УДК 621. 391

ANALYSIS OF NON-STATIONARY QUASI-PERIODIC INTERNET TRAFFIC BY THE METHOD OF INSTANT SPECTRA

Odegov N.A., Bahachuk D.H., Hordiienko O.O.

O.S. Popov Odesa National academy of telecommunications,
1 Kuznechna St., Odesa, 65029, Ukraine.
onick_64@ukr.net

АНАЛІЗ НЕСТАЦІОНАРНОГО КВАЗІПЕРІОДИЧНОГО ІНТЕРНЕТ-ТРАФІКА МЕТОДОМ МИТТЄВИХ СПЕКТРІВ

Одегов М.А., Багачук Д.Г., Гордієнко О.О.

Одеська національна академія зв'язку ім. О.С. Попова,
65029, Україна, м. Одеса, вул. Кузнечна,1.
onick_64@ukr.net

АНАЛИЗ НЕСТАЦИОНАРНОГО КВАЗИПЕРИОДИЧЕСКОГО ИНТЕРНЕТ-ТРАФИКА МЕТОДОМ МГНОВЕННЫХ СПЕКТРОВ

Одегов Н.А., Багачук Д.Г., Гордиенко А.А.

Одесская национальная академия связи им. А.С. Попова,
65029, Украина, г. Одесса, ул. Кузнечная, 1.
onick_64@ukr.net

Abstract. At large aggregation points, traffic growth processes are observed, which can be modeled by non-stationary quasi-periodic functions. A formal and informal analysis of traffic at a number of international aggregation points was performed. The traffic growth model was confirmed in the form of an algebraic sum of the main regular component and a number of cyclic components. The main cycles detected are in good agreement with the laws of rotation and rotation of the Earth and its interaction with neighboring celestial bodies. Not entirely obvious, but confirmed by the actual measurements of traffic, the dependence of human biological and social cycles on the rotation of all the same stars turned out to be. The daily, weekly, seasonal and annual traffic cycles turned out to be clearly pronounced. The nine-month and four-month cycles were unexpected. The latter are reflected only in the ancient calendars of the Slavs. In the daily dynamics, natural cycles prevail: day-night, the division of days into business and home traffic. Definitely unexpected was the discovered effect of desynchronization of Ukrainian traffic relative to the traffic of other European countries. The article concludes that this is due to incorrect determination of the decree time. A number of traffic dynamics models are proposed for solving subsequent prediction problems. It was found that in short time intervals the amplitude and phase components of the cyclic components are almost constant. An algorithm for selecting amplitudes of cyclical components and a regular trend component in a polynomial model of the latter is given. Software has been developed for solving similar problems of analyzing and predicting traffic using the instant spectra method.

Key words: traffic, random process, traffic aggregation point, prediction, Fourier transform, social and biological cycles.

Анотація. На великих вузлах агрегації спостерігаються процеси зростання трафіка, які можна моделювати нестаціонарними квазіперіодичними функціями. У статті зроблено формальний і неформальний аналіз трафіка на низці міжнародних точок агрегації. Підтверджено модель зростання трафіка у вигляді алгебраїчної суми основної регулярної компоненти і ряда циклічних складових. Основні виявлені цикли добре узгоджуються з закономірностями обертання Землі та взаємодії її з сусідніми небесними тілами. Не зовсім очевидною, але підтвердженою реальними вимірами трафіка виявилася залежність біологічних і соціальних циклів людини від обертання тих самих світил. Чітко вираженими стали добовий, тижневий, сезонний і річний цикли трафіка. Несподіваними стали дев'ятимісячні і чотиримісячні цикли. Останні мають відображення тільки в стародавніх календарях слов'ян. У добовій динаміці переважають природні цикли: день-ніч, ділення доби на діловий і домашній трафік. Безумовно несподіваним з'явився виявлений ефект розсинхронізації українського трафіка щодо трафіка інших європейських країн. У статті надано висновок про те, що це пов'язано з некоректним визначенням декретного часу. Запропоновано низку моделей динаміки трафіка для вирішення наступних завдань прогнозування. З'ясовано, що на малих часових відрізках амплітудні і фазові компоненти циклічних складових — практично константні. Дано алгоритм селекції амплітуд циклічних складових і регулярної складової тренда в поліноміальній моделі останнього. Розроблено програмне забезпечення для вирішення аналогічних завдань аналізу і прогнозування трафіка методом миттєвих спектрів.

Ключові слова: трафік, випадковий процес, точка агрегації трафіку, прогноз, перетворення Фур'є, соціальні та біологічні цикли.

Аннотация. На крупных узлах агрегации наблюдаются процессы роста трафика, которые можно моделировать нестационарными квазипериодическими функциями. В статье выполнен формальный и неформальный анализ трафика на ряде международных точек агрегации. Подтверждена модель роста трафика в виде алгебраической суммы основной регулярной компоненты и ряда циклических составляющих. Основные обнаруженные циклы хорошо согласуются с закономерностями вращения и вращения Земли и взаимодействия ее с соседними небесными телами. Не совсем очевидной, но подтвержденной реальными измерениями трафика оказалась зависимость биологических и социальных циклов человека от вращения все тех же светил. Четко выраженными оказались суточный, недельный, сезонный и годовой циклы трафика. Неожиданными оказались выявленные девятимесячные и четырехмесячные циклы. Последние имеют отражение только в древних календарях славян. В суточной динамике преобладают естественные циклы: день-ночь, деление суток на деловой и домашний трафик. Определенно неожиданным явился обнаруженный эффект рассинхронизации украинского трафика относительно трафика других европейских стран. В статье дан вывод о том, что это связано с некорректным определением декретного времени. Предложен ряд моделей динамики трафика для решения последующих задач прогнозирования. Выяснено, что на малых временных отрезках амплитудные и фазовые компоненты циклических составляющих — практически константны. Дан алгоритм селекции амплитуд циклических составляющих и регулярной составляющей тренда в полиномиальной модели последнего. Разработано программное обеспечение для решения аналогичных задач анализа и прогнозирования трафика методом мгновенных спектров.

Ключевые слова: трафик, случайный процесс, точка агрегации трафика, прогноз, преобразование Фурье, социальные и биологические циклы.

The article [1] concluded that in the near future there may be a need for radical technical re-equipment of fiber-optic transmission systems (FOTS). Reason: a sharp increase in the volume of processed traffic at the aggregation points. At the same time, this conclusion was based on data from various sources. Thus, the source [2] gives an estimate of about 20% of the annual increase in traffic volumes in world telecommunication networks. In [3], this indicator is taken at the level of 30%. The ITU report [4] shows a diagram from which it is clear that this indicator is approximately equal to 20%. (unfortunately, according to the citation rights of this source, it is impossible to provide this diagram here). In more recent research by Cisco [5], a forecast was made until 2022, from which it follows that, depending on the type of traffic, its volume may have an annual increase from 26% to 77%. Nokia's research [6] provides estimates of the annual increase in total traffic volume of about 25%, and peak values up to 40%.

