

Технологии Big Data и различные методы представления больших данных

© П.Д. Иванов, А.Г. Лопуховский

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Представлено исследование компании Cisco о внедрении больших данных. Обоснована перспективность применения технологий Big Data. Описаны современные методы визуализации больших данных. Показаны конкурентные преимущества при использовании технологий Big Data.

Ключевые слова: Big Data, большие данные, прогнозирование, визуализация, конкурентное преимущество, ИТ-технологии.

В настоящее время компании создают огромные объемы данных в формате, плохо соответствующем традиционному структурированному формату баз данных, таких как веб-журналы, видеозаписи, текстовые документы, машинные коды или геопространственные данные. Все это хранится во множестве разнообразных хранилищ, зачастую за пределами организации. В результате корпорации могут иметь доступ к огромному объему своих данных и не иметь необходимых инструментов, чтобы установить взаимосвязи между этими данными и сделать на их основе значимые выводы. С учетом того, что данные сейчас обновляются все чаще и чаще, возникает ситуация, в которой традиционные методы анализа информации не могут «угнаться» за огромными объемами постоянно обновляемых данных, что в итоге и открывает дорогу технологиям больших данных.

В современных компаниях различные датчики, видеокамеры, интеллектуальные счетчики и другие подключенные устройства генерируют огромные объемы данных, которые добавляются к уже хранящейся информации. Настоящий предприниматель может разглядеть во всей этой «лавине» данных полезную информацию, однако исследование, проведенное недавно по заказу компании Cisco, показало, что ИТ-специалисты и компании пока с трудом извлекают пользу из поступающей информации [1].

В ходе исследования под названием «Cisco® Connected World Technology Report», проведенного в 18 странах независимой аналитической компанией Insight Express, были опрошены 1800 студентов колледжей и такое же количество молодых специалистов в возрасте от 18 до 30 лет. Опрос проводился для того, чтобы выяснить уровень готовности ИТ-отделов к реализации проектов Big Data, получить представление о связанных с этим проблемах, технологических изъяс-
нах, выявить стратегическую ценность таких проектов.

Большинство компаний собирает, записывает и анализирует данные. Тем не менее, согласно отчету, многие компании в связи с Big Data сталкиваются с целым рядом сложных деловых и информационно-технологических проблем. Например, 60 % опрошенных признают, что Big Data могут усовершенствовать процессы принятия решений и повысить конкурентоспособность на мировом рынке, но лишь 28 % заявили о том, что уже получают реальные стратегические преимущества от накопленной информации.

Big Data: большой потенциал, высокий приоритет. Технологии Big Data может предоставить конкурентные преимущества тем, кто найдет новые, оригинальные способы использования данных.

Назовем 5 стран, где специалисты в большей степени склонны полагать, что технологии Big Data станут их конкурентным преимуществом:

- Китай (90 %);
- Мексика (85 %);
- Индия (82 %);
- Бразилия (79 %);
- Аргентина (78 %).

Более двух третей опрошенных ИТ-руководителей считают, что в ближайшие 5 лет Big Data станет стратегическим приоритетом в их компаниях. Наибольшую уверенность в этом выразили респонденты из следующих стран:

- Аргентина (89 %);
- Китай (86 %);
- Индия (83 %);
- Мексика (78 %);
- Польша (78 %).

Что для этого необходимо? 38 % опрошенных заявили, что хотя в их компаниях уже установлены решения Big Data, для раскрытия всех преимуществ им нужен стратегический план [2].

Ключевые проблемы. По мнению опрошенных ИТ-руководителей, внедрению Big Data мешает ряд проблем, и прежде всего — обеспокоенность по поводу информационной безопасности, а затем — ограниченные бюджеты и дефицит специалистов.

