УЧЕНЫЕ ЗАПИСКИ КАЗАНСКОГО УНИВЕРСИТЕТА.
СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ

2018, Т. 160, кн. 2 ISSN 2541-7746 (Print)
С. 392–398 ISSN 2500-2198 (Online)

UDK 519.257

AVIATION INDUSTRY STOCHASTIC MODEL BASED ON BIG DATA CONCEPT

S.S. Valeev, N.V. Kondratyeva

Ufa State Aviation Technical University, Ufa, 450008 Russia

Abstract

Modern aviation industry is a complex large-scale system that includes designing, testing, production, operation, maintains services, etc. Thus, it seems reasonable to provide information logistical support throughout the life cycle of single processes in this area. On the other hand, it allows building stochastic models to manage aviation industry in real-time mode under uncertainties and risks conditions. It is possible to apply this approach by using Big Data concept. The multi-agent method has been suggested to build the Markov decision process system model of the airline industry life cycle. For discovering the main control action to reduce costs of an airline company, the theoretic-information approach has been applied. The results of the analysis determine the agent action during the decision process to control the amount of freight related to fuel consumption.

Keywords: aviation industry, efficiency, stochastic model, decision making, Big Data

Introduction

Under the conditions of digital economy, aviation industry must be considered as an ultralarge-scale system. The ultra-large-scale system is a subclass of systems of the systems class. The system of system is a system that has: operationally independent subsystems; managerially independent components and subsystems; evolutionary development; emergent behavior; and geographic distribution [1].

The Northrop report [1] states that "the sheer scale of ultra-large-scale systems will change everything. Ultra-large scale systems will necessarily be decentralized in a variety of ways, developed and used by a wide variety of stakeholders with conflicting needs, evolving continuously, and constructed from heterogeneous parts. People will not just be users of a ultra-large-scale system; they will be elements of the system. The realities of software and hardware failures will be fundamentally integrated into the design and operation of ultra-large-scale systems. The acquisition of a ultra-large-scale system will be simultaneous with its operation and will require new methods for control. In ultra-large-scale systems, these characteristics will dominate. Consequently, ultra-large-scale systems will place unprecedented demands on software acquisition, production, deployment, management, documentation, usage, and evolution practices".

As it is known, industrial computer management systems is an important factor determining the success of the industry optimal functioning. The decision made by persons must manage a huge amount of information effectively in order to maximize its usefulness. Managing large amounts of information generated by computers, application of information from sensors and the data about the state of business processes of different industrial enterprises nowadays can be based on Big Data technologies. However, the main findings of the analysis undertaken in this area of research should be pointed as the lack of engineering management methods in regarding aviation industry. It disables considering aviation as a complex system. So, currently, it is impossible to manage and develop aviation industry integrally.

Modern aviation industry is a complex ultra-large-scale system that comprises designing, testing, production, operation, maintains services, computer services, and information technologies including hardware and software, etc. Thus, it seems reasonable to provide information logistical support throughout the life cycle of single processes in this area. On the other hand, it allows building stochastic models to manage aviation industry in real-time mode under uncertainties and risks conditions [2, 3]. It is possible to provide a concept for using Big Data [4–9].

The concept of Big Data combines techniques and technologies that extract knowledge from huge data streams in real time. An active interest from researchers and practitioners in the field of aviation in Big Data technology is due to the implementation of high-speed data channels for dynamic objects, as well as the implementation of NoSQL technologies and Cloud Computing. Generators of data streams in aviation are various sensors and devices of the aircraft. Considerable data stream is processed in solving the task of aircrafts group control. It is noted that a control system with gas turbine engines can generate an array of data of up to 20 TB per hour while flying.

As it is known, the idea of standardizing the exchange of information in the open system is based on the following principles:

• the principle of entropy reduction, i.e., bringing uniformity and standardization of a single form of objects and information about those objects;

• the principle of invariance, i.e., identification of requirements, which are necessary taking into account changes in the properties of standardized plants;

• the principle of preserving the viability of standardized requirements establishing reasonable limits of standardization.

