Approach to the Analysis and Processing of Data from IT-Services Monitoring System

Maksim A. Bolshakov
Saint Petersburg Information and Computing Centre JSC Russian Railways
Saint Petersburg, Russia
bolshakovm@yandex.ru

Sergei V. Pugachev
Emperor Alexander I St. Petersburg State Transport University
Saint Petersburg, Russia
nki-pugachev@yandex.ru

Igor A. Molodkin
Emperor Alexander I St. Petersburg State Transport University
Saint Petersburg, Russia
imolodkin@gmail.com

Nikolay N. Teslya
Laboratory of computer aided integrated systems, St.Petersburg Institute for Informatics and Automation of the RAS
Saint Petersburg, Russia
teslya@iias.spb.su

Abstract

Various instrumental IT infrastructure monitoring systems are considered and compared: Zabbix, Nagios, ManageEngine OpManager, Hewlett Packard Operations Manager, Naumen Network Manager and IBM Tivoli. The functions to be performed by the IT infrastructure monitoring and management system in its target state are specified. The current state of the IT infrastructure monitoring system in JSC Russian Railways is described. A mathematical formulation of the problem of determining control values of metrics and an example of developing a neural network to determine control values of metrics and recommendations for its improvement are given.

Introduction

The scope of the coverage of technological processes by automation systems in JSC Russian Railways is constantly growing. As of now, automation systems have accumulated and use in their work enormous quantities of data. New possibilities of data processing tools enable more comprehensive use of data being accumulated to solve a large variety of current problems [Rrw17]. Thus, Big Data technology is receiving ever-growing acceptance and demand for usage. It should be noted that a special feature of Big Data as applied to JSC Russian Railways consists of information coming from both external and internal sources. Provided data volumes fully conform to the currently accepted "7 Vs" of Big Data: Volume, Velocity, Variety, Veracity, Variability, Visualization, Value [Cuk14].

One of the most important ways of increasing the company operation efficiency is reducing failure rate and operating costs. Methods for solving these issues are generally similar both for railway infrastructure and IT infrastructure. Thus, one of the approaches to railway infrastructure condition monitoring implementation is using IRV concept (Instrumented Revenue Vehicles), which consists of the instrumentation of active cars with means for infrastructure condition monitoring. In locomotive facilities, information about the operation of traction equipment aggregates and parts is read directly by sensors located on the locomotive [Gol17]. In IT infrastructure, a similar approach is applied: to take readings about the operation of all assembly units in a number of different ways where the elements of computing resources employed in IT service creation are taken for assembly units.

When selecting a specific vendor of a monitoring system, you should decide according to the conformity between the functionality of the solution in question, problems and particularly the scale of problems to be solved on the IT landscape existing in the organization.

1 IT Infrastructure Monitoring Systems

Currently, the following instrumental systems are the most frequently used to solve the problem of full coverage by an IT infrastructure monitoring system.

- Zabbix is an open-source system featuring sufficiently high efficiency and readiness for scalability up to corporate-level data. Because of its open-source technology and wide applicability, there is a sufficiently active developers community with whose help even a novice administrator of monitoring system can quickly become familiar with its installation and maintenance.

In: B. V. Sokolov, A. D. Khomonenko, A. A. Bliudov (eds.): Selected Papers of the Workshop Computer Science and Engineering in the framework of the 5 th International Scientific-Methodical Conference "Problems of Mathematical and Natural-Scientific Training in Engineering Education", St.-Petersburg, Russia, 8–9 November, 2018, published at http://ceur-ws.org

However, the question of implementation itself (setting up and use of data collection metrics) can present serious difficulties for an administrator at the first stages of use as there are no ready-made agents of monitoring in this tool. Additionally, disadvantages include poor design of visualization tools.

- Nagios is an open-source tool with similar characteristics (both positive and negative) as Zabbix, excluding operability of using new settings of data-collection. While with Zabbix, data collection algorithms can be reorganized in online mode, in Nagios, the system should be rebooted after changes are made [Sha18].
- ManageEngine OpManager is a tool which allows you to work, among others, in terms of automated response to abnormal situations. It has a convenient and understandable interface but is limited by scalability in terms of data and fully justifies its use at small and medium-sized enterprises. On a higher level, it is significantly behind its competitors in terms of performance.
- Hewlett Packard Operations Manager is an example of a complete centralized monitoring system with high characteristics both in terms of user interface and data handling quality for various volumes.
- Naumen Network Manager is a Russian product with all the necessary parameters for a complete solution for the orientation of IT infrastructure monitoring processes. Under conditions where many state companies aim for import substitution of IT tools, such products should positively meet all client needs. At the present time, this tool has good characteristics both in terms of convenience of deployment, implementation and support, and in terms of collection and providing summary data after necessary processing.
- IBM Tivoli is a tool for centralized monitoring system creation and features simple installation, but initial configuration process requires the presence of highly qualified specialists. Generally, this is connected with its application on the corporate data level, when configuration and setup requires a large amount of work. In operation, the system is intuitive, with a vast set of functions ready-made and supported by the developer, it is possible to form monitoring agents at your own discretion using Agent Builder. The line of monitoring products from this vendor contains tools for all monitoring functions – from data collection by means of various agents and collected data processing, to automated response to failures and presenting necessary information via data displays for all user levels.

