

Conference Paper

Technological Support of Real-Time Interaction in Web Clients of Analytical Fraud Detection Systems

Radygin V. Y. and Kupriyanov D. Yu.

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse 31, Moscow, 115409, Russia

Abstract

The optimal tools selection for design of web-based visual mining client for real time fraud detection systems was discussed. The features of modern real time fraud detection software were analyzed. The necessity of transition to using of web-based technologies for client software design was shown. The market of web-frameworks and browser to web-server data exchange technologies were investigated. Basing on experimental research the most efficient toolset for design of web-client software for real time fraud detection systems was offered.

Keywords: fraud detection, Visual Mining, real time data exchange, web-visualization, webSockets, MessageBus.

Corresponding Author:

Radygin V. Y.

vyradygin@mephi.ru

Received: 11 December 2017

Accepted: 20 January 2018

Published: 13 February 2018

Publishing services provided by
Knowledge E

© Radygin V. Y. and Kupriyanov D. Yu.. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the FinTech and RegTech: Possibilities, Threats and Risks of Financial Technologies Conference Committee.

1. Introduction

There is a hyperactive evolution of information technology worldwide. This has led to automation of various fields of human activities, ranging from telecommunications environment to the law enforcement field. Not an exception is the issue of countering fraud. Databases, OLAP-systems, Data Mining and Visual Mining technologies have become powerful tools of modern financial monitoring institutions. Nevertheless, even the use of powerful mathematical instruments and software does not allow counter-acting the fraud without human control. Therefore, the final step in the fraud-detection process is always the decision to involve an operational law enforcement, which takes an expert analyst.

The effectiveness of protecting the financial structure from fraud and illegal activity depends on the speed and effectiveness of these analysts. Providing the ergonomics and simplicity of software, which financial monitoring specialists use, is the primary task of the management of any financial institution, such as commercial banks or state institutions that provides the economic security of the country.

 OPEN ACCESS

Improvement of hardware and data mining methods has led to the differentiation of modern tools of preventing fraud and rapid response to illegal actions. There are two main categories: post-factum analysis tools and real time tools. The second one is of greatest interest today, as it allows to quickly react to illegal activities and thus prevent fraud or minimize its effects in "hot pursuit". Such technologies are now widely used both in the banking sector and in many others fields, such as the field of electronic payments, Internet trading, etc. For example, the issues of real-time fraud detection among the Internet trading environment are devoted to a number of works. Today the most interesting of them were written by J. Akhilomen [1] and F.Carcillo [2]. J. Akhilomen in his work described the problem of rapid tracking facts of fraud with credit cards when paying using Internet. This solution based on the neural network methods. In turn, F. Carcillo solves a similar problem using Big Data methods and machine learning. W.N. Robinson [3] offers methods that solve the problem in a narrower field. He describes the creating of real-time system, which allows detecting the fraud when using prepaid cards. The solution is based on divergence analysis of the hidden Markov model.

New technologies, in turn, require new approaches to fraud protection. For example, the work of Y. Peng [4] describes the problem of fraudulent activities detection during online auctions. The solution based on the algorithm using bipartite cores. It is worth noting that none of the considered solutions excludes analyst work. This, in turn, determines the special requirements for these software products.

An increasing number of demanded real-time systems leads to the search for modern technological software to support their development and design. No less important is the role of the equipment that provides a visual representation of the identified knowledge and the critical need for developing complicated software for storing large amounts of data and all kinds of Data Mining.

Today there are a large number of visualizing analytical information methods, technologies for representing information on computers and protocols of information exchange between a client and a server. On the other hand, in the last decade the software and hardware market has undergone serious changes and has acquired a pronounced tendency to abandon client stand-alone applications in favor of web services, to abandon local storage in favor of cloud storage, to abandon desktops in favor of mobile devices. Today, the younger generation prefers not to work with a local application tied to a desktop, but with a web service available on any mobile device and providing the same user interfaces via a browser.

Thus, the most important problem of modern tools development for tracking fraud in real time is to provide effective technologies for creation of visual content in a web browser and mechanisms for the rapid exchange of data between the web browser and a remote server that performs Data Mining tasks. This problem is especially relevant when developing software for large geographically distributed companies that use a single security service.

