

**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ИНСТИТУТ  
ЭЛЕКТРОНИКИ И МАТЕМАТИКИ  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**ФОНД СОДЕЙСТВИЯ РАЗВИТИЮ МАЛЫХ ФОРМ ПРЕДПРИЯТИЙ  
В НАУЧНО-ТЕХНИЧЕСКОЙ СФЕРЕ**

**Научно-техническая  
конференция студентов, аспирантов  
и молодых специалистов НИУ ВШЭ**

**МАТЕРИАЛЫ КОНФЕРЕНЦИИ**



**Москва 2014г.**

**УДК 658.012; 681.3.06; 621.396.6.001.66(075); 621.001.2(031)**  
**ББК 2+3**  
**Н 34**

**Научно-техническая конференция студентов, аспирантов и молодых специалистов НИУ ВШЭ. Материалы конференции. - М. ~: МИЭМ НИУ ВШЭ, 2014. - 310.**

**ISBN 978-5-94768-062-1**

В материалах конференции студентов, аспирантов и молодых специалистов НИУ ВШЭ представлены тезисы докладов по следующим направлениям: прикладная математика; информационно-коммуникационные технологии; автоматизация проектирования, банки данных и знаний, интеллектуальные системы; компьютерные образовательные продукты; информационная безопасность; электроника и приборостроение; производственные технологии, нанотехнологии и новые материалы; современные технологии дизайн проектирования; информационные технологии в экономике, бизнесе и инновационной деятельности.

Материалы конференции могут быть полезны для преподавателей, студентов, научных сотрудников и специалистов, специализирующихся в области прикладной математики, информационно-коммуникационных технологий и электроники.

Редакционная коллегия: Тихонов А.Н., Азаров В.Н., Аристова У.В., Карасев М.В.,  
Кулагин В.П., Леохин Ю.Л., Львов Б.Г., Титкова Н.С.

Издание осуществлено с авторских оригиналов.

**ISBN 978-5-94768-062-1**

**ББК 2+3**  
**© Московский институт электроники и математики Национального исследовательского университета «Высшая школа экономики», 2014 г.**  
**© Авторы, 2014г.**

XML для представления данных, XDR для сжатия данных, RRDtool для хранения и визуализации данных. К сожалению, данная система лишена систем оповещения.

Zabbix — это открытое решение распределенного мониторинга корпоративного класса. [4] Zabbix обеспечивает мониторинг многочисленных параметров сети а также состояния и работоспособности серверов. Несмотря с возможностями масштабирования при больших объемах метрик возникают проблемы.

Nagios — это приложение мониторинга компьютерных систем и сетей с открытым кодом. Предназначена для наблюдения, контроля состояния вычислительных узлов и служб, а так же оповещения администратора в том случае, если какие-то из служб прекращают (или возобновляют) свою работу. [5] Nagios является популярной системой из-за своего мощного механизма оповещений.

OpenNMS — система мониторинга сетевой инфраструктуры уровня предприятия, распространяемая по модели свободного программного обеспечения (Open Source). [6,7] Система OpenNMS отвечает за мониторинг функционирующих в сетевой инфраструктуре сервисов, таких как Web, DNS, DHCP, сервисы СУБД. В системе упрощено добавление новых сетевых устройств благодаря автоматическому обнаружению.

Проект Graphite выполняет две довольно простые задачи: хранение данных, изменяемых со временем, и отображение их в виде графиков. Graphite предоставляет функции в виде сервиса, который прост в использовании и хорошо масштабируем. [8,9] Graphite прекрасно интегрируется с другим программным обеспечением и позволяют возможность собирать свою собственную мощную систему на базе Graphite.

