Conference Paper

Development of a Mathematical Model for Analyzing the Performance of Operators of Web-oriented Information Systems

V. Y. Radygin[1], D. Yu. Kupriyanov[1], A. S. Merkusheva[2], and D. A. Hotelov[2]

[1]Department of Financial Monitoring, National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse, 31, Moscow, 115409, Russia
[2]Moscow Polytechnic University, Bolshaya Semyonovskaya Street, 38, Moscow, 107023, Russia

Abstract

The issue of developing a mathematical model for estimating the dynamics of changes in the performance of operators of web-oriented information systems is considered. The relevance of these systems is shown both for the evaluation of the effectiveness of the units' activities and for the formation of the psychological climate of the collective. The generality of the solution of this issue and the issues of revealing abnormal behavior as a sign of harmful activity is described. The applicability of this approach for assessing the quality of interface updates of the system is underlined. The proposed mathematical model and the numerical algorithm for its building are thoroughly disassembled. A description of the system for evaluating the performance of operators of the ERP system of the National Research Nuclear University MEPhI, developed on the basis of this model, is given. Conclusions are drawn about the applicability of the proposed mathematical model.

Keywords: mathematical model, employee behavior, psychological climate of the team, detection of abnormal behavior, information system, Ruby on Rails, evaluation of the quality of interfaces, monitoring of operations, performance loss factors, criminal or illegal activity

Corresponding Author:
V. Y. Radygin
vyradygin@mephi.ru

Received: 22 July 2018
Accepted: 9 September 2018
Published: 8 October 2018

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the Breakthrough Directions of Scientific Research at MEPhI Conference Committee.

1. Introduction

Today, many human activities are associated with continuous work in information systems. Operators of customer service centers, call-centers and other departments of banks, insurance companies, online stores and many commercial and budgetary organizations perform thousands of operations on entering or selecting information daily. The effectiveness of their work can be determined by various factors, including the qualifications of employees, their psychological and physical condition, excessive or insufficient workload, the convenience of software (including the effectiveness of

the latest updates) and many other parameters. Each of these factors require management to have its own control, the loss of which can directly affect both the company's profits and its many potential customers.

Timely detection of the decline in the performance of both the individual employee and the team as a whole is an extremely important task for any successful leader. However, if operators perform thousands of similar actions per day, it is not enough to determine the fact of performance loss. It is also necessary to understand what the cause is and how to fix it. At the same time, we should not forget that excessive control can also lead to a drop in productivity due to the deterioration of the team's psychological state. The maximum importance of this factor for production efficiency is shown in the works of many authors. For example, Fernández-Aráoz C. in his book [1] shows the importance of not only the competence of employees, but also their contribution to the psychological climate of the team when creating a work team. It is important to understand that the emotional state can be violated both in the team as a whole and in individuals. The most vivid example of this situation can be professional burnout. The danger of these situations is described in detail in the works of N. Kind [2], L. Zaninotto [3] and J. Brearley [4].

An equally important problem of organizing the work of modern divisions related to work on computers and the Internet is the factor of distraction of employees for non-working activities. For example, many young professionals are characterized by a great loss of time for communication in social networks and instant messengers. Nevertheless, the activities of many organizations, for example, online stores do not allow you to block these services at the proxy server level. All this requires the introduction of delicate tools to monitor the effectiveness of employees, without direct intervention in their daily work.

It should be noted that there is a large number of software tools to monitor the actions of employees on the computer. These include, for example, the technologies KickDlr [5], HubStaff [6], Yaware [7] and StaffCop [8] and others. Nevertheless, these software have a number of significant drawbacks. First of all, all the specified applications require direct installation on the computers of employees. Thus, each employee understands that all his actions are controlled by management and system administrators. This leads to a sharp deterioration in the psychological state of the team, and often to the dismissal of key employees. In addition, huge amounts of data are generated by these applications, including screenshots of employee desktops and other large multimedia objects, which is extremely difficult to analyze from the management point of view.