Безопасность данных и управление рисками считают главной проблемой 27 % респондентов. По их мнению, трудности с защитой данных в проектах Big Data связаны с большими объемами информации, разными методами доступа к ней и с малыми бюджетами, выделяемыми на информационную безопасность. Более всех вопросами информационной безопасности обеспокоены респонденты из:

- Китая (45 %);
- Индии (41 %);
- США (36 %);
- Бразилии (33 %).

Нехватка бюджетных средств и времени на изучение технологии Big Data представляет собой главное препятствие с точки зрения трети опрошенных.

23 % респондентов главными препятствиями назвали дефицит ИТ-специалистов и нехватку опыта в области Big Data. Это особенно характерно для Японии (31 %) и Бразилии (30 %).

Проведенный компанией Cisco глобальный опрос показал: технология Big Data приносит конкурентные преимущества, вызывает рост трафика, увеличивает ИТ-бюджеты и помогает распространению облачных вычислений [3].

Анализируя данные, можно создавать новые сервисы и продукты, оптимизировать бизнес, повышать точность прогнозирования, следовательно, достигать нужного экономического эффекта. Технология Big Data не является чем-то принципиально новым. На современном этапе развития технологий стал меняться принцип подхода к анализу данных, вследствие того, что появились средства, позволяющие хранить, обрабатывать и анализировать большие объемы данных. Аналитики рассуждают следующим образом: «Мы не знаем, нужна ли нам информация, а если нужна, то какая, до тех пор, пока не проанализируем, насколько она взаимосвязана». Благодаря значительному снижению стоимости хранения информации появилась возможность собирать все больше данных и анализировать не связанные друг с другом факторы. Компьютер выявляет закономерности, которые не может обнаружить человеческий мозг, при этом выдавая совершенно неожиданные количественные взаимосвязи.

Технология Big Data предоставляет услуги, помогающие раскрыть коммерческий потенциал мегамассивов данных за счет поиска ценных закономерностей и фактов путем объединения и анализа больших объемов данных. Например, технологии Big Data могут связать цвет автомобиля с вероятностью его продаж в той или иной стране [4].

Консалтинговая компания Gartner выделяет 3+1 «V» Big Data (рис. 1):

1) Volume (объем): практически на всех предприятиях увеличиваются объемы данных. Создаются новые виды информации;

2) Variety (разнообразие): огромное разнообразие форматов и типов данных: фотографии, видео-, аудиозаписи, таблицы, документы, транзакции. Все это трудно хранить и анализировать;

3) Velocity (скорость): высокая скорость обработки и анализа данных;

4) Value (польза): хранение и обработка большого объема данных принесут большую пользу [5].

В настоящее время данные становятся важным фактором производства наряду с трудовыми ресурсами и капиталом. Все участники производственной цепочки опираются на технологии Big Data в своей работе. Теперь рассмотрим, как используют технологии Big Data промышленные предприятия.

