour is not in next layer
if neighbour not in nextLayer:
# adding node to the next layer
nextLayer.append(neighbour)
# marking node as parent
parent[neighbour] = node
# checking if this is last or penultimate layer
if layer < minDiameter - 2:
# checkinh if node has all connections
if getNeighbourNumber(graph, node) < len(graph[0]):
# answer cannot be determined
return UNKNOWN
# adding to path length sum
pathLengthSum += layer
# marking node as visited
visited.append(node)
# checking if all nodes were visited
if len(visited) == len(graph):
# checking if path length sum is minimal possible sum
if pathLengthSum == minPathLengthSum:
return PERFECT
else:
return WRONG
# changing layers
currentLayer = nextLayer
# checking id graph is finished
if isFinished(graph):
return WRONG
return UNKNOWN.
```

The above given operation principle of the program makes it possible to achieve the goal set by the authors of this paper, that is to find Referential Graphs.

In order to illustrate the study we will show an example depicting operation of this application.

It was assumed that the minimum spanning tree obtained from stage 1 has the form demonstrated in Figure 28.

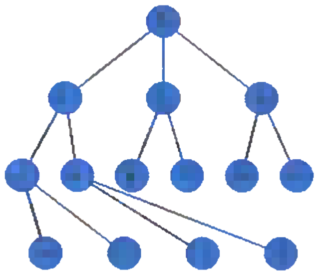

Fig. 28. Minimum spanning tree of a referential graph

The same tree seen from the point of view of exemplary nodes assumed to be start nodes is shown in Figure 29.

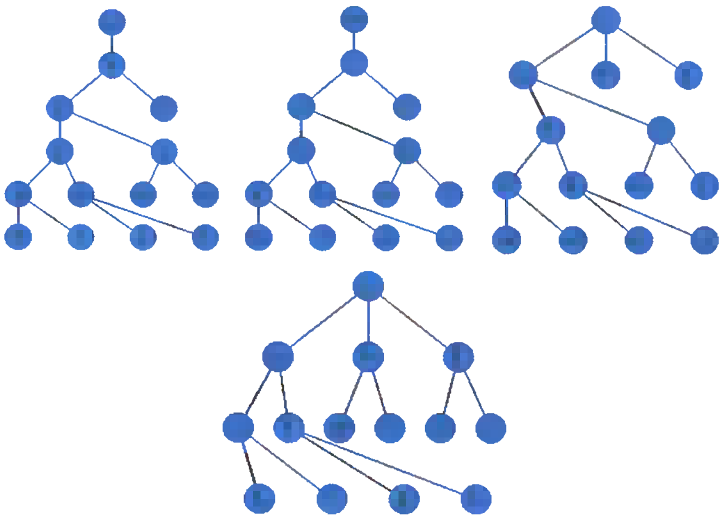

Fig. 29. Forms of the tree seen from exemplary source nodes

According to Figure 29, at the same stage of the analysis, in each case, graphs will produce the answer PERFECT.

After adding the successive four, out of eight, connections necessary to obtain the whole graph the structure has the following form (from the point of view of the chosen nodes) (Fig. 30).

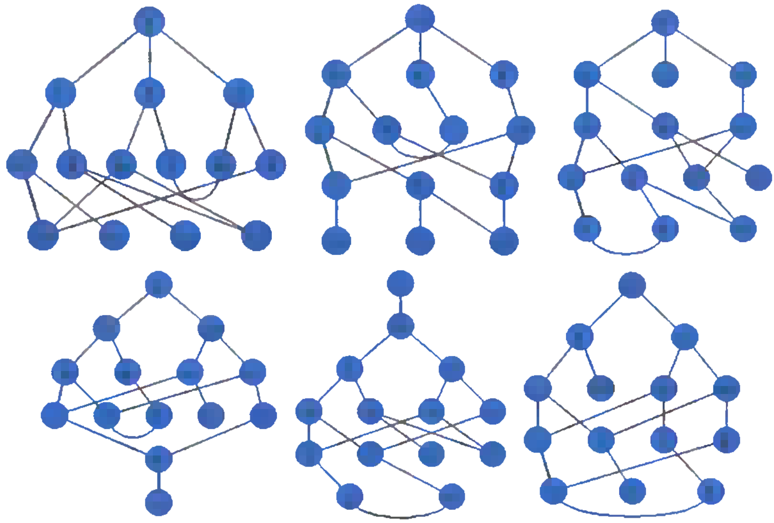

