УДК 004.056.5 + 004.7.056.5

МЕТОД ИДЕНТИФИКАЦИИ БОТНЕТОВ НА ОСНОВЕ МНОГОАГЕНТНОГО ПОДХОДА

М. Ю. Косенко, А. В. Мельников

Челябинский государственный университет

Поступила в редакцию 25.01.2015 г.

Аннотация. В работе представлен метод обнаружения ботнетов на основе многоагентного подхода. Данный метод основывается на работе распределенной системы обнаружения и блокирования атак типа «отказ в обслуживании» с последующим определением признаков атакующего бота и выявлением по этим признакам участников ботнета в сети Интернет.

Ключевые слова: атака типа «отказ в обслуживании», ботнет, идентификация ботнета, многоагентная система.

Annotation. In this paper presents a method for detecting botnets based on multi-agent approach. This method is based on the work of distributed system to detect and block denial of service attacks, followed by some signs of attacking bot detection and on these grounds participants botnet on the Internet.

Keywords: denial of service attacks, botnet, identification of botnet, multi-agent system.

ВВЕДЕНИЕ

Одной из самых опасных угроз безопасности в сети Интернет являются ботнеты. Ботнет – это сеть, состоящая из скомпрометированных хостов, управляемых некоторым вредоносным программным обеспечением (бот). Бот является частично автономной частью вредоносного программного обеспечения, которое контролируется удаленно. Ботнеты создаются путем инфицирования компьютера без ведома и согласия его владельца, например, рассылка вирусов прикрепленных к электронным письмам [1]. Применяются подобные сети для множества зловредных действий: абузоустойчивый хостинг (размещение порнографического, террористического контента), проведение всех атак типа «распределенный отказ в обслуживании», заражение шпионским вредоносным программным обеспечением, хищение конфиденциальных данных, широкомасштабная рассылка спама, «накручивание» кликов.

В настоящее время злоумышленниками отработаны техники построения ботнетов. Примеры известных зараженных сетей [2]:

- Ботнет Bredolab. Состоял примерно из 30 миллионов компьютеров. Использовался для организации DDOS-атак (в том числе на ряд российских новостных ресурсов, на сайт «Лаборатории Ксаперского»). В 2010 году принес своему создателю прибыль в размере 700 тысяч долларов. Методы распространения: эксплойт-пак, модификация страниц сайтов с ссылками на вредоносное программное обеспечение;
- Ботнет Storm. Использовался для рассылки спама, организации DDOS-атак. Около 2 миллионов зараженных компьютеров. Распространялся путем почтовых рассылок.
- Ботнет Spyeye. Использовался для банковского мошенничества. Состоял из 1,4 миллиона компьютеров.
- Ботнет Mariposa. Использовался для кражи конфиденциальной информации, организации DDOS-атак. Состоял из 12 миллионов компьютеров.
- Ботнет Zeroaccess. Использовался для «накручивания» кликов, рассылки спама, до-

бывания bitcoin. Состоял примерно их 9 миллионов компьютеров;

• Ботнет Zeus. Использовался для банковского мошенничества. Состоял примерно из 13 миллионов компьютеров;

• А также, DDOS-ботнет DRAGON, спам-ботнет Grum (около 1 млн. инфицированных компьютеров, 18 млрд. писем в месяц на пике работы), спам-ботнет Virut (более 300000 зараженных устройств) [3].

Такое многообразие различных ботнетов показывает, что создание и использование ботнетов – это одно из наиболее популярных направлений кибер-преступности. Таким образом, одной из актуальных задач на сегодняшний день становится задача идентификации участников ботнетов.

К наиболее распространенным методам идентификации ботнетов относится: анализ телеметрии, обнаружение аномалий, анализ журнала серверов DNS, система Honeypot.

Анализ телеметрии. Метод заключается в использовании сводной информации сетевого и транспортного уровня от сетевых устройств. К примеру, технология NetFlow часто применяется для обнаружения трафика DDOS атак, всплесков трафика SMTP, характерного для массовой рассылки спама и управляющего трафика контроллера ботнета. Проведено много исследований по использованию NetFlow для идентификации ботов, к примеруDISCLOSURE [4], Bottrack [5]. Методы представленные в этих исследованиях позволяют либо обнаруживать управляющие центры ботнетов, либо идентифицировать бота в рамках, пусть и достаточно больших, но корпоративных сетях. Метод анализ телеметрии не позволяет эффективно масштабировать систему идентификации ботнетов для крупных сетей.