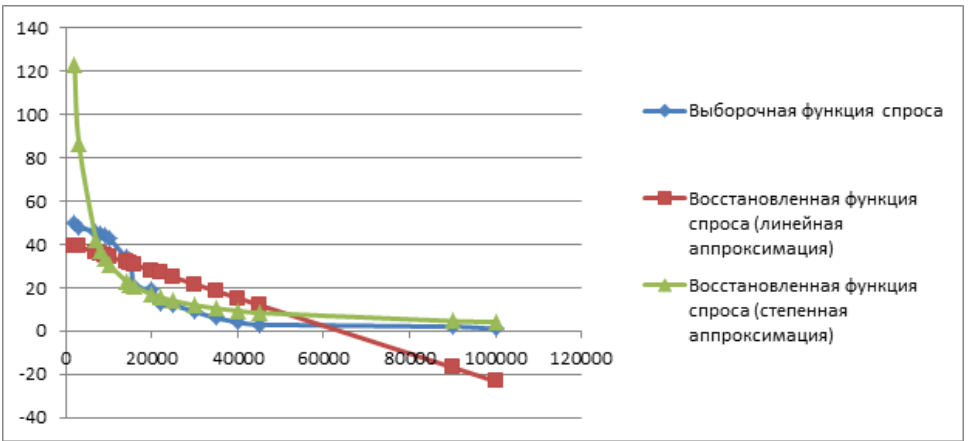


Рис. 3. График функций сброса

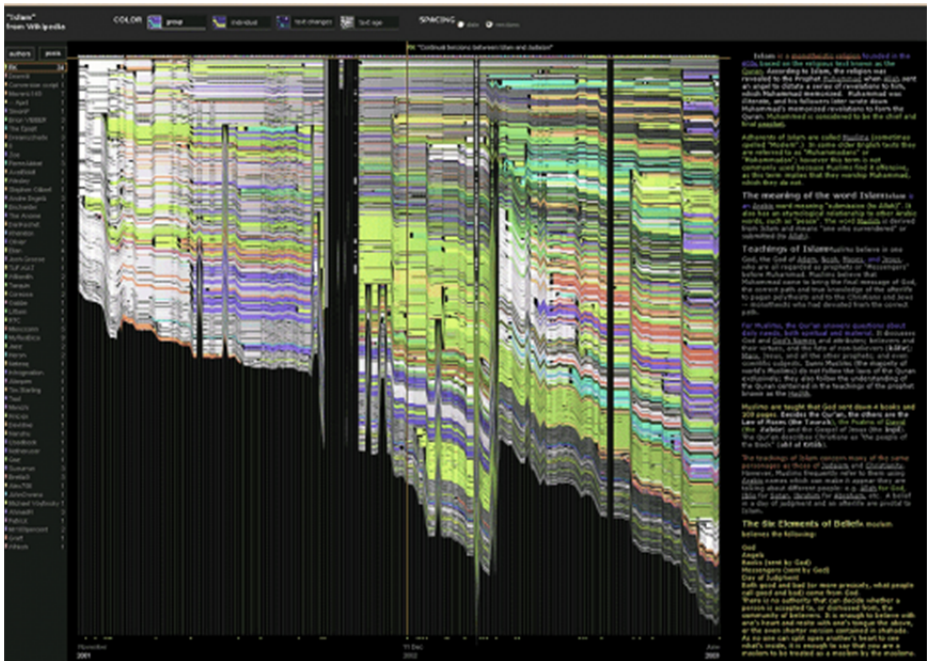


Рис. 4. Исторический поток

4. Пространственный поток (рис. 5). Эта технология предоставляет возможность пользователю следить за распределением и перемещением информации по всему миру. С помощью такого отображения данных можно выделить регионы, где данная информация наиболее востребована [9].

5. Семантическая сеть (рис. 6). Она представляет собой ориентированный граф, отображающий смысловые связи между объектами. Это средство семантического анализа, т. е. учитывающее смысловое сходство между объектами.

