

$$HNI = S_1^2 + S_2^2 + \dots + S_n^2.$$

где  $S_1, S_2, \dots, S_n$  – выраженные в процентах доли продаж фирм в отрасли, определяемые как отношение объема продаж фирмы к объему всех продаж отрасли.

В случае чистой монополии, когда отрасль состоит из одной фирмы, индекс Херфиндаля-Хиршмана будет равен 10000. Для двух фирм с равными долями  $HNI = 50^2 + 50^2 = 5000$ , для 100 фирм с долей в 1 %  $HNI = 100$ . Индекс Херфиндаля-Хиршмана реагирует на рыночную долю каждой фирмы в отрасли.

Таким образом, в результате расчетов с помощью *MS Excel* было получено значение индекса  $HNI = 611,2$ .

Коэффициент Херфиндаля-Хиршмана показывает, какое место, долю на данном рынке занимают продавцы, владеющие малыми долями. По значениям коэффициентов концентрации (*CR*) и индексов Херфиндаля-Хиршмана (*HNI*) выделяются три типа рынка:

- I тип – высококонцентрированные рынки: при  $70 \% < CR < 100 \%$ ;  $2000 < HNI < 10000$ ;

- II тип – умеренноконцентрированные рынки: при  $45 \% < CR < 70 \%$ ;  $1000 < HNI < 2000$ ;

- III тип – низкоконцентрированные рынки: при  $CR < 45 \%$ ;  $HNI < 1000$ .

Таким образом, полученное значение  $HNI < 1000$  свидетельствует о том, что рынок мясной продукции в Беларуси является низкоконцентрированным, т.е. с высоким уровнем конкуренции, и не требует антимонопольного вмешательства. Все это способствует честной конкурентной борьбе среди производителей.

<http://edoc.bseu.by>

*П.А. Плутов, В.И. Литвинец*  
(Беларусь, Минск)

## **ИНТЕГРАЦИЯ ТЕХНОЛОГИЙ BIG DATA И БИЗНЕС-АНАЛИЗА**

По оценкам IDC и McKinsey за период 2007-12 гг. отмечен рост цифровых данных в 20 раз, что составило  $> 50 \%$ /год, а прогнозы Oracle свидетельствуют, что к 2015 году мировой объем данных достигнет 7910 экзабайт [1, стр. 39]. Эксперты IDC сообщают о рекордном развитии технологий Big Data, рынок которых к 2016 году должен составить 23,8 млрд USD (рост на 31,7 %) [2].

В бизнес – аналитике особенно остро нуждаются крупные компании, работающие на высококонкурентных рынках. Им требуются технологии для обработки неструктурированных массивов данных (Big Data, BD) в области телекоммуникационных и финансовых услуг, а также в сфере межмашинного взаимодействия (M2M). Большинство корпораций используют аналитические службы на условиях аутсорсинга, что позволяет наращивать объем IT услуг.

*Big Data* – совокупность подходов, методов и IT-инструментов, которые консолидируют, управляют и анализируют данные, характеризующиеся 4 V:

*Volume* – объем данных выше границ вертикального масштабирования традиционных BI и рассматриваемых Big Data-технологий.

*Variety* – многообразие форматов представления данных удорожает интеграцию неструктурированной информации низкой плотности, например показания датчиков телеметрии, лог-файлы финансовых транзакций, поведение покупателей на веб-сайтах.

*Variability* – многовариантность интерпретации затрудняет анализ, время принятия решения невелико по сравнению со скоростью изменения данных.

*Velocity* – скорость поступления и необходимость быстрой обработки информации, а также скорость получения *воспринимаемых* результатов.

К методологии анализа *Big Data* следует относить: методы *Data mining*, кластерный и регрессионный анализ; *краудсорсинг*; прогнозную аналитику, интеграцию и сетевой анализ данных; распознавание образов и имитационное моделирование; пространственный и статистический анализ, генетические алгоритмы; машинное обучение и искусственные нейронные сети и т.д.

Технологии *Big Data*. В качестве базового принципа обработки больших данных принята архитектура *Shared Nothing (SN)*, обеспечивающая массивно-параллельную обработку, масштабируемую на сотни и тысячи узлов обработки. Сюда входит *Hadoop*, *MapReduce*, *R*, *NoSQL*. Применяемы также технологии BI и РСУБД с поддержкой языка SQL. Коротко о каждой составляющей:

*Hadoop* – проект фонда *Apache Software Foundation*, свободно распространяемый набор утилит, библиотек и программный каркас для распределенных программ, работающих на кластерах из сотен и тысяч узлов. Разработан в рамках вычислительной модели *MapReduce* и по состоянию на 2013 год *Hadoop* состоит из подпроектов – *Hadoop Common*, *Hadoop MapReduce*, *Hadoop YARN*. Используется для реализации поисковых и контекстных механизмов многих высоконагруженных веб-сайтов, в том числе, для *Yahoo!* и *Facebook*.

*MapReduce* – модель распределенных вычислений, представленная компанией *Google*, используемая для параллельных вычислений над очень большими, несколько петабайт, наборами данных в компьютерных кластерах.

*MapReduce* – фреймворк для вычисления набора распределенных задач с использованием большого количества компьютеров, образующих кластер.

Работа *MapReduce* состоит из двух шагов: *Map* и *Reduce*.

*R* – язык программирования для статистической обработки данных и работы с графикой, а также свободная программная среда вычислений с открытым исходным кодом в рамках проекта *GNU*.

*NoSQL* – ряд подходов, направленных на реализацию БД, имеющих существенные отличия от традиционных реляционных СУБД с доступом средствами языка SQL. Использование *NoSQL*-решений может осуществляться различными структурами данных: хеш-таблицы, деревья и др.