As can be seen, the estimates are very different. Cisco and Nokia companies can be suspected of some tendentiousness as telecommunications equipment manufacturers. In this case, any of these estimates makes you think at least.

To formalize the growth rate of traffic in the global networks, the well-known from economic theory index is used $CAGR$ (Compound Annual Growth Rate). In the calculations by the method of compound interest, the $CAGR$ index appears in the formula for the prediction:

$$V(nT) = V_0(1 + CAGR)^n, \tag{1}$$

where $V(nT)$ – the value of the investigated parameter in n periods of duration T; V_0 – base value of this parameter; $CAGR$ index is taken as a unit fraction.

In Fig. 1 given conditional predictions of traffic growth in accordance with the formula (1) for different $CAGR$ index values.

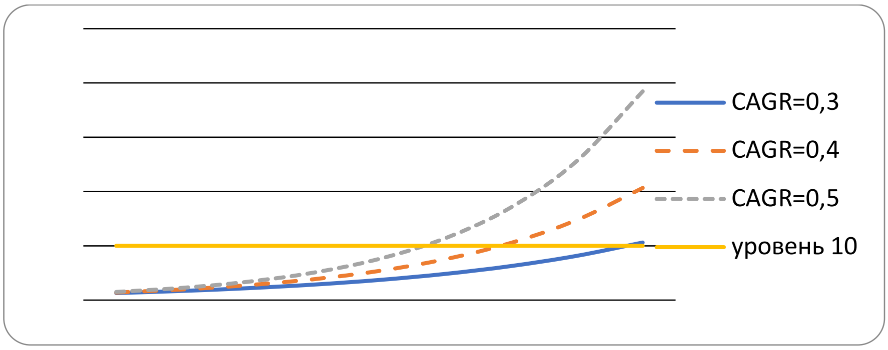

Figure 1 – Traffic growth prediction for various $CAGR$ index values

In this figure, the value of the quantity being investigated (traffic volume) in 2018 is taken as 1. As you can see, traffic increases by an order of magnitude when $CAGR = 0{,}3$ in about 8 years, when $CAGR = 0{,}4$ – in 6 years, and when $CAGR = 0{,}5$ – in less than 5 years. The rapid growth of traffic in the latter case may require not only upgrading the equipment of traffic aggregation nodes, but perhaps even replacing the backbone optical cable. The latter task is not solved in a short time and requires significant capital expenditures.

Therefore, adequate prediction of the growth of traffic volumes in telecommunication networks is a very topical issue from an economic and technical point of view. At the same time, predictions should not be too high, but not too low. In case of significant errors in either case, operators may incur unjustified losses due to incorrect planning of the development of their networks.

The positive aspect is that, in order to solve this problem, it is possible to involve not only (and not so much) a well-developed teletraffic theory, but a real metrological basis. As such a basis, can be taken directly the data of monitoring the traffic of aggregation nodes and data transmission channels. At the same time, even operators at the international and national levels often use the freely distributed software package MRTG (Multi Router Traffic Grapher) [7]. In particular, the results of measuring the averaged instantaneous values of traffic [8] using the MRTG for 5 years for the Brazilian exchange point IX.br are given in Fig.2.

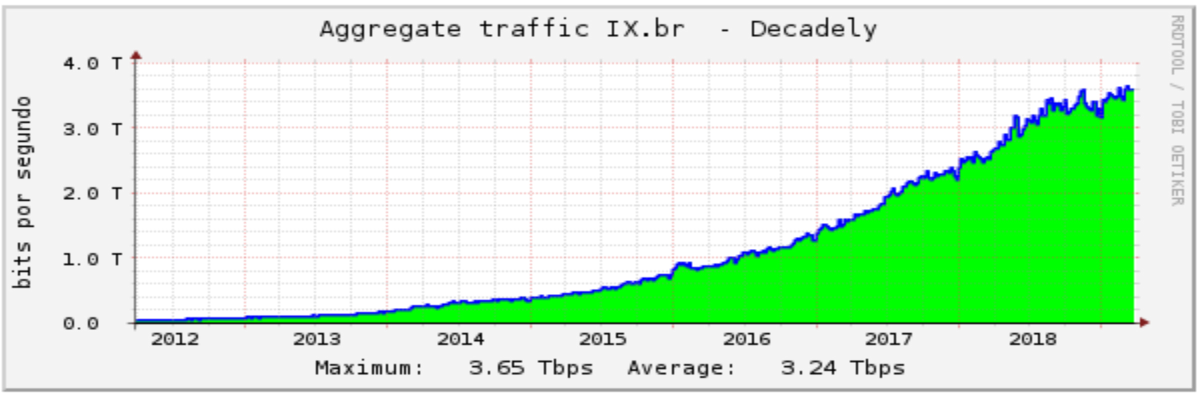

Figure 2 – Multi-year growth graph of instant traffic at the point of exchange IX.br

In accordance with Fig. 2 the instantaneous volume of traffic at the end of 2016 was about 1,4 Tbit/s, whereas at the end of 2017 it increased to 2,4 Tbit/s. The value per year was about 0,71. This value even exceeds all options of $CAGR$, presented in Fig. 1.

Thus, the solution of real problems of traffic prediction can and should be based on the data of monitoring the traffic of specific nodes and data transmission channels.

The purpose of this article is to develop a traffic analysis methodology for large aggregation sites, which requires solving the following tasks:

1. Synthesize the traffic model taking into account its non-stationarity.
2. Develop algorithms and software for traffic analysis.
3. Develop algorithms for parametric identification of predictive traffic models.

Synthesis of traffic patterns. In this case, the classical understanding of teletraffic [9], as a stream of elementary events (such as "call", "failure", "receiving a packet for processing", etc.), is not entirely constructive. Reason: significant amounts of data processed on aggregation sites. In this case, well-developed models such as Markov processes turn out to be unobservable, since the available measuring instruments do not allow separating elementary events from one another. In addition, the ideal solution of the practical problems of forecasting should be reduced not so much to extrapolating the statistical characteristics of the process being studied, as to obtaining an adequate point prediction.

