

# ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА И ВИЗУАЛИЗАЦИИ ДАННЫХ НА ПРИМЕРЕ ПРОГРАММНОГО ПРОДУКТА SPLUNK ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ УСТРОЙСТВ ИНТЕРНЕТА ВЕЩЕЙ

М. В. Захаров<sup>1\*</sup>, Р. В. Киричек<sup>1</sup>

<sup>1</sup> СПбГУТ, Санкт-Петербург, 193232, Российская Федерация

\* Адрес для переписки: [zaharov.spbgut@gmail.com](mailto:zaharov.spbgut@gmail.com)

## **Аннотация**

Четвёртая промышленная революция в скором будущем откроет для человечества невиданные ранее возможности. В различных областях знания — медицине, промышленности, сельском хозяйстве и т. д. внедрение автоматизированных комплексов и умных устройств, новых технологий виртуализации и автономной или управляемой робототехники позволит значительно повысить эффективность человеческой деятельности. Вместе с тем, использование передовых технологий предполагает отказ или пересмотр традиционных концепций построения, эксплуатации и управления производством, инфраструктурой и телекоммуникационными сетями. Одним из направлений дальнейшего развития может быть использование технологий анализа данных в режиме реального времени с применением технологий визуализации полученных результатов.

**Предмет исследования.** В статье рассмотрен пример использования программного продукта Splunk для случаев, которые могут происходить при взаимодействии различных типов устройств Интернета Вещей. **Метод.** Анализ предметной области. Синтез предложений по решению обозначенных в статье проблем. **Основные результаты.** Произведён критериальный подбор программного обеспечения для решения определённого в статье круга проблем. Сделан вывод о возможности решения поставленных в статье задач. **Практическая значимость.** Предложенные методы позволят значительно сократить накладные расходы при эксплуатации сетей связи, повысить качество обслуживания и надёжность сетей.

## **Ключевые слова**

Интернет Вещей, 5G, анализ данных, обработка данных, визуализация, Splunk.

## **Информация о статье**

УДК 004.725.7

Язык статьи — русский.

Поступила в редакцию 29.05.17, принята к печати 01.09.17.

**Ссылка для цитирования:** Захаров М. В., Киричек Р. В. Применение методов анализа и визуализации данных на примере программного продукта Splunk для повышения эффективности использования устройств Интернета Вещей // Информационные технологии и телекоммуникации. 2017. Том 5. № 3. С. 38–44.

# APPLICATION OF METHODS OF ANALYSIS AND VISUALIZATION OF DATA EXAMPLE, SOFTWARE SPLUNK TO IMPROVE THE EFFICIENCY OF USE OF DEVICES OF THE INTERNET OF THINGS

M. Zakharov<sup>1\*</sup>, R. Kirichek<sup>1</sup>

<sup>1</sup> SPbSUT, St. Petersburg, 193232, Russian Federation

\* Corresponding author: zaharov.spbgut@gmail.com

**Abstract**—The fourth industrial revolution in the near future will reveal to mankind unprecedented opportunities. In various fields of knowledge — medicine, industry, agriculture, etc. implementation of automated systems and smart devices, new technologies of virtualization and managed or Autonomous robotics will significantly increase the efficiency of human activity. However, the use of advanced technologies assumes the failure of or a revision of the traditional concepts of build, operate and manage manufacturing, infrastructure and telecommunication networks. One of the areas for further development could be the use of big data in real time with the use of technologies of visualization of the obtained results. **Research subject.** The article presents an example of the software using Splunk for events that can happen in the interaction of different types of devices in the Internet of Things. **Method.** Analysis of the subject area. Synthesis of proposals for the solution indicated in the article problems. **Core results.** Produce a criteria-based recruitment software for the solution defined in article range of problems. The conclusion is made about possibility of solution of the article task. **Practical relevance.** The proposed methods allow to considerably reduce overhead in the operation of communication networks, improve service quality and reliability of networks.