#### **Интеллектуализация мониторинга**

Существует один интересный open-source проект - Kale Stack, который состоит из двух частей: Skyline и Oculus. Цель данного проекта решить проблему мониторинга огромного числа графиков. [10]

Skyline - это система для обнаружения аномальных показателей. Она показывает все текущие показатели, которые были определены как аномальные. Skyline в режиме реального времени позволяет осуществлять мониторинг сотен метрик без необходимости настраивать пороги срабатывания для каждой метрики, как это требуется в Nagios. Skyline автоматически определяет отклонения от нормы для каждого показателя. Skyline опирается на консенсус ансамбля нескольких различных алгоритмов. Если большинство алгоритмов соглашается, что данная метрика является аномальной, то метрика будет классифицирована как аномальная.

На данный момент система имеет небольшой набор алгоритмов: среднее значение за час, стандартное отклонение от среднего, стандартное отклонение от скользящего среднего, тест Граббса и другие. Также система позволяет добавлять свои алгоритмы.

Оперируя данным набором алгоритмов можно подобрать оптимальный набор для различных данных, что позволит избежать ложных срабатываний системы.

Oculus позволяет сравнить подозрительную метрику с другими метриками. В результате можно сделать обоснованный анализ и исправить проблему.

#### **Заключение**

Каждая система имеет свои достоинства и недостатки. Для наиболее эффективного использования и решения поставленной задачи можно построить систему, комбинируя и используя различные системы мониторинга, которые позволяют наиболее эффективно выполнить поставленную задачу.

Ввиду постоянно растущих объемов данных, средства обычного мониторинга не так эффективны, так как постоянное наблюдение за десятками, а то и сотнями графиков является бесполезным или физически не возможным занятием. Решение данной проблемы лежит в основе интеллектуального мониторинга.

#### **Список литературы:**

1. Сравнение различных систем мониторинга. // Википедия - электронная энциклопедия. URL: [http://en.wikipedia.org/wiki/Comparison\\_of\\_network\\_monitoring\\_systems](http://en.wikipedia.org/wiki/Comparison_of_network_monitoring_systems) (дата обращения: 20.11.2013)

2. Андрей Бородийчук. Система мониторинга Cacti. 12.05.2006 URL: <http://hostinfo.ru/articles/773> (дата обращения 20.11.2013).

3. Ganglia. // Википедия - электронная энциклопедия. URL: <http://ru.wikipedia.org/wiki/Ganglia> (дата обращения 20.11.2013).

4. Универсальная система мониторинга Zabbix — введение. // Сообщество профессионалов habrahabr.ru, 25.10.2009. URL: <http://habrahabr.ru/post/73338> (дата обращения 20.11.2013).

5. Nagios. // Википедия - электронная энциклопедия. URL: <http://ru.wikipedia.org/wiki/Nagios> (дата обращения 20.11.2013).

6. Андрей Семенов. Устанавливаем и настраиваем систему мониторинга сети OpenNMS. База статей по компьютерной тематике OpenNET, 30.08.2008 URL: [http://www.opennet.ru/base/net/opennms\\_monitor.txt.html](http://www.opennet.ru/base/net/opennms_monitor.txt.html) (дата обращения 20.11.2013).

7. Андрей Семенов. Используем систему мониторинга сети OpenNMS. База статей по компьютерной тематике OpenNET, 24.09.2008 URL: [http://www.opennet.ru/base/net/opennms\\_monitor2.txt.html](http://www.opennet.ru/base/net/opennms_monitor2.txt.html) (дата обращения 20.11.2013).

8. Дмитрий Куликовский. Graphite — как построить миллион графиков. // Доклад на конференции Яндекс: Yet another Conference, 02.10.2013 URL: <http://tech.yandex.ru/events/yac/2013/talks/1122/> (дата обращения 20.11.2013).

9. Крис Девис. Graphite. // The Architecture of Open Source Applications. Под ред. Эми Браун и Грера Уилсона. - lulu.com, 2011. ISBN 978-1-257-63801-7. // Пер. Н. Ромоданова. URL: <http://rus-linux.net/MyLDP/BOOKS/Architecture-Open-Source-Applications/Vol-1/graphite-01.html> (дата обращения 20.11.2013).