As an example, we give the daily graph of instantaneous total traffic in all directions in the Ukrainian exchange point UA-IX [10] (Fig. 3). This graph gives the results of the preliminary processing of incoming requests for service, which can be expressed as follows:

$$V(t,\tau) = \frac{1}{\tau}\sum_{i=1}^{N} w_i = \frac{W(t,\tau)}{\tau}, \qquad (2)$$

where $V(t,\tau)$ – the result of averaging the flow of applications received for processing, starting from the moment t and ending with the moment $t+\tau$; N – total number of applications (for example, IP packets) received during this time; w_i – length of i-th packet (in bits or bytes); $W(t,\tau)$ – amount of traffic received during time τ. As you can see, the values of the process $V(t,\tau)$ have the dimension of the speed of transmission, reception or processing of data. At small values of the averaging interval τ "instantaneous" increment values of traffic volumes are obtained. Therefore, below, processes of the form (2) will be called instantaneous traffic or simply traffic, omitting the parameter τ.

Volume of total traffic

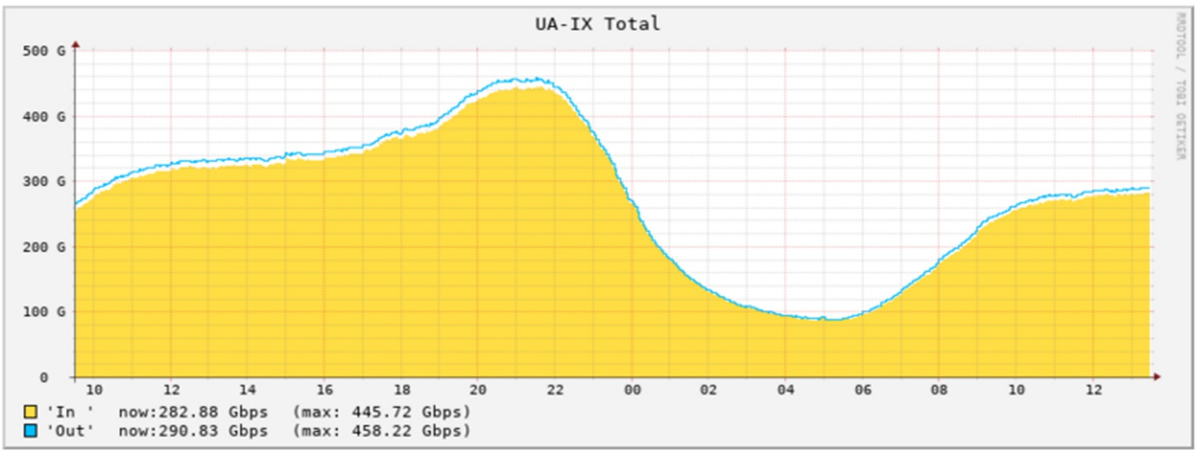

Figure 3 – Volume of total traffic per day at the exchange point UA-IX

For solving the problems of analyzing and predicting traffic at the data link layer, the representation (2) is sufficient. In the case of traffic analysis at the aggregation sites, it may still be necessary to represent it as characteristics of the flow of requests. For example, an important characteristic of routers is the maximum number of packets processed per unit of time. In this case, given the large number of applications, you can determine the average packet length and recalculate the instantaneous traffic values into the number of conditional packets. Errors at the same time will be minimal.

Thus, for our purposes, the representation of traffic in the form (2) is quite sufficient. Also, from Fig. 3 shows that traffic in this case is of a regular nature: there are no sharp outliers, which allows it to be approximated by smooth functions. In addition, the characteristic minima (about 5 o'clock in the morning) and maximums (about 9 pm o'clock) are noticeable. If we conditionally divide the time interval from 8 am to 11 pm into two components - “business” (approximately from 8 to 18 hours) and “home” (approximately from 16 hours), then the prevalence of “home” traffic is noticeable. It is important to note that about 20 years ago, in Ukraine, an inverse pattern was observed: business traffic prevailed. The reason is simple: at the cost of a round-the-clock package of $ 50 per month with dial-up access at a speed of 33.6 Kbps, Internet services turned out to be almost inaccessible for most users. And at work these services were paid by the employer.

Even from a separate graph in Fig. 3, it can be seen that the total incoming traffic on the exchange node is not just strongly correlated with outgoing traffic, but practically repeats it. Analysis of traffic for longer periods shows the presence of additional patterns. Thus, the analysis of graphs at the DE-CIX exchange point in Frankfurt [11] (Fig. 4 and 5) allows us to draw the following conclusions.

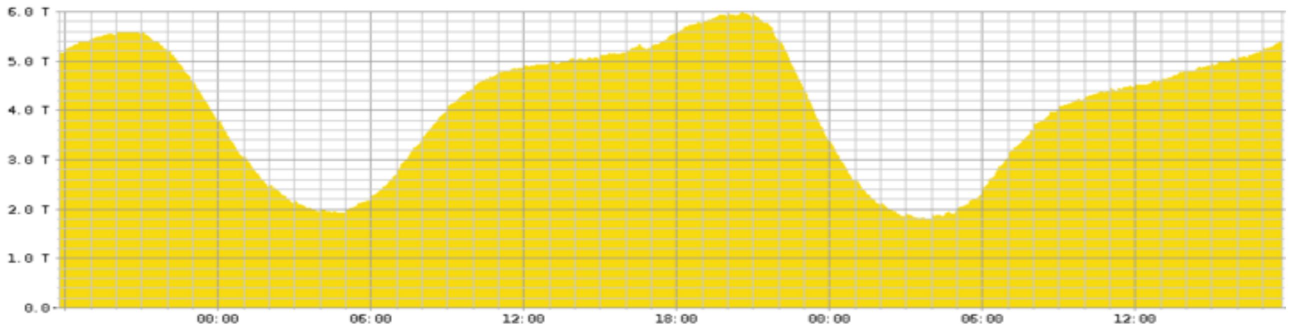

Figure 4 – Traffic at the exchange point of DE-CIX for two days

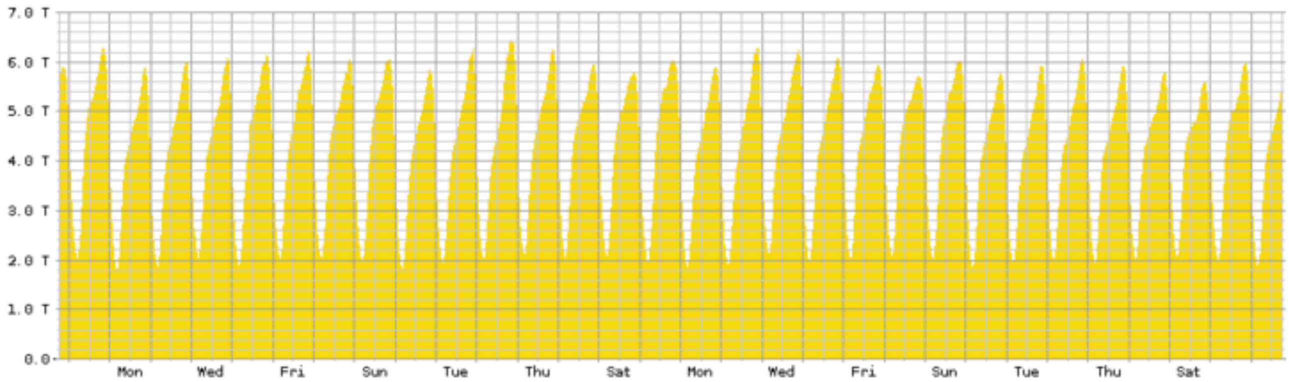

Figure 5 – Traffic at the exchange point DE-CIX for the month

As can be seen, there is a pronounced daily cyclical. In the strict sense, the observed process is not a periodic one, as can be seen from Fig. 5, where some pulsations are also noticeable: in the middle of the week, the process values are slightly higher than on weekends. In the same accounting system [11] a long-term chart is given (Fig. 6).

