Ссылка на статью:
// Ученые записки УлГУ. Сер. Математика и информационные технологии. УлГУ. Электрон. журн. 2021, № 2, с. 63–74.
Поступила: 30.11.2021
Окончательный вариант: 20.12.2021

УДК 004.056.5

Исследование файлов журнала веб-сервера на предмет активности ботнетов с целью совершенствования технологии защиты веб-серверов

Сутыркина Е. А.*, ***Бурмистров А. Н.***

*k@jcup.ru
УлГУ, Ульяновск, Россия

Авторами проведена работа по анализу log-файла веб-сервера, работающего на ОС Linux для выявления особенностей деятельности ботнетов, в автоматическом режиме осуществляющих атаки по подбору логина/пароля на сервера. Для формирования статистического отчета по данным, содержащимся в журнале сервера, разработан программный продукт, генерирующий выборку IP-адресов, портов и пар логин/пароль, соответствующих неудачным авторизациям на сервере, характерным для brute force атак. На основе полученной выборки IP-адресов, приложением формируется интерактивная карта обращений на сервер. В результате проведённого эксперимента составлена таблица IP-адресов ботнетов, выборка наиболее опасных для использования логинов и паролей, а также даны рекомендации по повышению степени уровня защищённости веб-сервера.

Ключевые слова: *безопасность веб-сервера, анализ log-файла, ботнет.*

Введение

В современном обществе, идущему по пути цифровизации, размещение информации в сети Интернет стало чем-то обыденным, а поддержание работоспособности веб-сервисов – одной из первостепенных задач. При этом остро стоит вопрос кибербезопасности, так как веб-ресурсы непрерывно подвергаются атакам злоумышленников, с целью получения несанкционированного доступа к размещенной на них информации, извлечения прибыли, установки вредоносного ПО и включения пользовательских устройств в ботнеты [12].

Примечательно, что первые этапы атак на веб-ресурсы – сканирование портов и перебор паролей [13] проходят в автоматическом режиме, как правило, с помощью «зомби-сетей», при этом, владельцы, в том числе и IOT устройств, включенных в распределённую систему, могут и не догадываться, что подверглись атаке и включены в бот-сеть, которая используется в деструктивных целях: рассылке спама, фишинге, организации DDoS-атак [14] или загрузке вредоносного ПО [3], [1].

В сложившихся условиях специалистам по информационной безопасности приходится анализировать всё больше векторов атак и быть готовыми распознать готовящееся наступление по второстепенным признакам, например, по растущему количеству неудачных авторизаций на сервере, а в случае прорыва эшелонированной защиты, - проанализировав логированные действия киберпреступников, выявить слабые места в выстроенной обороне веб-сервера и дать рекомендации по их устранению.

1. Ботнеты: разновидности и цели создания

Сеть роботов может быть создана для целенаправленных DDoS атак, проводимых с целью выведения из строя определенного веб-ресурса с последующим требованием выкупа для восстановления функционирования атакованного ресурса, для которого, доступность сервисов, в связи переходом в онлайн все большего количества бизнеса и работодателей, играет жизненно важную роль в поддержании функционирования всей компании.

Кроме того, ботнеты могут использоваться как промежуточное звено социальной инженерии - для организации спам-рассылок и рекламирования товаров и услуг, с последующим хищением личных данных пользователя на подложных сайтах социальных сетей или конфиденциальных данных сервисов интернет-банкинга [3].

И, наконец, ботнеты могут использовать взломанные ресурсы для несанкционированного обогащения, например, устанавливая майнеры на компьютеры [10].

В зависимости от целей, преследуемых злоумышленниками, ботсети могут различаться по своей архитектуре и функционалу.

1.1. Централизованные ботнеты

Такие ботнеты представляют собой иерархическую структуру, где во главе находится управляющий сервер, рассылающий компьютерам жертв конфигурационные файлы, содержащие инструкции для поддержания функционирования сети (рис. 1).

Недостатком такой архитектуры, очевидно, является наличие открытого доступа к адресу управляющего узла, вывод которого из строя, рушит всю сеть.