**Keywords**—Internet of Things, 5G, data analysis, data processing, visualization, Splunk.

## Article info

Article in Russian.

Received 29.05.17, accepted 01.09.17.

**For citation:** Zakharov M., Kirichek R.: Application of Methods of Analysis and Visualization of Data Example, Software Splunk to Improve the Efficiency of Use of Devices of the INTERNET of Things // Telecom IT. 2017. Vol. 5. Iss. 3. pp. 38–44 (in Russian).

## Введение

Возникновение и последующее повсеместное распространение концепции Интернета Вещей, а затем и Индустриального Интернета Вещей являются предвестниками и неотъемлемыми составляющими надвигающейся четвёртой промышленной революции, в рамках которой произойдёт резкое увеличение числа умных электронных устройств, окружающих человека в повседневной жизни [1, 2]. В настоящее время учёные и исследователи со всего мира ведут активные разработки в областях носимой и вживляемой биомедицинской электроники, автоматизации и роботизации промышленного производства и сельского хозяйства, автотранспорта и сферы услуг [3]. Целью данных разработок является

улучшение качества и увеличение продолжительности жизни людей. Многие умные устройства уже вошли в нашу жизнь и стали обыденностью. Это, например, спортивные браслеты, контролирующие показатели здоровья спортсменов, или бытовая электроника, которой пользователь может управлять удалённо со смартфона или планшета [4].

### **Проблематика**

Неотъемлемой частью сетей связи и автоматизированных систем управления всегда были и остаются технические специалисты, которые осуществляют контроль функционирования объектов инфраструктуры сети, производят ремонтно-восстановительные работы, участвуют в планировании расширения и развития телекоммуникационных сетей. Первоначально количество технических специалистов, необходимых для обслуживания сетей связи было достаточно велико. Со временем за счёт внедрения технологий автоматизации и информатизации потребность в большом количестве обслуживающего персонала исчезла - системы стали в значительной мере автономными.

Концепция Интернета Вещей изначально предполагала построение такого типа сетей, которые будут представлять собой автономные структуры, в максимально допустимой степени функционирующие без участия человека [5]. Однако даже в таких сетях в некоторых ситуациях без вмешательства человека не обойтись. Только хорошо подготовленный технический специалист может принять те или иные решения при возникновении неординарных ситуаций в сети и т. д. Без контроля и вмешательства человека сеть не может функционировать бесконечно долго — она может стать объектом атаки или подвергнуться другим воздействиям.

С ростом числа электронных устройств в сети и количества самих сетей, их объединяющих, придётся сделать выбор: или увеличивать количество обслуживающих сеть технических специалистов пропорционально росту устройств, или применять новые методы управления сетью (рис.). Первый вариант определённо является бесперспективным, поэтому стоит обратить внимание на различные методы повышения эффективности управления сетью. Одним из направлений повышения качества работы устройств в Интернете Вещей может стать применение в этой области методов анализа и визуализации данных в режиме реального времени [6].

### **Применение методов анализа данных**

Анализ данных это область математики и информатики, занимающаяся построением и исследованием общих математических методов и вычислительных алгоритмов извлечения знаний из экспериментальных данных. Также под анализом данных понимают процесс исследования, фильтрации, преобразования и моделирования данных с целью извлечения полезной информации и принятия решений [7]. Анализ данных имеет множество аспектов использования и подходов, охватывает различные области науки и деятельности человека.

В рамках Интернета Вещей и его приложений анализировать целесообразно журналы событий (лог-файлы), передаваемые данные и т. д. Основными задачами анализа данных можно считать контроль рабочих характеристик под-

ключённых к сети устройств, визуализацию полученных данных с целью упрощения их восприятия.