On the other hand, modern information systems in most cases are web-oriented applications that are operated on the company's server. At the same time, it is possible to maintain logs that can store not only changes made by the employees to the database objects but also all the operator actions in the browser on the information system site. Analyzing this information, it is possible to get an assessment of the employees' productivity without using additional software on their computers and excluding from the analysis all information not related to the direct task of the operator.

In addition, the analysis of information obtained from the log journals of information systems also allows you to evaluate the effectiveness of information systems updates, both in terms of work of individual employees and in terms of the work of a separate division or the entire organization as a whole. Particular difficulty in assessing the quality of information system updates is the analysis of changes in ergonomics of application's user interfaces. Today, the technology for assessing the convenience of interfaces that is named GOMS [9] is the most popular and requires the separation of different types of operations into categories, which is not always possible in complex integrated systems. On the other hand, the effectiveness of interface updates can be evaluated based on employee performance changes, which makes this task comparable with the task of assessing the performance of employees.

Thus, it is important to create an automated system that allows analyzing the effectiveness of employees or their groups based on the log journals of one or several information systems. An effective solution of this problem requires the construction of a mathematical model of the employee's behavior in the information system on the basis of multicomponent analysis. It is also worth noting that the introduction of a mathematical model of the employee's behavior will also help to identify situations that are characterized by a clear deviation from the norm in the actions of the employee, which may be a signal about criminal or illegal activity.

2. Mathematical Model of Operator Behavior of Web-oriented Information Systems

The data of information systems logs are used as input parameters of the mathematical model. This set of information includes the following components: the time of creation/modification/deletion of the object, including the values of its attributes; normal HTTP-requests to the system and their parameters; Ajax-requests to the system and their parameters; user actions on the system site using the mouse and keyboard.

Without preprocessing, such parameters cannot be used to build the model. Therefore, based on the business processes used by the organization, the most important components of the operator's production cycle are identified and considered in a time-aggregated state. For example, the following elements of the business process can be selected for the online store operator: order creation, adding products to the order, removing goods from the order, providing discounts or its analogues, changing the order status, changing the customer data and the delivery address. Based on the analysis of information system log journals, the following parameters, discretized by time, can be singled out: the number of orders created per hour; the number of goods added to the order per hour; the maximum number of products in one order per hour; the minimum number of products in one order per hour; the amount of ordered goods per hour; the amount of goods removed from orders per hour; average order price per hour; the number of discount facts; the maximum amount of the discount for an hour, etc. In General, the list of model parameters depends on the task and is determined by the analyst of the organization. The discretization of all parameters in time should be the same.

Let's introduce the following notation: $f_{i,j}$ – the aggregated value of the input parameter i at the step of discretization in time j; τ – the step of discretization the model in time; M – the number of input parameters. Then, the behavior of a person in an information system can be described using a discrete vector function:

$$f_j = \left\{ \begin{array}{c} f_{1,j} \\ f_{2,j} \\ \vdots \\ f_{M,j} \end{array} \right\}.$$

The employee's performance is largely determined by his chronotype [10]. Some employees work faster and more efficiently in the morning, while others in the evening. In this case, the organization of working time, determined by the company's business processes, can affect the amount of work that an employee can accomplish over a certain period of time. For example, the company's business processes can regulate the number of days off, determine the days of supply of new goods, banking days, etc. For this reason, under the unchanged operating conditions of the operator, it is always possible to distinguish a certain chronological cycle within which his behavior often has the same character. Let's denote this time interval by the symbol T.

We divide the domain of definition of the vector function f into segments of length T. We will estimate the efficiency of the employee's activity by using the deviation of his behavior on the current segment T_j from the model of his behavior constructed on the

basis of the previous segments $T_1 - T_{j-1}$. Such an approach will allow not only to assess the dynamics of changes in the efficiency of employees, but also to reveal extraordinary deviations in their behavior. The employee's behavior model is constructed by iteration, until the changes become insignificant within a given accuracy e, which is chosen empirically within 10–30%.