Fig. 30. Stage of searching for Referential Graph connections

After introduction of connections between nodes 12 & 8 and 9 & 11, which is definitely the most convenient stage to do it, obtained graph looks as follows (Fig. 31):

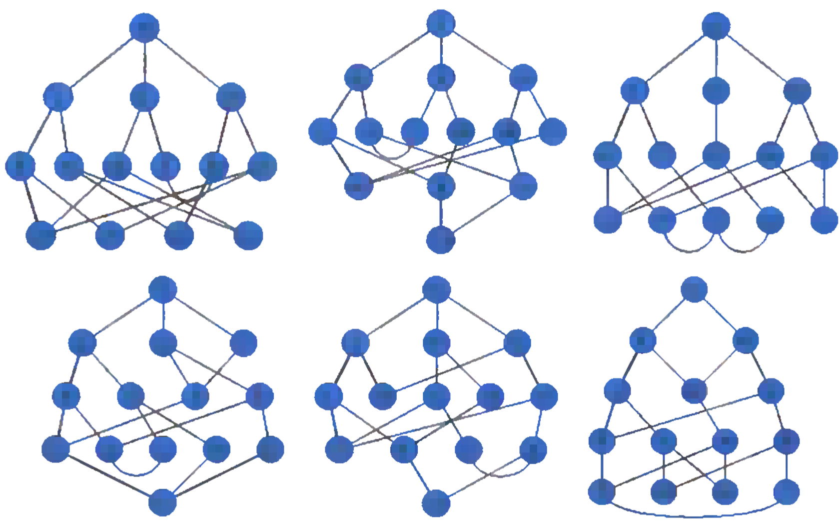

Fig. 31. Successive stage of seeking Referential Graph connections

Examination of the so obtained graph will result in the answer WRONG, that is, at this point, the process of Referential Graph seeking will be abandoned, as from the point of view of graphs no. 9, 10 and 11, there are two connections between the nodes of the first layer, and one node of the second layer, so there appeared a cycle connecting nodes 4, 9, 10 and 11.

However, the Referential Graph found thanks to the use of the discussed application is as follows (Fig. 32):

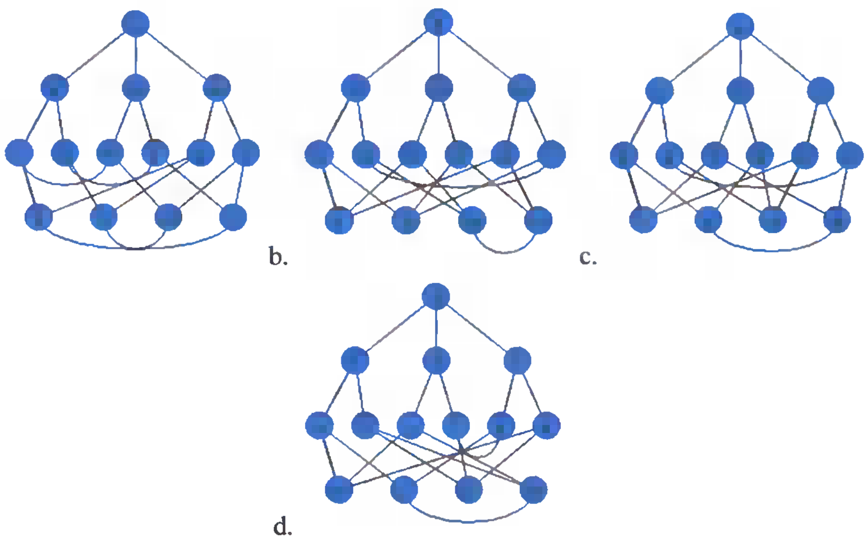

Fig. 32. Referential graph found thanks to GraphFinder4

Additional task to be performed by the discussed application is transformation of graphs shown in Figure 32 into the form of a chordal ring which is described below:

```
# function to finding Hamilton cycles in graph
def findHamilton(graph, path = [0]):
 # getting current node
 currentNode = path[-1]
 # checking that found path through all nodes
 if len(path) == len(graph):
 # checking if current node has connection to first node
 if path[0] in graph[currentNode]:
 # return found cycle
 return path
 return None
 # loop on neighbours
 for neighbour in graph[currentNode]:
 # checking if neighbour has been visited
 if neighbour not in path:
 # adding next node to path
 path.append(neighbour)
 # recursive call
 result = findHamilton(graph, path)
 # checking if cycle has been found
 if result is not None:
 return result
 # removing last element
 path.remove(neighbour)
 # cycle not found
 return None.
