

УДК 004:519.713

**В. И. ХАХАНОВ, Е. И. ЛИТВИНОВА, С. В. ЧУМАЧЕНКО,
И. ЕМЕЛЬЯНОВ, ТАМЕР БАНИ АМЕР***Харьковский национальный университет радиоэлектроники, Украина***ПРОЦЕССОРНЫЕ СТРУКТУРЫ ДЛЯ АНАЛИЗА BIG DATA**

Предлагается инфраструктура обеспечения параллельного анализа big data для поиска, распознавания и принятия решений на основе использования булеановой метрики измерения киберпространства. Она характеризуется наличием единственной логической хог-операции для определения кибер-расстояния путем циклического замыкания не менее одного объекта, что дает возможность на порядок повысить быстродействие анализа больших данных. Предложен новый подход векторно-логической обработки big data с полным исключением арифметических операций, влияющих на быстродействие и аппаратную сложность, который может быть эффективно реализован как на основе использования современных мультимикропроцессорных цифровых систем на кристаллах, так и с помощью виртуальных параллельных процессоров, функционирующих под управлением киберфизических систем или облачных сервисов-фильтров.

Ключевые слова: большие данные, киберинформационная система, киберфизический компьютеринг, мультимикропроцессор, киберпространство, векторно-логическая обработка больших данных.

1. Введение

Цель – создание метрико-семантической модели взаимодействия технологической культуры big data с интеллектуальными сенсорами мониторинга реального мира для управления киберфизическими процессами, повышения качества жизни людей и сохранения экосистемы планеты. Задачи из анализа источников:

1) определение актуальных рыночно-ориентированных направлений научно-образовательных исследований и дизрапторного развития киберэко-системы планеты [1-3];

2) использование технологической культуры big data для создания киберфизических систем мониторинга и управления [4-6];

3) рыночно ориентированные структуры киберфизических систем управления неприродными процессами [7-9].

Индексы числа научных публикаций [1] в IEEE Xplore формируют модные тренды в направлении создания киберфизической экосистемы: биотехнологии (13907), нанотехнологии (26176), Internet of Things (4200), Smart Everything (46439), Cloud Computing (19865), Big Data and Analytics (7709) [10], Cyber Physical System (CPS) (2348), мобильные технологии (220000) и социальные сети (19782), технологии киберзащит (124917), искусственный интеллект (149397). Интересен также рейтинг ссылок в IEEE Xplore для компьютерных направлений рынка научно-образовательных услуг: Program Engineering (79783), Computer Engineering (196748), System Engineering (376614), Electronics – 566496,

Social Engineering (10012), Radio Engineering (38546, 151094 – Radiosystems), Electronical Engineering (119856), Production Engineering (45293), Telecommunications (228278), Applied Mathematics (10034), Computer Science (181916), Media Devices and TV (28126), System Programming (159495), Computer Networks and Systems (44853), Internet (118157), Control Systems (563336), Computers (702123), Computer Science and Engineering (61369).

Прорывными системо-образующими дизрапторами для инвестиций временных, финансовых и людских ресурсов в ближайшие 8 лет будут:

1) Crowd-sourcing/open-sourcing of hardware development (419240),

2) Changes in educational structure/design (MOOCs) (387777),

3) Virtual/alternative currencies (Bitcoin) (71),

4) Smartphone for payment (216),

5) Cloud computing (20291),

6) Robots as source of labor (281),

7) Nonvolatile memory influencing big data accessibility and portability (2308),

8) Quantum/nondeterministic computing (7653) [11-13],

9) 3D printing (1335),

10) Green computing (5827),

11) New user interfaces (Siri, Kinect) (11051).

Киберфизическая система призвана сделать активной концепцию big data, рассматривая большие данные во взаимодействии с киберсистемами (облаками) управления, ориентированными на поиск, распознавание и принятие решений [14-16]. Структурное содержание CPS – совокупность коммуника-

ционно связанных реальных и виртуальных компонентов с выраженными функциями адекватного физического цифрового мониторинга и оптимального облачного компьютерного киберуправления для обеспечения качества жизни, продукции, процессов или сервисов в заданных условиях ограничений на время и ресурсы [17]. CPS включает компоненты: Cyber Control, Internet of Things или Cloud, Security, Intelligence, Big Data and Services, Digital Monitoring, Cyber Managing, Physical Smart Everything, Nature, Social, and Tech World. Регуляторные отношения (Relationship) между компонентами CPS формируются законами, уставами предприятий и организаций, морально-этическими правилами поведения внутри социальной группы. Направление движения RoadMap – Harmony of Human, Nature and Tech киберфизической системы человечества можно определить как достижение такого интегрального уровня развития киберфизических компонентов, который обеспечит гармонию жизни человека с природой и техникой (созданным миром – Created World).

Гармония предполагает создание киберинтеллекта, который к 2050 году должен позиционироваться как мозг человечества (Humanity Brain); цифровую идентификацию всех физических процессов, объектов и трехмерного пространства с помощью технологий Internet of Things, Smart Everything и Big Data.

Особое значение здесь приобретает достаточно новая «отчасти рекламная» парадигма Big Data, для критики которой существует не меньше специалистов, чем тех, кто считает ее дизраптором или прорывной технологией хранения и анализа данных 21 века. К сожалению, практически не существует приличного определения данного достаточно модного рыночного тренда.

Big Data – технологическая культура киберпространства, направленная на формирование динамически развивающейся инфраструктуры киберфизической экосистемы планеты путем метрико-семантической структуризации больших потоков (объемов) гетерогенных данных на основе использования интеллектуальных быстродействующих специализированных облачных фильтров параллельного мониторинга и метрического анализа извлекаемой информации для online управления физическими и виртуальными процессами.

Почему возникла парадигма больших данных как новый рыночный тренд, раскручиваемый учеными и компаниями? Короткий ответ содержит всего два тезиса:

1) большие данные невозможно просмотреть людскими ресурсами за приемлемое время – нужны фабрики для их хранения и многократного использования;

2) большие данные невозможно осмыслить даже совокупным мозгом всего человечества – нужны мощные интеллектуальные анализаторы для извлечения полезной информации, интегрируемой в знания.

Процесс получения знаний можно свести к четырем технологиям [14]:

- 1) экспериментальные измерения;
- 2) аналитические исследования;
- 3) численное моделирование;
- 4) инновационный компонент, замыкающий все процессы в систему – data-driven science for control.

Что касается эволюции киберпространства, то здесь наблюдается экспоненциальный рост объемов (удвоение каждые 1,5 года) и сложности данных, что приводит к трансформированию понимания процессов и структур:

- 1) от бедности данных к их перенасыщению;
- 2) от наборов данных к их потокам;
- 3) от статике данных к их динамическому развитию;
- 4) от обработки данных «время от времени» к их анализу в реальном масштабе времени;
- 5) от централизованного хранения данных к распределенным в пространстве big data;
- 6) от владения данными к обладанию их полезной экспертизы. Однако на пути становления киберэкосистемы планеты, возникают новые научные и практические проблемы, связанные с формированием киберфизических процессов:

- 1) data gathering – сбор данных от сенсоров и сетей;
- 2) data farming – создание фабрик и распределенных структур для хранения, архивирования, индексирования, поиска, объединения и разделения данных;
- 3) data mining – извлечение данных по шаблонам и их суперпозициям, кластерный анализ и классификация, поиск аномалий или выбросов, гиперпространственная визуализация;
- 4) data understanding – семантическое понимание данных;
- 5) new cyber knowledge – формирование новых знаний для управления киберфизическими процессами.

