TRANSPORT
2009
24(4): 339–344

THE RAILWAY INFORMATION SYSTEM AS A QUEUEING SYSTEM

Alla Melikyan

Transport and Telecommunication Institute, 1 Lomonosov str., LV-1019 Riga, Latvia
E-mail: melik@aircom.lv

Received 5 February 2009; accepted 3 November 2009

Abstract. The present article dwells upon the Information System (IS) of Latvian railways providing the analysis of the queuing system (QS) used to removing defects occurring in the infinite linear queuing system. The author of the article analyzes research data about the failures of system functioning in order to establish the parameters of the before mentioned system and other information systems used by Latvian railways.

The structure of accidents happening in the informative systems and the removal of possible defects are researched. The present investigation provides evidence that substantiates a Markov hypothesis about the peculiarities of IS service.

The author also examines the organizational structure of IS taking into consideration the level of the danger of possible accidents. The statistics of the failures of IS relates to the statistics of accidents that might happen to rolling stock.

Keywords: statistics, information systems, failures, queuing system, railway, frequency of failures.

1. Introduction

The author of the published article provides an insight into the process of exploiting the Information System (IS) by Latvian railways and analyzes the situation in terms of the queuing system (QS) functioning. Different approaches to investigating this problem, for example, the method of calculating the medium indicators of system functioning are developed. The applied method enables to analyze emergency situations such as a random increase in the intensity of a local stream of queries that involves using reserve possibilities of QS.

In the process of investigating the functioning of IS, the author has found out that different subsystems of IS randomly break and require the removals of nascent defects.

The defects happen quite rarely and are quickly removed. Thus, no necessity arises to make customers to wait for the service of queries.

The received query on repairing the functioning of IS has been immediately accepted and served. The author of the paper has chosen the infinite linear queuing system as the basis for research. The system will serve as the model for analyzing the exploitation of IS (Ganesh *et al*. 2004). The above mentioned source has stated that the received queries are immediately removed and a lack of service channels is never felt.

The infinite linear system always satisfies the needs of its clients. Thus, its relative carrying capacity, i.e. the part of the served queries from a number of receiving one is equals to 1.

The present research focuses on distributing the number of the busy channels of such system and the statistical results of its medium.

Therefore, the author analyzes IS as a source of the queries for removing defects in the infinite linear queuing system. For establishing the parameters of such system, the researcher should analyze the actual data about the failures of the functioning of some informational systems used by Latvian railways.

The author of the article goes into the structure of failures that may happen in the functioning of the informative systems and the removal of such defects. The undertaken research is meant for supporting the hypothesis proposed by Markov, exactly, the hypothesis about the peculiarities of IS service.

The author also examines the organizational structure of IS taking into consideration the level of the danger of possible accidents. The statistics of the failures of IS relates to the statistics of accidents that might happen to rolling stock.

2. Quantitative and Qualitative Analysis of the Incoming Queries

First, in order to investigate QS, it is necessary to define the qualitative and quantitative peculiarities of the incoming queries. Therefore, a book of recorded failures that

ISSN 1648-4142 print / ISSN 1648-3480 online
www.transport.vgtu.lt
DOI: 10.3846/1648-4142.2009.24.339-344

happened to IS in the period from 2004 to 2007 is used. The facts listed in the registration book contain information about the dates and times of accidents, the category of danger caused by failure, the time required for its removal and the name of a subsystem where failure occurred.

Considering the dates of incidents t, the series of time intervals $\tau_I = t_i - t_{i-1}$ existing between the queries about service containing observations $n = 223$ were constructed. The obtained data was grouped according to the duration of intervals between the queries and transferred to group 18. The results of grouping are illustrated in Fig. 1 and Table 1 (the borders of intervals are counted in accordance with time duration: 24 hours or it can be also defined as a day).