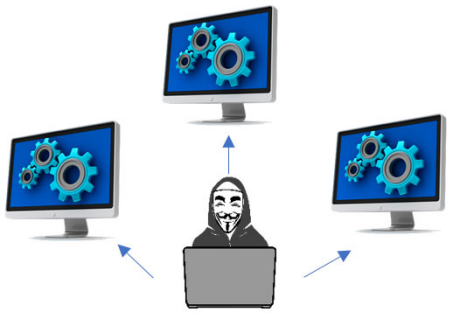

Рис. 1. Простая ботнет

1.2. Ботнеты с DGS

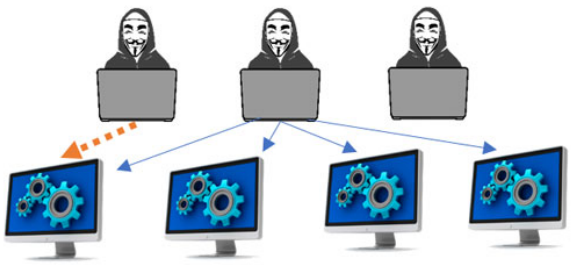

Рис. 2. Ботнет с DGS

Для того, чтобы успешно противостоять деятельности инженеров по компьютерной безопасности, злоумышленникам пришлось усовершенствовать принципы функционирования бот-сетей, результатом чего стало появления ботнетов, не имеющих жестко обозначенного управляющего узла, но получающих команды от сервера, адрес которого динамически генерируется по заранее заданной схеме. В случае блокировки доменного имени, киберпреступник попросту регистрирует новый, а ботнет перестраивается под новое управляющее звено (рис. 2).

1.3. Пиринговые ботнеты

Следующим витком эволюции бот-сетей стало появление распределённых Peer-To-Peer сетей, построенных по принципу однорангового взаимодействия, в которых роль управляющих серверов взяли на себя компьютеры пользователей, имеющие статический IP (рис. 3).

Естественным этапом развития ботнетов является объединение функционала централизованной и децентрализованной сетей.

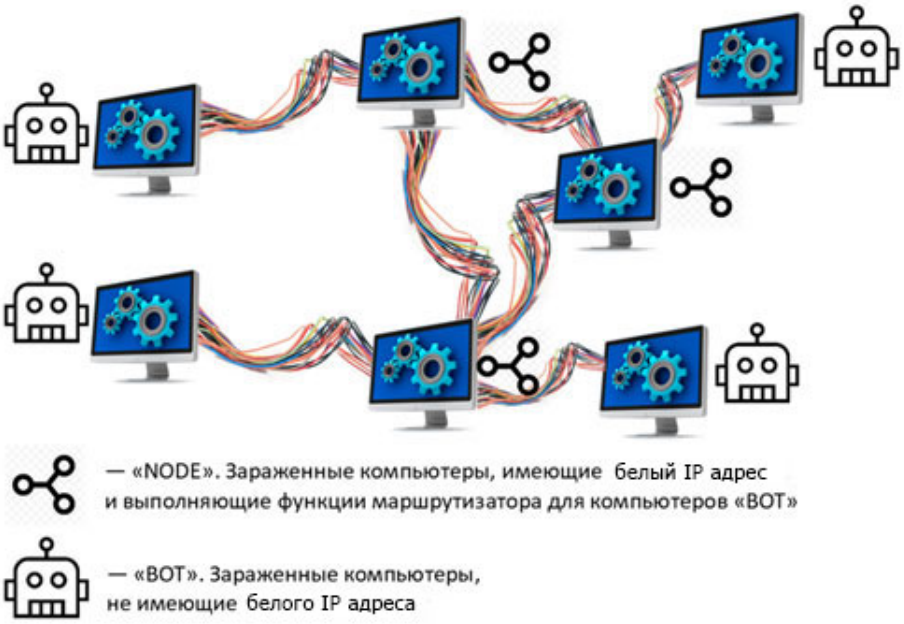

Рис. 3. P2P ботнет

Именно в такие сети может попасть устройство пользователя, уделяющего недостаточно внимания обеспечению конфеденциальности своих данных или пренебрегающего информацмонной гигиеной, используя ПО из непроверенных источников. При этом, ресурсы жертвы могут быть использованы как в качестве узла, так и в качестве управляющего устройства. Послднее представляется наиболее опасным, так как именно устройства со статическим IP являются потенциальными отправными точками для организации разного рода атак.

2. Постановка задачи

Обеспечение безопасности функционирования веб-сервера – это сложная комплексная задача (рис. 4), успешно выполнимая, только при слаженной работе специалистов разных уровней: от системных администраторов, до архитекторов ПО [2].

Несмотря на наличие широко известных общих рекомендаций по обеспечению функционирования веб-серверов [2], большое количество владельцев веб-ресурсов пренебрегает соблюдением минимальных требований к настройке системы. Более того, в связи с популяризацией и свободным распространением специального ПО, инструменты, позволяющие проводить автоматизированное тестирование на проникновение, становятся доступны злоумышленникам разной степени осведомленности, что неминуемо повышает вероятность успеха реализации возрастающих год от года атак на сервера. Таким образом, существующие стандарты противодействия киберугрозам и методы расследования киберпреступлений требуют постоянной доработки и актуализации и особого внимания здесь заслуживают веб-серверы – как основа функционирования веб-ресурсов в сети Интернет.