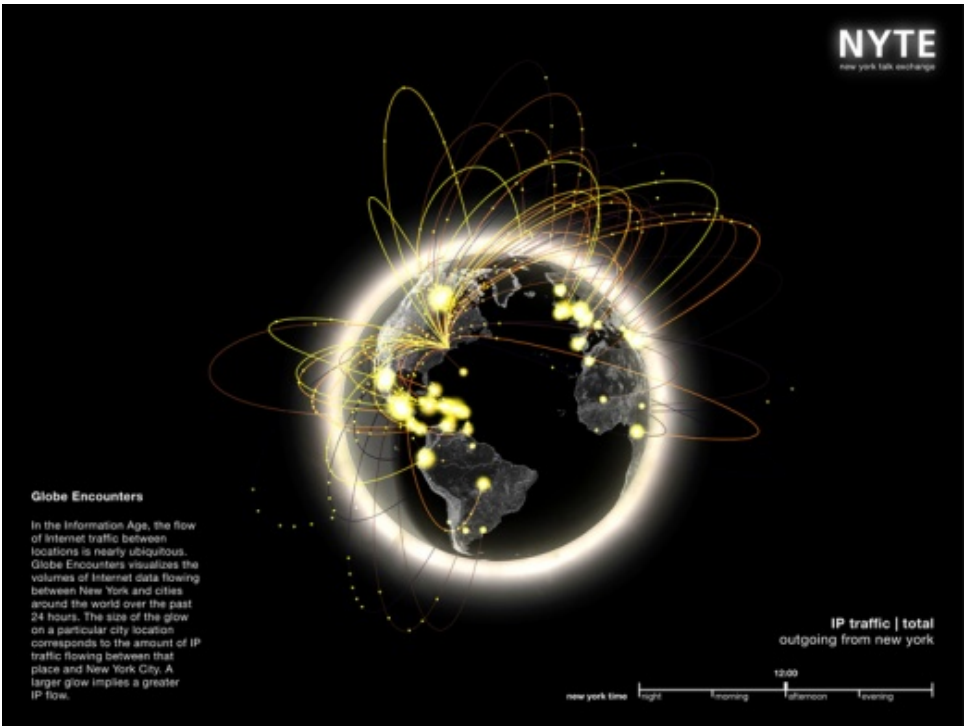


Рис. 5. Пространственный поток

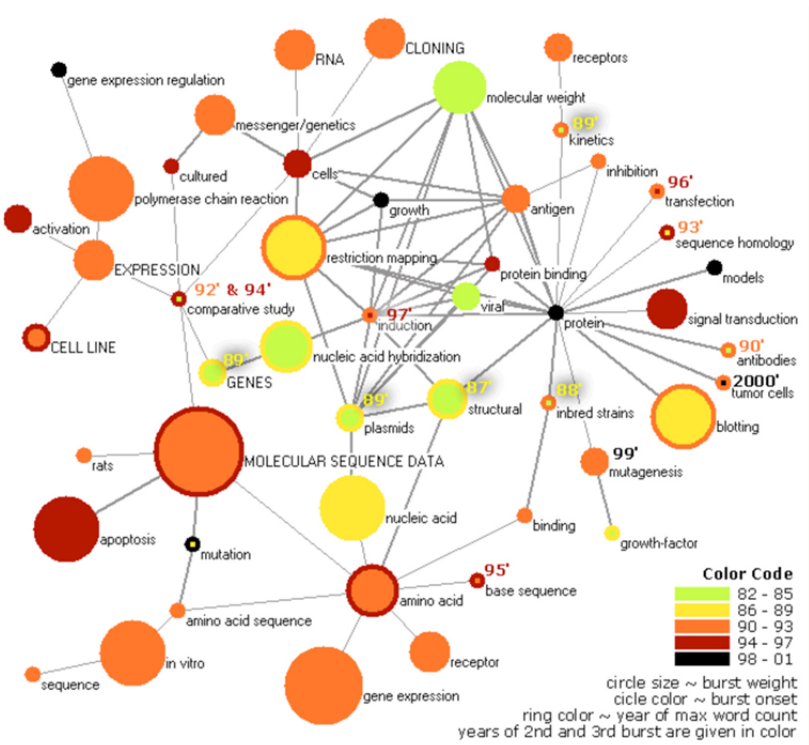


Рис. 6. Семантическая сеть

Выводы. В настоящее время предприятиям приходится работать с большими объемами информации, которая часто обновляется и приходит из разных источников. С помощью технологий Big Data предприятия могут анализировать огромные массивы данных и выявлять полезные закономерности, дающие им конкурентные преимущества.

Для более легкого восприятия и быстрого принятия управленческих решений необходимо представить результаты анализа данных визуально. На данный момент есть несколько видов представления массивов данных. Но существующие методы визуализации еще недостаточно развиты и требуют усовершенствования.

Компании, уже сейчас внедрившие технологии Big Data, в будущем получат большое конкурентное преимущество.