Today technological software tools for creation of visual content are widely represented. They include both low-level tools such as SVG, WebGL, Flash and high-level tools such as HighSharts or jqPlot that implements ready to use visualization templates using Java or Javascript. Thus, for example, in the works written by of D. Lv [5] and M. Pignatelli [6], the problem of scientific visualization based on SVG technology is described. A. Arbelaiz [7], in his work, demonstrated the possibilities of modern web-technologies using WebGL library for constructing three-dimensional scenes. The technology of using jqPlot tools for scientific visualization is widely represented in work written by K. Wang [8]. In particular, the problem of visualization of the results obtained in the Antarctic by analytical data set processing was discussed. Of particular interest is the work of L. Chi [9] describing the main features of modern tools for web visualization, oriented to work with mobile devices.

Thus, the problem of choosing the software for creating visual content in web applications is adequately viewed. Unfortunately, the problem of choosing the mechanisms for rapid data exchange between the browser and the web server with the local Data Mining system does not have ready to use solutions and requires detailed research.

2. Real-time interaction between the browser and the web server

It is worth noting that the real-time interaction technologies between the browser and the web server are in demand not only in the field of countering fraud, but in many other fields of human activity. For example, in work written by S. Mijovic [10], the problem of using the technology of real-time interaction between a web browser and a web server for developing the Internet of things elements is described. N.V. Lukyanova [11], in her work, shows the popularity of these technologies in the field of higher education.

Nevertheless, the tasks of financial monitoring in the choice of technologies for the rapid exchange of data between the browser and the web server have significant advantages over many other areas of human activity, consisting in a relatively small

number of simultaneous requests to the server per unit time. This advantage is due primarily to the fact that the constant users of such systems are the advantage of analysts whose staff in medium-sized companies is usually estimated in hundreds (and sometimes even tens) of people. Thus, when developing web applications for fraud monitoring systems, a wide range of protocols can be used between the browser and the web server, including protocols based on a persistent connection and regular polls.

Unfortunately, web-based information systems use basically a standard approach to organizing user interfaces. This approach consists in a one-way method of interaction between the client browser and the web server. When using it, interaction with the user is defined by a script that implements step-by-step execution of actions on the client side. Thus, the user initiates the transition to the next step of the script either by executing the HTTP request or by using the asynchronous request with Ajax. In normal mode, the server is not able to inform the client's web browser about the event itself (for example, about the detected suspicion of a fraudulent fact). In this case, visualization of information about a new event occurs only when the page is updated manually by the user or by the results of interactive JavaScript-routines. Thus, real-time visualization is not a part of ordinary web applications. Therefore, if necessary, additional communication tools and technologies are required.

There are three main technologies in the field of providing data exchange between the web server and the client browser today: polling technology, long polling technology (Ajax Long-Polling) and permanent connection technology. The technology of permanent connection can be implemented either on the basis of WebSockets (WebSockets) or on the basis of events sent by the server HTML 5 (HTML5 Server Sent Events - SSE). In detail, the basic communication technologies in web applications are described by K. Shuang [12]

The technology of polling is a periodical refreshing of the page using JavaScript, for example, every 5 seconds. This technology is the least demanding for the client browser version, but it has some significant disadvantages, such as a continual high load on the client browser due to the frequent page reloads and a continual high load on the network and the web server due to the large quantity of requests. The fact that there is a possible delay in time between the two next-in-time updates of the page can have negative consequences if fraud is detected. A long delay can lead to the loss of information by analysts, as well as to the difficulty in visualizing analytical content. In this regard, we can conclude that the technology of polling is not suitable for solving the problems of fraud tracking.

The technology of long polling is a sending ajax-message to the web server by the browser. The server ends the processing of ajax-message only if there are any special events. The advantage of this approach is that older browsers support this technology as well (within reasonable limits). The disadvantage is the complexity of the ajax-request to the web server, which is higher than the usual TCP / IP socket.