Этапы защиты Web-сервера

Этап	Наименование	Состав этапа
I этап	Начальный уровень безопасности	1. Обновление и модернизация установленного ПО. 2. Разделение задач по серверам. 3. Удаление лишних (ненужных) программ и приложений.
II этап	Защита от вторжения	1. Установка Firewall. 2. Удаленное администрирование системы. 3. Ограничение на использование скриптов. 4. Фильтрация пакетов с помощью маршрутизаторов. 5. Повышение квалификации сотрудников, разграничение прав.
III этап	Обнаружение и защита от атак	1. Разделение привилегий. 2. Аппаратные системы защиты. 3. Внутренний межсетевой экран. 4. Системы обнаружения атак на сеть. 5. Системы обнаружения атак на сервер.

Рис. 4. Основные стандарты, спецификации ИБ Веб-сервисов

Для проведения исследования, авторами был выбран типичный, ничем не выделяющийся среди прочих, веб-сервер, на котором настроено логирование действий, с операционной систе-

мой Linux, где единственным инструментом анализа и сортировки логов является командная строка.

Хотя, эффективной техникой сбора информации об активности на веб-сервере и являются специальные приманки [1], например некоммерческие ханипоты (SSH Honeypot, FTP Honeypot), авторами было принято решение, не устанавливать никакого дополнительного программного обеспечения на сервере так как их использование без должного статического и динамического анализа кода, могло привести к непредсказуемым последствиям.

Отметим, что мы рассматриваем реальный рабочий веб-сервер, а не локально настроенный стенд, поэтому для того, чтобы сервер в ходе исследования продолжал функционировать в штатном режиме, авторы сосредоточились на анализе лог-файла и разработке собственного программного продукта, который бы позволял ознакомиться со статистикой неудачных авторизаций, не нарушив при этом работоспособность ресурса.

3. Описание программного продукта

Для анализа логов сервера разработан программный продукт на языке Java в виде пользовательского приложения с графическим интерфейсом (рис. 5).

Приложение представляет собой форму, которая содержит поле подгрузки лог-файла, скачанного с веб-сервера и область вывода отчета по анализу содержимого загруженного файла в виде статистической сводки, на основе которой формируется выходной html-файл с размещённой на нём интерактивной картой запросов к веб-серверу.

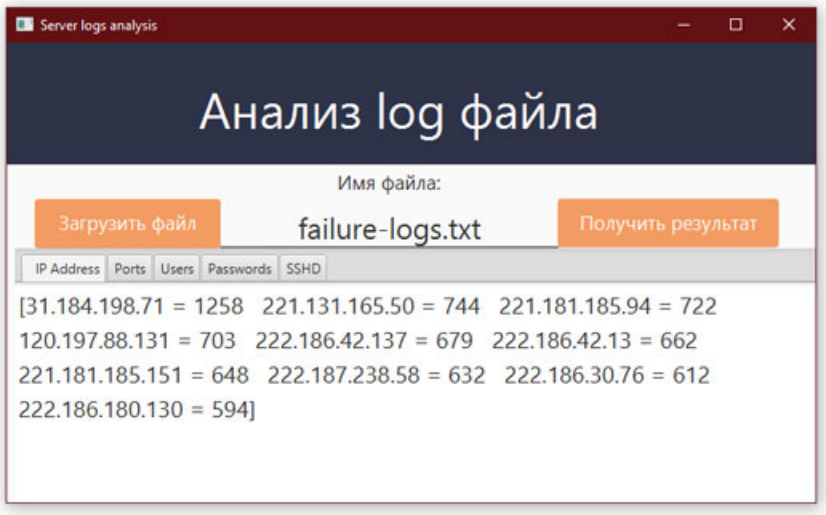

Рис. 5. GUI для анализа логов

Анализ содержимого лог-файла происходит в 2 этапа: сначала данные парсятся в соответствующие структуры данных, а затем сортируются «быстрой сортировкой» по убыванию количества вхождений.

После сортировки, на пользовательскую форму выводятся первые 10 элементов, соответствующих названию вкладки. Область вывода статистической сводки - кликабельна и позволяет скопировать полученную информацию для дальнейшей ручной обработки.

Кликая по вкладкам формы, пользователь может ознакомиться с результатами анализа загруженного файла, которые включают статистику:

- наиболее часто встречающихся IP адресов с количеством попыток входа;
- наиболее часто встречающихся портов;
- рейтинг наиболее часто подбираемых имен пользователей;
- рейтинг наиболее часто подбираемых паролей.