An important role in the framework of standardization of design solutions based on the idea of open systems is played by implementation of CALS technologies. It allows to form the basis for creation and usage of the unified information environment during the design, production, testing, and operation. These technologies are introduced at Sukhoi Aviation Holding Company, PJSC "Tupolev", Boeing Company, etc.

At the same time, Big Data application for design, testing and operation of control systems and aircraft power plants is only at the initial stage today. The reason is de to both the complexity of the design of control systems of power plants and aircraft and the heterogeneity of circulating information flows, as well as the lack of modern standards, regulatory processes of the development of advanced information management systems of power plants and aircraft.

1. Markov decision process system building for ultra-large-scale system management

As a decision making approach in the case of large-scale systems, Markov decision processes (MDP) can be applied. MDP is a mathematical model applied for decision making in situations where outcomes are partly random and partly under the control of a decision maker.

Definition 1. MDP means a stochastic system characterized by a 5-tuple $\langle \mathbf{S}, \mathbf{A}, A, p, g \rangle$ where the components are as follows: $\mathbf{S}$ is a finite set of discrete states and $\mathbf{A}$ is a finite set of control actions. Mapping $A : \mathbf{S} \to p(\mathbf{A})$ is the availability function that renders a set of actions available to each state where p denotes the power set. The transition function is given by $p : \mathbf{S} \times \mathbf{A} \to \Delta(\mathbf{S})$, where $\Delta(\mathbf{S})$ is the set of all probability distributions over $(\mathbf{S})$. Let p(y | x, a) denote the probability of arrival at state y after executing action $a \in A(x)$ in state x. The immediate-cost function is defined by $g : \mathbf{S} \times \mathbf{A} \to \mathbf{R}$, where g(x, a) is the cost of taking action a in state x.

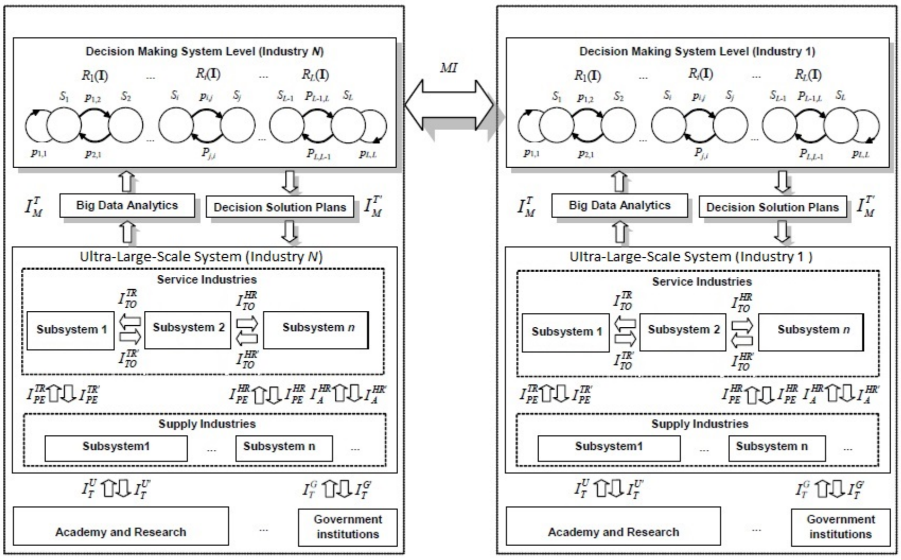

Fig. 1. Large Scale Model of Aviation Industry

Definition 2. A control policy determines the action to take in each state. A deterministic policy, $\pi : \mathbf{S} \to \mathbf{A}$, is simply a function from states to control actions. A randomized policy, $\pi : \mathbf{S} \to \Delta(\mathbf{A})$, is a function from states to probability distributions over actions. We denote the probability of executing action a in state x by $\pi(x)(a)$ or, for short, by $\pi(x, a)$.

In terms of logistics approach, models of large-scale system services can be represented as a Markov chain, where $p_{i,j}$ is the probability of transition from state S_i to state S_j relevant to processes of production to operation, $p_{j,i}$ is the probability of transition from state S_j to state S_i.