Сегодня также можно выделить несколько дифференцирующих принципов, характеризующих Big Data [10, 18-19]:

- 1) вместо причинно-следственных связей, предлагается использовать доминирование корреляции информационных объектов;
- 2) вместо выборки данных (максимум пользы из минимума информации) – предлагается использовать полное множество материалов;

3) вместо хранения данных – инновационно декларируется, что ценность данных заключается в уровне их многократного или массового использования вчера, сегодня и завтра для прогнозирования и/или управления действительностью;

4) вместо традиционных знаний для понимания прошлого предлагается приобретать способность прогнозировать будущее;

5) вместо структур данных с жесткими связями – адресная организация физических и виртуальных объектов и процессов;

6) вместо ручного ввода данных – использование Интернета как входа для киберсистемы: smart everything + internet of everything;

7) вместо вывода данных за пределы киберпространства – использование в качестве выхода киберсистемы Интернета и управляющих регуляторных воздействий для cyber physical systems;

8) вместо технологий пассивного отображения реального и виртуального мира – киберфизические системы мониторинга и анализа данных для управления физическими и виртуальными процессами;

9) вместо универсальных и тяжеловесных систем сбора и анализа информации – специализированные виртуальные параллельные мультипроцессоры мониторинга и управления физическими и виртуальными процессами;

10) вместо хаоса статических данных и знаний в киберпространстве Интернета – постепенная семантическая структуризация динамических потоков больших данных киберфизических процессов и явлений для их эффективного мониторинга, анализа и управления;

11) вместо неупорядоченных данных, трудных для понимания и использования человеком или киберсистемой – умные, метрически ранжированные информационные структуры, ориентированные на принятие оптимального решения;

12) вместо обособленного развития реального и виртуального пространств – постепенное создание замкнутой киберфизической экосистемы планеты для совместного гармонического развития реального и виртуального миров.

В дополнение к дифференцирующей метрике big data можно еще добавить совсем не уникальную характеристику VVVV: volume – большая размерность данных; velocity – высокое быстродействие предоставления сервиса; variety – мощная пространственно-временная семантика и онтология данных; veracity – высокая валидность и точность полезной информации.

Одним из возможных потребителей big data культуры является CPS, которая ориентирована на замену человека-менеджера кибероблачными сервисами управления социальными группами, биологи-

ческими, техническими и виртуальными объектами [14, 16, 17]. Цель киберфизической системы – Human Long Life Quality – определяется качеством жизни человечества, социальных групп и каждого человека в гармонии с зеленой планетой и искусственным миром. Цель, как функциональный выход CPS, зависит от точного цифрового мониторинга и оптимального киберуправления виртуальными и физическими ресурсами, включающими время, деньги, кадры и материалы. Главным отличием масштабируемой киберфизической системы является отсутствие человеческого фактора в блоке управления (Cyber), что делает ее, при конструктивном и гуманном законотворчестве, справедливой, эффективной, оптимальной, надежной и защищенной от субъективных ошибок менеджера.

Рыночно-привлекательные глобальные проекты сегодня выполняются под эгидой объединения физического и виртуального пространства в единое целое. Кибер-физическое пространство (Cyber Physical Space) – метрика телекоммуникационного взаимодействия физических, биологических и социальных объектов, процессов и явлений с виртуальными или облачными (компьютерными) технологиями мониторинга и управления на основе использования Internet of Things & Smart Everything для достижения социально значимых целей. Оно призвано интегрировать наиболее перспективные киберфизические технологические решения:

1) встроенный интерфейс непосредственной связи мозга человека с компьютером и/или киберпространством путем замены последовательных языковых интерфейсов на параллельные образные отношения;

2) создание искусственного интеллекта для самообучения и самосовершенствования киберфизических структур, программ и процессов;

3) нановыращивание компьютера путем аддитивного структурирования атомов;

4) самое интересное решение связано с неотвратимостью естественного отказа человечества от функций управления биологическими, социальными и техническими объектами и процессами в пользу беспристрастных киберфизических систем!

Девять технических лидеров IEEE Computer Society [1] объединили усилия для прогнозирования будущего планеты, в которое включены 23 компьютерные технологии 2022 года. Сформирована кибермода на ближайшие 8 лет: 3D printing, big data and analytics, open intellectual property movement, massively online open courses, security cross-cutting issues, universal memory, 3D integrated circuits, photonics, cloud computing, computational biology and bioinformatics, device and nanotechnology, sustainability, high-performance computing, internet of

things, life sciences, machine learning and intelligent systems, natural user interfaces, computer vision and pattern recognition, networking and interconnectivity, quantum computing, software-defined networks, multicore, and robotics for medical care.

Как итог сказанному выше можно констатировать, что киберфизические системы, большие данные и квантоподобные параллельные мультипроцессоры формируют сегмент рынка, предназначенный для поиска, распознавания и принятия эффективных управленческих решений.

Предлагается big data driven киберфизическая система (рис. 1) online управления физическими и/или виртуальными процессами, инвариантными по отношению к сферам человеческой деятельности.

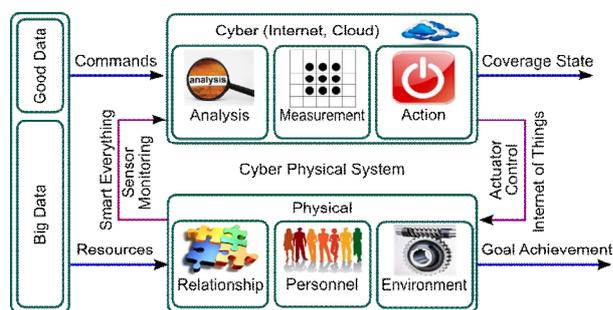


Рис. 1. Киберфизическая система управления процессами

Частные задачи, которые позиционируются как процессы, определенные во времени и пространстве, можно сформулировать в виде следующего:

- 1) списка: целеуказание при стрельбе, когда мишень управляет ракетой;
- 2) поиск контента по ключевым словам, которые отождествляются с булевыми примитивами;
- 3) управление светофором со стороны машин, формирующих в online-режиме матрицу актуатора;
- 4) учебный план управляет образовательным процессом студента;
- 5) взлетно-посадочная полоса руководит посадкой самолета; маршрут транспортного средства управляет его передвижением;
- 6) goadmap, отношения и умные сенсоры руководят социальными институтами, производственными предприятиями и компаниями.

Представленная big data driven киберфизическая система управления физическими процессами имеет инновационные отличия от существующих решений:

- 1) матричное представление цели, процесса, маршрута движения, компетенций объектов и субъектов,
- 2) online исполнение всех процессов,
- 3) использование нецифровой булеан-метрики для структурного и скалярного оценивания процессов и явлений,

4) использование метрического дискретного булевого пространства для идентификации процессов и явлений, применение big data технологической культуры для организации и активного управления киберфизическими процессами,

5) киберсистемное управление процессами на основе жесткого взаимодействия исполнительного и управляющего механизмов, сенсоров и актуаторов.

Взаимодействие интеллектуальных средств (фильтров) управления и big data в рамках киберфизической системы фильтров поиска полезной информации подобно процессу добычи золота из песка с помощью драги, настроенной на выделение тяжелого металла, когда легкие фракции вымываются водой. Следует заметить, что баланс экосистемы не нарушается, а добытое золото остается на планете, только в строго определенном месте. Аналогичные процессы протекают и в киберэкосистеме планеты, когда из хаоса киберпространства тяжелым трудом умных фильтров добывается полезная информация, чтобы ее затем положить в ячейку уже структурированной части киберпространства, которая становится reusable и может быть легко доступна всем желающим. Человечество стремится к порядку и структуризации, но оно также ответственно за мусор и хаос в киберпространстве, объемы которого еще в большей степени подчиняются закону Мура. Каждые 2 года информация удваивается, а к 2020 году ее объем составит 40 зеттабайт. Поэтому процесс упорядочения информации всегда будет отставать от мусоризации киберпространства. Сегодня человечество фильтрует и использует порядка 0,4 процента уже полезной информации. Дальше не будет больше. Это означает, что в ближайшие 100 лет актуальность технологической культуры big data для создания «правильной» инфраструктуры киберпространства будет только возрастать. Прогнозируется, что финансовые и кадровые инвестиции для создания инфраструктуры киберэкосистемы до 2020 года вырастут на 40%. Инвестиции в хранение и защиту информации, Big Data и Cloud Computing будут расти значительно быстрее [2].

