

Böyük Verilənlərdən Faydalı İnformasiyanın Aşkarlanmasında Mövcud Yanaşmalar

Rəna Qasımova

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
depart1@iit.ab.az

Xülasə — Yüzlərlə terabayt və ekzabayt həcmində böyük verilənlərin saxlanması, real-vaxt ərzində analitik emal, onlardan faydalı informasiyanın əldə edilməsi ciddi problemlərin yaranmasına səbəb olmuşdur. Məqalə böyük verilənlərin emal problemlərinə həsr olunur və böyük verilənlərdən faydalı informasiyanın aşkarlanmasında mövcud yanaşmalar analiz olunur. Eyni zamanda böyük verilənlərin əsas xarakteristikaları və xüsusiyyətləri tədqiq olunur.

Açar sözlər — *verilənlər xəzinəsi, bulud, strukturlaşdırılmamış informasiya, big data, NoSQL, MapReduce, Hadoop.*

I. GİRİŞ

Son zamanlar böyük verilənlər (Big Data) ideyası aktuallıq kəsb etmiş, bu sahəyə olan maraq daha da artmışdır. Big Data termini böyük həcmli verilənlərlə işləyən bütün sferalarda, iqtisadiyyatda, bank fəaliyyətində, istehsalatda, marketinqdə, telekommunikasiyada, veb analitikada, tibbdə və s. sferalarda istifadə edilir [1-3]. Hazırda IBM, Oracle, Microsoft, SAS, SAP, HP, Teradata, EMC və s. kimi nəhəng İT informasiya texnologiyaları istehsalçıları tərəfindən böyük vüerilənlərin (BV) emalını təmin edən startaplar, mürəkkəb informasiya sistemləri yaradılır [4-8]. Bütün aparıcı sahə analitikləri rəqabət qabiliyyətini təmin etmək üçün BV ilə iş bacarığının nə qədər vacib olmasına diqqətli cəlb etməyə çalışırlar. Bu bir daha böyük həcm, sürət və mürəkkəblilik kimi xarakteristikalarla təyin olunan BV sahəsindəki problemlərin nə qədər aktual olmasından xəbər verir. BV-nin saxlanması, idarə edilməsi, analitik emalı, təhlükəsizliyi və s. məsələlərinin həlli yeni texnologiyalar, idarəetmə metodları və daha mükəmməl intellektual analiz vasitələrini tələb edir.

Bununla belə, BV problemi hələ ilkin araşdırmalar səviyyəsindədir, yəni bu sahə hələ də tam olaraq təhlil olunmayıb. Aparılan bütün tədqiqatlar “Big Data” anlayışını, mahiyyətini, onun müxtəlif xarakteristikalarını təsnifləndirməyi, BV-nin mənbələrini, bu texnologiyanın imkanlarını, problemlərini, təhlükəsizlik məsələlərini hərtəfli tədqiq etməyə imkan verir. Eyni zamanda tədqiqatlar göstərir ki, BV-in emalı və analizi üçün mükəmməl analitik texnologiyalar və alətlər tələb olunur [9-11].

II. BİG DATA XÜSUSİYYƏTLƏRİ

Bu gün Big Data İT sahəsindəki nəhəng vendorların və beynəlxalq analitik agentlərinin istifadə etdikləri anlayışdır.

Qeyd edək ki, Big Data texnologiyasını və bu texnologiyanı digərlərindən fərqləndirən xüsusiyyətlərini araşdırarkən müxtəlif fikirlərə rast gəlmək olur. Mütəxəssislərin fikrincə Big Data texnologiyalar çoxluğudur və aşağıdakı xüsusiyyətlərə malikdir [12, 15]:

– **Çox böyük həcm.** Hazırda BV-nin miqyası terabaytlardan zətabaytlara qədər həcm ilə xarakterizə olunur. Bu gün müəssisə səviyyəsində verilənlərin belə həcmi artıq heç kəsi təəccübləndirmir.

– **Çoxlu sayda informasiya mənbələri.** Məlumdur ki, biznes analitika (Business intelligence, BI) emal olunmayan informasiyanın anlamlı, rahat formaya çevirmək üçün lazım olan metod və alətlərdir. Ənənəvi BI-də bir neçə informasiya mənbələri var, Big Data isə onlarla, minlərlə xarici mənbələrə malikdir.

