УДК 004.896

Ю. Х. Хамуков, Е. А. Ищукова, Л. З.-Г. Шауцукова

Проблема обеспечения информационной безопасности в условиях возрастания давления обволакивающего интеллекта

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 12–07–00744 а, № 13–07–01002 а

Представлен краткий обзор ситуации, сложившейся в проблеме обеспечения информационной безопасности вследствие быстрого внедрения технологий Big Data. На основе обзора делается вывод о бесперспективности традиционных технологий обеспечения информационной безопасности в новых условиях.

Ключевые слова: информационная безопасность, большие данные, экзабайты, аналитическая обработка, обволакивающий интеллект.

Ju. Kh. Khamukov, E. A.Ishchukova, L. Z.-G. Shautsukova

Problem of information security support in conditions of increasing pressure of coating intelligence

A brief review of the situation about the problem of information security support caused by fast introduction of the Big Data technologies is presented. Basing on the review we made a conclusion about the futility of traditional information security technologies in a new environment.

Keywords: information security, Big Data, exabyte, analytical processing, coating intelligence

Исследование проблем обеспечения информационной безопасности обретает содержательность при четком разделении понятий информация, информационный ресурс, данные и знания. До настоящего времени вопрос различия и взаимосвязи понятий «данные» и «информация» весьма дискуссионный [5]. В одних работах предлагаются существенно различные определения этих терминов, а в других они используются как синонимы. В частности, редко корректно используются термины «качество данных (quality data)» и «качество информации (quality of the information)»; управление (management) данными – управление информацией и т. п.

Профессор Рэй Р. Ларсон (Ray R. Larson) из Школы информационных технологий Калифорнийского Университета в Беркли [2] предлагает схему иерархического разделения понятий «данные», «информация», «знания» и «мудрость»:

- данные (data) – сырой, необработанный информационный материал;
- информация (Information) – структурированные и описанные данные;
- знания (Knowledge) – информация, которая кем-то воспринята, прочитана, услышана, или увидена и понята;
- мудрость (Wisdom) – единая совокупность понятых знаний.

Детально расписаны способы и методы обеспечения информационной безопасности в вузовском учебнике по защите информации [7].

Реализация представленных в учебнике требований, а также, к примеру, положений Доктрины национальной безопасности Российской Федерации, потребует оперирования сущностями понятий «данные», «информация», «знания». Эти сущности неотделимы от физических процессов в системах, которые генерируют данные, от процессов в гетерогенных коммуникационных системах и хранилищах информации. А в функционировании этих систем за последние два–три года произошли события, радикально меняющие тренды и перспективы развития информационных технологий в целом, и проблем обеспечения информационной безопасности в частности. Прежде всего, произошло то, что бывшее несколько десятилетий дежурной страшилкой околонаучной журналистики понятие «информационный взрыв», уже давно никого не пугавшее, вдруг приобрело качественно иное содержание. У этого феномена три основные составляющие – неожиданно быстрый рост генерации данных, переход информационных систем с «обработки данных» на анализ информации в приложениях, и, обу-

словленное вторым фактором, исключение человека из выработки решений по результатам анализа. Для нас важным является качественное изменение поведения такого феномена информатизации, как обволакивающий интеллект.

С 3 сентября 2008 г., после выхода специального номера журнала «Nature», посвященного вопросам влияния новых технологий работы с большими объемами данных на будущее науки, укрепилось новое понятие Большие Данные. Его придумал редактор номера Клиффорд Линч по аналогии с отражавшими новые парадигмы развития общества метафорами Большая Руда, Большая Нефть, Большая Химия и т. п. В последующие два года стало ясно, что новый термин отражает не только проблемы, связанные с ростом объемов и многообразием научных данных в академической среде, для которой термин предназначался. В 2009 г. им стала широко пользоваться деловая пресса, в 2010 появились первые специальные программные и технические продукты, предназначенные исключительно для обработки больших объемов данных, с 2011 года крупнейшие поставщики информационных технологий IBM, Oracle, Microsoft, Hewlett-Packard, EMC ориентируют стратегии развития на понятие о больших данных. В том же 2011 году специализирующаяся на рынках информационных технологий исследовательская и консалтинговая компания Gartner поставила Большие Данные на второе (после виртуализации!) место [3] в мировой информационно-технологической инфраструктуре впереди энергосбережения и мониторинга.