10. Abe, Jon. Introducing Kale. // Code as Craft, 2013. URL: <http://codeascraft.com/2013/06/11/introducing-kale> (дата обращения 20.11.2013).

## **ТУМАННЫЕ ВЫЧИСЛЕНИЯ И БЕСПРОВОДНЫЕ СЕНСОРНЫЕ СЕТИ**

*А.А. Дворников*

*НИУ ВШЭ,*

*Факультет информационных технологий  
и вычислительной техники МИЭМ НИУ ВШЭ*

#### **Аннотация**

В данной работе беспроводные сенсорные сети предложены как основа для туманных вычислений. В ней перечисляются функции, которыми не обладают современные беспроводные сенсорные сети для соответствия парадигме «туманных вычислений». В качестве решения приводятся механизмы, которые позволят приблизить беспроводные сенсорные сети к парадигме «туманных вычислений».

## Введение

Туманные вычисления — явление, появившееся в 2012 году [1]. Оно описывает распределённые вычисления, выполняющиеся на оконечных устройствах, имеющих ограниченные ресурсы и непосредственную связь с физическим миром.

Появление туманных вычислений стало возможным благодаря современной тенденции интеграции сетевых технологий во всё большее количество бытовых, промышленных, устройств, устройств личного пользования и иных устройств, которые раньше с сетевыми функциями не ассоциировались.

На практике туманные вычисления позволяют предоставить прозрачный доступ к сенсорам и актуаторам на различных оконечных устройствах, производить часть вычислений в самом «тумане», а также создавать вычислительные сети поверх наборов оконечных устройств. Подобные решения позволят повысить качество работы существующих специализированных систем, таких как системы класса «умный дом», «умный город», «умные производства» [2], позволяя им взаимодействовать с внешним миром и использовать различные интеллектуальные механизмы.

В данной работе автор обосновывает, почему именно беспроводные сенсорные сети он считает перспективной основой для туманных вычислений и приводит характеристики, которыми пока не обладают беспроводные сенсорные сети для соответствия данной парадигме, и существующие решения данных проблем.

### Беспроводные сенсорные сети

Технология беспроводных сенсорных сетей прочно заняла своё место в современной технике и науке. Известно немало сфер, где беспроводные сенсорные сети занимают одно из лидирующих положений — это системы интеллектуализации зданий и городов, интеллектуальные сети, системы мониторинга объектов и др.

Сегодня беспроводные сенсорные сети считаются наиболее перспективными, чем их проводные решения. Этот переход выделил энергосбережение и энергоэффективность [3] одним из наиболее приоритетных параметров для сенсорной сети и позволил использовать альтернативные источники энергии.

В настоящее время, по мнению автора, беспроводные сенсорные сети являются наиболее перспективной платформой для туманных вычислений.

### Парадигма Туманных вычислений

Туманные вычисления — это «платформа с высоким уровнем виртуализации, предоставляющая вычислительные службы, службы хранения информации, а также сетевые службы между оконечными устройствами и центрами облачных вычислений, не обязательно расположенных на крайних уровнях сети» (из статьи CISCO «Fog computing and its role in the internet of things»).

Из Рис. 1 видно, что «туман», это «облако», расположенное возле земли (физического мира). И, если облачные вычисления не имеют дела с физическими величинами и физическими действиями непосредственного, то «туман», наоборот, не только контактирует с физическим миром, но и имеет пространственную привязанность (знание расположения имеет смысл).

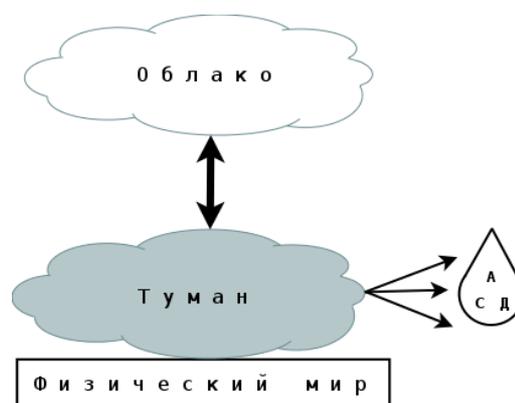


Рис.1. Сравнение облачных вычислений и туманных вычислений

Каждая составляющая («капля») «тумана» является самостоятельным вычислителем, обладающим вычислительной мощностью (часто довольно малой), имеет свой адрес, является сенсором и (или) актуатором (обладает набором действий).