```

After the Hamilton cycle was found and converted, the graph shown in fig. 31d assumed the following form (Fig. 33).

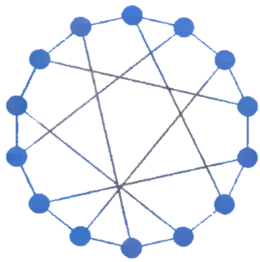

Fig. 33. An example of Referential Graph built with 14 nodes

Thanks to using the program a user can obtain the following information:

- Graph metrics, its basic features;
- Description of graph formation process;
- Topology of the obtained structure connections;
- Graph basic parameters;
- Length the above mentioned minimal cycle;
- Confirmation of appearance of the cycle covering all the graph nodes;
- Time of graph seeking.

After the searching process is finished, on the screen, apart from a description, there is an image of the obtained Referential Graph in two forms shown in Figure 34.

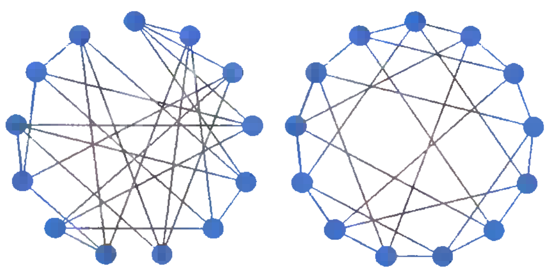

Fig. 34. Images of an exemplary Referential Graph

An additional option of this program is the possibility to turn the graphs which makes it easier to check whether the graphs using different chord lengths are isomorphic graphs.

While using the discussed program degree three and four graphs were surveyed and the effects of Referential Graphs seeking are demonstrated in figures 35 and 36.

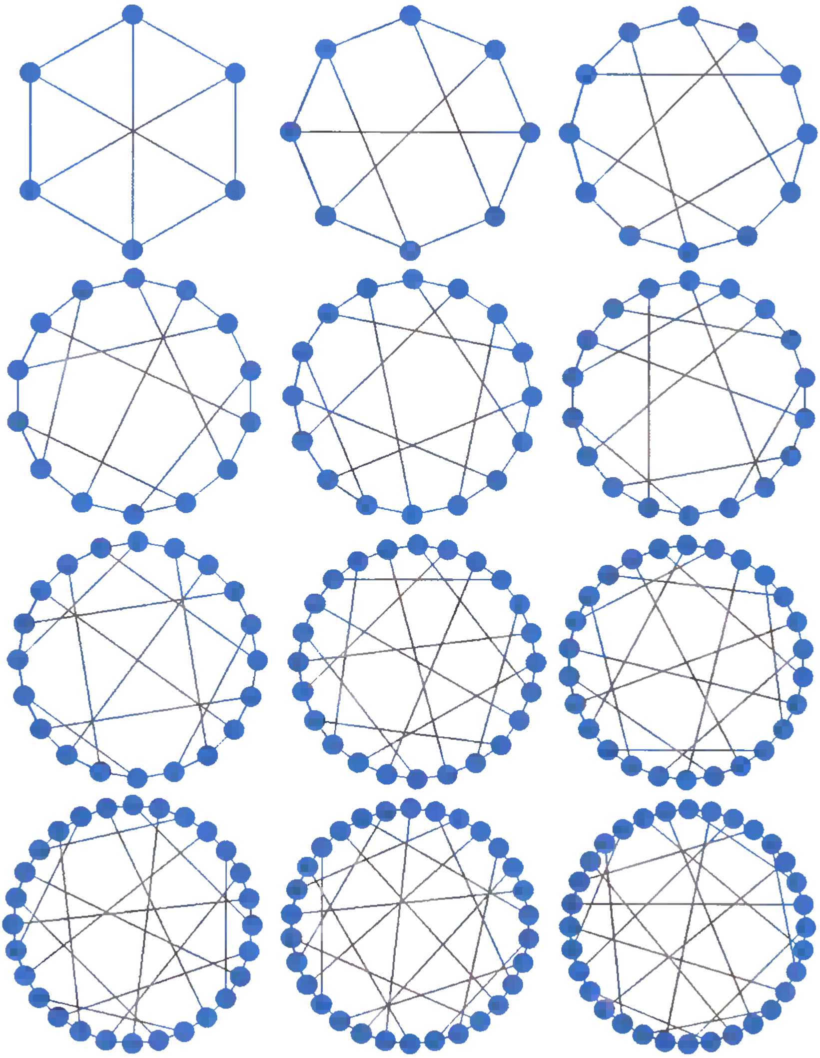

Fig. 35. Examples of degree three and four Referential Graphs

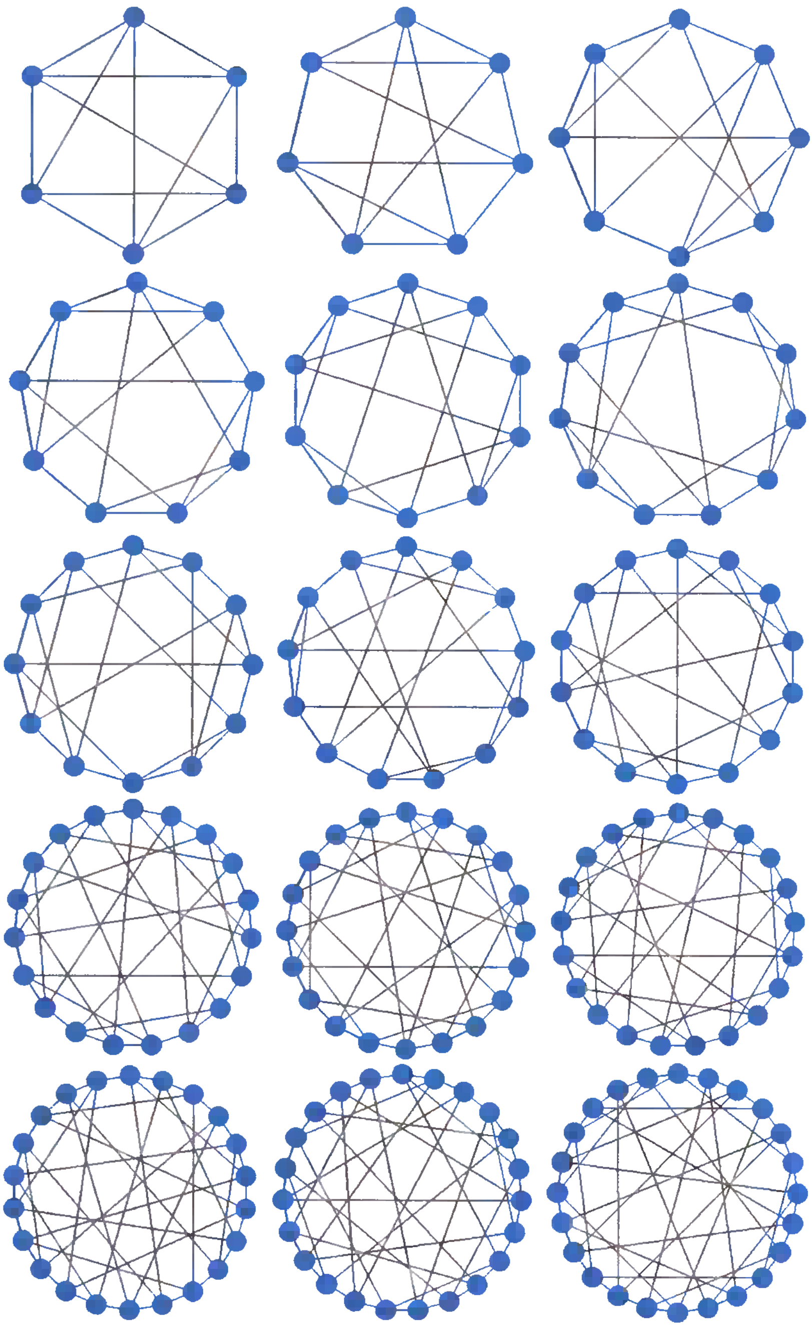

Fig. 36. Examples of found degree four Referential Graphs

4. CONCLUSIONS

In this paper, the notion of Referential Graphs has been defined in order to use it for optimization of a network connecting unified modules that can be used as telecommunication or specialized IT units. Using the studied chordal rings based networks for modeling makes it possible to reach a satisfying level of reliability, capacity, and the minimum information transmission delay thanks to minimization of the basic parameters – diameter and average path length. Parameters of Referential Graphs have been compared with the parameters of modified degree three and four structures proposed by authors. The article includes examples of Referential Graphs found thanks to using the program developed by the authors.

On the basis of obtained results they concluded that not for each number of nodes exists Referential Graphs. In case of degree three graphs, the only one optimal graph is NdR structure having 10 nodes (Pedersen graph), whereas there is no possibility to create a graph based on a single ring. There have been found no optimal chordal rings consisting of 22 nodes, which is of significance as it can indicate that it is not possible to build optimal Referential Graphs with more nodes than 10. Additionally, it would be advisable to check if it is possible to find an optimal graph for topologies with 46 nodes.