Сегодня более 60% компаний делают инвестиции в технологии Big Data, Cloud Services и аналитические продукты, чтобы иметь big data-driven киберуправление кадровыми и материальными ресурсами. Порядка 60% компаний в мире, по оценкам журнала «Форбс», готовы купить программные системы управления ресурсами. Компании замотивированы семью аргументами: продукты взрослеют и умнеют, их становится легко купить, появился удобный пользовательский интерфейс, системы способны интегрировать многочисленные программные средства компании, Big Data реально позволяет управлять кадрами – путем playing

“Moneyball” with their people data, облачные технологии позволяют легко переключаться на новые сервисы управления кадрами, а талант стал навсегда стратегическим товаром и главным вопросом каждого руководителя. Человеческий капитал, по оценке журнала «Форбс», имеет индекс важности для решения проблем компании, организации, государства – 2,44; управление и выполнение операций 2,10; инновации 1,99; остальные 7 имеют индексы: отношения с потребителями 1,72; глобальная политика 1,68; правительственное регулирование 1,55; глобальная экспансия 1,31; корпоративный бренд и репутация 0,92; устойчивость 0,82 и вера в бизнес 0,46. Дуализм управления на основе Big Data и Cloud Services включает: 1) детерминизм – технологии (правильные данные) управляют нами и 2) волюнтаризм – мы управляем технологиями. Оба варианта в своем комплексном развитии приводят рынок облачных технологий управления к детерминизму на основе использования концепции киберфизических систем, где фигурируют огромные массивы данных, не всегда достоверной информации. Но умная аналитика движков по Big Data пространствам должна научиться формировать правильное решение. Leon Trotsky: “Tell me anyway – maybe I can find the truth by comparing the lies”. Скажи мне что ты думаешь в любой форме, а я сумею найти правду сравнением даже ложных высказываний. По данным журнала Форбс технология Big Data сгенерирует в 2015 году 3,7 триллиона прибыли в продуктах и сервисах, что означает появление на рынке 4,4 миллиона новых рабочих мест. Если учесть, что во всех компаниях мира заработная плата составляет 40% доходов, то управление персоналом и ресурсами сегодня есть самая важная проблема бизнеса. Главный вывод из сказанного – человечество настолько гениально и одновременно несовершенно, что оно не может объективно управлять самим собой! Человек гениален в творчестве и бездарен в самоуправлении. Таким образом, мировой рынок безальтернативно приходит к необходимости использования кибероблачного управления ресурсами и кадрами без участия человека, но на основе извлечения правильных данных из киберинформационного «мусора».

Цель рыночного бренда big data – выращивание в киберпространстве культурного слоя инфраструктуры метрико-семантически упорядоченной легко доступной полезной reusable информации за счет разработки виртуальных облачных сервисов на основе параллельных мультипроцессоров, выполняющих роль быстродействующих интеллектуальных фильтров при поиске, распознавании и принятии решений.

Задачи технологической культуры big data:

1) хранение неструктурированной разнородной информации в надежных распределенных системах, обслуживаемых Hadoop-сервисами;

2) анализ big data в реальном времени с помощью низкоуровневых параллельных интеллектуальных быстродействующих процессоров среды Map-Reduce, помещенных в облачные сервисы;

3) создание новых метрик измерения расстояний между процессами и явлениями в киберпространстве для построения быстродействующих метрико-семантических фильтров поиска полезной информации;

4) разработка инфраструктурных решений для киберпространства в целях компактного хранения и быстрого семантико-метрического извлечения полезной классифицированной информации с помощью облачных сервисов и специализированных процессоров;

5) создание типового шаблона киберфизической системы управления big data analytics, использующей структуры облачных фильтров для извлечения полезной информации из больших объемов неструктурированных данных в целях получения прибыли путем big-data-driven управления физическими и виртуальными процессами;

6) построение масштабируемой киберинформационной системы массовой метрико-семантической переработки в реальном масштабе времени больших объемов данных в полезную структурированную информацию, используемую для управления физическим миром;

7) разработка big data-driven аналитической киберфизической системы прогнозирования (планирования) и управления неприродными (социальными) процессами и явлениями (катаклизмами) путем точного и исчерпывающего мониторинга общественного мнения для выработки регуляторных информационных управляющих воздействий в целях обеспечения качества жизни социальных групп и устранения конфликтов. Например, при желании власть имущих можно без финансовых и временных затрат устранить все конфликты в Украине путем законодательного учета интересов всех слоев населения во времени и пространстве на основе толерантного объединения исторических, языковых и национальных культур. Курс любой системы власти на пересечение интересов социальных групп, очевидно всем и не только знающим математику, непременно приводит к нарушению целостности и/или гибели государственных образований.

Структура киберинформационной системы (рис. 2) с сигналами мониторинга и управления включает компоненты:

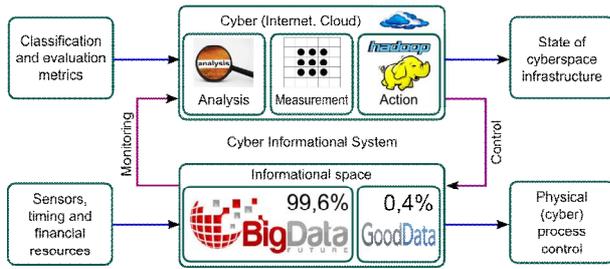


Рис. 2. Киберинформационная система трансформирования big data

- 1) «big data – good data», которые должны стремиться к процентному соотношению 90:10;
- 2) облачные сервисы-фильтры, предназначенные для формирования 10 процентов структурированной, легко доступной, полезной и многократно востребованной информации;
- 3) материальные и временные ресурсы для трансформирования киберпространства в семантико-метрическую инфраструктуру правильных данных;
- 4) метрики классификации и оценивания информационных объектов, необходимые для создания фильтров, анализа big data и синтеза структур правильных данных;
- 5) вектор состояния киберпространства, определяющий фактическое соотношение между «big data и good data» в реальном времени;
- 6) цель – достижение высокого уровня правильных данных (Good Data = PureData [18-19]) по отношению к «информационному мусору» – big data и последующего использования уже метрико-семантически-структурированных данных для оптимального управления киберфизическими процессами планеты.

По оценке журнала «Форбс» сегодня 36 процентов компаний вкладывают ресурсы в технологическую культуру big data. Однако только 13 процентов из них занимаются прогнозированием бизнесов в своем сегменте рынка. Тем не менее, уже 16 процентов компаний пытаются использовать добытые из big data правильные данные для управления киберфизическими процессами. Таким образом, можно сделать вывод, что взамен статистическому анализу данных по частичной представительной вы-

борке приходит точный и полный анализ больших данных по наперед заданной тематике, где экспертная формулировка проблемы есть искусство попадания в цель.