Обнаружение аномалий. В отличии от подхода на основе сигнатур, заключающегося в сопоставлении каждой атаки с имеющейся базой данных сигнатур, обнаружение аномалий заключается в обратном: описываются характеристики обычного трафика, а затем выполняется поиск отклонений. Использование такого подхода обеспечивает обнаружение и блокировку DDOS атак и попыток массового сканирования, предпринимаемых ботнетами. Данный метод хорош тем, что позволяет идентифицировать неизвестных ботов, но при этом допускается большое количество ложных срабатываний. В связи с чем, на методы, основанные на обнаружении аномалий, нельзя возложить решение задачи идентификации ботнетов в динамическом режиме.

Анализ журнала сервера DNS. Ботнеты часто используют бесплатные службы DNS, чтобы разместить адрес поддомена управляющих серверов, от которых можно принимать команды управления и получать обновления вредоносного кода. Часто код бота содержит жестко заданные ссылки на DNS-сервер, которые могут быть легко найдены любым средством анализа журнала DNS-запросов. При обнаружении таких служб администратор DNS-сервера может нейтрализовать ботнет путем переадресации зловредных поддоменовна несуществующий IP-адрес. Не смотря на эффективность этого метода, его сложно применять, потому что требуется сотрудничество со службами регистрации доменных имен и сторонними хостинг-провайдерами.

Система Honeypot. Ресурс-приманка, представляемый замкнутую, защищенную и контролируемую область, имитирующую уязвимую сеть, ресурс или сервис. Основная цель – приманить и обнаружить вредоносные атаки и попытки вторжения. Системы данного типа успешно могут использоваться для получения экземпляров вредоносного программного обеспечения используемого для организации ботнетов. Дальнейшее исследование полученных экземпляров позволяет получить характеристики работы бота, которые могут использоваться для идентификации ботнетов. Но сам по себе подход не является прямым методом идентификации ботнетов.

ИДЕНТИФИКАЦИЯ УЧАСТНИКОВ БОТНЕТОВ НА ОСНОВЕ МНОГОАГЕНТНОГО ПОДХОДА

Показанные проблемы использования существующих методов позволяют предло-

жить другой метод идентификации ботнетов, на основе многоагентного подхода. Использование многоагентного подхода позволит достичь основного преимущества – динамического решения задачи идентификации ботнетов. Также, данный подход позволит создать гибкий и масштабируемый метод.

Идея, вкладываемая в данный метод, состоит в следующем. Для того чтобы идентифицировать ботнет, в первую очередь необходимо обнаружить распределенную атаку типа «отказ в обслуживании», для осуществления которой чаще всего прибегают к использованию ботнетов. После обнаружения атаки необходимо блокировать её на стороне источника атаки, а атакующее средство взять под наблюдение для выявления характерных признаков работы бота. Далее, попытаться идентифицировать других участников ботнета путем поиска в различных сетях ранее обнаруженных признаков работы бота.

За основу при построении метода идентификации ботнетов была взята типовая структура сети Интернет основанная на взаимодействии между автономными системами. Предлагаемый метод идентификации ботов базируется на средстве защиты от распределенных атак типа «отказ в обслуживании» с возможностью обнаружения атаки в сети цели атаки, а предотвращения генерации атаки в сети источника. Проведя декомпозицию решаемой задачи, можно выделить перечень известных задач, решение которых приведет к требуемому результату:

- задача обнаружения атаки типа «распределенный отказ в обслуживании»;
- задача блокирования атаки;
- задача выявления характерных признаков работы бота;
- задача идентификации бота;
- задача координации агентов системы;
- задача контроля и мониторинга работы агентов;
- задача накопления информации;
- задача визуализации атак и ботнетов.

Целью данной работы является описание метода идентификации участников ботнетов на основе многоагентного подхода. Многоагентный подход фактически избавляет от проблем масштабирования при росте системы идентификации. Выявление набора одинаковых признаков взаимодействия ботов с контролерами ботнетов могут решить проблему автоматизации обнаружения ботов.

Признаки идентификации ботов. В качестве общих признаков ботов можно выделить:

- IP-адрес или доменное имя контролирующего центра ботнета;
- характеристики HTTP или IRC пакетов с определенными командами управления;
- размерность сетевых пакетов;
- временные интервалы сетевых взаимодействий;
- трафик злонамеренной активности, к примеру, сканирование, рассылка спама, загрузка бинарных файлов [6, 7];
- информацию протоколов DNS, SMTP;
- используемый протокол обмена данными и порты транспортного уровня.