In case of degree four graphs, did not find any optimal graph, apart from complete graph constructed of 5 nodes. Despite long lasting tests, there have been found no structure with 16, 17 and 18 nodes.

Due to the time required to perform calculations, examination of degree three graphs could be carried out for structures consisting of 32 nodes, whereas for degree four graph – to 24. Thus, the main problem to be solved at the present stage is a successive modification of the program or use of higher performance computers (Cloud Computing) which should result in shortening the time of searching for nodes with higher numbers of nodes, thus enabling optimization of connections for more developed network.

Such an application provides the possibility of operation in a network and this will allow to use significantly larger computing potential for seeking graphs. The proposed architecture of this application is shown in fig. 37.

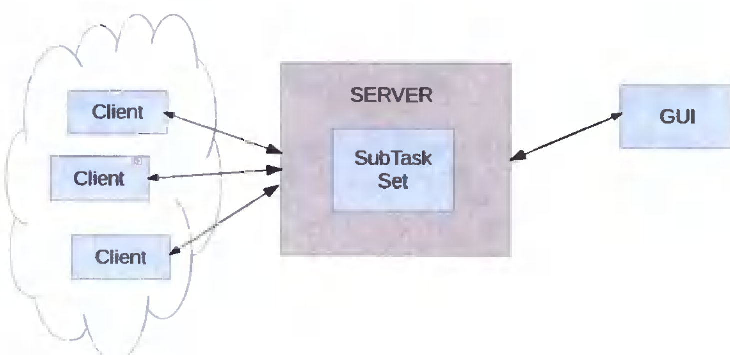

Fig. 37. A scheme of the proposed application: Server – application running on a machine available for the system remaining elements; SubTaskSet – a set of subsets which are assigned to clients. Client-application operating on a machine designed for computing; GUI – application of the user's interface

The server observes calls coming from clients and GUI, it accepts tasks from GUI, decomposes them, and assigns the tasks to particular clients and saves the results. The content of SubTaskSet is obtained from decomposition of the task which is performed by Server (e.g. search of Referential Graph with defined degree and number of nodes). A client receives tasks from the server and returns the obtained results, using all available cores of the processor. GUI enables reception and assignment of new tasks as well as collection and survey of the results gathered by the server. Thanks to implementation of such a solution it is possible to find Referential Graphs with simultaneous use, in our case, more than 80 cores. This is expected to enable disperse many doubts, and solve problems that occur while examining chordal rings which, though being important for modeling of real-world data communications network, have not been accounted for in this paper.

BIBLIOGRAPHY

[1] Bhuyan L.N., 1987. Interconnection Networks for Parallel and Distributed Processing. IEEE Computer 20(6), 9-12.

[2] Bermond J.C., Comellas F., Hsu D., 1995. Distributed loop computer networks: A survey. Journal of Parallel and Distributed Computing 24, 2-10.

[3] Blazewicz J., Ecker J.K., Plateau B., Trystram D., 2000. Handbook on parallel and distributed processing. Springer Berlin. Heidelberg.

[4] Kotsis G., 1992. Interconnection Topologies and Routing for Parallel Processing Systems. ACPC, Technical Report Series, ACPC/TR92-19.

[5] Mans B., 1999. On the Interval Routing of Chordal Rings. ISPAN '99 IEEE – International Symposium on Parallel Architectures, Algorithms and Networks, Fremantle, Australia, 16-21.

[6] Kiedrowski P., 2009. Easy Applicable Algorithm for Accelelerate Reading Process in AMR Systems based on WSN Solutions, Image Processing & Communications Challenges (ed. R.S. Choraś, A. Zabłudowski), Academy Publishing House EXIT, 482-487.

[7] Kiedrowski P., 2012. Media Independent Protocol Suite for Energy Management Systems Based on Short Range Devices. Prace Instytutu Elektrotechniki LIX, z. 258, 189-199.

[8] Kiedrowski P., Boryna B., Marciniak T., 2013. Last-mile smart grid communications based on hybrid technology as a reliable method of data acquisition and distribution, Rynek Energii 1(104), 127-132.

[9] Narayanan L., Opatrny J., Sotteau D., 2001. All-to-All Optical Routing in Chordal Rings of Degree 4. Algorithmica 31, 155-178.

[10] West D.B., 2000. Introduction to Graph Theory, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.