2. Критерии и структуры для оценивания качества взаимодействия объектов в киберпространстве

Цель – разработка универсального критерия и структур оценивания качества взаимодействия объектов в детерминированном киберпространстве на основе использования только логических операций для определения расстояний между процессами и явлениями путем построения высокопроизводительных параллельных процессоров, ориентированных на эффективное решение задач поиска, распознавания и принятия решений. Задачи: 1) разработка универсальной аналитической модели для определения бинарного взаимодействия между процессами и явлениями в многозначной логике; 2) разработка неарифметической вычислительной параллельной структуры для метрического оценивания детерминированных процессов, явлений и выбора оптимального решения. Входной вектор $m = (m_1, \dots, m_j, \dots, m_k)$, $m_j \in \{0,1,x\}$ и анализируемый объект $A = (A_1, \dots, A_j, \dots, A_k)$, $A_j \in \{0,1,x\}$, который также представлен вектором, имеют одинаковую размерность k [14]. Степень принадлежности m -вектора к A обозначается как $\mu(m \in A)$. Процедура вычисления векторного критерия качества зависит от значности алфавита:

$$Q' = \begin{cases} m \oplus A \leftarrow \{m_j, A_j\} \in \{0,1\}; \\ m \Delta A \leftarrow \{m_j, A_j\} \in \{0,1,x\}. \end{cases}$$

При рассмотрении практически ориентированных задач необходимо найти лучшее решение из конечного числа вариантов взаимодействия $m \oplus A_i$ для его последующей реализации без арифметических операций, которые на один два порядка снижают быстродействие. Пусть, например, имеются два вектора A B , относительно которых необходимо выполнить операции: $a = m \oplus A$, $b = m \oplus B$, чтобы оценить близость каждого из них к вектор-запросу m , как показано в табл. 1.

Таблица 1

Поиск лучшего решения

$A =$	1 0 1 0 0 1 0 1	$A =$	0 0 0 1 0 0 0 1	$A =$	1 0 1 0 0 1 1 1
$B =$	0 1 0 0 1 0 1 1	$B =$	0 1 0 0 1 1 0 0	$B =$	0 1 0 0 1 0 0 1
$m =$	1 1 1 1 1 1 0 0	$m =$	1 1 1 1 1 1 0 0	$m =$	1 1 1 1 1 1 0 0
$\mu_a(m \in A) = m \oplus A$	0 1 0 1 1 0 0 1	$\mu_a(m \in A) = m \oplus A$	1 1 1 0 1 1 0 1	$\mu_a(m \in A) = m \oplus A$	0 1 0 1 1 0 1 1
$\mu_b(m \in B) = m \oplus B$	1 0 1 1 0 1 1 1	$\mu_b(m \in B) = m \oplus B$	1 0 1 1 0 0 0 0	$\mu_b(m \in B) = m \oplus B$	1 0 1 1 0 1 0 1
$a = slc(\mu_a)$	1 1 1 1 0 0 0 0	$a = slc(\mu_a)$	1 1 1 1 1 1 0 0	$a = slc(\mu_a)$	1 1 1 1 1 0 0 0
$b = slc(\mu_b)$	1 1 1 1 1 1 0 0	$b = slc(\mu_b)$	1 1 1 0 0 0 0 0	$b = slc(\mu_b)$	1 1 1 1 1 0 0 0

Далее предлагается простая и доступная для понимания и реализации структура параллельного вычисления лучшего варианта на основе нечисленного сравнения двух альтернативных векторов a и b , полученных на основе использования одноктактной операции slc – сдвиг всех единиц влево с уплотнением [15]. После параллельного сдвига за один такт всех единиц в регистрах векторного критерия качества, оценивающих взаимодействия объектов в киберпространстве, теоретически возможны три варианта соотношения единиц, представленных в табл. 2 (взаимодействие ранее полученных векторов a и b).

Таблица 2

Взаимодействие векторов

$a =$	1 1 1 1 0 0 0 0	$a =$	1 1 1 1 1 1 0 0
$b =$	1 1 1 1 1 1 0 0	$b =$	1 1 1 0 0 0 0 0
$a \wedge b$	1 1 1 1 0 0 0 0	$a \wedge b$	1 1 1 0 0 0 0 0
$q^b = (a \wedge b) \oplus b$	0 0 0 0 1 1 0 0	$q^b = (a \wedge b) \oplus b$	0 0 0 0 0 0 0 0
$q^a = (a \wedge b) \oplus a$	0 0 0 0 0 0 0 0	$q^a = (a \wedge b) \oplus a$	0 0 0 1 1 1 0 0
$Q^b = \bigvee_{i=1}^k q_i^b = 1$		$Q^b = \bigvee_{i=1}^k q_i^b = 0$	w i n n e r
$Q^a = \bigvee_{i=1}^k q_i^a = 0$	w i n n e r	$Q^a = \bigvee_{i=1}^k q_i^a = 1$	

$a =$	1 1 1 1 1 0 0 0	$Q^b = \bigvee_{i=1}^k q_i^b = 0$	w i n n e r
$b =$	1 1 1 1 1 0 0 0	$Q^a = \bigvee_{i=1}^k q_i^a = 1$	w i n n e r
$a \wedge b$	1 1 1 1 1 0 0 0		
$q^b = (a \wedge b) \oplus b$	0 0 0 0 0 0 0 0		
$q^a = (a \wedge b) \oplus a$	0 0 0 0 0 0 0 0		

Нулевое значение Q -критерия означает лучшую альтернативу из рассмотренных, которая используется далее для сравнения с другими оценками или в качестве конечного варианта решения проблемы. Регистровые переменные a и b , обозначающие векторы сжатых влево единичных значений, объединяются и инвертируются для одновременного выполнения хог-операций. Затем результаты в виде состояний регистров подаются на входы двух логических элементов $ог$, которые формируют уже состояния двух булевых переменных, которые создают три сочетания: 00, 01, 10. Нулевое значение одной из двух переменных означает лучшее решение, которое необходимо выбрать. Два нулевых состояния означают, что оба решения равнозначны по уровню предпочтения. Единичная комбинация булевых переменных невозможна. Схемная реализация кибер-выбора из двух альтернатив, соответствующая приведенной выше логической структуре, представлена на рис. 3.

Если учесть, что выбор лучшего решения должен идентифицироваться логическим сигналом 1 (вместо 0), что соответствует максимальному значению функции принадлежности, определенной ранее, тогда схемная структура для определения лучшего варианта из двух альтернатив может быть незначи-

тельно модифицирована с помощью дополнительных инверторов на выходах логических $ог$ элементов, формирующих $\{Q^a, Q^b\} \in Q$.

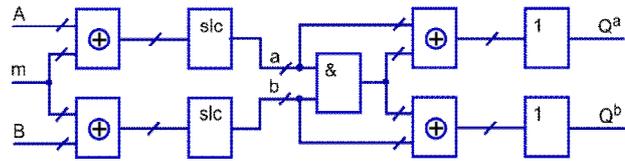


Рис. 3. Схемная реализация выбора оптимального решения

В этом случае на выходах цифровой структуры возможны следующие пары состояний: 10, 01 и 11, где выбор решения (a или b) осуществляется по единичному значению одного из выходов. Пара сигналов 00 на внешних выходах схемы с инверторами невозможна. Таким образом, точный поиск заказанной информации в big data можно и нужно выполнять на основе только логических операций and , or , not , xor , slc без использования арифметических функций, что позволяет проектировать быстродействующие векторно-логические физические и/или виртуальные мультипроцессоры для существенного уменьшения времени исполнения сервисов облачными приложениями. Вычислительная процедура поиска лучшего варианта из двух возможных сводится к параллельному выполнению 4-х векторных операций и одной логической, в результате которых на одном или нескольких выходах формируется значение нуля, которое идентифицирует лучшее решение:

$$Q^a = \bigvee_i \{ [slc(m \oplus A) \wedge slc(m \oplus B)] \oplus A \},$$

$$Q^b = \bigvee_i \{ [slc(m \oplus A) \wedge slc(m \oplus B)] \oplus B \}.$$

Векторно-логический SIMD-мультипроцессор характеризуется отсутствием арифметических операций, параллельным вычислением расстояния между запросом и информационными квантами, а также одновременным определением лучшего из возможных n -решений по минимуму функции принадлежности, что дает возможность на порядок повысить быстродействие максимально точного поиска данных в big data. Его структура изображена на рис. 4, где представлены только логические примитивы для выполнения векторных и булевых (битовых) операций. Процессор работает следующим образом: вектор запрос m , состоящий из k -битов, взаимодействует по хог-функции с матрицей M , имеющей n строк или векторов. В результате выполнения хог-операций формируются n функций принадлежности, определяющие степень близости