Fig. 1 demonstrates that the most efficient type of distributing the duration of intervals between the queries is exponential distribution,

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0, \end{cases} \quad (1)$$

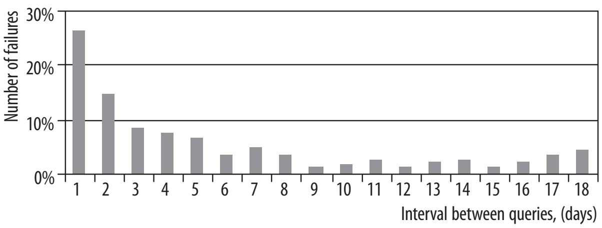

Fig. 1. The amount of accidents considering the intervals between the queries (time duration: twenty four hours)

Table 1. The borders of intervals in days (time duration: twenty four hours)

No	Interval	n_i	h_i	p_i
1	0–1	59	0.265	0.169
2	1–2	33	0.148	0.140
3	2–3	21	0.094	0.117
4	3–4	18	0.081	0.097
5	4–5	15	0.067	0.081
6	5–6	8	0.036	0.067
7	6–7	11	0.049	0.056
8	7–8	8	0.036	0.046
9	8–9	3	0.013	0.038
10	9–10	4	0.018	0.032
11	10–11	6	0.027	0.027
12	11–12	3	0.013	0.022
13	12–13	4	0.018	0.018
14	13–14	6	0.027	0.015
15	14–15	3	0.013	0.013
16	15–17	5	0.022	0.019
17	17–20	6	0.027	0.018
18	>20	10	0.045	0.025

where: λ – the distribution of parameter; t – time.

The parameter λ of this distribution can be achieved according to the following statement: the medium result is exponential to the distributed casual result (Gray 1988).

$$M\left[T\right] = \frac{1}{\lambda}, \quad (2)$$

where: $M[T]$ denotes the amount of time intervals between the queries.

Furthermore, the statistics of research provides the estimation of the medium length of intervals between the queries:

$$\frac{1}{n}\sum_{i=1}^{n}\tau_i = 5.41 \;(\text{twenty four hours}), \quad (3)$$

where: τ_i – time interval between the queries; n – the number of observations.

Thus, we obtain the following result:

$$\lambda \approx \frac{1}{5.41} = 0.185\,\text{failures/day}. \quad (4)$$

Thus, a hypothesis about the type of distributing the intervals between the queries is proved and the frequency of this distribution is counted. Exactly, the frequency of using this figure in the chosen intervals was investigated:

$$h_i = \frac{n_i}{n}, \quad (5)$$

where: n_i – the frequency of observations used in grouping intervals; n – the number of observations used in grouping intervals

A theoretical possibility of agreement between the exponentially distributed casual result and parameter 0.185 and its appropriateness for p_i are calculated to select the criterion.

$$\chi^2 = \sum_{i=1}^{18}\frac{\left(h_i - p_i\right)^2}{p_i}\cdot n, \quad (6)$$

where: p_i – the calculated possibility of agreement between random variable T and interval grouping.

Data on the made calculation is presented in Table 1.

The value of selective criterion χ^2 equals 30.5 which is less than the result of distribution, $\chi^2_{0.99}(16) \approx 32$. Thus, when the level of significance is 0.01, the following hypothesis is proved: time interval between the queries is distributed according to the model law with parameter 0.185. Therefore, the process of the incoming queries from IS can be described by the term *poisson*, i.e. random ordinary incoming queries without any consequences.

The conducted research analyzes the frequency of failures occurring in IS in order to investigate the invariance of the incoming queries (Fig. 2). The records about the failures happening in IS were grouped according to the months and levels of complications caused by them.

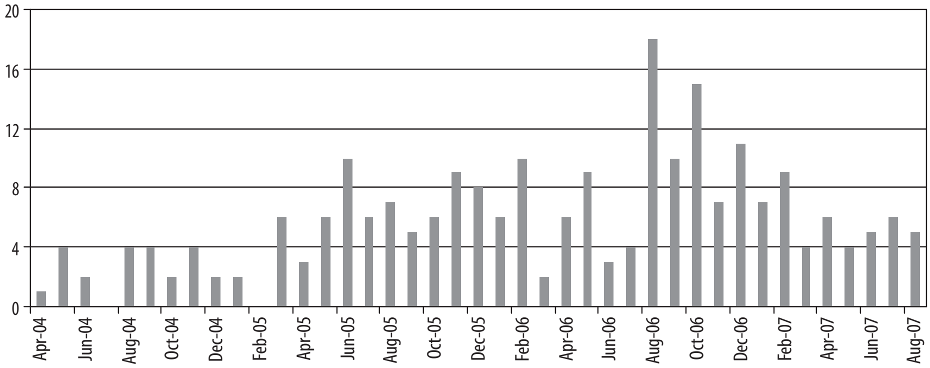

Fig. 2. The series of failures in IS considering time periods (number of failures by month, 2004–2007)

The majority of accidents were categorized as the first level ones, and therefore it was not necessary to conduct differentiation considering levels.