Let's describe the algorithm of constructing the model, which consists of seven steps, and introduce some additional notation:

1. Let's set the value of discretization step number j to 1.
2. Then calculate the value of vector function F_j:

$$F_j = \left\{ \begin{matrix} F_{1,j} \\ F_{2,j} \\ \vdots \\ F_{M,j} \end{matrix} \right\}.$$

   To do this, for each component $f_{i,j}$ of the vector function f_j is approximated by b-splines [10] on the segment T_j of the least-squares line calculated for it (Figure 1).

3. Then, for each $F_{i,j}$ it is required to calculate the following evaluation parameters: $Fmax_{i,j} = max_{k \in T_j} f_{i,k} - F_{i,k}$; $Fmin_{i,j} = -min_{k \in T_j} f_{i,k} - F_{i,k}$;

$$Fsq_{i,j} = \sum_{k \in T_j} {f_{i,j}}^2 - {F_{i,j}}^2; \; Ffre_{i,j} = \sum_{k \in T_j} g_k,$$

$$\text{in which } g_k = \begin{cases} 1 & if \; |f_{i,k} - F_{i,k}| > \dfrac{e\left(Fmax_{i,j} + Fmin_{i,j}\right)}{100} \\ 0 & else \end{cases};$$

$$Fwaf_{i,j} = \frac{1}{2} \sum_{k \in T_j} p_k, \text{ in which } p_k = \begin{cases} 1 & if \; f_{i,k} = F_{i,k} \\ 0 & else \end{cases}.$$

4. Using the aforementioned calculated parameters, it is required to calculate the estimated parameters for f_{j+1} and F_j.
5. If none of the estimated parameters of the $(j + 1)$-st step exceeds more than e percent of the corresponding parameter of the j-th step, then it shall be assumed that the vector function F_j describes model of human behavior and stop the algorithm.

6. Otherwise it is necessary to set the value of the variable j to $j + 1$.
7. And replace the values of the vector function f_j in the k-th step of the discretization of the time interval T_j ($f_j{}^k$) according to the following principle:

$$\tilde{f}_j^k = \left\{ \begin{array}{c} \frac{\sum_{p=1}^{j} f_{1,p}^k}{j} \\ \frac{\sum_{p=1}^{j} f_{2,p}^k}{j+1} \\ \vdots \\ \frac{\sum_{p=1}^{j+1} f_{M,p}^k}{j+1} \end{array} \right\}.$$

8. After that it is necessary to return on step (ii) and repeat steps (iii–viii).

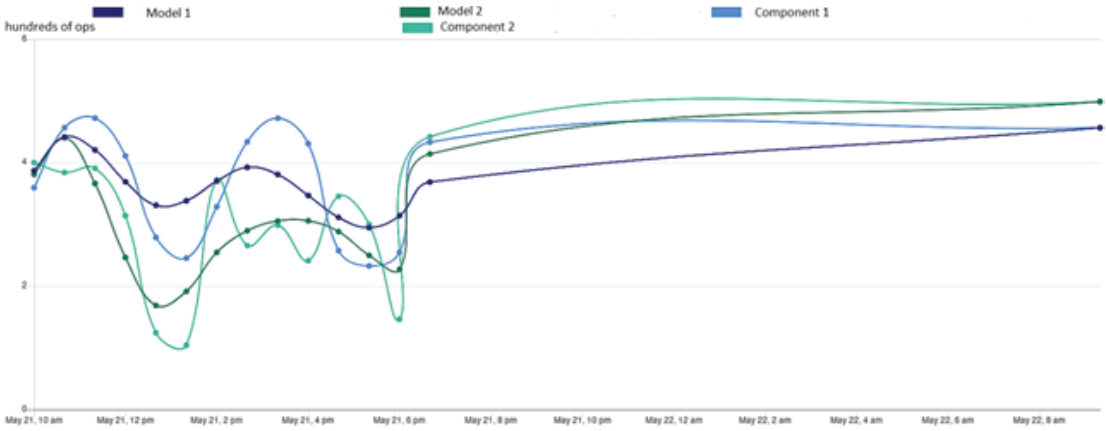

Figure 1: Graphs of two components and their models on a time interval T = 1 day (with 8 hours of employees working day).