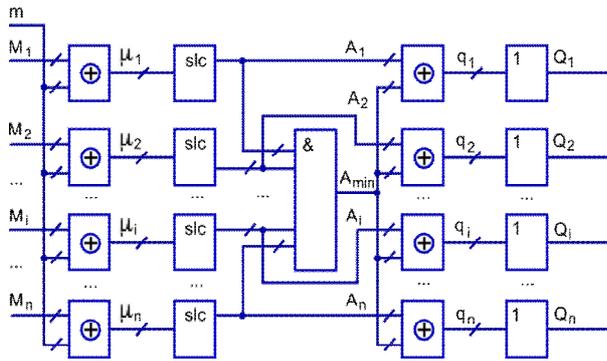


Рис. 4. Векторно-логический мультипроцессор

или расстояние между запросом и каждой вектор-строкой матрицы M . Для оценивания расстояний и выбора лучшего (минимального) взаимодействия выполняется регистровая операция slc , которая осуществляет уплотнение всех единиц со сдвигом влево за один автоматный такт, что дает возможность оценивать минимальное расстояние $m \oplus M_i$ номером бита, в котором находится правая крайняя единица. Для определения номера вектор-строки, формирующей минимум функции принадлежности, осуществляется параллельная поразрядная операция логического умножения над всеми векторами, содержащими сдвинутые влево единичные значения, что дает возможность вычислить вектор с минимальным числом единиц A_{min} . Последний используется для определения номера или индекса вектор-строки матрицы M , имеющей лучшее значение функции принадлежности, путем выполнения векторной хог-операции между A_{min} и всеми сдвинутыми функциями принадлежности A_i ($i=1, n$). В результате формируются векторы q_i ($i=1, n$), биты которых определяют входные значения каждого из n логических элементов $ог$. Выход каждого $ог$ -элемента равен единице, если существует хотя бы одно единичное значение в результатах сравнения $A_i \oplus A_{min}$. Если таких единиц нет, то минимальное расстояние между $m \oplus M_i$ идентифицируется 0-состоянием одного или, возможно, нескольких выходов Q_i ($i=1, n$). Аналитическая модель поиска оптимального решения в киберпространстве

по вектор-запросу оперирует пятью параллельными логическими операциями, выполняемыми последовательно:

$$Q^i = \bigvee_{j=1,k} \{ \bigwedge_{p=1,n} slc(m \oplus M_i) \} \oplus M_i.$$

Структурная модель векторного логического процессора, соответствующая аналитической модели формирования оптимального решения, имеет вид, представленный на рис. 4.

Пример вычисления состояний выходов векторно-логического процессора, как реакции на запрос m для матрицы M , состоящей из двух вектор-строк, представлен в табл. 3 и иллюстрирует формирование векторов единичных значений, сдвинутых влево: $\{A_1, A_2\} \in A$.

Не существует принципиальных ограничений для оценивания взаимодействия объектов в киберпространстве путем использования нечисленных метрик, исключающих арифметические операции. Более того, все расстояния в информационном мире можно измерять с помощью хог-операции или симметрической разности, которые обеспечивают выбор лучшего решения на основе векторно-логических критериев качества взаимодействия.

Следующий пример иллюстрирует работу мультипроцессора в многозначном алфавите описания логических переменных, ориентированном на компетентностное рейтинговое учение молодежи. Пусть имеется группа студентов, получившая сессионные оценки по восьми экзаменам в метрике A, B, C, D , которые кодируются соответствующими векторами: 1000, 1100, 1110, 1111. Необходимо определить лучшего студента, который интегрально получил максимальные баллы за сессию. Результаты вычислений представлены в табл. 4, 5.

В табл. 4 вектор-эталон m приведен к лучшим фактическим оценкам, полученным студентами по каждому экзамену. Табл. 5 оперирует вектором-эталонном с теоретически возможными лучшими ($A=1000$) значениями тестирования знаний. В обоих случаях интегральный критерий качества сессии определяет лучшим – студента под номером 3.

Таблица 3

Формирование векторов

$M_1 =$	1 0 1 0 0 1 0 1	$M_1 =$	0 0 0 1 0 0 0 1	$M_1 =$	1 0 1 0 0 1 1 1
$M_2 =$	0 1 0 0 1 0 1 1	$M_2 =$	0 1 0 0 1 1 0 0	$M_2 =$	0 1 0 0 1 0 0 1
$m =$	1 1 1 1 1 1 0 0	$m =$	1 1 1 1 1 1 0 0	$m =$	1 1 1 1 1 1 0 0
$\mu_1(m \in M_1) = m \oplus M_1$	0 1 0 1 1 0 0 1	$\mu_1(m \in M_1) = m \oplus M_1$	1 1 1 0 1 1 0 1	$\mu_1(m \in M_1) = m \oplus M_1$	0 1 0 1 1 0 1 1
$\mu_2(m \in M_2) = m \oplus M_2$	1 0 1 1 0 1 1 1	$\mu_2(m \in M_2) = m \oplus M_2$	1 0 1 1 0 0 0 0	$\mu_2(m \in M_2) = m \oplus M_2$	1 0 1 1 0 1 0 1
$A_1 = sl(\mu_1)$	1 1 1 1 0 0 0 0	$A_1 = sl(\mu_1)$	1 1 1 1 1 1 0 0	$A_1 = sl(\mu_1)$	1 1 1 1 1 0 0 0
$A_2 = sl(\mu_2)$	1 1 1 1 1 1 0 0	$A_2 = sl(\mu_2)$	1 1 1 0 0 0 0 0	$A_2 = sl(\mu_2)$	1 1 1 1 1 0 0 0

Таблица 4

Приведенный интегральный критерий сессии

$M_1 =$	A	B	D	D	B	A	C	C
$M_2 =$	C	C	D	A	B	B	A	D
$M_3 =$	B	C	C	B	A	A	C	D
$m =$	A	B	C	A	A	A	A	C
$\mu_1(m \in M_1) = m \oplus M_1$	0000	0000	0001	0111	0100	0000	0110	0000
$\mu_2(m \in M_2) = m \oplus M_2$	0110	0010	0001	0000	0100	0100	0000	0001
$\mu_3(m \in M_3) = m \oplus M_3$	0100	0010	0000	0100	0000	0000	0110	0001
$A_1 = \text{sl}(\mu_1)$	1111	111						
$A_2 = \text{sl}(\mu_2)$	1111	111						
$A_3 = \text{sl}(\mu_3)$	1111	11						
$M_3 =$	w	i	n	n	e	r		

Таблица 5

Интегральный критерий сессии

$M_1 =$	A	B	D	D	B	A	C	C
$M_2 =$	C	C	D	A	B	B	A	D
$M_3 =$	B	C	C	B	A	A	C	D
$m =$	A	A	A	A	A	A	A	A
$\mu_1(m \in M_1) = m \oplus M_1$	0000	0100	0111	0111	0100	0000	0110	0110
$\mu_2(m \in M_2) = m \oplus M_2$	0110	0110	0111	0000	0100	0100	0000	0111
$\mu_3(m \in M_3) = m \oplus M_3$	0100	0110	0110	0100	0000	0000	0110	0111
$A_1 = \text{sl}(\mu_1)$	1111	1111	1111	1				
$A_2 = \text{sl}(\mu_2)$	1111	1111	1111	1				
$A_3 = \text{sl}(\mu_3)$	1111	1111	111					
$M_3 =$	w	i	n	n	e	r		

Для кодирования всех пяти градаций болонской метрики оценивания знаний предлагается использовать и нулевую комбинацию: A=0000, B=1000, C=1100, D=1110, E=1111. В этом случае табл. 6, 7 дают аналогичный предыдущему результат выбора лучшего студента по итогам сессии.