Архитектура многоагентной системы. Полученные в процессе декомпозиции задачи можно отнести к различным классам функциональности: {Обнаружение, Блокирование, Исследование, Идентификация, Координация, Интерфейс}. Каждому классу может соответствовать свой тип агента, решающий задачи класса. Таким образом, многоагентная система идентификации ботнета имеет вид

$$MAS = \{A_{detection}, A_{blocking}, A_{discrovery},$$

$$A_{identification}, A_{coordination}, A_{interface}\},$$

где $A_{detection} = \left\{ A^1_{detection}, \ldots, A^n_{detection} \right\}$ – множество агентов обнаружения атаки типа «распределенный отказ в обслуживании». Агенты данного класса решают задачу обнаружения атак и реагируют на неё определенным в сценарии реагировании образом. В каждой автономной системе сети Интернет располагается как минимум один агент данного класса $A^i_{detection}$, где $i = 1..n$ – номер автономной системы сети Интернет.

$A_{blocking} = \left\{ A^1_{blocking}, \ldots, A^n_{blocking} \right\}$ – множество агентов решающих задачу блокирования обнаруженной атаки. В каждой автономной системе сети Интернет располагается как минимум один агент данного класса $A^i_{blocking}$, где

$i = 1..n$ – номер автономной системы сети Интернет.

$A_{discovery} = \left\{ A^1_{discovery}, \ldots, A^n_{discovery} \right\}$ – множество агентов выявления признаков бота. Класс агентов решающий задачу определения характерных признаков работы бота. В каждой автономной системе сети Интернет располагается как минимум один агент данного класса $A^i_{discovery}$, где $i = 1..n$ – номер автономной системы сети Интернет.

$A_{identification} = \{A^1_{identification}, \ldots, A^n_{identification}\}$ – множество агентов идентификации работы бота в рамках автономной системы. Агенты данного класса анализируют трафик сети на наличие признаков функционирования ботов. В каждой автономной системе сети Интернет располагается как минимум один агент данного класса $A^i_{identification}$, где $i = 1..n$ – номер автономной системы сети Интернет.

$A_{coordination}$ – множество агентов сети решающих задачу распространения информации об активных агентах.

$A_{interface}$ – множество агентов сети решающих следующие задачи: контроль и мониторинг работы сети агентов, визуализация атак, хранение информации.

Таким образом, структура многоагентной системы идентификации ботов состоит из следующих элементов:

1. Агент обнаружения атаки типа «распределенный отказ в обслуживании» $a_i \in A_{detection}$.

2. Агент выявления признаков бота $a_i \in A_{discovery}$.

3. Агент идентификации ботов $a_i \in A_{identification}$.

4. Агент блокирования атак $a_i \in A_{blocking}$. Функционирует, когда его расположение является сетью источника атаки. В частности, осуществляет реагирование на основе информации полученной от агентов обнаружения атак согласно профилю сетевой безопасности (блокирование систем задействованных в реализации атаки, оповещение ответственных лиц по электронной почте, SMS).

5. Агент координации $a_i \in A_{coordination}$. Распространяет информацию о местонахождении различных агентов с целью осуществления взаимодействия между ними.

6. Интерфейсный агент $a_i \in A_{interface}$. Устанавливается в любой точке глобальной сети Интернет. Предназначен для контроля и мониторинга работы сети агентов, предоставления графического интерфейса визуализации обнаруженных атак, хранения и обеспечения доступа к истории обнаруженных атак.

Концептуальный алгоритм функционирования системы Botnet MultiAgent Recognition (BNMAR) заключается в следующем (рис. 1):

1. Агент обнаружения атаки типа «распределенный отказ в обслуживании» обнаруживает атаку на подконтрольную ему сеть.

2. Агент обнаружения атаки сообщает агенту координации информацию о сетях источнике обнаруженной атаки.

3. Агент координации передает агентам блокирования, находящимся в соответствующих источникам атаки автономных системах, информацию об атакующем узле.

4. Агент координации передает агенту выявления признаков бота, контролирующего сеть источника атаки, информацию об атакующем узле.

5. Агент координации передает интерфейсному агенту информацию об атаке.

6. Агент блокирования прекращает злонамеренную активность узлов находящихся в контролируемой им сети.