Let us consider a more detailed representation of the scheme (Fig. 1). Traditionally, the production and operation stages in the life cycle of large-scale system products have a week information logistics interconnection I_i. The responsibility for production to operation stages transition, as a rule, is problem of the large-scale system (transitions $p_{i,j}$) and $p_{j,i}$. It is possible to decompose each transition to some more detailed stages. For instance, testing of new production units can be also considered as separate processes. Thus, it is possible to describe the large-scale system as the following. Let $S = \{s_1, \dots, s_L\}$ $(j = 1, \dots, L)$ be the state space which is the same for every step of different processes presenting the life cycle of large-scale system objects. One-dimensional array of probability distribution on the set of indices i for the n-th step (transition) is denoted as $\pi_1 = (\pi_n^{(1)}, \dots, \pi_n^{(L)})$. Herewith, $\pi_n^{(j)}$ is the probability of the event when a large-scale system object is in state j at n-th step. Let us determine state probability π_n distribution at every single step.

Let us define T as an event corresponding to the situation when a large-scale system object is in a state j at $(n+1)$-th step. In accordance with the greed notation, the event probability will be $P(T) = \pi_{n+1}^{(j)}$. Let us define S_i as an event when an aviation object is in a state i at n-th step, therefore $P(S_i) = \pi_n^{(i)}$. Thus, p_{ij} is the probability that the aviation object is in state j at $(n+1)$-th step if it was in a state i at n-th step, i.e.,

$p_{ij} = p(T|S_i)$. We can find the event probability using the total probability equation:

$$\pi_{n+1}^{(j)} = \sum_{i=1}^{L} \pi_n^i p_{ij}, \quad j = 1, \ldots, L. \tag{1}$$

Equation (1) allows to sequentially (step by step) determine the change in the probability distribution of states of the aviation object while providing aviation industry service. The initial distribution of probabilities should be a priori given. The probabilities are directly dependent on the risk of failure $R_i(I)$ at each step of the product life cycle (Fig. 1).

For example, in aviation industry, service industries and supply industries may be allocated. Service industries interact with a product directly on the base of different organizational technical system. Supply industries are related with product indirectly and include power engineering and so on. All these subsystems are connected with material and information flows. In this case, we need to minimize the risk of service failure $R_i(I)$:

$$\sum_{i=1}^{L} R(I) \to \min. \tag{2}$$

Interrupting the information flows or information lack or invalidation lead to increasing the risks $R_i(I)$. Therefore, in our opinion, a Markov decision process system should be provided for aviation industry services, as well as system information support through all the aviation industry life cycle elements.

2. Big Data processing for Markov decision process system building

As it can be seen, this approach requires a huge information content to be processed at the national and worldwide levels. Herewith, the following urgent tasks may be allocated for processing the flows:

- data mining as a tool of information flow analysis;
- building a large amount data warehouse;
- forecasting and optimizing processes in aviation industry;
- high-technology tools and equipments to provide an information support of aviation industry services.

A multi-agent approach is suggested to build the Markov decision process system model of the industrial service life cycle [4, 7]. To date, it has been applied effectively for business process optimization, as well as complex social and technical systems management and control. Let us present the example of a strategy for agent in MDP of airport management. Airports are large-scale systems that operate in unique and evolving physical, financial, and regulatory environments. For estimation of the airport efficiency, statistics of total movement, air transport movements, terminal passengers, and freight are applied. The UK Civil Aviation Authority collects statistics from more than 60 UK Airports for the period of 1981–2015 years. So, we can say that some kind of Big Data technologies has been already applied. Let us find the basic strategy for agent in MDP on the basis of these data. Data of total movement, air transport movements, terminal passengers, freight for 1981-2015 years are presented in Fig. 2.