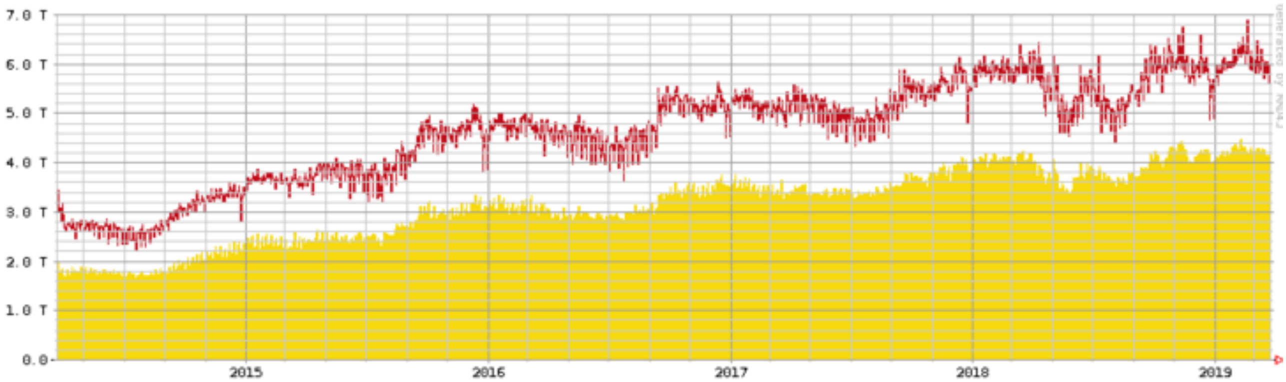

Figure 6 – Traffic at the DE-CIX interchange point for five years

In Fig. 6, patterns of long periods are observed: the presence of seasonal effects (in winter, user activity increases slightly) and the tendency to a gradual increase in traffic.

Taking into account the visible traffic trends at large aggregation sites (non-stationarity and the presence of cyclical patterns), the processes under consideration can be attributed to non-stationary quasi-periodic processes. To synthesize models of such processes, unfortunately, it is difficult to use models of rare events, including models of self-similar traffic [12, 13] and models derived from queuing theory [14]. At the same time, given the almost regular nature of such processes, autoregressive models or moving averages, or composite models, may be productive [15]. The advantage of such models in this case is their obvious connection with the differential equations that describe the observed processes on the substantive level. A certain complexity of using them for large traffic aggregation nodes is the substantial dependence of the obtained solutions on the order of equations and the discretization step. Some information on the order of the equations is provided by the analysis of autocorrelation functions [16], but this method is applicable for stationary processes. In our case, this is not true. The review article [17] considers, among other things, linear models of factor analysis. But what are the real factors involved in processing the analysis of dispersions is far from obvious.

In this article, we accept the traffic model as follows:

$$V(t) = a_0(t) + \sum_{k=1}^{N} a_k(t)\cos(k\omega t + \varphi_k(t)) + \xi(t), \quad \omega = 2\pi / T, \tag{3}$$

where $a_0(t)$ – non-harmonic trend component; $a_k(t)$ – some real functions (hereinafter - harmonic coefficients); N – the number of explicitly taken into account cyclical components of the quasi-harmonic component of the trend; $\varphi_k(t)$ – constituent phases; T – baseline observation period; $\xi(t)$ – implementation of a random process on the studied time interval. Depending on the problem being solved, the parameter T may be days, weeks, years, etc. In the most general form, the functions entering into equation (3) must meet the only formal requirement: to form a system of linearly independent functions.

An expression similar to equation (3) for discrete time is obtained by simply indexing the argument.

Model (3) in general form can describe any implementation of any random process of continuous time. We emphasize that it is precisely one single implementation that the researcher has only in the case of a non-stationary process with non-stationary increments. This model is phenomenological and does not follow from any theory. It is based on numerous observations of traffic schedules, similar to those shown in Fig. 2-5. With the proper choice of the type of free functions, it takes into account the laws of rotation and rotation of the Earth, cycles caused by biological and social factors, geographical features, and many others. For example, the comparison of graphs in Fig. 2 and 5 shows that in countries close to the equator, seasonal effects are practically absent. Another observation: the only European country in which the minimum traffic phase occurs at 5 am is Ukraine (in all other countries the typical minimum phase is 4 am). Or we have a different biological entity, or the decree time does not correspond to human biorhythms.

Instant traffic spectra analysis. P Let us consider the result of the fast Fourier transform (FFT) traffic per week at the node [11] in the form of the amplitude spectrum (Fig. 7). As you can see, the characteristic periodicity reflects the main cycles of human life: day; half day; eight hours (a third of the day to work, a third of the day to rest, a third of the day to sleep); six o'clock (morning, afternoon, evening, night). Note that cycles with shorter periods are less pronounced. The peak at a frequency of 1/7 days (position 5 in Fig. 6) is also noticeable - a weekly cycle. It is due to the social schedule of the week, and the week, in turn, due to the phases of the moon with a period of $7_{1/4}$ days.

Over longer periods, additional patterns appear. In Fig. 8 are graphs of the amplitude spectra of traffic at the same exchange point for 4 weeks. In this case, the FFT was performed separately for each week without overlapping the readings. Here it can be seen that the week shift has practically no effect on the amplitudes of the first four harmonics. Non-stationarity manifests itself in low stability of the amplitude of the weekly cycle.

Definitely, an unexpected result is obtained by analyzing the phase spectrum on the same sample of data for 4 weeks (Fig. 9). At the points of local maxima of the amplitude spectrum, the phases of the cyclic components of model (3) are stationary. It follows that in some cases in the model (3) the constituent phases $\varphi_k(t)$ can be considered constant.

Spectral analysis on samples that include annual intervals provides information on slow cycles, which is required in real forecasting problems. In Fig. 10 shows the traffic graph at the MSK-IX Moscow exchange point [18], and Fig. 11 – its amplitude spectrum. Two obvious peaks are visible here: with a frequency of 1 period per year and approximately 53 cycles per year (weekly cycle).