Таблица 6

Приведенный интегральный критерий сессии по Болонской метрике

$M_1 =$	A	B	D	D	B	A	C	C
$M_2 =$	C	C	D	A	B	B	A	D
$M_3 =$	B	C	C	B	A	A	C	D
$m =$	A	B	C	A	A	A	A	C
$\mu_1(m \in M_1) = m \oplus M_1$	0000	0000	0010	1110	1000	0000	1100	0000
$\mu_2(m \in M_2) = m \oplus M_2$	1100	0100	0010	0000	1000	1000	0000	0010
$\mu_3(m \in M_3) = m \oplus M_3$	1000	0100	0000	1000	0000	0000	1100	0100
$A_1 = \text{sl}(\mu_1)$	1111	111						
$A_2 = \text{sl}(\mu_2)$	1111	111						
$A_3 = \text{sl}(\mu_3)$	1111	11						
$M_3 =$	w	i	n	n	e	r		

Таблица 7

Интегральный критерий сессии по Болонской метрике

$M_1 =$	A	B	D	D	B	A	C	C
$M_2 =$	C	C	D	A	B	B	A	D
$M_3 =$	B	C	C	B	A	A	C	D
$m =$	A	A	A	A	A	A	A	A
$\mu_1(m \in M_1) = m \oplus M_1$	0000	1000	1110	1110	1000	0000	1100	1100
$\mu_2(m \in M_2) = m \oplus M_2$	1100	1100	1110	0000	1000	1000	0000	1110
$\mu_3(m \in M_3) = m \oplus M_3$	1000	1100	1100	1000	0000	0000	1100	1110
$A_1 = \text{sl}(\mu_1)$	1111	1111	1111					
$A_2 = \text{sl}(\mu_2)$	1111	1111	1111					
$A_3 = \text{sl}(\mu_3)$	1111	1111	111					
$M_3 =$	w	i	n	n	e	r		

Если примитивы оценок неравнозначны по весу в метрике сравнения, то их нельзя обозначать унитарными кодами, которые выполняют лишь роль идентификаторов элементов в универсальном множестве равных примитивов. Поэтому для обозначения весов в кодах оценок был использован фактор количества единиц. Тем не менее, последующие действия, ориентированные на определение интегрального качества знаний студентов для выбора лучшего из них путем сравнения с идеальным результатом, не связаны с какими-либо арифметическими операциями, а используют только логические процедуры.

Стратегически задача линейной вычислительной сложности решалась в рамках следующей формулировки – найти студента, который имеет минимальное расстояние к заранее известному идеальному результату, в виде вектор-эталона m . Альтернативная стратегия предполагает поиск лучшего из n студентов путем последовательного сравнения каждого из них друг с другом, на что будет затрачено существенно больше времени, поскольку вычислительная сложность такой процедуры $-(n^2/2)-n$. Для ранжирования всех студентов относительно идеального результата необходимо выполнить $(n^2/2)-n$ векторно-логических операций. После определения на каждом шаге лучшего студента соответствующую ему строку матрицы M (экзаменационной компетенции академической группы учащихся) следует исключить из последующего рассмотрения путем занесения в ее разряды единичных значений сигналов.

Для эффективной работы логического мультипроцессора необходимо сформировать M -матрицу (рис. 5) возможных вариантов решения проблемы, которая, в частности, может быть продуктом применения поисковой системы Google (Hadoop) к киберпространству Internet (big data), используемой для грубой и широкой выборки, когда количество найденных информационных фрагментов достигает сотен или тысяч вариантов. Затем наступает очередь функционирования мультипроцессора, формирующего точное решение по запросу m , которое должно быть сохранено в структурированной, специализированной части киберпространства для последующего многократного использования. Поэтому входом и выходом логического мультипроцессора следует считать формы киберпространства: Internet of Things, Big Data, Cyber-Physical Systems. Рыночная привлекательность предложенного мультипроцессора заключается в возможности его использования для: повышения качества и быстродействия поисковых процедур в big data, создания встроенных автоматических, автономных систем диагностирования

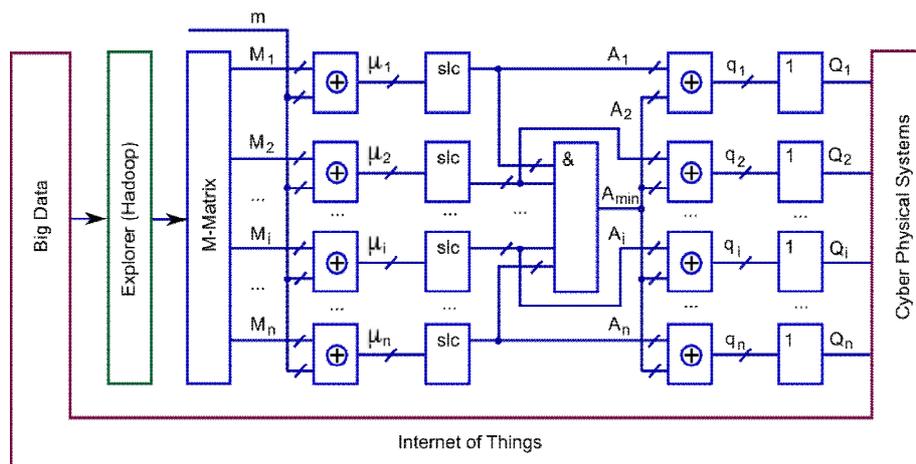


Рис. 5. Взаємодія мультипроцесора з кіберпростором

и восстановления работоспособности, средств целеуказания и распознавания образов.

Типичной для киберфизических систем, использующих информационное пространство, является функциональность, когда по запросу возникают многоальтернативные варианты его исполнения в векторно-логической форме описания компонентов киберпространства (субъектов, процессов или явлений), которые необходимы для управления социальными, биологическими и неприродными производственно-технологическими процессами без участия человека.

Заключение

Существующие программные продукты и публикации практически не предлагают ассоциативно-логических технологий поиска, распознавания и принятия решений в дискретном информационном пространстве [17], состоящем из big data. Практически все они используют универсальную систему команд современного дорогостоящего процессора с математическим сопроцессором. В то же время, аппаратные специализированные средства логического анализа, являющиеся их прототипами [15], как правило, ориентированы на побитовую или неекторную обработку информации. Предложенный новый подход векторно-логической обработки больших данных с полным исключением арифметических операций, влияющих на быстродействие и аппаратную сложность, может быть эффективно реализован на основе использования как современных мультипроцессорных цифровых систем на кристаллах, так и с помощью виртуальных параллельных процессоров, функционирующих под эгидой киберфизических систем или облачных сервисов-фильтров. Фактическая реализация подхода основана на предложении инновационных моделей и методов, использующих идею векторно-логической мет-

рики киберпространства.

Метрика анализа киберпространства (big data), которая характеризуется использованием единственной логической хог-операции для определения кибер-расстояния путем циклического замыкания не менее одного объекта, что дает возможность на порядок повысить быстродействие анализа big data и подсчет нечисленных структурных критериев качества взаимодействия информационных объектов на основе использования векторных логических операций для точного поиска, распознавания образов и принятия решений.