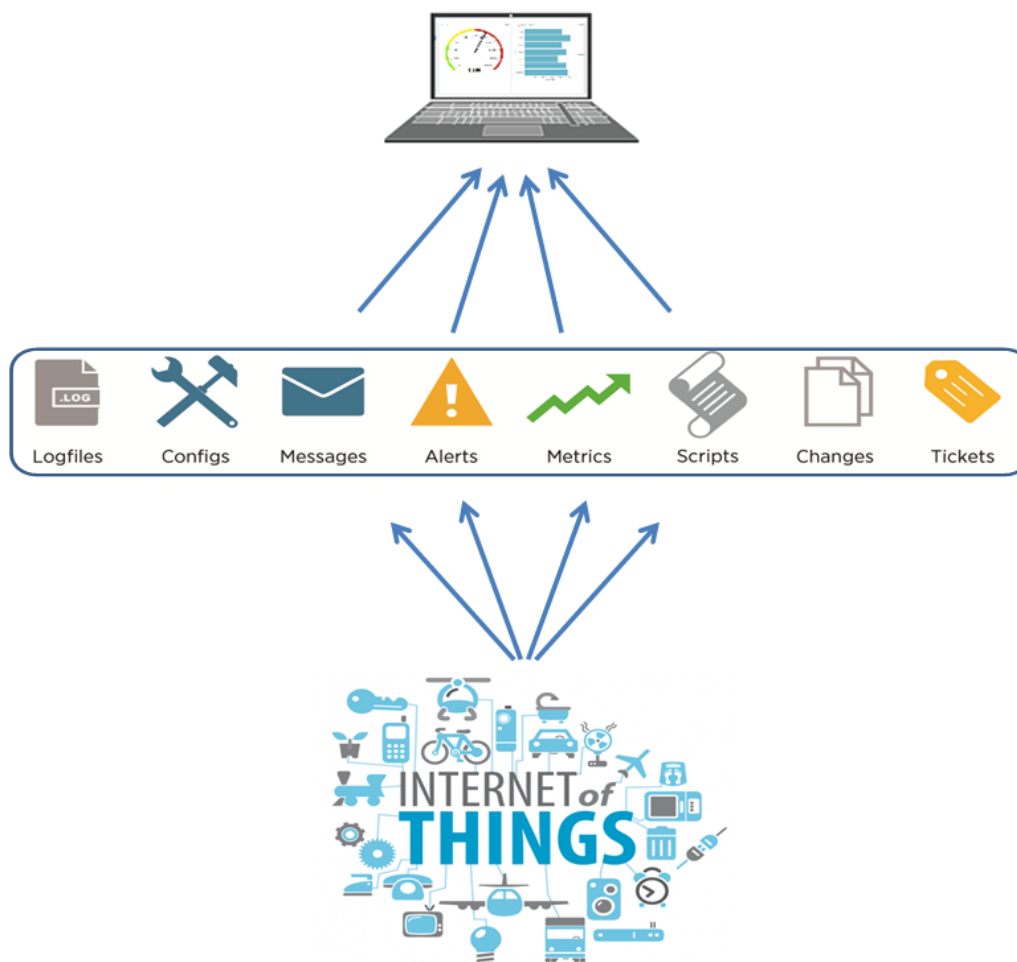


Рис. Разнообразие типов данных и необходимость их анализа

### **Инструменты анализа данных в режиме реального времени**

Одним из популярных инструментов анализа данных можно считать программный продукт Splunk от компании Splunk Inc. Эта платформа позволяет производить агрегацию, анализ и взаимодействие с большими объёмами данных, которые генерируются технологической инфраструктурой, системами безопасности и бизнес-приложениями. В результате пользователь получает наглядные статистические данные, которые значительно упрощают процесс принятия решений и управления в технически-сложных средах и производствах.

Кроме базового функционала Splunk имеет целый ряд модулей, позволяющих увеличить эффективность использования данного продукта. Стоит обратить внимание на модули категории IoT & Industrial Data, которые доступны на сайте производителя, а именно Splunk Machine Learning Toolkit и Fabrix Data Visualization for Splunk.