The analysis of time series has shown the absence of a seasonal component of accident frequency and autocorrelation has been discovered within the limits of the level of white noise in all lags (Fig. 3).

However, the analysis of daily fluctuations in the intensity of failures has shown a significant interrelation between the frequency of failures in IS and daytime. The author of research has prepared statistics on daily fluctuations in the frequency of failures having grouped the records provided in the registration book according to the hours the accidents happened within the time period from 2004 to 2007 (see Fig. 4).

Fig. 4 illustrates that the majority of failures occur in the morning and afternoon hours and make 53% of all failures happening from 7.00 a.m. to 2.00 p.m. The highest intensity of IS load is observed at 8.00 a.m. and at the midday. A local increase in the intensity of the load in the evening and night hours is observed at midnight.

Thus, the queries about the removal of failures that appear in IS vary during daytime. Designating through the *h(t)* part the numbers of failures during an hour from t till $t+1$ in the general amount of failures, one can define the intensity of the queries during daytime as $\lambda(t) = \lambda \times h(t)$.

The researcher may analyze the instability of the queries from two different angles which should be done to evaluate the functioning of IS. On the one hand, daily fluctuations in intensity can be ignored. In this case, there should be counted the average intensity of the queries occurring at different daytime. This approach allows getting though a simple but approximate model of the queuing system. On the other hand, it is possible to divide daytime into several temporal intervals that are to be characterized by the permanent intensity of the queries. Then, the functioning of the system should be analyzed separately in accordance with each of these intervals. This approach is recommended in the case of applying the flexible regime of managing resources used for the queuing system.

Unexpected result was obtained analyzing the registration book records on the subsystems of IS that cause different failures. The carried out research discovered that the majority of failures was caused by three subsystems of railway IS – the automated control system of stations SAVS, the automated system of the operative management of transportations APOVS as well as by the system analyzing the incomes of freight traffic and calculations with clients APIKS (Fig. 5). The subsystems EXPRESIS and INTERREPORT also cause a number of problems.

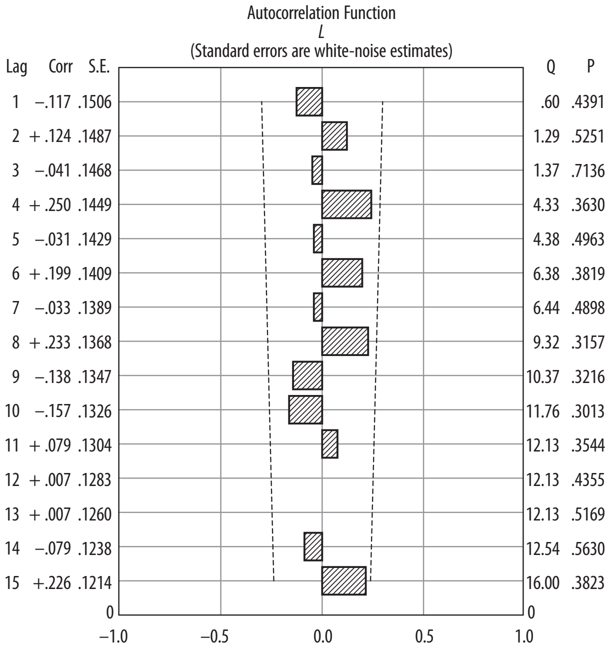

Fig. 3. A correlogram of the time series of failures in IS (Melikyan 2007)

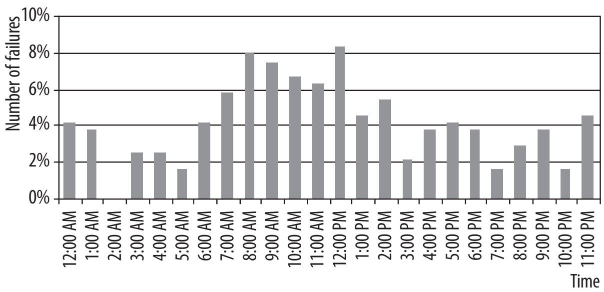

Fig. 4. The frequency of failures in IS considering hours (%)