– **Strukturlaşdırılmamış informasiya.** Ənənəvi verilənlər xəzinəsi (VX) relyasion baza əsasında qurulmuşdur və strukturlaşdırılmış informasiyanın axtarışı üçün nəzərdə tutulmuş alətdir. Adi relyasion bazalarının emal edə bilmədiyi yüzlərlə terabayt və ekzabayt həcmində müxtəlif tip strukturlaşdırılmamış informasiyanın toplanması və saxlanması, axtarışı, təhlükəsizliyi, analizi kimi məsələlərin həlli üçün yeni alətin yaradılmasına ehtiyac yaranır. Big Data mənbələrinin əsas hissəsini də strukturlaşdırılmamış və qismən strukturlaşdırılmış verilənlər təşkil edir. Bu halda Big Data yanaşması prinsiplial olaraq daha mükəmməl həll olan şablonlar əsasında axtarışı təklif edir, bu isə öz növbəsində relyasion verilənlər bazasından fərqli olan informasiyanın saxlanması strukturunu nəzərdə tutur.

– **Dinamik dəyişən informasiya.** Böyük verilənlər çox böyük informasiya massivlərinin toplanması ilə xarakterizə olunur ki, bu da hazırda istənilən təşkilatın qarşılaşdığı problemdir. Böyük həcm problemi ilk növbədə saxlama problemi yaradır ki, bu da genişmiqyaslı saxlama və paylanmış emal tələb edir. Eyni zamanda bütün verilənləri, hətta filtrlənmiş verilənləri də saxlamaq getdikcə daha baha başa gəlir. Məlumdur ki, saxlama texnologiyaları bahadır və onların uçuzlaşması yeni verilənlər mənbəyinin yaranmasına nisbətən daha yavaş baş verir. Bununla bağlı təşkilat səviyyəsində bu və ya digər verilənlərin hansı vaxta qədər saxlanması dəqiq müəyyən olunmalıdır. Məsələn, bəzi verilənlər təşkilatda uzun illər ərzində tələb oluna bilər, digərləri isə artıq bir neçə saatdan sonra, analitiklər onlardan lazım olanı götürdükdən sonra faydasız ola bilər.

Son zamanlar isə verilənlərin emalı və yadda saxlanılmasına xidmət edən “grid” və “cloud computing” texnologiyalarının tətbiqi saxlama sahəsindəki problemləri demək olar ki, aradan qaldıra bilmişdir. Beləliklə, Big Data BI-nin ənənəvi alətləri ilə həll olunan oxşar məsələləri həcm, mənbə, struktur və paylanma mərhələsində daha geniş kontekstdə həll edir. Nəticədə Big Data və BI texnoloji səviyyədə bir-birindən ciddi fərqlənirlər (cədvəl 1.) [16-19].

CƏDVƏL 1. Big Data və biznes analitika arasındakı fərqli xüsusiyyətlər

Ənənəvi BI	Big Data
Vahid korporativ VX	Xəzinə paylanmış fayl sistemində yerləşə bilər
Verilənlərin formatı emal funksiyalarının tələbinə uyğunlaşdırılır	Emal funksiyaları müxtəlif verilənlərin formatına uyğunlaşdırılır
Verilənlərin formatı strukturlaşdırılmışdır	Həm strukturlaşdırılmış, həm də strukturlaşdırılmamış informasiya ilə işləmək nəzərdə tutulur
Tarixi informasiya	Ən yeni verilənlərlə iş
Verilənlərin əldə edilməsi, ötürülməsi və emalı əsasən ardıcıl, dəqiq müəyyən olunmuş prosedurlara uyğun baş verir	Verilənlərin kütləvi paralel emalı tətbiq olunur (Massively Parallel Processing, MPP)

