Resource management of a distributed stream data processing system in safety systems of infrastructure objects

S.S. Valeev[1], N.V. Kondratyeva[1], A.S. Kovtunenko[1], M.A. Timirov[2], R.R. Karimov[1]

[1]Ufa State Aviation Technical University, K. Marx str. 12, Ufa, Russia, 450008
[2]Bashkir Production Association "Progress", Kommunisticheskaya 23 Ufa, Russia, 450076

Abstract. The solution of the problem of resource management in distributed computing systems of processing stream data in safety systems of distributed objects is considered. The tasks of streaming data processing in a multi-level multi-agent evacuation system in an infrastructure object are considered. The features of the mathematical model of a distributed stream data processing system are discussed.

1. Introduction

Modern infrastructure facilities include various building complexes that have a large area and the possible presence of a large crowd of people [1-3]. The purpose of risk management in the security system of an infrastructure facility is reduced to minimizing collective risk in the event of a critical situation in the cyber-physical system or in conditions of infrastructure degradation. Different systems for collecting and processing information are used to analyze the state of these objects: sensors of the state of physical parameters, video monitoring systems and telecommunication systems. It should be noted that part of the information is transmitted and processed as stream data [4, 5]. In the event of a critical situation, the possible loss of current information about the state of the information field due to factors affecting the elements of the computing system that affect the physical parameters of the internal space of the infrastructure facility.

The task of reallocating the resources of a computing system in a technical security system in which information about the state of an infrastructure object is collected from various sensors in real time is considered. This system can be attributed to the class of self-learning systems of stochastic distributed objects having a variable structure. The observed information space is the internal space of the infrastructure facility.

The state of the information field may change depending on the internal and external factors affecting the physical flows, which in turn affects the distribution of risk levels.

The solution to the problem of risk management is based on the elements of self-learning of a multi-level management system at the planning and coordination levels [6, 7]. The formation of scenarios for the evolution of a critical situation and the formation of possible ways to minimize risks is provided by a self-learning multi-agent data collection system.

A multi-agent system can be implemented in different ways: we can use a multi-agent software platform, such as JADE [5], or implement a software agent based on different programming languages: Python, Java, etc. Databases can be implemented on different concepts, for example, IRIS [8]. In any case, it is necessary to solve the tasks of reallocation of real-time computing power to solve problems of collecting and processing streaming data in security systems of a distributed infrastructure object.

2. Multiagent system of infrastructure object safety

A set of agents of the coordination level use the information received by the system planner to solve the task of minimizing risks through interaction mechanisms in multi-agent systems. At the level of coordination, among other things, the problem of ensuring the integration of information on the state of the control object using multi-agent sensor systems must be solved, which makes it possible to minimize errors of the first kind and of the second kind due to the effect of self-organization of sensors of various nature. At this level, the problem of developing control actions on the ventilation system, fire protection system, information support for evacuation must be also solved.
The architecture of multiagent system of cyber physical system safety is presented in Fig. 1. The structure of this system is based on hierarchical architecture.

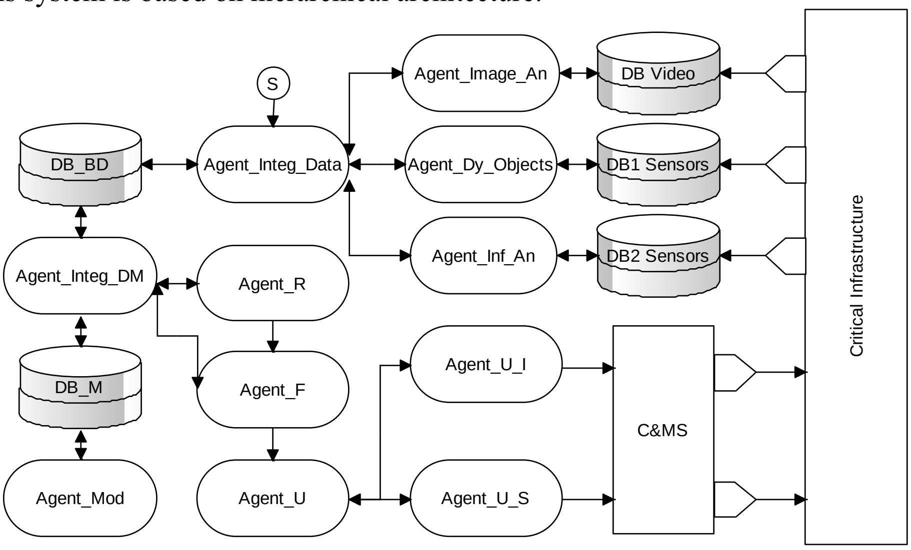

Figure 1. Multiagent system of infrastructure object safety.

In the sense of Big Data concept here we have different flows of information and different formats of data. We need to gather information from different types of sensors and store these data in Databases: DB_Video – Data Base of Video Sensors Data, DB_Sensor1 – Data Base of Infrastructure Sensors, DB_Sensor2 – Data Base of Dynamic Objects, DB_BD – Data Base of Mix Content, DB_M – Data Base of Simulation Results Data.
Multigent system of safety of cyber-physical system includes: Agent_Image_An – agent of Video Stream Flow Analyzer, Agent_Dy_Objects – agent of Dynamic Objects trajectory Identification, Agent_Inf_An – agent of Infrastructure State Identification, Agent_Integ_Date – agent of Integration of different flows from sensors, Agent_Mod – Agent of Simulation of the different Scenarios of Critical Situations, Agent_Integ_Dm – Agent of Analysis the state of Cyber-Physical System, Agent_R – agent of risk field identification, Agent_F – agent of integrated risk identification, Agent_U – agent of control and management of minimizing risk, Agent_U_I – agent of control signal generation for actuators, Agent_U_S – agent of management decisions realization, C&MC – implementation of control and management system.
So, here we have massive application of different streaming data flows and in case of degradation of information system we need to solve the problem resource allocation problem.