Основная цель подхода – расширить возможности БД там, где SQL недостаточно гибок, и не вытеснять его там, где он справляется со своими задачами. В основе *NoSQL* лежат тезисы: – нереляционная модель данных; открытый исходный код; высокая горизонтальная масштабируемость.

Также существуют большое количество не рассматриваемых здесь аппаратных решений, например – *Oracle Big Data appliance, Greenplum appliance*.

Целевая аудитория пользователей Big Data в экономике распределена по следующим направлениям:

- Интернет-проекты (*Google, Amazon, Twitter, Facebook*) для анализа, обработки и хранения данных различных web-ресурсов в целях улучшения поисковых ВІ-алгоритмов, выявления предпочтений, анализа и персонализации действий пользователей, маркетинга и т.д.

- Телекоммуникационный сектор (*AT&T; Verizon Communications*). Многофакторный анализ звонков клиентов.

- Ритейл (*Ikea, WalMart, Publix*). Интернет или традиционные бизнесы с большой нагрузкой на вебсайты с потребностью анализа лог-покупок.

- Банковский сектор (*HSBC, Bank of America Merrill Lynch, Commercial Bank of Qatar*). Основными пользователями являются инвестиционные банки, работающие в реальном масштабе времени с большими потоками информации различных торговых площадок.

Перспективы применения Big Data в РБ выявляют следующие проблемы белорусского рынка, препятствующие его использованию.

*Во-первых*, это уровень готовности белорусских компаний, потенциальных пользователей Big Data: телекоммуникационного сектора (МТС, *Velcom, Life*), ритейла (*Евроопт, Белмаркет, Гиппо*), банковской сферы (БПС-Сбербанк, Беларусбанк, Белагропромбанк). Уровень подготовки персонала и высокая стоимость решений не позволяет самостоятельно внедрять Big Data. Непрерывно изменяются данные и подходы к их анализу. Внедрение проекта, достижение высокой достоверности данных и аналитических моделей является требованием времени.

*Во-вторых*, отсутствие квалифицированных специалистов – аналитиков в РБ и в СНГ (*Data Mining Engineers, Manager Data Analytics, Data Scientists*).

Возможная реализация. Платформы *Oracle Big Data Appliance, EMC Greenplum* могут найти эффективное применение в потребительском секторе. Дополнительное использование индивидуальных карточек клиентов с RFID, позволит создавать потребительские профили. Это позволит анализировать CLV клиента, его пожизненную ценность для ритейлера. Грамотное использование Big Data позволит фокусироваться на наиболее прибыльных клиентах, их удержании, расширении корзины покупок, позволит сформировать индивидуальную программу лояльности.

Анализ потребительского профиля позволит создавать индивидуальные рекомендации, подсказки к продуктам, акции, магазины в сети, которые ему наиболее удобно посещать. Аналитика Big Data позволит отследить изменения в поведенческих моделях, и более точно на это реагировать. Используя большие объемы детальных данных о продажах и профиля конкретного покупателя позволит улучшить эффективность маркетинга. Уменьшается время на отслеживание реакции рынка, с момента запуска продукта в ритейлерскую сеть, а также другие накладные расходы (проведение опросов, интервьюирование покупателей и т.д.).

Моделирование в платформе *Greenplum* с помощью библиотеки *MADlib* позволяет быстро анализировать данные, проведение регрессии с целью определения самых важных переменных, а также их выборку. Правильное использование позволит повысить коэффициент связи потребительского спроса со множественными переменными. Скорость построения модели и выполнения скоринга в десятки раз выше чем у традиционных BI систем.

Очевидно, что та компания, которая сможет внедрить и использовать технологии Big Data, принимать решения, а также получит конкурентные преимущества и позволит оказывать более качественные услуги клиентам.

#### Литература

1. Explosion in big data causing data centre crunch. // Information Age – March/April 2012. URL: <http://goo.gl/JW6QL> (дата обращения 20.01.2013)
2. IDC-Press Release, FRAMINGHAM, Mass. January 8, 2013. // IDC Corporate portal. URL: <http://goo.gl/mZy16> (дата обращения 20.01.2013)

*И.И. Станкевич*  
(Беларусь, Минск)

### **АВТОМАТИЗАЦИЯ СОСТАВЛЕНИЕ БИЗНЕС-ПЛАНА РАЗВИТИЯ ОРГАНИЗАЦИИ С ПОМОЩЬЮ ПК «НИВА-СХП: БИЗНЕС-ПЛАН»**

Бизнес-планы, как правило, составляются в следующих случаях: для получения кредита, при получении государственного финансирования, для привлечения инвестора, для текущего планирования, для обоснования инвестиционных проектов.

В нашем случае, бизнес-план – это подробный план деятельности сельскохозяйственной организации на определенный период, устанавливающий показатели, которые необходимо достичь предприятию, с учетом собственных возможностей, а также при необходимости привлечения сторонних организаций.

Алгоритм составления бизнес-плана коммерческих организаций должен учитывать современные тенденции в развитии национальной и мировой экономике:

- рост производительности и стоимости ресурсов, вовлекаемых в процесс производства продукции с целью определения более точного их расхода на единицу продукции, оптимального их соотношения, повышения репрезентативности расчетных данных;

- потенциальную возможность альтернативного использования инвестиционных, производственных, финансовых ресурсов в процессе производства и реализации продукции, предполагающую оптимальное обоснование значений всех плановых параметров бизнес-плана коммерческой организации в процессе его практической реализации;