В статье из CISCO «Fog computing and its role in the internet of things» приводятся следующие характеристики парадигмы туманных вычислений, которые отличают вычислительное «облако» от вычислительного «тумана»:

1. Крайнее положение, осведомлённость о своей позиции и малые задержки в сети.

К сожалению, проблема больших задержек в сети неразрешима для существующих радиостандартов, кроме семейства протоколов IEEE 802.11 (WiFi).

2. Географическая распределённость.

Географическая распределённость достижима современными протоколами для беспроводных сенсорных сетей за счёт включённых в них протоколов ретрансляции.

3. Большое количество сенсоров.

4. Очень большое количество узлов.

5. Поддержка мобильности.

6. Поддержка реального времени.

Большинство существующих операционных систем для беспроводных сенсорных сетей являются операционными системами реального времени.

7. Предпочтительным считается беспроводное взаимодействие.

8. Гетерогенность.

9. Взаимодействие с другими типами сетей и их объединение.

За счёт использование мостов беспроводные сенсорные сети хорошо объединяются, но из-за отсутствия единого стандарта протоколов сети данная функция на настоящие момент не используется полностью.

10. Поддержка вычислений на узлах и взаимодействие с облаком.

На данный момент есть определённые трудности с унификацией распределённых вычислений на узлах и их взаимодействием с облаком.

### Задержки в беспроводных сенсорных сетях

Проблема больших задержек в беспроводных сенсорных сетях связана, в первую очередь, с низкой пропускной способностью сетевых стандартов. Приоритетным направлением в БСС является высокая энергоэффективность и как плата за это - невысокая пропускная способность. Исключением из данного правила являются беспроводные сенсорные сети на стандартах семейства 802.11 (WiFi), потребляющие при этом больше электроэнергии, чем иные сетевые стандарты для

беспроводных сенсорных сетей.

### **Географическая распределённость беспроводных сенсорных сетей**

Большинство протоколов беспроводной сенсорной сети поддерживает ретрансляцию данных по сети, что позволяет покрывать сетью большие площади, несмотря на небольшой радиус действия одного узла.

Это достаточно эффективный способ увеличения зоны покрытия узла, но существуют дополнительные алгоритмы, позволяющие оптимизировать зону покрытия в беспроводных сенсорных сетях.

Одно из семейств подобных алгоритмов носит название роевых, так как в них используются механизмы роевого взаимодействия из средств искусственного интеллекта. Примером такого алгоритма является алгоритм поведения пчелиного роя, где мобильные агенты стремятся обеспечить максимальное покрытие вокруг точки интереса [4].

### **Взаимодействие беспроводных сенсорных сетей с другими сетями и их объединение**

Большинство протоколов беспроводных сенсорных сетей не обладают прозрачной совместимостью с сетями TCP/IP и друг с другом.

TCP/IP достаточно сложен в реализации, из-за чего многие производители оборудования не берутся за подобные решения. Тем не менее, существует свободная реализация стека TCP/IP под устройства с ограниченными ресурсами — uIP, которая поставляется с некоторыми операционными системами. Стек прозрачен по отношению к TCP/IP сетям и в нём реализован механизм сокетов. Существуют и другие реализации TCP/IP под устройства с ограниченными ресурсами. Для взаимодействия с сетями TCP/IP также существуют сетевые мосты.

Из-за отсутствия стандартов на протоколы для беспроводных сенсорных сетей каждый производитель стремится выпустить свой собственный стек протоколов, что приводит к отсутствию совместимости с устройствами других производителей. Тем не менее, сети можно объединять с помощью сетевых мостов.