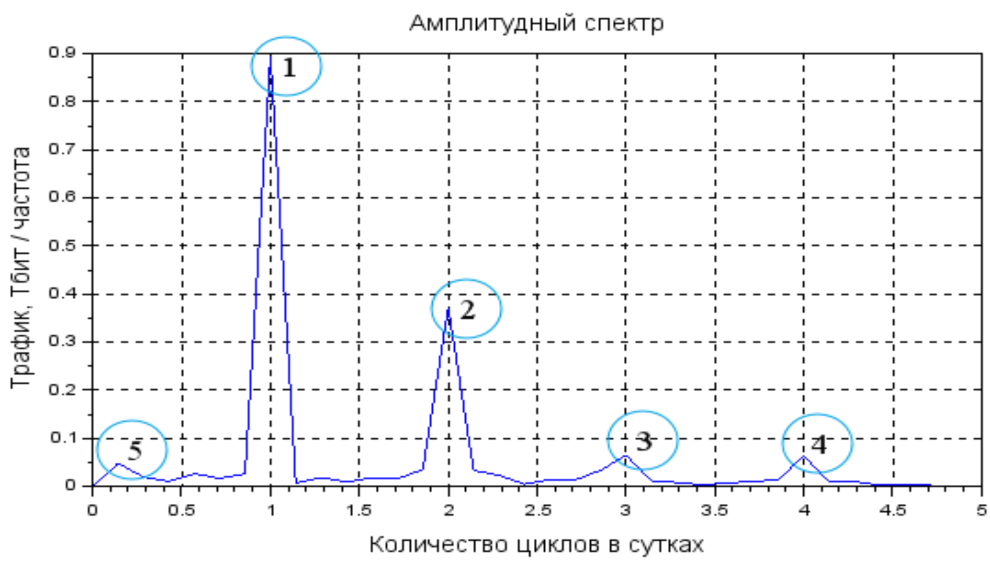

Figure 7 – The amplitude spectrum of traffic at the DE-CIX exchange point for the week: 1 – daily cycle; 2 – day-night cycle; 3 – eight hour cycle; 4 – six hour cycle; 5 – cycle of 7 days

Figure 8 – Traffic amplitude spectra for 4 weeks at DE-CIX exchange

Figure 9 – Phase traffic spectra for 4 weeks at DE-CIX exchange

Figure 10 – Annual schedule of daily traffic maximums for the year at the exchange point MSK-IX

At a different time scale (Fig. 12) the semi-annual cycle (winter - summer) and the three-month cycle (seasons of the year) are clearly visible. There is also a blurred extremum in the vicinity of 13-18 cycles per year (20-28 days), which is consistent with the lunar month and the biological cycles of a part of humanity.

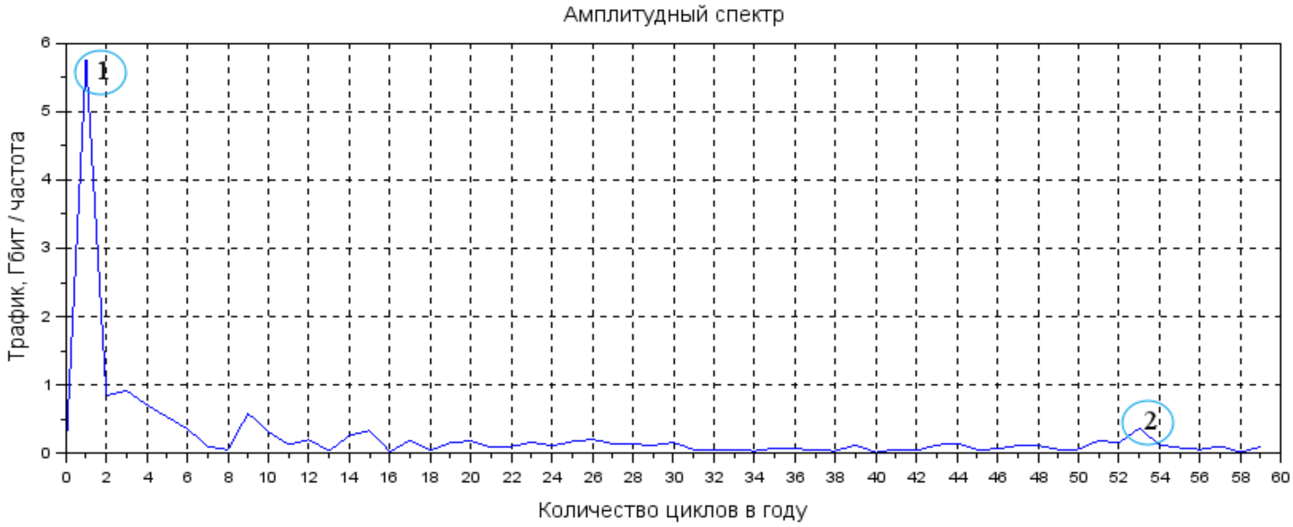

Figure 11 – The amplitude spectrum of the graph in Fig. 9

The four-month cycle (position 1 in Fig. 11) and the forty-day cycle (position 2 in Fig. 11) deserve special attention. In this case, the division of the year into three, rather than four seasons, and the division of the year into nine rather than twelve months is perhaps more appropriate for the person's biorhythms than the periods of the current calendar. Indeed, it can be seen that the amplitude of the 9-month cycle prevails over the amplitude of the 10-month cycle, and the 12-month cycle is generally lost against this background. It can be assumed that for a man, due to the not quite obvious patterns of planetary motion, a year that includes 9 months is more natural. The ten-month calendar is a tribute to the decimal number system. And the 12-month annual cycle is generally complete nonsense to please two narcissistic Roman emperors. From this graph it follows that our ancestors were still wiser: they subordinated their lives to natural laws, and not to social conventions [19].