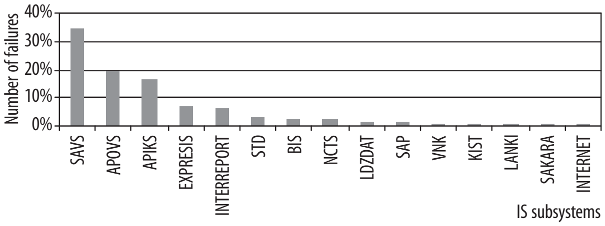

Fig. 5. The amount of failures considering subsystems

On the one hand, this result is important because it emphasizes the weak links of IS (Домарев 2001). Thus, it is necessary to pay special attention to these links during the exploitation of IS. On the other hand, it allows differentiation when describing QS that serves IS. Besides, it is possible to choose several sources of queries, i.e. the subsystems of IS. These items have different quantitative descriptions such as the intensity of the queries.

3. Time Distribution for the Analysis of the Queries

Another important characteristic of QS is time distribution for query satisfaction.

The author of the carried out research did certain calculations meant to study the following issues: time distribution for repairs according to the dates the failure occurred and the query was received; time distribution for the staff dealing with IS and asking a repair brigade to remove the failure. The following items are involved in the process of calculation: t_1 denotes time distribution for the staff dealing with IS, t_2 denotes time distribution for necessary repairs. Thus, the author did the following calculation meant to discover the time of repairs: $\tau_I = t_{2i} - t_{1i}$ containing $n = 224$ supervisions.

The data was analyzed according to the duration of the intervals between the queries forming 11 groups. The results of grouping are presented in Table 2.

It is necessary to mention a possibility of examining different subsystems of IS as the independent sources of the queries for service. Removing the failures of different subsystems requires different periods of time which should be done prior to making any conclusions according to the data shown in Table 2. Furthermore, Fig. 6 illustrates that removing the failures of APIKS subsystem requires the largest amount of time.

Table 2. The distribution of intervals considering hours

No	Interval	n_i	h_i	p_i
1	0–1	67	0.299	0.217
2	1–2	49	0.219	0.170
3	2–3	23	0.103	0.133
4	3–4	17	0.076	0.104
5	4–5	13	0.058	0.082
6	5–6	9	0.040	0.064
7	6–7	8	0.036	0.050
8	7–8	8	0.036	0.039
9	8–9	3	0.013	0.031
10	9–10	5	0.022	0.024
11	>10	22	0.098	0.087

Taking into consideration the analysis of statistics revealing the amount of time needed for repairing different subsystems in this case, the author examines data on time distribution (see Figs 7 and 8).

In accordance with the theory, a sum of a large number of independent services, the exponential distribution of time for which is not full enough, allows to make a conclusion on the *poisson* nature of this phenomenon. The result of grouping the data according to time spent for repairing all subsystems of IS (Table 2) is illustrated in Fig. 9 showing similarities existing between the processes described above and the exponential law.

The author of the article did certain calculations in order to prove the following hypothesis: time needed for removing the failure is distributed according to the exponential law. The author also calculated the average time limit for repair. Thus,

$$f_T(t) = \begin{cases} \mu e^{-\mu t}, & t \geq 0, \\ 0 & , t < 0. \end{cases} \quad (7)$$