Приложение Splunk Machine Learning Toolkit предоставляет возможность использовать дополнительные команды языка SPL (*Splunk Processing Language* — специализированный язык запросов, использующий для работы в Splunk), дополнительные инструменты для визуализации, а также наглядные

примеры, предназначенные для обучения пользователя основам Machine Learning. Приложение включает в себя обучающие примеры с подборками тестовых данных, которые позволяют приступить к изучению возможностей приложения непосредственно сразу после установки.

Приложение позволяет:

- предсказывать значения числовых полей на основе линейной регрессии;
- предсказывать значения категориальных полей на основе логистической регрессии;
- обнаруживать числовые выбросы с помощью функции статистического распределения;
- обнаруживать категориальные выбросы;
- создавать временные ряды для прогнозирования событий;
- анализировать кластерные числовые события.

Приложение Fabrix Data Visualization for Splunk предназначено для визуализации данных в области Индустриального Интернета Вещей (IIoT) и различных типов промышленного производства. Визуализация обеспечивает панорамное обозрение происходящих производственных и информационных процессов. Это позволяет повысить качество оперативного управления предприятием и решать возникающие проблемы в режиме реального времени.

Приложение позволяет:

- создавать интерактивные шкалы, счётчики, индикаторы, указатели, карты и графики;
- отображать как информацию об уже собранных данных, так и отображать значения величин, измеряемых в реальном времени;
- индивидуально настраивать рабочее пространство приложения, размещая необходимые объекты визуализации и создавая между ними связь.

## **Заключение**

В большинстве случаев анализ данных предполагает апостериорное исследование массивов данных для выявления различных типов закономерностей и явлений. Эта область по-прежнему остаётся перспективным полем для исследователей и в наше время. С другой стороны, в сложных конвергентных системах, которыми становятся телекоммуникационные сети, количество параметров и характеристик работы телекоммуникационных устройств, которые необходимо контролировать, чрезвычайно возросло, и будет увеличиваться и дальше [7, 8, 9, 10, 11]. В такой ситуации для обеспечения эффективного управления в условиях развития и распространения Интернета Вещей и сетей 5G необходимо использовать технологии анализа различных типов служебных данных в режиме реального времени с применением технологий визуализации. Это позволит без увеличения количества обслуживающих работников из числа технического персонала обеспечить требуемую надёжность сетей и приемлемое качество предоставления телекоммуникационных услуг.

## **Литература**

1. Кучерявый А. Е. Интернет Вещей // Электросвязь. 2013. № 1. С. 21–24.
2. Ястребова А. А., Выборнова А. И., Киричек Р. В. Обзор концепции тактильного интернета и технологий для его реализации // Информационные технологии и телекоммуникации. 2016. Т. 4. № 4. С. 89–96.

3. Koucheryavy A. State of Art and Research Challenges for USN Traffic Flow Models // 16th International Conference on Advanced Communication Technology (ICTACT). 2014. pp. 336–340.
4. Кучерявый А. Е., Маколкина М. А., Киричек Р. В. Приложения дополненной реальности в летающих сенсорных сетях // Связь в революциях. Революции в связи. 2017. С. 168–171.
5. Бородин А. С., Кучерявый А. Е. Сети связи пятого поколения как основа цифровой экономики // Электросвязь. 2017. № 5. С. 45–49.
6. Севостьянов В. Н. BIG Data как инновационный путь развития компании // Молодой ученый. 2014. № 8.2. С. 55–57.
7. Герасименко Е. М. Интеллектуальный анализ данных. Алгоритмы data mining. Ростов-на-Дону. 2017. 84 с.
8. Борисов Е. Г., Владыко А. Г., Парамонов А. И., Киричек Р. В. Самоорганизующиеся сети связи мультиагентных робототехнических систем // XIX Всероссийская научно-практическая конференция «Актуальные проблемы защиты и безопасности». 2016. С. 210–217.
9. Кучерявый А. Е., Владыко А. Г., Киричек Р. В., Маколкина М. А., Парамонов А. И., Выборнова А. И., Пирмагомедов Р. Я. Перспективы научных исследований в области сетей связи на 2017–2020 годы // Информационные технологии и телекоммуникации. 2016. Т. 4. № 3. С. 1–14.
10. Владыко А. Г., Киричек Р. В., Великоречин М. А., Думин Д. И. Комплексная методика тестирования фрагмента программно-конфигурируемой сети // Информационные технологии и телекоммуникации. 2015. № 2 (10). С. 20–29
11. Vybornova A., Koucheryavy A. Traffic Analysis in Target Tracking Ubiquitous Sensor Networks // Lecture Notes in Computer Science. 2014. Vol. 8638. pp. 389–398.