III. BIG DATA EMALININ TEXNOLOJİ MƏRHƏLƏLƏRİ

Hazırda aparılan elm-tədqiqat işlərinin əksəriyyəti BV-nin texnoloji problemlərinin həllinə yönəlmişdir. Yüzlərlə terabayt və ekzabayt həcmində BV-nin saxlanması, real-vaxt ərzində analitik emalı, onlardan faydalı informasiyanın əldə edilməsi ciddi problemlərin yaranmasına səbəb olmuşdur. Big Data-nın əsas mənbələrini sensor və sosial şəbəkələr, müxtəlif sahələr haqqında məlumat ötürücüləri, bank əməliyyatları, coğrafi informasiya sistemləri (GIS), Qlobal Mövqəyəyinetmə Sisteminin siqnailləri (GPS), elmi eksperimentlər, elektron poçt, smartfonlar vasitəsilə alınmış rəqəmsal foto və videolar, böyük şirkətlərin bazarları, böyük satış mərkəzləri, domen adları serverləri (Domain Name Server, DNS) və s. təşkil edir. Bu mənbələr tərəfindən yaradılan böyük həcmli verilənlərin məhsuldar istifadəsi üçün uyğun texnologiyalar lazımdır. Bunlar verilənlərin toplanması, saxlanması və analitik emalı sistemlərindən ibarətdir. Onun müxtəlif hissələrində aşağıdakı texnologiya qruplarından istifadə edilə bilər [20]:

- analitik alqoritmlər;
- paralel proqramlaşdırma metodları;
- bulud hesablamaya resursları;
- fərdi kompüterdən strateji təyinatlı superkompüterlərə qədər hesablamaya sistemləri;
- saxlama sistemləri;
- şəbəkələr;
- mürəkkəb teleskoplardan və tomoqraflardan sadə radiotezlikli identifikasiya (Radio Frequency Identification, RFID) texnologiyalarına qədər müxtəlif növ daxiletmə qurğuları və s.

Sadalanan siyahıda səviyyələr yüksəldikcə verilənlərin biliklərə yaxınlığı da artır. Analitika verilənlərin biliyə çevrilməsi prosesini başa çatdırır. Siyahıda aşağı səviyyələr yoxlanmış elmi və mühəndis həllərinə əsaslanır, yuxarı səviyyələr isə yeni olduqlarına görə zəif inkişaf etmişlər və kifayət qədər elmi təminatla malik deyillər. Mütəxəssislər bilavasitə IT-dən Big Data üçün texnologiyalar kateqoriyasına verilənlərin kütləvi paralel emalı platformalarında analitik sistemləri, verilənlərin bulud servisini, Hadoop və MapReduce texnologiyalarını və NoSQL tipli paylanmış verilənlər bazasının idarəetmə sistemlərini (VBIS) və s. daxil edirlər. Apache Software Foundation-un layihəsi olan Hadoop daha geniş yayılmış texnologiyadır, paylanmış hesablama mühitində böyük həcmli verilənlərin (petabayt miqyasında) emalı və analizi üçün əsas platformadır, MapReduce modelinin açıq kodlu (open access) sistemidir. Hadoop iki əsas komponentdən təşkil olunmuşdur: Hadoop MapReduce və Hadoop Distributed File System (HDFS). Burada MapReduce paralel hesablamalara, HDFS paylanmış fayl sistemi isə verilənlərin idarə edilməsinə cavab verir [21-25].

Böyük verilənlərin növbəti dalğasını insanın kompüterlə qarşılıqlı əlaqəsini təmin edən yeni qurğular yarada bilər. Onlara elektron kağız, taktill əks əlaqə texnologiyaları (haptics), müxtəlif videoshəkillər (video visor), həmçinin xüsusi yaddaşlı məhsul və əşyalar, Vikipediya oxşar açıq kontekstlə işləmək üçün sistemlər və s. aiddir. Məlumdur ki, geniş mənada texnologiya, əmtəə və xidmətlərin istehsalı üçün zəruri bilik kapitalı, tətbiqi mənada isə məhsulun hazırlanması prosesində maddənin, enerjinin, informasiyanın çevrilməsi, materialların emalı və yenidən emalı, hazır məhsulların yığılması, keyfiyyət və idarəetməyə nəzarət üçün lazım olan metodlardır. Texnoloji ardıcılıqdakı maşın və mexanizmlər ilkin xammal istehlaka hazır olan məhsula çevirir. Hazırkı inqilab amorf, qeyri-müəyyən “informasiya texnologiyaları”nı rədd edərək daha aydın “verilənlər texnologiyaları”nı təsdiq edir. Texnoloji ardıcılığın girişində xam verilənlər, çıxışda isə insan tərəfindən istifadəyə hazır olan verilənlər olmalıdır. Bu aspektdə müasir texnologiyalar verilənlərdəki informasiyanın yeni biliklərə çevrilməsinə və ya biliklərin əldə edilməsinə imkan verməlidirlər [26, 27].