3. Formulation of the optimal resource allocation problem

The model of the processing and storing of streaming data (PSSD) based on the computer network includes resource model of the computer network *N*, model of data processing data processing system (DPS) and the relation *D* describing system deployment in network.

$$IS = \langle N, DPS, D \rangle \,. \tag{1}$$

Here:

- N– the computational network,
- DPS – the PSSD model,
- D – the deployment relation that shows $\forall p \in P$ on what node it physically deployed

$$D = \{(w, p) : w \text{ executes } p;\ w \in W,\ p \in P\}. \qquad (2)$$

Let $B : W \times R \to \mathbf{R}^+$ – function that shows $\forall w \in W$ amount of the involved resource. Values of function can be calculated from resource capacities of data processing model and network model as follows.

$$B_{cpu}(w) = \sum_{\forall p \in D(w)} \frac{V(p)}{T(p)}. \qquad (3)$$

$$B_{mem}(w) = \sum_{\forall a \in (D \circ C)(w)} M(a). \qquad (4)$$

$$B_{net}^{in}(w) = \sum_{\forall p \in D(w)} \frac{\sum_{\forall a \in I(p) \cap A \setminus (D \circ C)(w)} M(a)}{T(p)}. \qquad (5)$$

$$B_{net}^{out}(w) = \sum_{\forall p \in D(w)} \frac{\sum_{\forall a \in I(p) \cap (D \circ C)(w)} M(a)}{T(p)}. \qquad (6)$$

The introduced notation makes it possible to formulate the problem of optimal allocation of computational resources in PSSD as the problem of combinatorial optimization. It is necessary to find such deployment relation D^*, that gives maximum to predefined efficiency function F.

$$D^* = \arg \max_{D \subset W \times P} F(IS\langle D \rangle). \qquad (7)$$

The following constraints apply

$$B_r(w) \le S_r(w). \qquad (8)$$

Let $rem(w)$ – the unitless function, that shows average free resources for every node as

$$rem(w) = \frac{1}{|R|} \sum_{\forall r \in R} \frac{S_r(w) - B_r(w)}{S_r(w)}. \qquad (9)$$

and the average total of free resources as

$$\langle rem \rangle_W = \frac{1}{|W|} \sum_{\forall w \in W} rem(w). \qquad (10)$$

The mathematical problem of HCNUL resource allocation can be formulated as follows. Subject to constraints (14) the optimal deployment at every point of time is

$$D(t) = \arg \min_{D \subset W \times P} \left(\max_{w \in \delta(D)} \left(\sigma_\tau^2 [rem(w, \tau)] \right) \right). \qquad (11)$$

Here:

- $\sigma_\tau^2[\xi(\tau)]$ denotes the dispersion of stochastic process ξ on interval $\tau \in (0, t)$,
- $\delta(D)$ denotes the domain of relation D – the set of nodes where system is deployed

4. Conclusion

The task of risk management in cyber-physical systems on the basis of multilevel management system is considered. The principals of designing and architecture of multilevel self-learning multiagent system for risk management are discussed. Self-learning ability of propose multiagent system is based on application of Big Data algorithms an machine learning algorithms.

5. References

[1] Curry, D.M. A system-of-systems approach to improving intelligent predictions and decisions in a time-series environment / D.M. Curry, W.W. Beaver, C.H. Dagli // SoSE, 2018. – P. 98-105.

[2] Gosavi, A. Predicting response of risk-seeking systems during project negotiations in a system of systems / A. Gosavi, S. Agarwal, C.H. Dagli // IEEE Systems Journal. – 2017. – Vol. 11(3). – P. 1557-1566.

[3] Agarwal, S. Multi-agent based architecture for acknowledged system of systems / S. Agarwal, L.E. Pape, N.H. Kilicay-Ergin, C.H. Dagli // CSER, 2014. – P. 1-10.

[4] Kondratyeva, N. Distributed information and control system for emergencies in critical infrastructures / N. Kondratyeva, S. Valeev // Proceedings of 10th conference on Application of Information and Communication Technologies (AICT), 2016. – P. 1-4.

[5] Kovtunenko, A. Distributed streaming data processing in iot systems using multi-agent software architecture / A. Kovtunenko, A. Bilyalov, S. Valeev // 18th International Conference on Next Generation Wired/Wireless Networking (NEW2AN), and 11th Conference on Internet of Things and Smart Spaces (ruSMART) conference proceedings, 2018. – P. 572-583.

[6] Valeev, S. Technical safety system with self-organizing sensor system and fuzzy decision support system / S. Valeev, N. Kondratyeva // Proccedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2015. – P. 1-6.

[7] Valeev, S. Infrastructure object management in the conditions of critical situations based on situational models / S. Valeev, N. Kondratyeva // Proceedings of 10th International Conference Management of Large-Scale System Development, 2017. – P. 1-4.