Technologies that use both WebSockets and server-side HTML 5 events are in fact based on establishing a connection using a protocol that is an add-on over TCP / IP between clients and the server. The advantage of this approach is the simplicity of the connection, achieved due to the renunciation of the HTTP protocol. The disadvantage is that older browsers do not support it. The most evolving among the technologies using webSockets or sent HTML5 events to the server are webSockets.

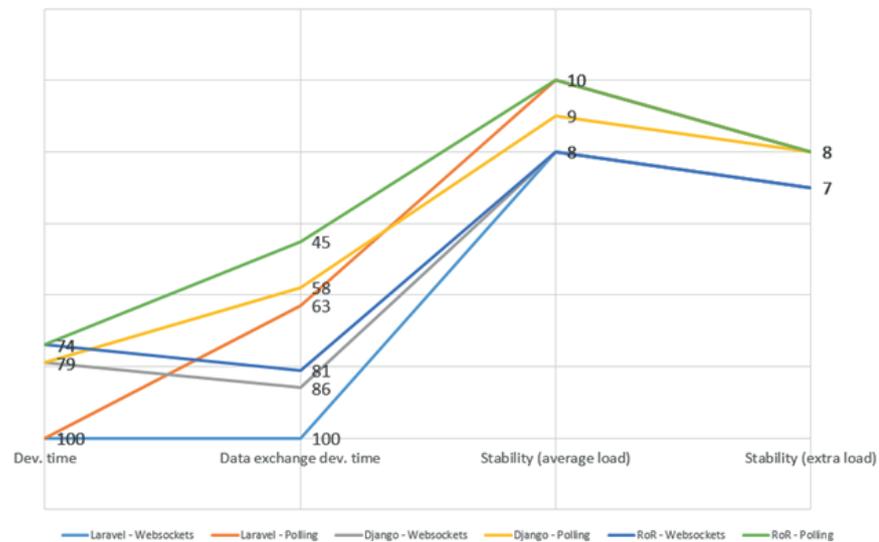
Thus, today there are only technologies using webSockets or long polling that can be used for fraud detection visualization tasks in the field of financial monitoring. Therefore, the choice of real-time communication technology between the browser and the web server for such tasks can be reduced to a choice between them. In this paper, a study has been performed to determine the most effective of these tools in the framework of modern free software technologies for web development.

3. Selection of web-client design technology for fraud detection software

The research on modern market of web frameworks using webSockets or long polling data exchange has been done to find out best combination of technologies for design of fraud detection software.

According to the research by independent Australian Company BuiltWith [13] the most popular free web-frameworks are frameworks built on the following programming languages: PHP, Java, Ruby and Python. Because of high complexity of development the Java language was excluded from analysis. For each of the remaining programming languages the most popular frameworks have been selected. The selection was based on analytical data of GoogleTrends and rating of the most popular web frameworks represented by HotFrameworks service. As a result, the next frameworks have been selected: Larvel (PHP programming language), Django (Python programming language) and Ruby on Rails (Ruby programming language).

For all selected frameworks, the experimental research has been conducted. The research had two main ways: effectiveness of development and stability in use. For assessing these indicators simple combination of web-server and web-client with the



Picture 1: The results of the experimental research on base technologies of real-time web-services development.

real-time data exchange has been created. This software solves the task of analytical database investigation of transactions by electronic cards.

Thus, six test applications have been created. Each of them was solving identical tasks by own technical tools including next combinations: Laravel-WebSockets, Laravel-Polling, Django-WebSocket, Django-Polling, Ruby on Rails-WebSocket, Ruby on Rails-Polling. For implementation of the real time data exchange, the most popular packages or built-in modules have been used.

The results of the experimental research are assessed by four indicators. The first indicator is the average development time of the base core. The second indicator is the average development time of the real time data exchange. The third indicator is the system stability with average load by user queries (in ten points). The fourth indicator is the system stability with extraordinary load by user queries (in ten points). The generalized results are shown in Pic. 1.

Combination of the framework Ruby on Rails and gem-package MessageBus has the best result by research estimation. The MessageBus gem-package is the implementation of long polling technology for Ruby programming language. Its advantages are caused by simplicity of the Ruby language and easiness of the realization of the MessageBus technology in comparison with webSockets. This advantage becomes more noticeable if the fact of building webSockets in Ruby on Rails by framework developer is taken into account.