[11] Arden W., Lee H., 1981. Analysis of Chordal Ring Network. IEEE Transactions on Computers 30(4), 291-295.

[12] Gavoille C. (n.d). A Survey on Internal Routing. http://deptinfo.labri.ubordeaux.fr/~gavoille/article/ survey /node28.html

[13] Bujnowski S., 2003. Analysis & Synthesis Homogeneous Structure Networks Connecting Communications Modules. PhD Thesis, UTP Bydgoszcz.

[14] Pedersen J. M., Gutierrez J.M., Marciniak T., Dubalski, B. Zabłudowski A., 2009. Describing N2R Properties Using Ideal Graphs. Advances in Mesh Networks, MESH 2009. The Second International Conference on Advances in Mesh Networks. ISBN: 978-0-7695-3667-5, 150-154.

[15] Azura R.N.F., Othman M., Selamat M.H., Hock P.Y., 2008. Modified Degree Six Chordal Rings Network Topology. Proceedings of Simposium Kebangsaan Sains Matematik Ke-16, 3-5, 515-522.

[16] Azura R.N.F., Othman M., Selamat M.H., Hock P.Y., 2010. On properties of modified degree six chordal rings networks. Malaysian Journal of Mathematical Sciences 4(2), 147-157.

[17] Farah R.N., Othman M., Selemat H., Hock P.Y., 2008. Analysis of Modified Degree Six Chordal Rings and Traditional Chordal Rings Degree Six Interconnection Network. International Conference of Electronics Design. Penang, Malaysia.

[18] Farah R.N., Othman M.., 2010. In Modified Chordal Rings Degree Six Geometrical Representation Properties. Proceedings of Fundamental Science Congress, Kuala Lumpur, Malaysia.

[19] Farah R.N., Othman M., Selamat M.H., 2010. Combinatorial properties of modified chordal rings degree four networks. Journal of Computer Science 6(3), 279--284.

[20] Farah R. N., Othman M., Selamat M. H., 2010. An optimum free-table routing algorithms of modified and traditional chordal rings networks of degree four. Journal of Material Science and Engineering 4(10), 78-89.

[21] Bujnowski S., Dubalski B., Pedersen J.M., Zabłudowski A., 2009. Struktury topologiczne CR3m oraz NdRm. Przegląd Telekomunikacyjny LXXXI, 8/9, 1133--1141.

[22] Pedersen J.M., Patel A., Knudsen T.P., Madsen O.B., 2004. Generalized Double Ring Network Structures. Proc. of SCI 2004. The 8th World Multi-Conference On Systemics. Cybernetics and Informatics 8. Orlando, USA, 47-51.

[23] Pedersen J.M., Knudsen T.P., Madsen O.B., 2004. Comparing and Selecting Generalized Double Ring Network Structures. Proc. of IASTED CCN 2004. The Second IASTED International Conference Communication and Computer Networks. Cambridge, USA, 375-380.

[24] Zabłudowski Ł., Dubalski B., Kiedrowski P., Ledziński D., Marciniak T., 2012. Modified NDR structures. Image Processing and Communications 17(3), ISSN 1425-140X, 29-47.

[25] Pedersen J. M., Riaz M.T., Madsen O.B., 2005. Distances in Generalized Double Rings and Degree Three Chordal Rings. Proc. of IASTED PDCN 2005. IASTED International Conference on Parallel and Distributed Computing and Networks. Innsbruck, Austria, 153-158.

[26] Ledziński D., Bujnowski S., Marciniak T., Pedersen J.M., Gutierrez Lopez J., 2013. Network Structures Constructed on Basis of Chordal Rings 4^{th} Degree. Image Processing and Communications Challenges 5. Advances in Intelligent Systems and Computing 233, 281-299.

GRAFY REFERENCYJNE

Streszczenie

W artykule zdefiniowano pojęcie Grafu Referencyjnego, za pomocą którego można modelować sieci teleinformatyczne w celu ich optymalizacji. Zdefiniowano parametry takiego typu grafów, które zostały porównane z parametrami modyfikowanych pierścieni cięciwowych trzeciego i czwartego stopnia. Do modelowania poszczególnych przypadków opracowano program symulacyjny. Zaprezentowano opis tego programu oraz wyniki otrzymane przy wyszukiwaniu grafów referencyjnych dla szerokiego spektrum danych wejściowych.

Słowa kluczowe: grafy, grafy referencyjne, projektowanie sieci telekomunikacyjnych

ISSN 1899-0088