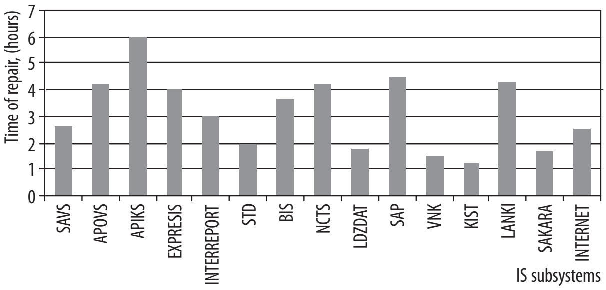

Fig. 6. The average time limit needed for repairing subsystems

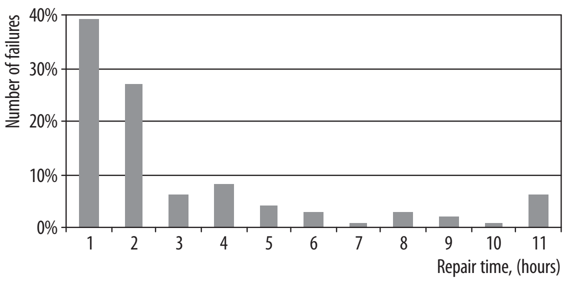

Fig. 7. Time distribution needed for repairing SAVS

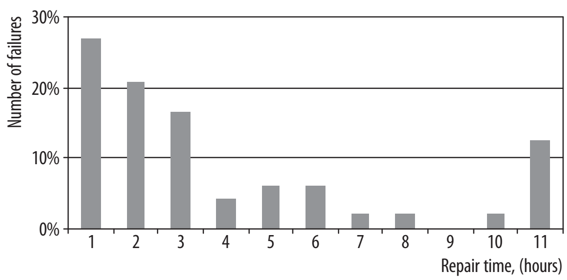

Fig. 8. Time distribution needed for repairing APIKS

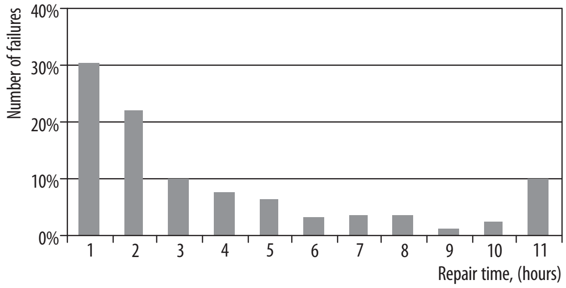

Fig. 9. Time distribution needed for repairing all subsystems

According to the suggested statistics, the amount of time is as follows:

$$\frac{1}{n}\sum_{i=1}^{n}\tau_i \approx 4.09 \text{ hours}. \quad (8)$$

So, the estimation of the exponential distribution parameter is:

$$\mu \approx \frac{1}{4.09} = 0.245 \text{ failures/hour} \quad (9)$$

or 5.87 inquiries a day.

The author of the study also took into consideration the data compiled in Table 2 and calculated the selective criterion – χ^2 the value of which appeared to be 20.4, that is less than the distribution of the obtained result $\chi^2_{0.99}(9) \approx 21.7$. Thus, a hypothesis that time needed for removing the failure is distributed according to the exponential law is proved and can be evidenced by the fact when the level of signification is 0.01 and the parameter is 5.87. Therefore, service provided for the queries (QS) and the queries of IS can be described as having *poisson* nature.

4. Calculation of QS Efficiency Indicators

The above described calculations allow analyzing QS as the infinite linear system proposed by Markov. The state of QS should be defined as constantly changing due to the influence and service of these queries. These processes have *poisson* nature. The following graph illustrates QS having stationary queries (Fig. 10).

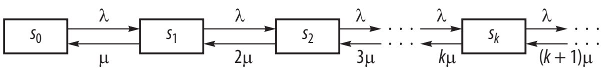

Fig. 10. The graph of infinite linear QS

The graph illustrates the process of calculating s_k, where k denoting the queries are served by the channels of system k, in particular, s_0 is the state when the system is free and does not serve any query. The arrows represent the possible transformations of QS from one state into another. The transitions from the left to the right are initiated by the queries having intensity λ and from the right to the left – by servicing these queries with k working channels having intensity $k\mu$.

Thus, the following system of Kolmagorov's differential equations is constructed due to this graph:

$$\begin{cases} \dfrac{dp_0}{dt} = -\lambda p_0 + \mu p_1 \\ \dfrac{dp_1}{dt} = \lambda p_0 - (\lambda + \mu) p_1 + 2\mu p_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dp_k}{dt} = \lambda p_{k-1} - (\lambda + k\mu) p_k + (k+1)\mu p_{k+1} \\ \dots\dots\dots\dots\dots\dots\dots\dots \end{cases} \quad (10)$$

A solution to it provides the possibility of discovering the distribution of the number of QS busy channels (Papadopoulos *et al.* 1993). Hence,

$$p_k = \frac{\rho^k}{k!} e^{-\rho}, \quad (11)$$

where: $\rho = \dfrac{\lambda}{\mu}$ and $k = 0..\infty$.

The probability that the system is going to be stopped is p_0=96.9 in the case when λ = 0.185 and μ = 5.87 queries a day which corresponds to the average intensity of the amount of the queries during the day. Furthermore, it is possible that one query is served with index p_1 = 3.1, whereas two queries are served with index p_2=0.05. In this case, emergency work is inevitable. So, the necessity to remove more than one failure at the same time arises which is an unusual phenomenon of IS.