Verilənlərlə işləyən texnologiyaların əsas fərqləndirici xüsusiyyətlərindən biri ondadır ki, onlar insan üçün işləyirlər. Avtomatik quraşdırılmış sistemlər istisna olmaqla qalan kompüter sistemləri son nəticədə insanın istifadə etdiyi verilənlərin hazırlanmasına xidmət edir, yalnız bu sistemlər verilənləri informasiyaya və sonra biliyə çevirə bilər. Bu cür texnologiyaların paradoksalığı ondadır ki, texnoloji ardıcılığın girişində verilənlərin həcmi daim artırsa, buradan da böyük verilənlərin problemi yaranır. Giriş və çıxışdakı verilənlərin həcmi arasındakı disproporsiya mahiyyətcə verilənlərlə işləyən texnologiyaların inkişafında əsas istiqaməti müəyyən edir: daxil edilən verilənləri itirməməklə hansısa bir yolla giriş axınına cilovlamaq, sonra bütün verilənlərdən daha əhəmiyyətli olanı seçmək və insanın qavraması üçün asan şəkildə təqdim etmək lazımdır. Hazırkı iqtisadi böhran exafloodla (exaflood, petabaytdan sonra gələn, 10¹⁸ bayta bərabər olan verilənlərin ölçü vahidi exabyte və flood, yəni, daşqın sözlərindən yaranmışdır) bağlı sahələrdə inkişafı dayandırmamış, lazımı verilənləri o qədər də mühüm

olmayanlardan ayırmaq üçün filtr rolunu oynayan mürəkkəb hadisələrin emalı texnologiyalarının (Complex Event Processing, CEP) inkişafını sürətləndirmişdir. Bu qrup texnologiyalar BV-lərlə bağlı əsas məsələlərdən birini həll edir, yəni, xam verilənlərdən yeni, gizli biliklərin və faydalı məlumatların əldə olunmasına keçidi asanlaşdırır [28].

Bu gün böyük həcmli verilənlərdən faydalı informasiyanın aşkarlanması üçün ən qabaqcıl metodlar təmin edilir, belə məsələləri effektiv həll etməyə imkan verən texnoloji alətlər yaradılır. Mütəxəssislər tətbiq sahəsindən asılı olmayaraq böyük verilənlər emalının texnoloji ardıcılığını yeddi əsas mərhələyə bölürlər [29-31]:

– **Verilənlərin toplanması.** Xam verilənlər xəzinələrdən, qurğuların vericilərindən, şəbəkə mənbələrindən toplanma bilər. Bu mərhələ mühəndis nöqtəyi-nəzərindən daha əhəmiyyətdir, lakin burada da verilənlərin tipini ayırmaq lazımdır, məsələn, mətn verilənlərini ədədi verilənlərlə səhv salmaq olmaz.

– **Verilənlərin sintaksis və qramatik təhlili.** Burada verilənlərin strukturlaşması, onların kateqoriyalar üzrə paylanması həyata keçirilir, təhlil isə bir neçə səviyyədə yerinə yetirilir. Aşağı səviyyəyə fiziki siqnalların emalı, sıxılmış faylların açılması və deşifrəlməsi daxildir. Bit səviyyəsində mətn, media və digər faylların ayrılması, mətn səviyyəsində qramatik və struktur təhlili aparılır.

– **Filtrasiya.** Burada verilənlərin analizi metodlarına, məsələn, mürəkkəb hadisələrin emalı texnologiyaları vasitələrinə müraciət etmədən giriş axınına azaldaraq yalnız faydalı verilənləri saxlamaq lazımdır.

– **Verilənlərin əldə edilməsi.** Verilənlərin əldə edilməsi təsvir və verilənləri ayırmağa imkan verən statistik və digər metodlardan ibarətdir, verilənləri uyğun riyazi kontekstə həll edir.

– **Təqdimat.** Verilənlər üçün daha yaxşı təqdimat formasını müəyyən etmək tələb olunur (diaqramlar, siyahılar, ağaclar və s.). Verilənlərin vizuallaşma üsulu sadə cədvəl və qrafiklərdən tutmuş mürəkkəb iki və üç ölçülü təsvirlərə qədər dəyişir.