### **Поддержка вычислений на узлах и взаимодействие с облаком в беспроводных сенсорных сетях**

На сегодняшний день не существует единых стандартов и универсальных решений для вычислений на узлах и связи с облаком. Как правило, для каждого решения используется свой собственный протокол.

Тем не менее, существует реализация распределённой базы данных на узлах беспроводной сенсорной сети — TinyDB [5], которая выступает основой для распределённых вычислений на узлах.

#### **Заключение**

В работе приведено определение туманных вычислений, перечислены аргументы, согласно которым именно платформа беспроводных сенсорных сетей должна стать основой «туманных вычислений».

Проанализированы характеристики беспроводных сенсорных сетей, которых не хватает для соответствия парадигме туманных вычислений. Найдены решения, способные восполнить недостатки платформы беспроводных сенсорных сетей, мешающие их применения в парадигме «туманных вычислений».

Данное научное исследование (проект № 14-05-0064) выполняется при поддержке Программы «Научный фонд НИУ ВШЭ» в 2014г.

#### **Список литературы:**

1. Bonomi F. и др. Fog computing and its role in the internet of things // Proceedings of the first edition of the MCC workshop on Mobile cloud computing. , 2012. С. 13–16.

2. «Умные» среды, «умные» системы, «умные»

производства. Санкт-Петербург: Фонд «Центр стратегических разработок «Северо-Запад», 2012.

3. Восков Л. С., Ефремов С. Г. Задача увеличения времени автономной работы беспроводных сенсорных сетей в системах сбора данных и способ ее решения // Датчики и системы. 2013. № 4(167). С. 2-9.

4. Ozturk C., Karaboga D., Gorkemli B. Artificial bee colony algorithm for dynamic deployment of wireless sensor networks // Turk J Electr Eng Comput Sci. 2012. Т. 20. № 2. С. 1–8.

5. Qela B. Adaptive Systems for Smart Buildings Utilizing Wireless Sensor Networks and Artificial Intelligence // 2012.

## **ПРИМЕНЕНИЕ MRMR КРИТЕРИЯ ОТБОРА ПРИЗНАКОВ В ЗАДАЧЕ КЛАССИФИКАЦИИ СПАМА В ИНТЕРНЕТЕ**

*И.О. Карбачинский*

*НИУ ВШЭ,*

*Факультет прикладной математики  
и кибернетики МИЭМ НИУ ВШЭ*

Ежедневно в сети появляются тысячи новых сайтов, содержащих спам и некачественный контент. Существует множество различных видов спама, но в данной работе под спамом будут подразумеваться web страницы, содержащие вредоносные, некачественные и бесполезные для пользователя данные. Антиспам система как правило представляет собой бинарный классификатор, обученный на некотором множестве web документов. Для обучения классификатора, вне зависимости от используемого алгоритма, необходимо из каждого элемента обучающего множества выделить некоторый набор информативных признаков. Данные признаки позволяют численно охарактеризовать каждый документ в обучающем множестве. Можно придумать очень много таких признаков. Однако, не все они будут полезны для обучения классификатора. Поэтому существует множество различных способов, позволяющих осуществить отбор наиболее релевантных признаков. Одним из таких способов является применение критерия MRMR (Maximum Relevance Minimum Redundancy) отбора признаков. В данной работе показывается эффективность данного критерия применительно к задаче классификации спама и показывается, что критерий работает даже в условиях относительно малого числа информативных признаков, позволяя значительно повысить качество классификации.