However, selection of one vendor of monitoring tools does not result in use of drastically different approaches in its implementation. A working IT infrastructure monitoring and management system should perform the following functions in its target state:

- Function of detection of IT components and determination of dependencies between them –including automatic monitoring agent installation and subsequent information collection, among others, in order to search for relations between these metrics characterizing the work of IT components.
- Function of performance management – data collection in order to forecast utilization of resources and form proposals on load redistribution/balancing.
- Function of correlation and event management – providing message reception from all sources of data monitoring and its subsequent analysis in order to enrich data about an event, identify data similarity and determine root cause.
- Function of automated response to events in order to restore operability of a component or entire service.

2 Current State of IT Infrastructure Monitoring System in JSC Russian Railways

Presently, in JSC Russian Railways, an umbrella solution is implemented as a centralized monitoring system, where data are transmitted by means of various “probes” at the lower level of information collection (monitoring agents from different vendors) and then processed centrally by IBM Tivoli. Because of the wide variety of ready-for-service monitoring agents from IBM, the portion of Tivoli “probes” at the lower level of monitoring makes up about 90% of the total quantity of information collection tools throughout JSC Russian Railways the next largest tools in terms of coverage are Zabbix tools. It is the umbrella-type structure of the monitoring system that allows coverage of absolutely all elements of IT infrastructure and processing of this data according to a single logic avoiding various local and not interrelated monitoring systems, for instance for each type of equipment or geographical location.

Thus, the main factors affecting the decision on vendor selection are the readiness to work with existing volumes of client’s IT infrastructure and the cost of this decision. It is the transition from simple IT infrastructure monitoring to IT infrastructure management and IT services monitoring that requires the highest expenditures, because for virtually all vendors the elements of the functions of the product line implementing data processing and forecasting are the most expensive. That is why (among other causes) companies often stop developing their monitoring system at the level of equipment information collection and generation of events relating to failures of this equipment, without considering the relations between these elements.

Where as in the target state, the resource and service model is determined as a result of taking inventory of all resources employed in IT service provisioning and subsequent determination of the influence of elements on each other and on key parameters of the performance of the whole service. This concept is the key concept for the evaluation of possible consequences of failures/faults in the operation of individual elements of infrastructure affecting final IT service provisioning feasibility. In this case, the condition of each element is characterized by metrics taken automatically [Oht06].

Obtained values of metrics are compared with reference ones, and when they exceed acceptable limits

the monitoring system informs appropriate technical support personnel about the abnormal state of the service. This creates the problem of correct determination of these reference or marginal values of normal condition for each metric and their mutual interrelations.

Initially, these values are determined with the help of experts, however, the risk of subjectivity in this case can not be totally excluded. That is why the problem of maintaining marginal values and their mutual interrelations in their actual condition for large number of metrics is rather complex, and additionally, when a group of experts is permanently engaged this task is cost intensive [Spr16, Aua07, Mar09].

Within the framework of the current level of IT monitoring system development, the data volume in MCC JSC Russian Railways is 11 terabytes, and this value will only increase as solving the problem of automated prioritizing of metrics taken from IT infrastructure and creating resource and service models bigger than current data storage horizon is required. At the present time, when aggregated data are stored, for most metrics the storage horizon resides in the interval of 1-3 months. Raw data is actually stored for a substantially shorter period.

In this case, the specialists responsible for IT infrastructure monitoring and support have at their disposal 1318 unique metrics, the combinations of which are imposed on selected IT services. Special attention should be paid to data heterogeneity for specified metrics. For example, for each IT service the following metrics should be analyzed: numerical values of processor utilization expressed in percent, remaining free space on the virtual server in megabytes, response time of network equipment via ICMP in milliseconds, and text values of responses for Blade chassis operation status.

Specified heterogeneity together with a large volume of data is the cause of impossibility of problem solving based just on the knowledge, skills and competence of experts involved in the support of IT infrastructure on which MCC JSC Russian Railways IT services are deployed [Ort91].