– **Təqdimatın təkmilləşdirilməsi.** Burada verilənlərin təqdimatı formalarının redaktəsi aparılır.

– **Qarşılıqlı əlaqə.** Verilənlərlə fəal işləmək üçün verilənlərlə manipulyasiya fəndləri və daha əyani təqdimatı təmin edən üsullar işlənir.

Verilənlərin toplanması və onların təhlili əhəmiyyətli texnologiyalar vasitəsilə reallaşır, filtrasiya və əldə edilmə verilənlər haqqında yeni elmin predmetidir, verilənlərin təqdimatı və dəqiqləşdirilməsi qrafiki dizayn sahəsinə daxildir.

NƏTİCƏ

Big Data sadə düşüncə deyil, gələcək texniki inqilabın simvoludur. BV-lərlə analitik iş zəruriliyi İT-sənayesini tamam dəyişəcək yeni proqram və aparat platformalarının yaranmasını stimullaşdırır. Artıq bu gün böyük həcmli verilənlərin analizi üçün ən qabaqcıl metodlar təmin edilir: süni neyron şəbəkələri – bioloji neyron şəbəkələrinin təşkili və funksiyası prinsipi üzrə qurulmuş modellər, prediktiv

analitika, statistika, Natural Language Processing metodları və s. Həmçinin kraudsorsinq, A/B testləşmə, sentiment analiz və ekspertləri cəlb edən metodlardan istifadə edilir.

Beləliklə, dünyada gedən proseslərə və aparıcı ölkələrin təcrübəsinə nəzər saldıqda aydın olur ki, müxtəlif mənbələrdən avtomatik və fasiləsiz olaraq generasiya olunan verilənlərin real-vaxt ərzində analitik emalı və analizində fərqli modellər, konseptual yanaşmalar təklif olunur. Bu mühüm faktı nəzərə alaraq, beynəlxalq təcrübədə BV-nin analitikası istiqamətində aparılan tədqiqatların araşdırılması olduqca zəruridir. Aydındır ki, çox böyük həcmdə rəqəmsal verilənlərin toplanması cəmiyyətin bütün sferalarının qarşılaşdığı problem olduğu üçün, onunla bağlı elmi-tədqiqat işləri də bir çox elm sahəsi üçün aktualdır. Çünki, bu xam, strukturlaşdırılmamış verilənlər cəmiyyətin bütün sahələrini kökündən dəyişə biləcək təsirə malik bilik mənbəyidir. Bu baxımdan, Big Data texnologiyasının bütün aspektlərinin elmi-nəzəri əsaslarının işlənməsi böyük əhəmiyyət kəsb edir. Baxılan problemin araşdırılması, elmi-nəzəri tədqiqatların təhlili əsasında alınan nəticə göstərir ki, BV-dən yeni biliklərin və faydalı məlumatların əldə olunması verilənlərin idarə edilməsində, onlardan dəyər yaradılmasında, mərkəzləşdirmə, interaktivlik, onlayn rejim və digər spesifik xüsusiyyətləri sayəsində ictimai sferanın yeni, unikal modeli formalaşır.