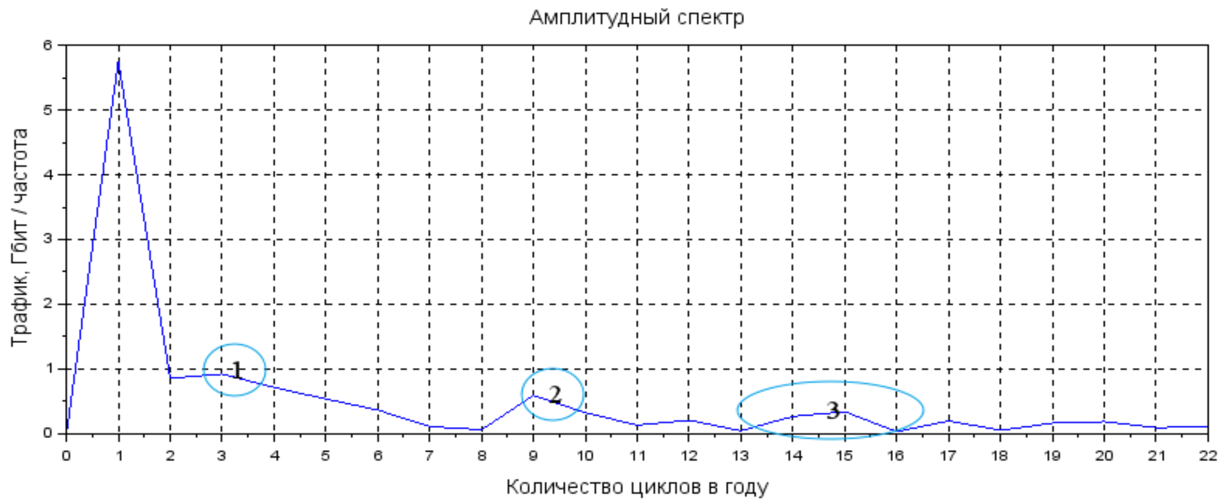

Figure 12 – Additional traffic cycles:

1 – four months; 2 – forty days; 3 – lunar

In all cases presented in Fig. 6 – 12, performed by the FFT was a discrete analogue of the instantaneous spectrum [20]:

$$S(\omega,T)=\int_0^T V(t)\,e^{-j\omega t}dt\,, \tag{4}$$

where $S(\omega,T)$ – traffic spectral density; $V(t)$ – traffic; T – baseline observation period (day, week, month, year). In Fig. 8 and Fig. 9 are given the images of instant traffic for several base periods (week) with a shift by the value of the base period. In this case, the results of instantaneous spectral transformations can be represented as a sequence of functions:

$$S_n(\omega,T)=\int_{(n-1)T}^{nT} V(t)\,e^{-j\omega t}dt,\quad n=1,2...\ . \tag{5}$$

Discrete analogs of transformations (4-5) suggest that at relatively short time intervals (when the observation interval slightly exceeds the base period) in model (3), the phases and harmonic coefficients can be considered approximately constant values. Then the non-harmonic component of the trend $a_0(t)$ makes the greatest contribution to the nonstationarity of the observed processes. Therefore, it is important to parameterize the model of this dependency.

Taking into account the obvious correlation of traffic at large aggregation sites with the growth of the economies of individual countries, in some cases, the compound percentage model is productive (1). A definite lack of an analytical representation of this model is an obvious dependence on the period T. Imagine this relationship in general:

$$a_0(n\tau)=a_0(1+\alpha(\tau))^n\,,$$

where $\alpha(\tau)$ – *COGR* coefficient, represented as an explicit function of period τ. If you run the limit on $\alpha(\tau)\to 0$ in the identity $[1+\alpha(\tau)]^n=[1+\alpha(\tau)]^{\frac{n\alpha(\tau)}{\alpha(\tau)}}$ and use the second remarkable limit, then you can go to the model of continuous time:

$$a_0(n\tau)=a_0(t_0)\exp(n\alpha(\tau))\,. \tag{6}$$

Such a model of “infinite growth” is in good agreement with traffic at the Brazilian exchange point (Fig. 2) at a certain time interval. In Fig. 13 given the original graph at this point of exchange and its approximation using the model (6). The model identification was performed in the area from the last quarter of 2012 to the middle of 2018. As you can see, in this area the model is in good agreement with the traffic measurement data. At the same time, from the end of 2018 to the current moment, there is an imbalance of the forecast. The growth rate of traffic is slowing. The growth rate of traffic is slowing.

$$a_0(t)=A_0(1-\exp(\beta(t))),\quad \beta(t)<0\quad \forall t>0\,, \tag{7}$$

where A_0 – saturation level; $\beta(t)$ – monotone function of t.

The combined use of functions (6) and (7) at different time intervals allows us to simulate typical phases of the development of real systems: growth, growth slowdown, plateau, growth, etc.

The task of separating and predicting these phases is not formally solved. At the same time, a simple and universal model of dependence of anything on time is a polynomial function. The disadvantage of such a model is the sensitivity to the depth of the prediction: the highest components provide “candles” for relatively long-term extrapolations. For short-term predictions, such a model is probably the most adequate.

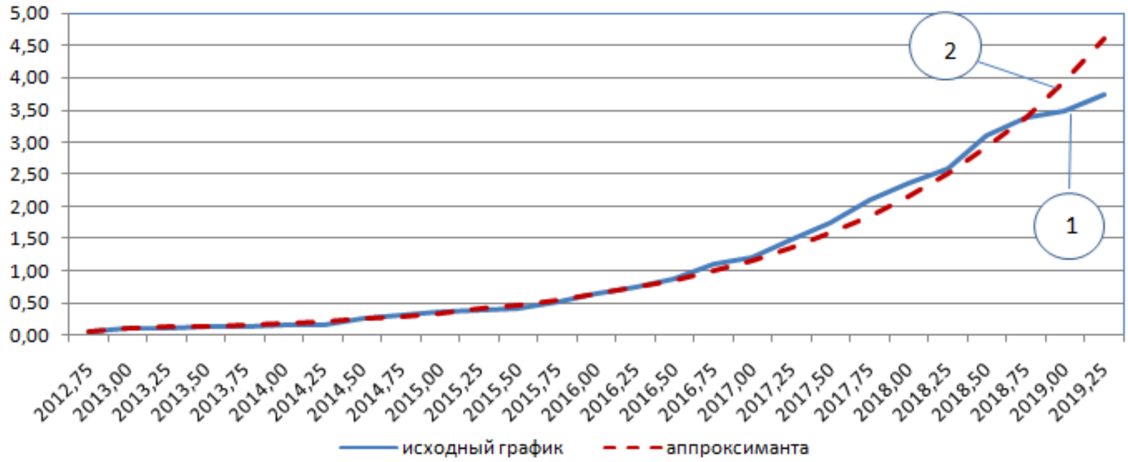

Figure 13 – Approximation of traffic in the Brazilian exchange point:
1 – original traffic graph; 2 – compound percentage approximation

In simple cases, using this type of dependency allows you to select a regular trend component and cyclical components. So, from Fig. 9 it follows that in the short sections of the forecast, the phases of the cyclic components can be considered almost constant. So, from Fig. 8 follows that the amplitudes of the cyclic components are also almost constant over short periods of time. Suppose that the main trend has the form of a linear dependence (1st order polynomial) : $a_0(t) = a_{00} + a_{01}t$, then the expression (3) takes the simple form:

$$V(t) = a_{00} + a_{01}t + \sum_{k=1}^{N} a_k \cos(k\omega t + \varphi_k) + \xi(t), \quad \omega = 2\pi / T. \tag{8}$$

When digitally processing "signals" of the form (8) at frequencies $k\omega t$ additional components arise due to the FFT function $a_{00} + a_{01}t$. For the selection of "true" coefficients $a_k, k > 0$, cyclic components depending on (8) we differentiate twice the right and left parts of this expression:

$$\frac{d^2V(t)}{dt^2} = -\sum_{k=1}^{N} a_k \cos(k\omega t + \varphi_k) + \frac{d^2\xi(t)}{dt^2}, \quad \omega = 2\pi / T. \tag{9}$$

As can be seen, in the representation (9) the main trend is completely absent. Of course, the second derivative of the noise component $\xi(t)$ may also affect the identification of coefficients $a_k, k > 0$. Moreover, this component is due, as a rule, to sudden emissions of observed processes. Its amplitude derivatives can thus significantly exceed the derivatives of the regular components. In this case, to identify the "true" amplitudes of cyclical components, the following identification algorithm turned out to be productive.