3 Mathematical Formulation of Problem

The mathematical formulation of reference problem metrics values can be represented in the following form: for each metric M_i, ($i \in [1, m]$) it is necessary to determine reference values K_i, which makes it possible to unambiguously split the whole array metric values M_i into normal and abnormal subsets (the boundary condition can be derived).

To do this it necessary to find vector $\overline{K} = (K_1, ..., K_i, ..., K_m)$ under condition:

$$\overrightarrow{K} \times M \to \overrightarrow{F},$$

where $M = \left(M_{ij}\right)$ – is the matrix of i-th metric values ($i \in [1, m]$) in j-th period of time ($j \in [1, n]$);

$\overrightarrow{F} = (F_1, ..., F_j, ..., F_n)$ – is resulting condition of service, F_j is condition of service for each j-th point of time.

$$F_j = \begin{cases} 0; \\ 1, \end{cases}$$

where 0 is normal condition, and 1 is a fault.

The problem under consideration can be solved using classic methods for the system of solving simultaneous linear equations:

$$\begin{cases} K_1 \times M_{11} + \cdots + K_i \times M_{1i} + \cdots + K_m \times M_{1m} = F_1; \\ \ldots \\ K_1 \times M_{j1} + \cdots + K_i \times M_{ji} + \cdots + K_m \times M_{jm} = F_j; \\ \ldots \\ K_1 \times M_{n1} + \cdots + K_i \times M_{ni} + \cdots + K_m \times M_{nm} = F_n. \end{cases}$$

At the same time, this problem can be solved using apparatus of artificial neural networks [Aya07]. These networks have various structures and each one has advantages and disadvantages for solving different kinds of problems. Assuming that at the j-th point of time the value of metrics is given for all previous measurements and there is a problem of the resulting condition of the IT service (failure/degradation) change probability estimation then it is possible to speak about a forecasting problem, i.e. about a special case of a regression problem. For such problems, the use of forward propagation neural networks (perceptrons) is the most justified [Naz03].

In this case, the number of layers for forward propagation networks and the number of neurons in each layer are the values upon which, on the one hand, the speed and, on the other hand, the quality of proposed neural network learning depends. The degree of network architecture complication and the increase in the number of neurons, in turn, depend on existing computational capability limitations.

Under the conditions of MCC JSC Russian Railways operations, the speed of these calculations is not less important than K_i, coefficient calculations quality because of this problem it is absolutely necessary to organize periodic model recalculation in order to maintain its actual condition.

To achieve this goal, and to mitigate possible negative influence on IT services provided by MCC JSC Russian Railways to consumers, the process of periodic learning of the neural network should be executed at the time of minimal utilization of the computing system, for instance, in daily relearning mode at night (according to the Moscow time zone).

The learning process itself should be built in the format of learning by instruction – learning by means of the presentation of multiple available examples of input data M and reference solutions $\overrightarrow{F}$. As stated above, this problem, in substance, is a forecasting problem which is most often solved using a back propagation learning algorithm. Its main disadvantage consists in the learning process being too long which makes it essentially unusable for the given problem. At the present time, there are enough faster algorithms such as: conjugate gradient method, RProp method of Levenberg–

Marquardt, etc.

The optimal choice for solving the problem is the RProp (Resilient Propagation) method known as the method of resilient error propagation. It outperforms the standard back propagation method in terms of learning time length, particularly with regard to the heterogeneity of available data [Nov16].

At the beginning of learning, all weight factors K_i are set in a random manner (as small values close to zero), and further when examples are input the network error is minimized.

In the learning process using the RProp algorithm, partial derivative signs are used to trim weight factors. For each K_i weight factor in the chain for k-th neuron, the separate modifier value entered Δ_{ik}, is used to calculate the size of correction for each relevant weight factor.

To determine the correction value the following convention is used in each step:

$$\Delta^{j}{}_{ik} = \begin{cases} \eta^{+}\Delta^{(j-1)}{}_{ik}, если \dfrac{\partial E^{j-1}\partial E^{j}}{\partial K_{ik}\partial K_{ik}} > 0; \\ \eta^{-}\Delta^{(j-1)}{}_{ik}, если \dfrac{\partial E^{j-1}\partial E^{j}}{\partial K_{ik}\partial K_{ik}} < 0, \end{cases}$$

Where $0 < \eta^{-} < 1 < \eta^{+}$.

Specific values of modifiers can be different but most often the values proposed in [Rdm93] and tested on multiple examples are used:

$$\eta^{-} = 0.5; \quad \eta^{+} = 1.2;$$

$\dfrac{\partial E^{j}}{\partial K_{ik}}$ – partial derivative of activation function by weight factor at j-th point of time.