Наиболее существенно в этом процессе то, что происходит интеграция приложений. Решение задач обмена данными между приложениями приводит к исключению человека из технологий работы с информацией, и корпоративные системы превращаются в технические (вернее – телекоммуникационные) системы управления. Человек реально оказывается «включенным» в быстро развивающиеся отношения между приложениями ПО информационных систем. Наибольшие изменения произойдут, по прогнозам специалистов, в промышленном производстве, здравоохранении, торговле, госуправлении, и, естественно, в финансовой сфере. Сообразно развитию событий в хозяйственной сфере жизнедеятельности уже с этого года Большие Данные вошли как академический предмет в вузовские программы по науке о данных [1] и вычислительным наукам и инженерии [4].

В новое понятие попадают давно известные продукты. Специалисты Gartner еще в 2011 г. ввели в отчете Hype Cycle, содержащем анализ состояния и перспектив новых технологий, новую позицию Big Data and Extreme Information Processing and Management с оценкой перспективы массового внедрения до 2015–2016 гг.

На протяжении предыдущих 65 лет истории компьютеров так и не было понято, что же такое данные, что с ними происходит в результате обработки и как они связаны с информацией. Эти основополагающие понятия были предметом только интуитивного восприятия. При этом невероятными темпами развивались собственно технологии работы с данными. А кибернетика и теория информации остались на уровне 50-х гг., когда только осваивались технологии расчетов на ламповых компьютерах. А теперь Облака, Большие Данные и Аналитика определяют характер и роль в жизни человечества современных IT. Пришло осознание того, что то, что мы привычно называем теорией информации по Клоду Шеннону, на самом деле является не более чем статистической теорией передачи сигналов. А информация, необходимая и воспринимаемая человеком – что-то совсем иное. При этом важнейшим условием развития стали результаты аналитических исследований, что, в свою очередь, резко подняло требования к масштабированию систем хранения данных. Только масштабирование – горизонтальное и вертикальное – позволяет одновременно сильно загружать системы для выполнения аналитических задач, и, при этом, обеспечить постоянную доступность всех сервисов, приложений и самих данных.

Редактор журнала «Web 2.0 Journal» Дайон Хинчклиф [6] предпринял попытку классификации Больших Данных сообразно с результатом, который ждут от их обработки.

Согласно предложению Хинчклифа, Big Data можно поделить на три группы: Быстрые Данные (Fast Data) с характерным объемом в терабайтах; Большая Аналитика (Big Analytics) с петабайтными данными, и Глубокое Проникновение (Deep Insight) с экзабайтами и зеттабайтами данных. Но главное различие между группами – в качестве результатов обработки данных.

Новые проблемы обеспечения информационной безопасности в создавшихся условиях отразились в СУБД, переданной Агентством национальной безопасности (АНБ) США фонду Apache Software Foundation. СУБД назвали Accumulo. Она основана на системе доступа на метках. Идеология системы отражена в пояснительной записке к проекту: «Необходимо создать гибкую, высокопроизводитель-

ную, распределенную систему хранения пар ключ-данные, с эффективно выражающими смысл метками, обеспечивающими детальный контроль доступа к данным. За последние три года мы добились немалых успехов в работе над этим проектом и полагаем, что общедоступность и открытая разработка пойдут на пользу как самому проекту, так и всем, кто в нем заинтересован».

В отличие от аналогичных продуктов, созданных по тому же принципу, что и Google BigTable, и способных, благодаря распределенной архитектуре, работать в нескольких серверах одновременно, Accumulo позволяет присвоить каждой ячейке данных свою уникальную метку. В результате становится возможным обеспечить детальный контроль доступа к данным, в виде, например, предоставления внешнему серверу доступа только к определенным ячейкам хранилища, метки на которых снабжены своими метками на основе набора правил. По мнению разработчиков, такие системы хранения могли бы стать основой безопасных хранилищ данных для организаций со строгими требованиями к безопасности и защите личных данных – учреждений здравоохранения и правительственных органов.

Наряду с ранее известным подходом к проблеме обеспечения безопасности на основе меток – Security Enhanced Linux (SE Linux) – реализованный специалистами АНБ подход к безопасности на основе меток дает возможность администраторам систем создавать наборы правил, детально описывающих действия, которые может выполнять каждая программа.

Учитывая, что над Accumulo трудятся сотни специалистов, нельзя обойти вниманием эффект или, скорее, «феномен Сноудена». Во всяком случае, независимо от того, благо для общества привнес своими действиями Сноуден, или, наоборот, открыл некий «ящик Пандоры» с непредсказуемыми угрозами и опасностями, проблема обеспечения информационной безопасности приняла качественно иной характер. Первым делом обществу были представлены реальные масштабы контроля жизнедеятельности индивидуума и новые возможности, предоставляемые технологиями обработки метаданных – данных, описывающих другие данные. Стало возможным манипулировать с фрагментами перехватываемой информации (данных), не раскрывая их содержания, а только оперируя с меткой, присвоенной данному фрагменту. Обойдены юридические ограничения на работу с персональными данными.