The intensity of the queries increases up to 1.9 in the rush hour exactly between 8.00 a.m. and the midday. Thus, in comparison with the average result it reaches λ_{krit} = 0.355 queries a day, and therefore during these hours, QS does not function with full intensity having index p_0 = 94.1, one failure is removed with index p_1 = 5.7 and two failures are simultaneously removed with index p_2 = 0.17. The latter result exceeds an analogical result of the average intensity of failures for 3.5.

Therefore, the above calculations testify to the fact that one repair brigade is able to manage the problems of IS. Thus, QS may have only one channel. It is possible to manage the emergency cases happening quite rarely. If the failure occurs, the queuing system is used and the average length of the queries should be quite short.

5. Interrelation Between the Failures of IS and Railway Accidents

The author of the carried out investigation has also made another observation including interrelation that exists between the frequency of accidents in IS and the frequency of accidents and the rolling stock of Latvian railways. The author has analyzed statistics on accidents occurring on Latvian railways. The interrelation of the failures of IS and railway accidents was investigated. Furthermore, it was discovered that no significant correlation existed between these phenomena. Having calculated time needed for repairing IS, the author has found out the interrelation between time needed for repairing IS denoted as L and the amount of accidents denoted as A. Thus,

$$L = \sum \tau_i \qquad (12)$$

where: calculation was conducted for all subsystems each month and the interrelation between time L and the number of accidents A exists (Fig. 11).

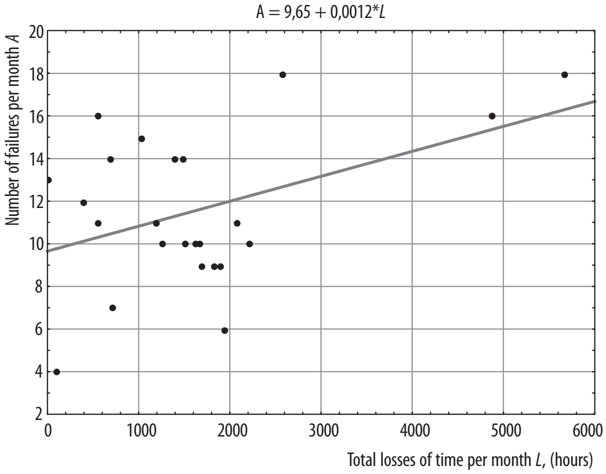

Fig. 11. The interrelation between the amount of accidents and time needed for repairing IS

According to the data obtained, the factor of correlation between these phenomena makes 0.43. The factor of determining the model of regression is low and makes only 0.18. The author of research has worked on the earlier studies on the influence of natural factors on the frequency of accidents happening to different kinds of transport and come to the conclusion it is impossible to specify the reason of the accident to full extent. The failures occurring in IS can be characterized as indirect factors.

6. Conclusions

To sum up, the analysis of the collected statistics on the failures of IS in Latvian railways provides evidence that substantiates the following hypothesis: the process of repairing needed to remove failures is to be defined as the Markov process used for QS. The author of the article has investigated the structure and parameters of IS that can be described as the source of the queries. Research covered the area of temporal needs for repairing IS. The basic characteristics of IS functioning were stated.

The undertaken study has evaluated the degree of the influence on the frequency of failures, the amount of time needed for removing these failures and the frequency of accidents occurring to the rolling stock of Latvian railways.

References

Ganesh, A.; O'Connell, N.; Wischik, D. 2004. *Big Queues*. Springer-Verlag. 254 p. doi: 10.1007/b95197

Gray, R. M. 1987. *Probability, Random Processes and Ergodic Properties*. Springer. 295 p.

Papadopoulos, H. T.; Heavey, C.; Browne, J. 1993. *Queuing Theory in Manufacturing Systems Analysis and Design*. Springer. 420 p.

Melikyan, A.; Misnevs, B.; Seregin, R. 2007. Statistical research of weather conditions influence on the railway transport accident rate, in *Proceedings of the 7th International Conference Reliability and Statistics in Transportation and Communication (RelStat'07)*, 24–27 October 2007 Riga, Latvia, 120–126.

Домарев, В. В. 2001. *Безопасность информационных технологий. Методология создания систем защиты* [Domarev, V. V. Safety of Information Technologies. The Methodology for the Establishment of Protection Systems]. Киев: ДиаСофт. 688 с. (in Russian).