ƏDƏBİYYAT

- [1] Clifford L. Big data: How do your data grow?, Nature, 2008, vol. 455, p. 28-29.
- [2] The digital universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. Study report, IDC, December 2012. www.emc.com/leadership/digital-universe
- [3] Wei Fan, Albert Bifet. Mining big data: current status, and forecast to the future, ACM SIGKDD Explorations Newsletter, 2012, vol. 14, No. 2, p. 1-5.
- [4] InfoSphere Platform: Big Data Analytics, 2013, <http://www-01.ibm.com/software/>
- [5] Oracle and Big Data: Big Data for the Enterprise, 2013, <http://www.oracle.com/>
- [6] Big Data, 2013, <http://www.microsoft.com/>
- [7] Big Data - What Is It? 2013, <http://www.sas.com/big-data/>
- [8] SAP HANA integrates predictive analytics, text and big data in a single package, 2013, <http://www54.sap.com/>
- [9] Big Data Solutions, 2013, <http://www8.hp.com/>
- [10] Madden S. From Databases to Big Data, IEEE Internet Computing, 2012, vol.16, issue 3, p. 4-6.
- [11] What is big data? - Bringing big data to the enterprise, 2013. <http://www-01.ibm.com>
- [12] Laney D. 3D Data Management: Controlling Data Volume, Velocity and Variety. Technical report, META Group, Inc (now Gartner, Inc.), February 2001. <http://blogs.gartner.com>
- [13] Birke R., Björkqvist M., Chen L. Y., Smirni E., Engbersen T. (Big) data in a virtualized world: volume, velocity, and variety in cloud datacenters, Proceedings of the 12th USENIX conference on File and Storage Technologies (FAST'14), USENIX Association Berkeley, CA, USA, 2014, p. 177-189.
- [14] Richard Price. Volume, velocity and variety: key challenges for mining large volumes of multimedia information, Proceedings of the 7th Australasian Data Mining Conference (AusDM '08), Australia, 2008, vol. 87, p. 17-17.
- [15] Chiang R.H.L., Goes P., Stohr E.A. Business Intelligence and Analytics Education, and Program Development: A Unique Opportunity for the Information Systems Discipline, Journal ACM Transactions on

- Management Information Systems, 2012, vol. 3, No. 3, Article 12 (p. 12:1-12:13).
- [16] Chen H., Chiang R.H. L., Storey V.C. Business intelligence and analytics: from big data to big impact, *Journal MIS Quarterly*, 2012, vol. 36, No. 4, p. 1165-1188.
- [17] Omar El-Gayar, Prem Timsina. Opportunities for Business Intelligence and Big Data Analytics in Evidence Based Medicine, *HICSS '14 Proceedings of the 2014 47th Hawaii International Conference on System Sciences(HICSS '14)*, USA, 2014, p. 749-757.
- [18] Statchuk C., Iles M., Thomas F. Big data and analytics, *Proceedings of the 2013 Conference of the Center for Advanced Studies on Collaborative Research (CASCON '13)*, USA, 2013, p. 341-343.
- [19] Foster Y., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations, *Intern. J. of High Performance Computing Applications*, 2001, vol. 15, No. 3, p. 200-222.
- [20] Черняк Леонид. Большие Данные – новая теория и практика, *Открытые системы*, 2011, № 10, с. 18-25.
- [21] Dean J., Ghemawat S. MapReduce: simplified data processing on large clusters, *Communications of the ACM*, 2008, vol. 5, No.1, p. 107-113.
- [22] Lee K-H., Lee Y-J., Choi H., Chung Y.D., Moon B. Parallel data processing with MapReduce: a survey, *ACM SIGMOD Record*, 2011, vol. 40, No. 4, p.11-20.
- [23] Brunozzi Simone. Big Data and NoSQL with Amazon DynamoDB, *Proceedings of the 2012 workshop on Management of big data systems (MBDS '12)*, 2012, USA, p. 41-42.
- [24] Weiyi Shang, Zhen Ming Jiang, Hadi Hemmati, Bram Adams, Ahmed E. Hassan, Patrick Martin. Assisting developers of big data analytics applications when deploying on hadoop clouds, *Proceedings of the 2013 International Conference on Software Engineering (ICSE '13)*, NJ, USA, 2013, p. 402-411.
- [25] Чак Лэм. Hadoop в действии, Издательство: ДМК Пресс 2012, 424 с.
- [26] Черняк Леонид. Вычисления с акцентом на данные, *Открытые системы*, 2008, № 8, с. 36-39.
- [27] Wu X., Zhu X., Wu G., Ding W. Data Mining with Big Data, *Journal IEEE Transactions on Knowledge and Data Engineering*, USA, 2014, vol. 26, No.1, p. 97-107.
- [28] Wang Y.H., Cao K., Zhang X.M. Complex event processing over distributed probabilistic event streams, *Computers & Mathematics with Applications*, 2013, vol. 66, No. 10, p. 1808-182131.
- [29] Черняк Леонид. Смутное время СУБД, *Открытые системы*, 2012, № 2, с. 16-21.
- [30] Черняк Леонид. Что делать с хаосом данных?, *Открытые системы*, 2013, № 9, с. 16-20.
- [31] Дубова Наталья. Большие Данные крупным планом, *Открытые системы*, 2011, № 10, с. 30-33.