

## Программное обеспечение для обработки и анализа данных

Разорёнова Т.Р. Разорёнов Н.А.

Белорусский национальный технический университет

Интерес к инструментам сбора, обработки, управления и анализа больших данных (Big Data) проявляют едва ли не все ведущие ИТ-компании.

Объектом исследования являются большие данные, которые накапливаются при осуществлении деятельности различных организаций, а на основании их анализа принимаются управленческие решения. Предметом исследования являются компьютерные технологии, позволяющие качественно и своевременно обрабатывать накопленную информацию.

Программное обеспечение Business Intelligence (BI) и Data Mining как раз и призваны собирать такого рода информацию, аккумулировать и поставлять ее туда, где эта информация сейчас необходима.

Современные системы BI состоят из компонент:

- 1) корпоративные хранилища данных (ХД), консолидирующие в себе всю существенную информацию о работе организации, выверенную, очищенную, подготовленную для последующего удобного использования;
- 2) средства ETL (Extraction, Transformation, Loading), обеспечивающие извлечение данных из многочисленных систем-«источников», преобразование, очистку и загрузку этих данных в ХД;
- 3) средства доступа к данным и их визуализации (это те инструменты, с которыми непосредственно работают конечные пользователи).

Кроме этих основных компонент, важными составляющими BI-систем являются: модели данных ХД, (оптимизированные типовые структуры базы данных ХД для той или иной отрасли бизнеса); средства управления метаданными BI-системы; сервера OLAP-анализа и др.

Существует множество разнообразных методик анализа массивов данных, в основе которых лежит инструментарий, заимствованный из статистики и информатики. В отличие от других технологий поддержки принятия решений Data Mining обладает более высокой степенью интеллектуальности, позволяет автоматизировать анализ данных и работать в любых масштабах. Использование больших данных ведет к существенному улучшению процессов принятия решений, положительно отражается на конкурентоспособности компаний и упрощает управление рисками. К прогрессивным методам визуализации анализа больших данных относятся облако тегов, кластерграммы, исторический и пространственный потоки.

Технологии Big Data успешно реализуются в различных индустриях: энергетика, медицина, банки, страхование, маркетинг, ЖКХ и др.

УДК 681.513.5

## **Работа с протоколом HTTP на языке PHP**

Русак Ю.П., Новиков С.О.

Белорусский национальный технический университет

При создании web-приложений на языке PHP зачастую требуется работать напрямую с удаленным сервером по протоколу HTTP для получения и отправки данных.

HTTP (HyperText Transfer Protocol, протокол передачи гипертекста) – это протокол прикладного уровня; для отправки HTTP запросов и обработки ответов язык PHP предусматривает как встроенные функции, так и более низкоуровневый подход с генерацией HTTP заголовков и отправкой их через tcp-сокеты.

Протокол HTTP реализует принцип запрос/ответ. Запрашивающая программа–клиент инициирует взаимодействие с отвечающей программой–сервером и посылает запрос, содержащий заголовки и непосредственно тело запроса.

С помощью сокетов на языке PHP можно реализовать как клиент, так и сервер HTTP. Реализация GET-запроса выглядит следующим образом:

```
$sock = socket_create(AF_INET, SOCK_STREAM, SOL_TCP) or  
exit("Unable to create socket");  
socket_connect($sock, 'example.com', 80);  
socket_write($sock, "GET / HTTP/1.1\r\n");  
socket_write($sock, "Host: example.com\r\n");  
socket_write($sock, "Connection: Close\r\n\r\n");  
$buff = socket_read($sock, 1024);  
socket_close($sock);
```

Здесь создается TCP-сокеты, открывается соединение с сервером на порту 80 и отсылается запрос. HTTP-заголовки при отправке разделяются переводами каретки. После отправки запроса происходит считывание ответа в буфер размером 1024 байт, размер буфера устанавливается произвольно. Затем сокет закрывается и соединение с сервером разрывается.

С помощью сокетов можно реализовать GET, POST, HEAD запросы, а также другие более редкие запросы как PUT и DELETE. Также, задав сокету режим прослушивания входящих соединений, мы можем принимать входящие подключения от удаленных клиентов и реализовать HTTP-сервер.