7. Агент выявления признаков бота анализирует активность узлов замеченных в атаке. В результате чего выявляет характерные признаки работы бота.

8. Агент выявления признаков бота сообщает характерные признаки работы бота агенту координации.

9. Агент координации рассылает информацию о работе ботов агентам идентификации ботов.

10. Агенты идентификации анализируют трафик в своей сети, пробуя обнаружить полученные признаки работы бота. В случае удачной идентификации, передают информацию о боте агенту координации, который направляет её интерфейсному агенту для дальнейшего принятия решения.

Описание моделей системы идентификации. В работе выделены следующие модели: модели всех упомянутых выше агентов, а так

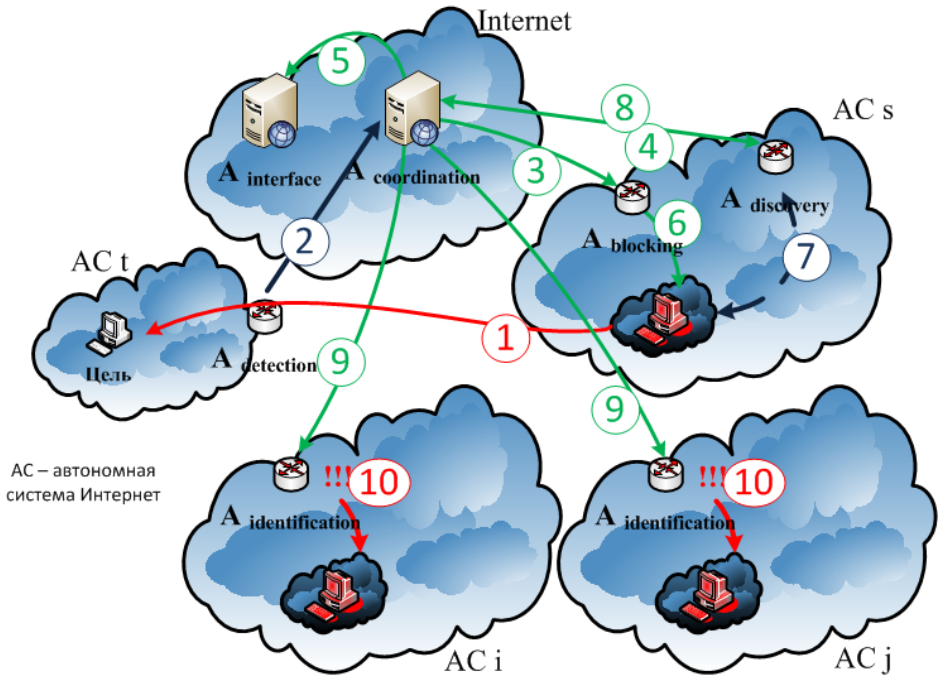

Рис. 1. Схема многоагентной системы идентификации ботнетов

Рис. 2. Представление основных моделей системы идентификации ботнетов

же модель взаимодействия агентов (рис. 2). Модели агентов предназначены для представления процессов решаемых агентами задач. Они включают частные онтологии агентов, базовые функции и специальные функции агентов, протоколы их взаимодействия, сценарии их поведения.

Предлагаемый перечень базовых функций агентов включает следующие функции: функции инициализации, окончания работы, доступ к частной онтологии агента, контроль списка активных агентов, базовая работа с модулями транспортного уровня (создание соединения, посылка сообщения, закрытие соединения). Также для некоторых агентов предполагаются специализированные функции, основанные на базовых. Для агентов обнаружения атаки их реализация будет зависеть от используемого метода обнаружения, для агентов выявления признаков бота – от используемых методов анализа деятельности бота, для агентов блокирования – от политики реагирования на атаку, для агентов идентификации – от метода анализа сетевого трафика.

Протоколы взаимодействия агентов представляются в виде последовательности команд с определенными параметрами. Протоколы взаимодействия агентов основываются

на транспортном уровне, предоставляемом коммуникационной средой. В работе для обмена сообщениями между агентами будут использоваться протоколы TCP и UDP. Выбор протокола будет основываться на уровне затрачиваемых ресурсов коммуникации.

В работе предполагаются различные сценарии поведения агентов. В некоторых случаях, сценарии конкретных агентов будут зависеть от политики безопасности принятой в системе идентификации.