Новая структурно-цифровая модель точного поиска информации, которая характеризуется использованием векторных функций принадлежности запроса к информационным компонентам big data и последующего их сравнения для определения минимального расстояния, дающего возможность выбрать оптимальное решение при существенном уменьшении времени поиска за счет полного исключения арифметических операций.

Новая модель вычислительного дискретного автомата, которая характеризуется транзакционным взаимодействием компонентов памяти, исполняющих роль комбинационных и последовательностных элементов, реализованных в форме кубитных или «квантовых» примитивов, что дает возможность создавать параллельные виртуальные компьютеры для эффективного решения задач анализа big data без наличия арифметических команд и обеспечивать высокое быстродействие облачно-ориентированных процессоров.

Усовершенствованное определение адресно-автоматной модели виртуального компьютера <MQT> как организация структуры функциональных примитивов памяти без гальванических или проводных связей, на которых определены адресные транзакции данных во времени и пространстве для достижения поставленной цели – анализа объектов в

киберпространстве big data. Предложен Q-метод синтеза вычислительных устройств, который характеризуется использованием векторно-кубитной формы компактного описания структурных компонентов, что дает возможность существенно повысить быстродействие адресно-ориентированного моделирования логических элементов и уменьшить размерность памяти для хранения структуру данных цифрового устройства.

Новая киберинформационная модель анализа big data, использующая средства облачных сервисов, киберфизические системы, параллельные виртуальные мультипроцессоры с минимальным набором векторно-логических операций для точного поиска информации на основе предложенной булевой метрики и векторно-логических критериев качества, что дает возможность постепенной классификации и упорядочения хаотических данных big data в масштабах киберэкосистемы планеты.

Практическая значимость предложенных моделей заключается в необходимости реструктуризации киберпространства путем замены концепции аморфных big data на семантически классифицируемую информационную инфраструктуру полезных данных, предназначенных для управления киберфизическими процессами. В связи с этим предложены направления формирования технологической культуры big data для постепенного повышения уровня полезной информации от 0,4% до 10% путем компетентностной реструктуризации киберпространства больших данных.

Новая модель векторно-логического SIMD-мультипроцессора, который характеризуется отсутствием арифметических операций, параллельным вычислением расстояния между запросом и информационными квантами, а также одновременным определением лучшего из возможных n-решений по минимуму функции принадлежности, что дает возможность на порядок повысить быстродействие максимально точного поиска данных в big data.

Литература

1. *IEEE CS 2022 Report* / H. Alkhatib, P. Faraboschi, E. Frachtenberg, H. Kasahara, D. Lange, P. Laplante, A. Merchant, D. Milojicic, K. Schwan // *Computer*. – 2015. – Vol. 48, Iss. 3. – P. 68-76.
2. *Рост объема информации - реалии цифровой вселенной* [Электронный ресурс]. – Режим доступа: <http://www.tsonline.ru/articles2/fix-corp/rost-obema-informatsii--realii-tsifrovoy-vselennoy#sthash.rpNodQLF.dpuf>. – 16.04.2016.
3. *Майер-Шенбергер, В. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим* / В. Майер-Шенбергер, К. Кеннет. – Манн, Иванов и Фербер, 2013. – 240 с.

4. *Demchenko, Y. Defining architecture components of the Big Data Ecosystem* / Y. Demchenko, C. de Laat, P. Membrey // *International Conference on Collaboration Technologies and Systems (CTS)*. – 2014. – P. 104 – 112.
5. *Challenges for MapReduce in Big Data* / K. Grolinger, M. Hayes, W.A. Higashino, A. L'Heureux, D.S. Allison, M.A.M. Capretz // *IEEE World Congress on Services (SERVICES)*. – 2014. – P. 182 – 189.
6. *Zhang, Lichen. A framework to specify big data driven complex cyber physical control systems* / Lichen Zhang // *International Conference on Information and Automation (ICIA)*. – 2014. – P. 548 – 553.
7. *Zhang, Lichen. Designing big data driven cyber physical systems based on AADL* / Lichen Zhang // *International Conference on Systems, Man and Cybernetics (SMC)*. – 2014. – P. 3072 – 3077.
8. *Michalik, P. Concept definition for Big Data architecture in the education system* / P. Michalik, J. Stofa, I. Zolotova // *12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. – 2014. – P. 331 – 334.
9. *Munoz, M. Space systems modeling using the Architecture Analysis & Design Language (AADL)* / M. Munoz // *International Symposium on Software Reliability Engineering Workshops (ISSREW)*. – 2013. – P. 97 – 98.
10. *PureData System for Analytics* [Электронный ресурс]. – Режим доступа: http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=SP&infotype=PM&apname=SWGE_WA_UZ_USEN&htmlfid=WAS12385USEN&attachment=WAS12385USEN.PDF#loaded. – 1.04.2016.
11. *Nielsen, M. A. Quantum Computation and Quantum Information*. / M. A. Nielsen, I. L. Chuang. – Cambridge University Press, 2010. — 676 p.
12. *Stenholm, S. Quantum approach to informatics* / S. Stenholm, S. Kalle-Antti. – John Wiley & Sons, Inc, 2005. – 249 p.
13. *Qubit Model for solving the coverage problem* / V. I. Hahanov, E. I. Litvinova, S. V. Chumachenko, B. A. Abbas, E. A. Mandefro // *Proc. of IEEE East-West Design and Test Symposium*. – Kharkov, 2012. – P. 142 – 144.
14. *Big Data Driven Cyber Analytic System* / V. Hahanov, E. Litvinova, Wajeb Gharibi, S. Chumachenko // *2015 IEEE International Congress on Big Data, New York, USA*. – New York, 2015. – P. 615–622.
15. *Бондаренко, М. Ф. Структура логического ассоциативного мультипроцессора* / М. Ф. Бондаренко, В. И. Хаханов, Е. И. Литвинова // *Автоматика и телемеханика*. – 2012. – № 10. – С. 71–92.
16. *Hellinger, A. Cyber-Physical Systems. Driving force for innovation in mobility, health, energy and production*. / A. Hellinger, H. Seeger. – Springer, 2011. – 48 p.
17. *Cyber Physical Social Systems – Future of Ukraine* / V. Hahanov, W. Gharibi, A.P. Kudin, I. Hahanov, C. Ngene, Y. Tiekura, D. Krulevska, A. Yerchenko, A. Mishchenko, D. Shcherbin, A. Priymak //

Proceedings of 12th IEEE EWDT Symposium. – Kiev, 2014. – P. 67–81.

18. *Toward Scalable Systems for Big Data Analytics: A Technology Tutorial [Электронный ресурс] / Hu Han, Yonggang Wen, Tat-Seng Chua, Xuelong Li.* – Режим доступа: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6842585>. – 1.04.2016.

19. Gebara, F. H. IBM Research–Austin. *Second-Generation Big Data Systems / F. H. Gebara, H. P. Hofstee, K. J. Nowka // IEEE Computer magazine.* – 2015. – Vol. 48, no. 1. – P. 36–41.

References

1. Alkhatib, H., Faraboschi, P., Frachtenberg, E., Kasahara, H., Lange, D., Laplante, P., Merchant, A., Milojevic, D., Schwan, K. IEEE CS 2022 Report. *Computer*, 2015, vol. 48, iss. 3, pp. 68–76.

2. *Rost ob'ema informatsii - realii tsifrovoy vselennoi* [The growth of information is the realities of the digital universe]. Available at: <http://www.tsonline.ru/articles2/fix-corp/rost-obema-informatsii--realii-tsifrovoy-vselennoy#sthash.rpNodQLF.dpuf> (accessed 1.04.2016).