For obtaining the main action to reduce costs of airline company, the theoretic-information approach based on the notion of transfer entropy was applied. The transfer entropy extends the concept of mutual information to provide a direction sensitive measure of information flow between two time series. Formally, the transfer entropy

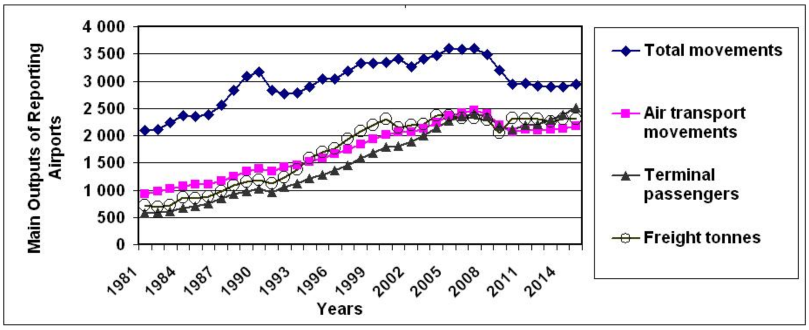

Fig. 2. UK Airports Statistics

Table 1. Transfer Entropy from Y to X

Y to X	Total Movements	Air Transport Movements	Terminal Passengers	Freight Tonnes
Total Movements	0.0	0.2	0.3	0.0
Air Transport Movements	0.2	0.0	0.0	0.2
Terminal Passengers	0.3	0.0	0.0	0.7
Freight Tonnes	0.0	0.0	0.0	0.0

from time series Y to X is given by

$$T_{Y\to X} = \sum p(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log \frac{p(x_{n+1}|x_n^{(k)}, y_n^{(l)})}{p(x_{n+1}|x_n^{(k)})}, \tag{3}$$

where x_{n+1} is the value of X at time $n+1$ and $x_n^{(k)}$ is the $k(l)$ lagged values of $X(Y)$ at time n.

The definition of transfer entropy assumes X as Markov process. The transfer entropy measures the additional amount of information Y contains about X over the information contained in the Markov embedding. The results of calculation of transfer entropy for statistics data sets are presented in Table 1.

The results of the analysis of transfer entropy calculations determine the agent action in MDP is control amount of freight or related to fuel consumption. That is why aircrafts of low-cost airlines companies are operating with a minimum set of optional equipment, further reducing costs of acquisition and maintenance, as well as keeping the weight of the aircraft lower and, thus, saving fuel.

Conclusions

The scope of Big Data technologies is very wide. In particular, it covers a complex of organizational and technical systems, such as aircraft systems, aircraft building enterprises, etc. Data on the functioning of these systems are characterized by high volume, heterogeneous structure, and a high frequency of updating.

The tendency of ultra-large-scale systems development can be worldwide integration of their elements on the basis of the common information space. Under these circumstances, ICT tasks become even more important. These tasks are data mining, data warehouse building, creating high technology tools and equipments, and forecasting and optimizing processes in large-scale systems on the basis of Big Data analytics.

The tasks may be solved with the help of interconnected information and communication systems based on innovative technologies.

Big Data approach is suggested to build the MDP model of large scale system. The MDP model implementation will improve the information structure of the large scale system, which, in turn, will reduce costs and increase revenues in the area.

Acknowledgements. The study was supported in part by the Russian Foundation for Basic Research (project no. 15-08-04877-a).

References

1. Feiler P., Gabriel R.P., Goodenough J., Linger R., Longstaff T., Kazman R., Klein M., Northrop L., Schmidt D., Sullivan K., Wallnau K. *Ultra-Large-Scale Systems: The Software Challenge of the Future.* Pittsburgh, Carnegie Mellon Univ., Software Eng. Inst., 2006. xi+134 p.

2. Hu T.-C., Cabrera M.O., Volodin A. Almost sure lim sup behavior of dependent bootstrap means. *Stochastic Anal. Appl.*, 2006, vol. 24, no. 5, pp. 939–952. doi: 10.1080/07362990600869969.

3. Anatolyev S., Vasnev A. Markov chain approximation in bootstrapping autoregressions. *Econ. Bull.*, 2002, vol. 3, no. 19, pp. 1–8.

4. Valeev S.S. Big data technologies in aviation. *Proc. 2nd Int. Conf. Information Technologies for Intelligent Decision Making Support*, 2014, pp. 150–152.