Для того, чтобы определить, что некоторый информативный признак лучше другого, необходимо определить некоторую численную меру, характеризующую качество произвольно выбранного признака или множества признаков в контексте решаемой задачи. Интуитивно может показаться, что чем больше значения признака отличаются на документах разных классов, тем лучше этот признак. Однако, это не совсем так. Во-первых признаки могут коррелировать друг с другом, и не имеет никакого смысла использовать в процессе обучения избыточные признаки. Примером таких признаков могут являться, например, частоты некоторых фиксированных слов на странице. Во-вторых, значения различных признаков могут варьироваться на совершенно различных подмножествах обучающего множества. И даже если признак хорошо разделяет лишь малое подмножество документов из разных классов, он может быть очень полезен. Также немаловажную роль играет выбранный алгоритм обучения. Необходимо отбирать признаки с учетом особенностей работы выбранного

|   |        |
|---|--------|
| <b>Кузьмина Ю.О.</b> Компьютерное моделирование процесса всестороннейковки стали AISI304  | 32-33  |
| <b>Цикина Н.В.</b> Методы оценки энтропии печатного текста  | 33-34  |
| <b>Гайдуков Р.К.</b> Существование решения уравнения типа Релея   | 35     |
| <b>Малкин Р.Е.</b> Метод обработки информации для мониторинга лесных пожаров  | 35-36  |
| <b>Тарасов Н.А. Тарасов И.Н.</b> Новая методика статистического доказательства логнормальности биржевых колебаний с помощью фрактальной бинарной сети | 37-38  |
| <b>Дорофеев П.А.</b> Исследование задачи о пропускной способности регулируемого перекрестка методами математического моделирования                    | 38-39  |
| <b>Захарьев И.Ю.</b> Интерпритация данных эксперимента по газовой формовки сплава АМГ6 для определения условий его сверхпластичности                  | 39-40  |
| <b>Ефремова Е.В.</b> Моделирование некомпланарного многоимпульсного перелета за орбиту Луны   | 41     |
| <b>Деркач А.В.</b> Новая математическая модель дисконтной облигации   | 42     |
| <b>Ермаков П.Д.</b> Выделение ключевых понятий из неструктурированных текстов на примере выделения навыков и требований из текстов резюме и вакансий  | 43     |
| <b>Семёнов В.П.</b> Исследование резонансных явлений в расширенном уравнении Матье  | 43-45  |
| <b>Калашников Г.А.</b> Регуляризация уравнений Дарси в случае микроскопической гидродинамической неустойчивости Релея-Тейлора                         | 45-46  |
| <b>Быков Ю.В.</b> Решётки полного ранга, инвариантные относительно конечных кватернионных групп   | 46-48  |
| <b>Тимербулатов Д.Р.</b> Анализ информации о дальностях для решения навигационных задач   | 48-49  |
| <br>  |        |
| <b><u>Секция 2 "Информационно-коммуникационные технологии"</u></b>  | 50-165 |
| <b>Ролич А.Ю.</b> Проектирование энергоэффективных пользовательских интерфейсов интернет-вещей  | 50-51  |
| <b>Колесов С.Н.</b> Эффективный open-source мониторинг  | 51-52  |
| <b>Дворников А.А.</b> Туманные вычисления и беспроводные сенсорные сети   | 52-54  |
| <b>Карбачинский И.О.</b> Применение MRMR критерия отбора признаков в задаче классификации спама в интернете   | 54-55  |
| <b>Гераськин П.Б.</b> Анализ систем управления кредитными рисками на основе ансамблей методов машинного обучения                                      | 55-56  |
| <b>Громов И.Ю.</b> Метод синтеза систем обеспечения тепловых режимов радиоэлектронной аппаратуры  | 57     |
| <b>Калачёв Я.Б.</b> Метод проверки отчетов о проделанной работе по техническому заданию на полноту  | 58-60  |
| <b>Луцкий В.А.</b> Исследование возможностей интеллектуального анализа данных при адаптивном передвижении шестиногого шагающего робота                | 60-61  |
| <b>Николаев П.В. Лышов С.М.</b> Математическое моделирование процесса вибродиагностики латентных дефектов в конструкциях электронных средств          | 62     |
| <b>Горохова-Алексеева А.В.</b> Алгоритм автоматической коррекции цветовой температуры завершенных видеопотоков  | 63     |