3. Maier-Shenberger, V., Kennet, K. *Bol'shie dannye. Revolyutsiya, kotoraya izmenit to, kak my zhivem, rabotaem i myslim* [Big Data: A Revolution that Will Transform How We Live, Work and Think]. Mann, Ivanov i Ferber Publ., 2013. 240 p.

4. Demchenko, Y., de Laat, C., Membrey, P. Defining architecture components of the Big Data Ecosystem. *International Conference on Collaboration Technologies and Systems (CTS)*, 2014, pp. 104–112.

5. Grolinger, K., Hayes, M., Higashino, W. A., L'Heureux, A., Allison, D. S., Capretz, M. A. M. Challenges for MapReduce in Big Data. *IEEE World Congress on Services (SERVICES)*, 2014, pp. 182–189.

6. Zhang, L. A framework to specify big data driven complex cyber physical control systems. *International Conference on Information and Automation (ICIA)*, 2014, pp. 548–553.

7. Zhang, L. Designing big data driven cyber physical systems based on AADL. *International Conference on Systems, Man and Cybernetics (SMC)*, 2014, pp. 3072–3077.

8 Michalik, P., Stofa, J., Zolotova, I. Concept definition for Big Data architecture in the education

system. *12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*, 2014, pp. 331–334.

9. Munoz, M. Space systems modeling using the Architecture Analysis & Design Language (AADL). *International Symposium on Software Reliability Engineering Workshops (ISSREW)*, 2013, pp. 97–98.

10. *PureData System for Analytics*. Available at: http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?sutype=SP&infotype=PM&appname=SWGE_WA_UZ_USEN&htmlfid=WAS12385USEN&attachment=WAS12385USEN.PDF#loaded. (accessed 1.04.2016).

11. Nielsen, M. A., Chuang, I. L. *Quantum Computation and Quantum Information*. Cambridge University Press Publ., 2010. 676 p.

12. Stenholm, S., Kalle-Antti, S. *Quantum approach to informatics*. John Wiley & Sons, Inc. Publ., 2005. 249 p.

13. Hahanov, V. I., Litvinova, E. I., Chumachenko, S. V., Abbas, B. A. A., Mandefro, E. A. Qubit Model for solving the coverage problem. *Proc. of IEEE East-West Design and Test Symposium*, Kharkov, 2012, pp. 142–144.

14. Hahanov, V., Litvinova, E., Gharibi, W., Chumachenko, S. Big Data Driven Cyber Analytic System. *IEEE International Congress on Big Data*, New York, 2015, pp. 615–622.

15. Bondarenko, M. F., Hahanov, V. I., Litvinova, E. I. Struktura logicheskogo assotsiativnogo mul'tiprotssora [The structure of the logical association multiprocessor]. *Avtomatika i telemekhanika*, 2012, no. 10, pp. 71–92.

16. Hellinger, A., Seeger, H. *Cyber-Physical Systems. Driving force for innovation in mobility, health, energy and production*. Springer Publ., 2011. 48 p.

17. Hahanov, V., Gharibi, W., Kudin, A.P., Hahanov, I., Ngene, C., Tiekura, Y., Krulevska, D., Yerchenko, A., Mishchenko, A., Shcherbin, D., Priymak, A. Cyber Physical Social Systems – Future of Ukraine. *Proc. of 12th IEEE EWDT Symposium*, Kiev, 2014, pp. 67–81.

18. Han, Hu, Wen, Yonggang, Chua, Tat-Seng, Li, Xuelong. *Toward Scalable Systems for Big Data Analytics: A Technology Tutorial*. Available at: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6842585> (accessed 1.04.2016).

19. Gebara, F. H., Hofstee, H. P., Nowka, K. J. IBM Research–Austin. *Second-Generation Big Data Systems. IEEE Computer magazine*, 2015, vol. 48, no. 1, pp. 36–41.

Поступила в редакцію 1.04.2016, рассмотрена на редколлегии 14.04.2016

ПРОЦЕСОРНІ СТРУКТУРИ ДЛЯ АНАЛІЗУ BIG DATA

*В. І. Хаханов, Є. І. Литвинова, С. В. Чумаченко,
І. Ємельянов, Тамер Бані Амер*

Пропонується інфраструктура забезпечення паралельного аналізу big data для пошуку, розпізнавання та прийняття рішень на основі використання булеаної метрики вимірювання кіберпростору. Вона характеризується наявністю єдиної логічної хог-операції для визначення кібер відстані шляхом циклічного замикання не менше одного об'єкта, що дає можливість на порядок підвищити швидкодію аналізу великих даних. Запропоновано новий підхід векторно-логічної обробки big data з повним виключенням арифметичних операцій, що впливають на швидкодію і апаратну складність, який може бути ефективно реалізований як на

основі використання сучасних мультипроцесорних цифрових систем на кристалах, так і за допомогою віртуальних паралельних процесорів, що функціонують під управлінням кіберфізичних систем або хмарних сервісів-фільтрів.

Ключеві слова: великі дані, кіберінформаційна система, кіберфізичний комп'ютинг, мультипроцесор, кіберпростір, векторно-логічна обробка великих даних.

PROCESSOR STRUCTURES FOR BIG DATA ANALYSIS

*V. I. Hahanov, E. I. Litvinova, S. V. Chumachenko,
I. Yemelianov, Tamer Bani Amer*

An infrastructure for parallel analyzing big data in order to search, pattern recognition and decision-making based on the use of the Boolean metric of cyberspace measurement is proposed. It is characterized by using only logical xor-operation for determining the cyber-distance by means of cyclic closing at least one object, which allows significantly increasing the speed of analysis of large data. A new approach to vector-logical big data processing based on the total removal of arithmetic operations, which impact on the performance and hardware complexity, is offered; the approach can be efficiently implemented based on modern multiprocessor digital systems-on-chips and virtual parallel processors, operating under the control of cyber-physical systems or cloud service-filters.

Key words: big data, cyber information system, cyber physical computing, multiprocessor, cyberspace, vector-logical processing big data.

Хаханов Владимир Иванович – д-р техн. наук, декан факультета комп'ютерної інженерії та управління, професор кафедри автоматизації проектування вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина, e-mail: hahanov@icloud.com

Литвинова Евгения Ивановна – д-р техн. наук, професор кафедри автоматизації проектування вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина, e-mail: litvinova_eugenia@mail.ru.

Чумаченко Светлана Викторовна – д-р техн. наук, заведуюча кафедрой автоматизації проектування вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина, e-mail: ch_s_v@mail.ru

Емельянов Игорь – аспірант кафедри автоматизації проектування вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина, e-mail: apot@kture.kharkov.ua.

Тамер Бани Амер – аспірант кафедри автоматизації проектування вычислительной техники Харьковского национального университета радиоэлектроники, Харьков, Украина, e-mail: hahanov@kture.kharkov.ua.

Hahanov Vladimir – D. Sc., Dean of Computer Engineering Faculty, Professor of Design Automation Department of Kharkov National University of Radio Electronics, Kharkov, Ukraine, e-mail: hahanov@icloud.com.

Chumachenko Svetlana V. – D. Sc., Head of Design Automation Department of Kharkov National University of Radio Electronics, Kharkov, Ukraine, e-mail: ch_s_v@mail.ru.

Litvinova Eugenia I. – D. Sc., professor of Design Automation Department of Kharkov National University of Radio Electronics, Kharkov, Ukraine, e-mail: litvinova_eugenia@mail.ru.

Yemelianov Igor – PhD student of Design Automation Department of Kharkov National University of Radio Electronics, Kharkov, Ukraine, e-mail: apot@kture.kharkov.ua.

Tamer Bani Amer – PhD student of Design Automation Department of Kharkov National University of Radio Electronics, Kharkov, Ukraine, e-mail: hahanov@kture.kharkov.ua.