The advantages of MessageBus technology over webSockets technology at average load is explained by the difference between WebSocket protocol and HTTP protocol.

Many users have faced blocking or missing some packages of webSocket queries at local networks of their companies. Nevertheless, minimization of restrictions on webSocket protocol by administration of the local networks has not given any big advantages for webSockets over MessageBus. Taking into account features of the fraud detection software makes it possible to consider them of being of low importance.

4. Conclusion

As a result of experimental research the most efficient toolset for design of web-client software for real time fraud detection systems was offered. The optimal free software combination is the symbiosis with Ruby on Rails framework and long polling technology based on gem-package MessageBus.

Acknowledgements

This work was supported by Competitiveness Growth Program of the Federal Autonomous Educational Institution of Higher Education National Research Nuclear University MEPhI (Moscow Engineering Physics Institute).

References

- [1] J. Akhilomen, "Data Mining Application for Cyber Credit-Card Fraud Detection System" in *Proceedings of the 13th Industrial Conference – Advances in Data Mining: Applications and Theoretical Aspects*, New York, NY, USA, 2013, (LNAI 7987) pp. 218–228.
- [2] F. Carcillo, A. Dal Pozzolo, Y. Le Borgne, O. Caelen, Y. Mazzer, G. Bontempi, "SCARFF : A scalable framework for streaming credit card fraud detection with spark" in *Information Fusion*, #41 (2018), Elsevier, 2017, pp. 182–194.
- [3] W.N. Robinson, A. Aria, "Sequential fraud detection for prepaid cards using hidden Markov model divergence" in *Expert Systems With Applications*, #91 (2018), Elsevier, 2017, pp. 235–251.
- [4] Y. Peng, L. Zhang, Y. Guan, "Detecting Fraud in Internet Auction Systems" in *Expert Systems With Applications*, #91 (2018), Elsevier, 2017, pp. 187–199.
- [5] D. Lv, M. Yu, J. Song, K. Qian, Y. Cui, "Study on Police Graphic Plotting Technology based on Web" in *Proceedings of the 6th International Conference on Software and*

- Computer Applications*, (ICSA 2017), ACM, 2017, pp. 137–143.
- [6] M. Pignatelli, "TnT: A set of libraries for visualizing trees and track-based annotations for the web" in *Bioinformatics*, Vol. 32, Iss. 16, 2016, pp. 2524–2525.
- [7] A. Arbelaz, A. Moreno, L. Kabongo, A. García-Alonso, "X3DOM volume rendering component for web content developers" in *Multimedia Tools and Applications*, Vol. 76, Iss. 11, 2017, pp. 13425–13454.
- [8] K. Wang, D. Neudegg, C. Yuile, M. Terkildsen, R. Marshall, M. Hyde, G. Patterson, C. Thomson, A. Kelly, Y. Tian, "Antarctic space weather data managed by IPS radio and space services of Australia" in *Data Science Journal*, Vol. 13, 2014, pp. PDA44–PDA50.
- [9] S. Mijovic, E. Shehu, C. Buratti, "Comparing Application Layer Protocols for the Internet of Things via Experimentation" in *Proceedings of the 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow*, (RTSI 2016), IEEE, 2016, pp. 1–5.
- [10] L.-J. Chi, C.-H. Huang, K.-T. Chuang, "Mobile-friendly and streaming web-based data visualization" in *Proceedings of the Conference on Technologies and Applications of Artificial Intelligence*, (TAAI 2016), IEEE, 2016, pp. 124–129.
- [11] V.Y. Radygin, N.V. Lukyanova, D.Yu. Kupriyanov, "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), AIP Publishing, 2017, pp. 020015-1 – 020015-8.
- [12] K. Shuang, K. Feng, "Research on server push methods in web browser based instant messaging applications" in *Journal of Software*, Vol. 8, Iss. 10, IEEE, 2013, pp. 2644–2651.
- [13] <http://trends.builtwith.com/framework> – statistical analysis of using different frameworks in the Internet from BuiltWith Pty Ltd Company (last visit date 14.11.2017).