Общая модель агента. Существует целый ряд математических моделей многоагентных систем, в каждом из которых делается акцент на каком-либо аспекте системы. Согласно [8], выделяют следующие модели многоагентных систем: модели, являющиеся развитием понятия алгебраической системы по А. И. Мальцеву, «Искусственный рой» [9], модель, предложенная К. Цетнаровичем, основанная на идее трехступенчатого определения основных понятий. Наиболее адекватной для поставленной задачи является модель, основанная на понятии алгебраической системы по А. И. Мальцеву. Данная модель удачна в связи со следующими аспектами:

• Открытость [10]. Возможность агентов интегрироваться в системы, совместно решающие сложные задачи.

• Позволяет разделить уровни описания отдельных агентов и многоагентной системы как целого.

• Ориентирована на описание конечного множества действий.

• Модель ориентирована на искусственных агентов.

Таким образом, МАС можно выразить следующим образом [8]: $MAS = (A, E, R, ORG)$, где A – множество агентов; E – среда, в которой находится данная МАС; R – множество взаимодействий между агентами; ORG – множество базовых организационных структур, соответствующих конкретным функциям (ролям) агентов и установившимся отношениям между ними.

Для описания введенного множества R взаимодействий между агентами и между агентами и окружающей средой вводится три языка разного уровня со следующими коммуникационными функциями: язык составления общих планов и взаимодействия с другими агентами (L2), язык локального планирования (L1), язык исполнительно уровня (L0). Это позволит создать многоуровневую архитектуру агента, что приведет к разбиению функциональных возможностей агента на несколько иерархических уровня. Каждый такой уровень взаимодействует с остальными в порядке иерархии. Примером такой архитектуры является InteRRaP (INTErgation of Reactive behavior and Rational Plannig – объединение реактивного поведения и рационального планирования). Акт взаимодействия с использованием некоторого языка Lx обозначим через $r(\mathrm{Lx})$. Тогда $R = (\{r(\mathrm{L}2)\}, \{r(\mathrm{L}0)\})$. Язык L1 предназначен для построения планов агента в рамках множества ORG.

Отдельный агент же, в рамках выбранной модели, может быть описан как четверка: $A_i = (E_i, R_i, ORG_i, C)$, где E_i – элементы коммуникационной среды, включая источники информации $(E_i \subseteq E)$; R_i – подмножество связей данного агента с другими $(R_i \subseteq R)$; ORG_i – подмножество, описывающее организационную структуру агента (или множество его функций выполняемых в общей структуре МАС, $ORG_i \subseteq ORG$); C – внутренняя структура агента.

Внутренняя функциональная структура отдельного агента может быть представлена пятеркой $C = (K, F, I, G, B)$, где K – подсистема – ядро, отвечающее за динамическую реализацию ORG, F – подсистема, отвечающая за выполнение конкретных функций агента, I – подсистема, отвечающая за взаимодействия с источниками информации, G – подсистема, отвечающая за взаимодействие с другими агентами, B – база знаний агента. Метамодель агента представлена на рис. 3. Центральный блок метамодели описывает структуру базового агента, на основе которого будут строиться основные агенты системы. Дополнительные блоки описывают структуру основных агентов, отражая специальные функции, зависящие непосредственно от роли агента в системе.