Следующее действие Сноудена – раскрытие службы SIGAD US-984XN, больше известной по кодовому имени PRISM, и изымавшей у девяти ведущих интернет компаний все пересылаемые цифровые фотографии, файлы, электронные письма и видеопереговоры. Причем, по обнародованным Сноуденом схемам, АНБ использует средства слежки, работающие в реальном времени. И, наконец, в июле (2013 г.) Сноуден обнародовал сверхсекретный доклад с описанием программы для просмотра содержимого сотен баз данных на 700 серверах, разбросанных по всему миру.

Таким образом, прежние представления об информационной безопасности как обеспечения целостности, конфиденциальности, сохранности и т. п. в эпоху Больших Данных теряют содержание. Big Data в большей степени, чем прежние феномены информатизации, олицетворяют сущность начавшейся технологической эпохи. Необходимость в аналитической работе с большими данными заметно изменит лицо ИТ-индустрии и стимулирует появление принципиально новых программных и аппаратных платформ. Учитывая, что IT-технологии переходят в область фундаментальных процессов, описываемых в категориях статистической физики и термодинамики, при разработке способов обеспечения информационной безопасности, неотделимой от безопасности жизнедеятельности вообще, следует пользоваться фундаментальными характеристиками физических систем – энтропия, параметр порядка системы и т. п.

Библиографический список

1. Магистратура в области обработки Больших Данных [Электронный ресурс]. – Режим доступа: http://iacs.seas.harvard.edu/master-of-science-in-cse.

2. Мельников, В. П. Информационная безопасность и защита информации [Текст]/ Мельников В. П., Клейменов С. А., Петраков А. М. – 3-е изд., стер. – М. : Академия, 2008. – 336 с.

3. Методики работ с данными [Электронный ресурс]. – Режим доступа: http://www.computing.dundee.ac.uk/study/postgrad/degreedetails.asp?17.

4. Проблема Больших данных и новые методы счета [Электронный ресурс]. – Режим доступа: http://www.osp.ru/nets/2011/04/13010795/.

5. Технологии управления данными [Электронный ресурс] / Jim Harris. Data, Information, and Knowledge Management, Information Management, 2011. – Режим доступа: http://www.tsi.lv/ResTech/2011/vol6_2/section4.pdf.

6. Хамуков, Ю. Х. Феноменологическое описание проблемы обеспечения информационной безопасности на основе семантизации сообщений и принципа обволакивающего интеллекта [Текст] / Ю. Х. Хамуков // Международный конгресс по интеллектуальным системам и информационным технологиям IS&IT'13. – 2–9 сентября. – Россия, Краснодарский край, пос. Дивноморское.

7. Черняк, Леонид. Большие Данные – новая теория и практика [Текст] / Леонид Черняк // Открытые системы. СУБД. – М. : Открытые системы, 2011. – № 10.

Bibliograficheskij spisok

1. Tehnologii upravlenija dannymi [Jelektronnyj resurs] / Jim Harris. Data, Information, and Knowledge Management, Information Management, 2011. – Rezhim dostupa: http://www.tsi.lv/ResTech/2011/vol6_2/section4.pdf.

2. Mel'nikov, V. P. Informacionnaja bezopasnost' i zashhita informacii [Tekst]/ Mel'nikov V. P., Klej-menov S. A., Petrakov A. M. – 3-e izd., ster. – M. : Akademija, 2008. – 336 s.

3. Chernjak, Leonid. Bol'shie Dannye – novaja teorija i praktika [Tekst] / Leonid Chernjak // Otkrytye siste-my. SUBD. – M. : Otkrytye sistemy, 2011. – № 10.

4. Metodiki rabot s dannymi [Jelektronnyj resurs]. – Rezhim dostupa: http://www.computing.dundee.ac.uk/study/postgrad/degreedetails.asp?17.

5. Magistratura v oblasti obrabotki Bol'shih Dannyh [Jelektronnyj resurs]. – Rezhim dostupa: http://iacs.seas.harvard.edu/master-of-science-in-cse.

6. Problema Bol'shih dannyh i novye metody scheta [Jelektronnyj resurs]. – Rezhim dostupa: http://www.osp.ru/nets/2011/04/13010795/.

7. Hamukov, Ju. H. Fenomenologicheskoe opisanie problemy obespechenija informacionnoj bezopasnosti na osnove semantizacii soobshhenij i principa obvolakivajushhego intellekta [Tekst] / Ju. H. Hamukov // Mezhdunarodnyj kongress po intellektual'nym sistemam i informacionnym tehnologijam IS&IT'13. – 2–9 sentjabrja. – Rossija, Krasnodarskij kraj, pos. Divnomorskoe.