### References

1. Koucheryavy A. Internet of Things // *Electrosvyaz*'. 2013. No. 1. pp. 21–24.
2. Yastrebova A., Vybornova A., Kirichek R.: Review the Concept of the Tactile Internet and Technology for its Realization // *Telecom IT*. 2016. Vol. 4. Iss. 4. pp. 89–96. (in Russian)
3. Koucheryavy A. State of Art and Research Challenges for USN Traffic Flow Models // 16th International Conference on Advanced Communication Technology (ICTACT). 2014. pp. 336–340.
4. Koucheryavy A., Makolkina M., Kirichek R. Augmented Reality Applications in a Flying Sensor Networks // *Communication in revolutions. Revolutions in communication*. 2017. pp. 168–171.
5. Borodin A., Koucheryavy A. Fifth Generation Networks as a Base to the Digital Economy // *Electrosvyaz*'. 2017. No. 5. pp. 45–49.
6. Sevostyanov V. BIG Data as an Innovative Way of the Company's Development // *Molodoi uchenyi*. 2014. No. 8.2. pp. 55–57.
7. Gerasimenko E. Intelligent Data Analysis. Data Mining Algorithms. Rostov-na-Donu. 2017. 84 p.
8. Borisov E., Vladyko A., Paramonov A., Kirichek R. Self-Organizing Communication Network of Multi-Agent Robotic System // XIX All-Russian scientific practical conference "Actual Problems of Protection and Safety". 2016. pp. 210–217.
9. Koucheryavy A., Vladyko A., Kirichek R., Makolkina M., Paramonov A., Vybornova A., Pirmagomedov R.: The prospects for research in the field of communications networks on the 2017–2020 years // *Telecom IT*. 2016. Vol. 4. Iss. 3. pp. 1–14 (in Russian).
10. Vladyko A., Kirichek R., Velikorechin M., Dumin D. Benchmarking Methodology of Software-Defined Networks // *Telecom IT*. 2015. Vol. 2 (10). pp. 20–29 (in Russian).
11. Vybornova A., Koucheryavy A. Traffic Analysis in Target Tracking Ubiquitous Sensor Networks // *Lecture Notes in Computer Science*. 2014. Vol. 8638. pp. 389–398.

**Захаров Максим Валеович**

– аспирант, СПбГУТ, Санкт-Петербург, 193232,  
Российская Федерация, zaharov.spbgut@gmail.com

**Киричек Руслан Валентинович**

– кандидат технических наук, доцент, СПбГУТ,  
Санкт-Петербург, 193232, Российская Федерация,  
kirichek@sut.ru

***Zakharov Maxim***

– Postgraduate student, SPbSUT, St. Petersburg,  
193232, Russian Federation,  
zaharov.spbgut@gmail.com

***Kirichek Ruslan***

– Candidate of Engineering Sciences, Associate  
Professor, SPbSUT, St. Petersburg, 193232,  
Russian Federation, kirichek@sut.ru