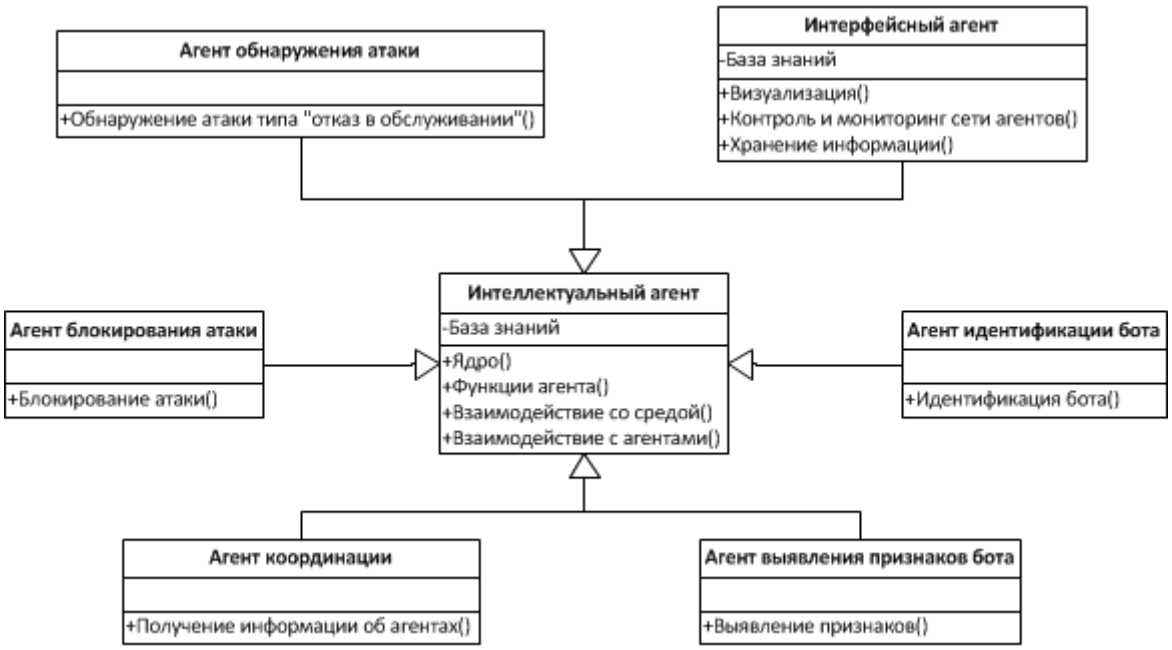

Рис 3. Метамодель агента

ЗАКЛЮЧЕНИЕ

Предложенный метод позволяет обнаружить ботнетов на основе анализа функционирования ботов участвующих в конкретной атаке. Особенность метода заключается в возможности идентифицировать ботов не принимавших участия в атаке за счёт работы распределенной сети интеллектуальных агентов. В работе предложено:

- архитектура многоагентной системы идентификации ботнетов;
- концептуальный алгоритм работы подобной системы.

СПИСОК ЛИТЕРАТУРЫ

1. *Ianelli N. and Hackworth A.* «Botnets as a vehicle for online crime.» CERT Coordination Center, 2005, pp. 1–28.

2. *Трегубенко В.* «Топ-10 ботнетов», Хакер № 188, 2014.

3. Group-IB Threat Intelligence Report 2012 – 2013 H1.

4. *Bilge L., Balzarotti D., Robertson W., Kirda E., Kruegel C.* Disclosure: detecting botnet command and control servers through large-scale NetFlow analysis. In: ACSAC, ACM (2012), p. 129-138.

5. *Francois J., Wang S., State R., and Engel T.* Bottrack: Tracking botnets using netflow and pagerank. In IFIP Networking, 2011.

6. *Zhuge J., Holz T., Han X., Guo J., Zou W.* «Characterizing the ircbased Botnet phenomenon.» Peking University& University of Mannheim Technical Report, 2007.

7. *Collins M., Shimeall T., Faber S., Janies J., Weaver R., Shon M. D., Kadane J.* «Using uncleanliness to predict future Botnet addresses» in Proceedings of ACMIUSENIX Internet Measurement Conference (IMC'07),2007.

8. *Тарасов В. Б.* «От многоагентных систем к интеллектуальным организациям: философия, психология, информатика.» – М. : Эдиториал УРСС, 2002. – 352 с.

9. *Адамацкий А. И., Холланд О.* «Роящийся интеллект: представления и алгоритмы», Информационные технологии и вычислительные системы. –1998. – № 1. – С. 45–53.

10. *Городецкий В. И., Грушинский М. С., Хабалов А. В.* «Многоагентные системы (обзор)», Журнал «Новости Искусственного Интеллекта», № 2, 1998.

11. *Никишова А. В.* Многоагентная система обнаружения атак на информационную систему предприятия: диссертация кандидата технических наук, Волгоград, 2013, 109 с.

Косенко Максим Юрьевич – преподаватель кафедры информационных технологий института информационных технологий Челябинского государственного университета.
Тел. 8-351-799-72-88.
E-mail: kosenko@csu.ru

Мельников Андрей Витальевич – доктор технических наук, профессор, директор института информационных технологий Челябинского государственного университета.
Тел. 8-351-799-74-11.
E-mail: mav@csu.ru

Kosenko Maxim U. – Lecturer, Department of Information Technology, Institute of Information Technologies, Chelyabinsk State University.
E-mail: kosenko@csu.ru

Melnikov Andrey V. – Doctor of Technic Science, Professor, Head of Institute of Information Technologies, Chelyabinsk State University.
E-mail: mav@csu.ru