If in the current step the partial derivative with respect to corresponding weight K_{ik} has changed its sign, then it follows that the last change was too large and the algorithm has exceeded the local minimum. Consequently, the amount of change should be decreased and the previous weight factor value should be returned, in other words, the "rollback" should be performed. When the derivative retains its sign, the modifier value should be additionally increased to accelerate convergence.

After the values of modifiers have been updated, the change of factors themselves is made according to the convention:

$$\Delta K_{ik} = \begin{cases} -\Delta^{j}{}_{ik}, если \dfrac{\partial E^{j}}{\partial K_{ik}} > 0; \\ \Delta^{j}{}_{ik}, если \dfrac{\partial E^{j}}{\partial K_{ik}} < 0; \\ 0, иначе. \end{cases}$$

$$K^{j+1}{}_{ik} = K^{j}{}_{ik} + \Delta K^{j}{}_{ik}.$$

The discussed example is a suitable variant of network generation and its subsequent learning for set problem, and in its terms the key effect of the use of the considered technologies becomes clear – initial resource and service model building in online mode without expert engagement.

Conclusion

The key effect of the use of the considered technologies consists of initial resource and service model building in online mode without engagement from experts.

In our opinion, it is practical to continue further studies in the direction of detailed configuring of the neural network architecture to solve the problem in question and the use of available computational capabilities for periodic neural network relearning based on constantly updated teaching selections.

In this case, the question of effective use of computing system, namely its dynamically distributed resources, should be built on the principles of parallel processing calculation tasks while using algorithms employed in the distribution of works for multiprocessor computing systems [Mld19].

This will make it possible to provide development of adequate models not requiring substantial debugging by a group of experts in terms operation of IT services, and keeping them in their current state which ultimately provides obvious improvement in the real-time evaluation quality of MCC JSC Russian Railways IT services.

Further development of monitoring system should be performed in relation to the results obtained.

References

[Aya07] N. Ayachitula. IT Service Management Automation - a Hybrid Methodology to Integrate and Orchestrate Collaborative Human Centric and Automation Centric Workflows. / N.Ayachitula, M. J. Buco, Y. Diao, M. Surendra, R. Pavuluri, L. Shwartz, C. Ward – In IEEE SCC, 2007. 574–581 p.

[Cuk14] K. Cukier. A Revolution That Will Transform How We Live. / K. Cukier, V. Mayer-Schonberger NY: Mariner Books, 2014. 240 p.

[Gol17] A.S. Golubev. Digital Railway is Reality / A.S. Golubev, A.V. Skryabin – Russia, Eurasia News. 2017, №12.

[Mar09] P. Marcu. Managing Faults in the Service Delivery Process of Service Provider Coalitions. / P. Marcu, L. Shwartz, G. Grabarnik, D. Loewenstern – In IEEE SCC, 2009. Pp. 65–72.

[Mol19] I.A. Molodkin, S.G. Svistunov. Comparative Analysis of Scheduling Algorithms in Multiprocessor Systems, Intellectual

Technologies on Transport. 2018, № 2. Pp. 41–46.

[Naz03] A.V. Nazarov. Neural Network Algorithms of Forecasting and Optimization of Systems / A.V. Nazarov, A.I. Loskutov, Saint-Petersburg: Science and technique. 2003. 384 p.

[Nov16] P.A. Novikov. Software for Mobile Indoor Navigation Using Neural Networks / P.A. Novikov, A.D. Khomonenko, E.L. Yakovlev. Information Management System. 2016, №1. P.32-39.

[Ort91] J. Ortega. Introduction to Parallel and Vector Solutions of Linear Systems / J. Ortega – Russia, Moscow: World. 1991. 367 p.

[Oht06] M.Yu. Ohtilev. Intelligent Technologies for Monitoring and Control of Structural Dynamics of Complex Technical Objects / M.YU. Ohtilev, B.V. Sokolov, R.M. Yusupov, Moscow: Science, 2006. 410 p.

[Rie93] M.A. Riedmiller. Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. / M. Riedmiller, H.Braun – In IEE, Conf. on Neural Networks. San Francisco, 1993. Pp. 586-591.

[Rrw17] The Concept of Implementation of the Complex Scientific and Technical Project "Digital Railway" – Russia, Moscow, 2017. 92 p.

[Sha18] K. S. Shardakov. Comparative Analysis of the Popular Monitoring Systems for Network Equipment Distributed Under the GPL License / K.S. Shardakov, V.P. Bubnov, Intellectual Technologies on Transport. 2018, №1. Pp. 44–48.

[Sup16] M. Supriya. Monitoring and Evalution in adaptation. / M.Supriya, S.Truck, P.Davies – Darwin, 2016. 56 p.