ЛИТЕРАТУРА

- [1] Большие данные помогут наращиванию ценности ИТ-департаментов. *Открытые системы*, 2013, № 4. URL: <http://www.osp.ru/news/2013/0403/13018290/> (дата обращения 16.07.2014)
- [2] Глобальное исследование Cisco выявило растущую роль сетей в центрах обработки данных. *Новости «Русского переплета»*, 2010, 10 января. URL: <http://www.pereplet.ru/news/index.cgi?id=25434> (дата обращения 16.07.2014)
- [3] Cisco, *Big Data: большой потенциал, высокий приоритет*. URL: <http://www.cisco.com/web/RU/news/releases/txt/2013/04/040113b.html> (дата обращения 16.07.2014)
- [4] Найдиц Андрей. *BigData: проблема, технология, рынок*. URL: <http://www.compress.ru/article.aspx?id=22725&iid=1044> (дата обращения 16.07.2014)
- [5] Pettey Christy. *Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data*. URL: <http://www.gartner.com/newsroom/id/1731916> (дата обращения 16.07.2014)
- [6] Зейников Александр. *Большие данные — неотъемлемая часть нашей жизни*. URL: <http://habrahabr.ru/post/197462/> (дата обращения 16.07.2014)
- [7] Rivera Janessa. *Gartner Survey Finds 42 Percent of IT Leaders Have Invested in Big Data or Plan to Do So Within a Year*. URL: <http://www.gartner.com/newsroom/id/2366515> (дата обращения 16.07.2014)
- [8] Артемов Сергей. *BigData: новые возможности для растущего бизнеса*. URL: <http://www.jet.msk.su/upload/iblock/eec/BigData.pdf> (дата обращения 16.07.2014)
- [9] TADVISER. *Большие данные (BigData)*. URL: [http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%91%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D0%B5_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5_\(Big_Data\)](http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%91%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D0%B5_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5_(Big_Data)) (дата обращения 16.07.2014)

Статья поступила в редакцию 28.08.2014

Ссылку на эту статью просим оформлять следующим образом:

Иванов П.Д., Лопуховский А.Г. Технологии Big Data и различные методы представления больших данных. *Инженерный журнал: наука и инновации*, 2014, вып. 9. URL: <http://engjournal.ru/catalog/it/asu/1229.html>

Иванов Павел Дмитриевич — аспирант, ассистент кафедры предпринимательства и внешнеэкономической деятельности МГТУ им. Н.Э. Баумана.

e-mail: ivanovpd@bmstu.ru

Лопуховский Антон Геннадьевич — студент кафедры предпринимательства и внешнеэкономической деятельности МГТУ им. Н.Э. Баумана.

e-mail: zond4@mail.ru

Big Data technologies and different methods of their presenting

© P.D. Ivanov, A.G. Lopukhovskiy

Bauman Moscow State Technical University, Moscow, 105005, Russia

Currently, companies are creating huge amounts of data, which are presented in a format consistent with the traditional poorly structured format databases such as weblogs, videos, text documents, computer codes or geospatial data. All of this is stored in a variety of different repositories, often outside the organization. As a result, corporations are able to have access to an enormous amount of their data and do not have the necessary tools to establish the relationship between these data and to make meaningful conclusions based on them. Given the fact that the data is now updated more and more often, there is a situation in which the traditional methods of data analysis can not keep up with the huge volume of constantly updated data, which ultimately paves the way for big data technologies. This paper presents a study on the implementation of Cisco big data. The usefulness of Technology Big Data is substantiated. Modern methods of visualizing Big Data are described. Competitive advantage by using Big Data technologies is shown.

Keywords: *Big Data, forecasting, visualization, competitive advantage, IT technology.*

Ivanov P.D., postgraduate, assistant lecturer of the Department of Entrepreneurship and Foreign Economic Activities of the Bauman Moscow State Technical University.
e-mail: ivanovpd@bmstu.ru

Lopukhovskiy A.G., a student of the Department of Entrepreneurship and Foreign Economic Activities of the Bauman Moscow State Technical University.
e-mail: zond4@mail.ru