Under the condition that the maximum values of the components of the vector function f_j are bounded above and below, this algorithm converges in a finite number of steps to a certain solution. If only the restriction on the lower boundary of the vector function is guaranteed, then for the convergence of the algorithm it is necessary to consider the trend of each component during the model creation.

3. Approbation of the Model

As an approbation of the developed model, a system for assessing the effectiveness and identifying deviations in the work of ERP-system operators of the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute) was created. The algorithm for model creation was written using the Python programming language and the SciPy mathematical library. It should be noted that university ERP-system based on the Ruby on Rails. In a number of early works of the authors, the

high efficiency of this framework for solving the problems of various areas of human activity has been shown [11, 12]. The high efficiency of this web development platform is also indicated by its relevance in the world. According to a study by the Australian independent agency BuiltWith [13], more than 11% of the most visited sites are implemented using Ruby on Rails technology.

The system for assessing the effectiveness and identify deviations in the work of the operators was used for the detailed evaluation of a number of departments, as well as individual employees, working with the University of ERP-system. The results can be used as an effective component KPI contract. In addition, with this mathematical model has been identified and eliminated a number of inefficient decisions made when updating the interface of the University of the ERP-system.

4. Conclusion

Thus, as a result of the research, a mathematical model of the employee's behavior in the information system based on multicomponent analysis is proposed. The model allows to evaluate the productivity of employees and divisions. This helps to detect abnormal behavior with the aim of suppressing criminal or illegal activity. Also it helps to assess the quality of updates to the interface of web systems using the analysis of log journals data.

Approbation of the work showed its qualitative applicability for solving real applied problems.

References

[1] Fernández-Aráoz, C. (2014). *It's Not the How or the What but the Who: Succeed by Surrounding Yourself with the Best*, p. 244. Harvard Business Review Press.

[2] Kind, N., Eckert, A., Steinlin, C., et al. (2018). Verbal and physical client aggression – A longitudinal analysis of professional caregivers' psychophysiological stress response and burnout. *Psychoneuroendocrinology*, vol. 94, pp. 11–16.

[3] Zaninotto, L., Rossi, G., Danieli, A., et al. (2018). Exploring the relationships among personality traits, burnout dimensions and stigma in a sample of mental health professionals. *Psychiatry Research,* vol. 264, pp. 327–333.

[4] Brearley, J. (2018). Personal experience of professional burnout. *Veterinary Record*, vol. 182, no. 16, pp. i–ii.

[5] The official website of monitoring system KickDler, https://www.kickidler.com/ (last accessed June 22, 2018).

[6] The official cite of monitoring system HubStaff, https://hubstaff.com/employee_monitoring (last accessed June 22, 2018).

[7] The official cite of Yaware application, https://yaware.ru (last accessed June 22, 2018).

[8] The official cite of StaffCop application, www.staffcop.ru (last accessed June 22, 2018).

[9] Raskin, J. (2000). *The Humane Interface: New Directions for Designing Interactive Systems*, p. 233. Addison-Wesley Professional.

[10] Alberg, J. (1972). *The Theory of Splines and Its Applications*, p. 319. Moscow: Mir.

[11] Radygin, V. Y., Lukyanova, N. V., and Kupriyanov, D. Yu. (2017). LMS in university for in-class education: Synergy of free software, competitive approach and social networks technology, in *Proceedings of the International Scientific-Practical Conference Information Technologies in Education of the XXI Century*, #91 (ITE-XXI), pp. 020015-1–020015-8. AIP Publishing.

[12] Radygin, V. Y., Kupriyanov, D. Yu. (2018). Technological support of real-time interaction in web clients of analytical fraud detection systems, in *Proceedings of the III Network AML/CFT Institute International Scientific and Research Conference "FinTech and RegTech"*, pp. 450–457. KnE Social Sciences.

[13] Statistical analysis of using different frameworks in the Internet from BuiltWith Pty Ltd Company, http://trends.builtwith.com/framework (last accessed June 22, 2018).