Step 1. Over the original data set is performed FFT.

Step 2. The procedure of screening high-frequency components is performed - a low-pass filter is applied.

Step 3. The inverse FFT is performed.

Step 4. A numerical differentiation of order n+1 is performed, where n – maximum degree of polynomial dependence $a_0(t)$.

Step 5. FFT is performed.

Step 6. Band-pass filters are allocated amplitude coefficients of cyclic components.

Step 7. The inverse FFT is performed taking into account the phases and the main trend and noise component is distinguished by the difference of functions. Next, the noise component is eliminated by statistical data processing.

The application of this algorithm in the cases under consideration gives quite adequate analysis results. So, in Fig. 14 shows the results of the identification of a two-week traffic model. Individual emissions in positions 1 and 2 do not significantly affect the accuracy of the approximation. The relative error of the model in this case is only 3%, which can be considered a negligible value.

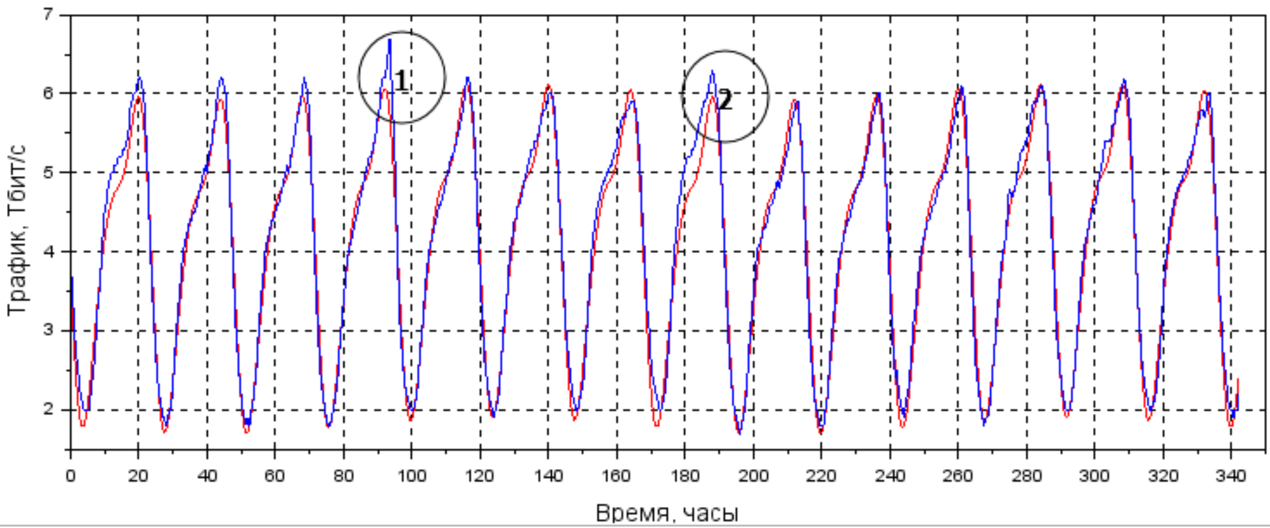

Figure 14 – Approximation of traffic at the DE-CIX Frankfurt exchange point

At the same time, in model (8) there was a limitation of the number of cyclic components $N = 4$. Thereby, an additional effect of information compression is achieved. The initial data set contains 685 samples, and for the approximation 4 amplitude values are used, 4 phase values and one constant is a constant level. Compress almost 2 orders of magnitude!

Conclusions

1. The solution of traffic prediction problems is relevant for any operator.

2. On large aggregation sites, there is a mass effect of events, when traffic can be considered an almost regular function of time.

3. In the short time sections, the studied traffic has approximately constant phases of the main cyclic components and approximately constant amplitudes.

4. Long-term trends can be modeled as a function of increase in the form of a complex percentage, models with saturation and polynomial functions.

5. Model in the form of a non-stationary quasi-periodic process adequate to the problem of traffic prediction at large aggregation sites.

6. Typical periodicity, clearly following from the analysis of numerous graphs at international exchange nodes: daily, weekly, seasonal. The article also noted some unexpected periodicities, which are still to be understood.

REFERENCES:

1. Odegov N.A Osnovy teorii approksimacii opticheskih signalov gaussovymi impul'sami / N.A. Odegov // Visnik Hmel'nic'kogo nacional'nogo universitetu. Tekhnichni nauki. – 2018. – № 5. – P. 42–72.
2. BigData shagaet po planete. Kolichestvo dannyh v mire udvaivaetsya kazhdye dva goda [Elektronnyj resurs]. – Rezhim dostupa: https://rg.ru/2013/05/14/infa-site.html.
3. Informacionnyj vzryv [Elektronnyj resurs]. – Rezhim dostupa: https://ru.wikipedia.org/wiki/Informacionnyj_vzryv.
4. Izmerenie informacionnogo obshchestva / Mezhdunarodnyj soyuz elektrosvyazi // Place des Nations

CH1211 Geneva Switzerland, 2014 [Elektronnyj resurs]. – Rezhim dostupa: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2014/MIS_2014_Exec-sum-R.pdf.

5. Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper. [Elektronnyj resurs] – Rezhim dostupa: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html.
6. Internet-trafik (mirovoj rynok). Prognoz Nokia Bell Labs [Elektronnyj resurs]. – Rezhim dostupa: http://www.tadviser.ru/index.php/ Stat'ya: Internet-trafik_(Mirovoj rynok).
7. Multi Router Traffic Grapher [Elektronnyj resurs] – Rezhim dostupa: https://oss.oetiker.ch/mrtg/.
8. Total traffic (all IX.br) [Elektronnyj resurs] – Rezhim dostupa: http://ix.br/cgi-bin/all.
9. Kornyshev Yu.N. Teoriya teletrafika: uchebnoe posobie / Yu.N. Kornyshev, A.P. Pshenichnikov, A.D. Harkevich. – M.: Radio i svyaz', 1996. – 272 s.
10. Statistika: ob"em summarnogo trafika v tochke obmena UA-IX [Elektronnyj resurs]. – Rezhim dostupa: http://www.tadviser.ru/index.php/ https://www.ix.net.ua/ru/o-kompanii/statistika.
11. DE-CIX Frankfurt statistics [Elektronnyj resurs]. – Rezhim dostupa: https://www.msk-ix.ru/traffic/.
12. Lozhkovskyi A.G. Povyshenie tochnosti rascheta koefficienta samopodobnosti trafika paketnoj seti svyazi / A.G. Lozhkovskyi, E.V. Levenberg // Naukovi praci ONAZ im. O.S. Popova, 2017, № 2. – S. 63-68.
13. Strelkovskaya I. Estimation of qos characteristics of self-similar traffic for the wb/m/1/ queuing system / Strelkovskaya I., Siemens E., Solovskaya I., Fedotova I. // Naukovi praci ONAZ im. O.S. Popova, 2018, № 1. – S. 27-33.
14. Lozhkovskyi A.G. Teoriya masovogo obslugovuvannya v telekomunikaciyah / Lozhkovskyi A.G. – Odesa: ONAZ im. O.S. Popova, 2010. – 112 s.
15. Nikityuk L.A. Cognitive monitoring in an information and communication network operation system / Nikityuk L.A. // Naukovi praci ONAZ im. O.S. Popova, 2018, № 2. – S. 48-58.
16. Ovchinnikov K.A. Prognozirovanie setevogo trafika pri pomoshchi avtoregressionnyh modelej / Ovchinnikov K.A., Bushmanov V.S. Pervaya Mezhdunarodnaya nauchno-prakticheskaya konferenciya «Problemy infokommunikacij. Nauka i tekhnologii». – Har'kov, 9-11 sentyabrya 2013 g. – S. 177-179.
17. Rukkas K.M. Sravnitel'nyj analiz metodov prognozirovaniya trafika v telekommunikacionnyh sistemah / Rukkas K.M., Solyanik Yu.V., Ovchinnikov K.A., Olotu Oluvatosin David // Problemy telekommunikacij № 1 (13), 2014. – S. 84-95.
18. MSK-IX / Trafik [Elektronnyj resurs]. – Rezhim dostupa: https://www.msk-ix.ru/traffic/.
19. Slavyanskij kalendar': mesyac 40 dnej, nedelya 9 dnej [Elektronnyj resurs]. – Rezhim dostupa: http://derzhavarus.ru/kalendar-leto-mesyac-nedelya.html.
20. Harkevich A.A. Spektry i analiz / Harkevich A.A. – M.: Knizhnyj dom «LIBKOM», 2009. – 240 s.

ЛИТЕРАТУРА:

1. Одегов Н.А Основы теории аппроксимации оптических сигналов гауссовыми импульсами / Н.А. Одегов // Вісник Хмельницького національного університету. Технічні науки. – 2018. – № 5. – С. 42–72.
2. BigData шагает по планете. Количество данных в мире удваивается каждые два года [Электронный ресурс]. – Режим доступа: https://rg.ru/2013/05/14/infa-site.html.
3. Информационный взрыв [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/ Информационный_взрыв.
4. Измерение информационного общества / Международный союз электросвязи // Place des Nations CH-1211 Geneva Switzerland, 2014 [Электронный ресурс]. – Режим доступа: https://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2014/MIS_2014_Exec-sum-R.pdf.
5. Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper. [Электронный ресурс] – Режим доступа: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html.
6. Интернет-трафик (мировой рынок). Прогноз Nokia Bell Labs [Электронный ресурс]. – Режим доступа: http://www.tadviser.ru/index.php/ Статья: Интернет-трафик_(Мировой рынок).
7. Multi Router Traffic Grapher [Электронный ресурс]. – Режим доступа: https://oss.oetiker.ch/mrtg/.
8. Total traffic (all IX.br) [Электронный ресурс]. – Режим доступа: http://ix.br/cgi-bin/all.
9. Корнышев Ю.Н. Теория телетрафика: учебное пособие / Корнышев Ю.Н., Пшеничников А.П., Харкевич А.Д. – М.: Радио и связь, 1996. – 272 с.
10. Статистика: объем суммарного трафика в точке обмена UA-IX [Электронный ресурс]. – Режим доступа: http://www.tadviser.ru/index.php/ https://www.ix.net.ua/ru/o-kompanii/statistika.
11. DE-CIX Frankfurt statistics [Электронный ресурс]. – Режим доступа: https://www.msk-ix.ru/traffic/.
12. Ложковский А.Г. Повышение точности расчета коэффициента самоподобности трафика пакетной сети связи / А.Г. Ложковский, Е.В. Левенберг // Наукові праці ОНАЗ ім. О.С. Попова, 2017. – № 2. – С. 63-68.

13. Strelkovskaya I. Estimation of qos characteristics of self-similar traffic for the wb/m/1/ queuing system / Strelkovskaya I., Siemens E., Solovskaya I., Fedotova I. // Наукові праці ОНАЗ ім. О.С. Попова, 2018. – № 1. – С. 27-33.
14. Ложковський А.Г. Теорія масового обслуговування в телекомунікаціях / Ложковський. А.Г. – Одеса: ОНАЗ ім. О.С. Попова, 2010. – 112 с.
15. Nikityuk L.A. Cognitive monitoring in an information and communication network operation system / Nikityuk L.A. // Наукові праці ОНАЗ ім. О.С. Попова, 2018. – № 2. – С. 48-58.
16. Овчинников К.А. Прогнозирование сетевого трафика при помощи авторегрессионных моделей / К.А. Овчинников, В.С. Бушманов. // Первая Международная науч.-практ. конф. [«Проблемы инфокоммуникаций. Наука и технологии»], (Харьков, 9-11 сентября 2013 г.). – Харьков,2013. – С. 177-179.
17. Руккас К.М. Сравнительный анализ методов прогнозирования трафика в телекоммуникационных системах / Руккас К.М., Соляник Ю.В., Овчинников К.А., Олоту Олуватосин Давид // Проблемы телекоммуникаций, 2014. – № 1 (13). – С. 84-95.
18. MSK-IX / Трафик [Электронный ресурс]. – Режим доступа: https://www.msk-ix.ru/traffic/.
19. Славянский календарь: месяц 40 дней, неделя 9 дней [Электронный ресурс]. – Режим доступа: http://derzhavarus.ru/kalendar-leto-mesyac-nedelya.html.
20. Харкевич А.А. Спектры и анализ / Харкевич А.А. – М.: Книжный дом «ЛИБКОМ», 2009. – 240 с.

DOI 10.33243/2518-7139-2019-1-1-121-134