5. Valeev S., Kondratyeva N. Large scale system management based on Markov decision process and big data concept. *Proc. 2016 IEEE 10th Int. Conf. on Application of Information and Communication Technologies (AICT)*, 2016, pp. 6–9. doi: 10.1109/ICAICT.2016.7991829.

6. Valeev S., Kondratyeva N. Distributed information and control system for emergencies in critical infrastructures. *Proc. 2016 IEEE 10th Int. Conf. on Application of Information and Communication Technologies (AICT)*, 2016, pp. 58–61. doi: 10.1109/ICAICT.2016.7991653

7. Kondratyeva N., Valeev S. Fatigue test optimization for complex technical system on the basis of lifecycle modeling and big data concept. *Proc. 2016 IEEE 10th Int. Conf. on Application of Information and Communication Technologies (AICT)*, 2016, pp. 75–78. doi: 10.1109/ICAICT.2016.7991656.

8. Kondratyeva N., Valeev S. Technical safety system with self-organizing sensor system and fuzzy decision support system. *Proc. 2015 IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE)*, 2015. pp. 1–4. doi: 10.1109/FUZZ-IEEE.2015.7337962.

9. Valeev S.S., Taimurzin M.I., Kondratyeva N.V. An adaptive data acquisition system in technical safety systems. *Autom. Remote Control*, 2013, vol. 74, no. 12, pp. 2137–2142. doi: 10.1134/S0005117913120151.

Recieved
December 19, 2017

Valeev Sagit Sabitovich, Doctor of Technical Sciences, Professor, Head of the Department of Informatics

Ufa State Aviation Technical University
ul. Karla Marksa, 12, Ufa, 450008 Russia
E-mail: *vss2000@mail.ru*

Kondratyeva Natalia Vladimirovna, Candidate of Technical Sciences, Associate Professor of the Department of Informatics

Ufa State Aviation Technical University
ul. Karla Marksa, 12, Ufa, 450008 Russia
E-mail: *knv24@mail.ru*

УДК 519.257

Стохастическая модель авиационной индустрии на основе концепции больших данных

С.С. Валеев, Н.В. Кондратьева

Уфимский государственный авиационный технический университет, г. Уфа, 450008, Россия

Аннотация

Современная авиационная индустрия представляет собой сложную крупномасштабную систему, которая включает в себя проектирование, производство, эксплуатацию, техническое обслуживание авиационной техники и т. д. Таким образом, представляется актуальным осуществлять информационную логистическую поддержку на протяжении всего жизненного цикла индустрии. С другой стороны, это позволяет строить стохастические модели для для управления авиационной индустрии в режиме реального времени. Такой подход становится возможным за счет использования концепции больших данных. Предлагается многоагентная технология построения марковской модели принятия решений для оптимизации жизненного цикла авиакомпании. Для выявления основного управляющего воздействия по снижению издержек авиакомпании был применен теоретико-информационный подход. Результаты анализа показали, что основным воздействием агента в процессе принятия решений является контроль количества груза, связанного с расходом топлива.

Ключевые слова: авиационная индустрия, эффективность, стохастическая модель, марковский процесс принятия решений, большие данные

Поступила в редакцию
19.12.17

Валеев Сагит Сабитович, доктор технических наук, профессор, заведующий кафедрой информатики

Уфимский государственный авиационный технический университет
ул. К. Маркса, д. 12, г. Уфа, 450008, Россия
E-mail: *vss2000@mail.ru*

Кондратьева Наталья Владимировна, кандидат технических наук, доцент кафедры информатики

Уфимский государственный авиационный технический университет
ул. К. Маркса, д. 12, г. Уфа, 450008, Россия
E-mail: *knv24@mail.ru*

For citation: Valeev S.S., Kondratyeva N.V. Aviation industry stochastic model based on big data concept. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2018, vol. 160, no. 2, pp. 392–398.

Для цитирования: *Valeev S.S., Kondratyeva N.V.* Aviation industry stochastic model based on big data concept // Учен. зап. Казан. ун-та. Сер. Физ.-матем. науки. – 2018. – Т. 160, кн. 2. – С. 392–398.