Далее, для большей наглядности процесса обращений к веб-серверу, IP-адреса, полученные после обработки лог-файла, автоматически преобразуются в пул координат объектов. Затем автоматически генерируется html-файл, визуализирующий запрос соединения от полученных объектов к рассматриваемому серверу (рис. 6).

Рис. 6. Визуализация запросов авторизации на web-сервере

На html-странице располагается построенная на основе открытой JavaScript библиотеки Leaflet [8], интерактивная карта, функционал которой включает: Зумирование, перетаскивание, обработку клика.

На карте кружками показаны источники обращений к веб-серверу,

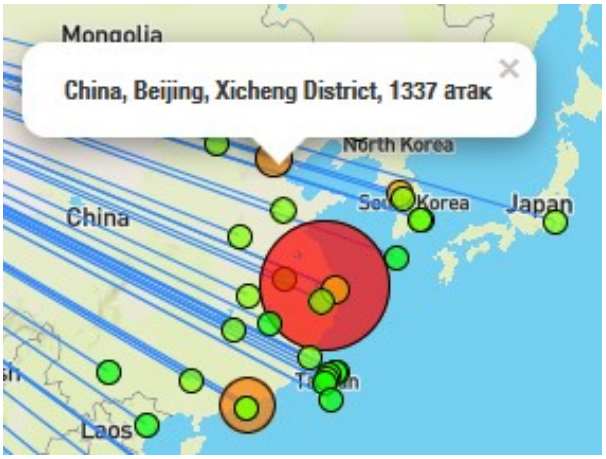

Рис. 7. Всплывающая подсказка

координаты которых были получены путём реализации API, обеспечивающим геолокацию, предоставленным ip-api [7].

Кружки и соединительные линии, изображенные на карте, кликабельны, по клику на них всплывает подсказка, содержащая информацию о количестве обращений и их источнике с указанием города и страны (рис. 7).

Чем больше радиус кружка, тем больше обращений к серверу зарегистрировано с данными координатами.

Цвет кружка варьируется в диапазоне
$hsl(x_0, 100\%, 50\%, .5)$ - $hsl(x_{max}, 100\%, 50\%, .5)$,
где тон x_j определяется углом поворота цветового спектра, замкнутого в круг, соответствующего нормализованному значению количества обращений от одного объекта с заданными координатами к веб-серверу (рис. 8).

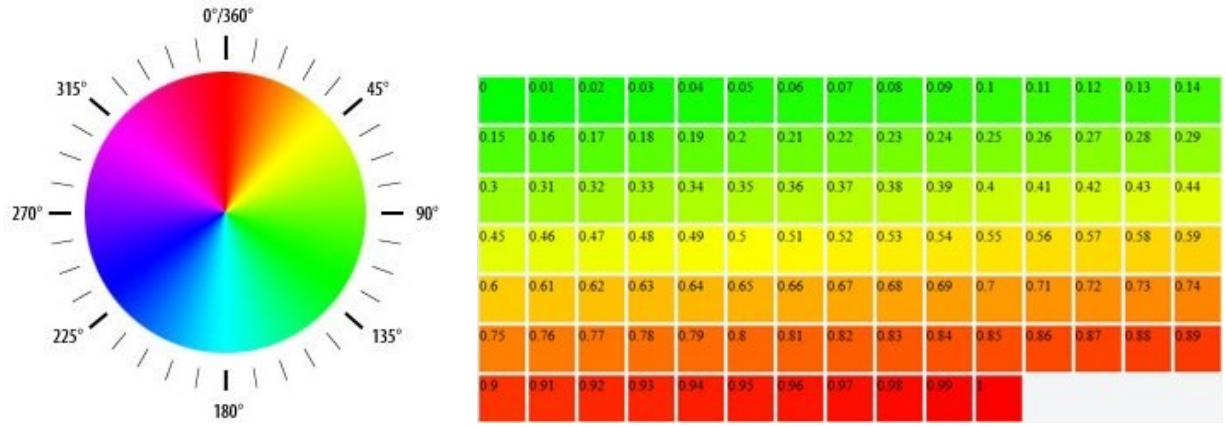

Рис. 8. Цветовой круг и нормализованная шкала значений

Например, зелёный цвет определяется углом поворота в 120°, что соответствует 0 обращений к серверу на нормализованной шкале. Максимальному количеству запросов, (11112 в нашем случае), на нормализованной шкале будет соответствоть 1, что будет определять угол поворота в 0° на цветовом круге.

Для взаимодействия с приложением от пользователя не требуется специальных знаний в программировании, при этом, скромный функционал приложения позволяет отслеживать динамику обращений к серверу и делать заключения о вероятных атаках на веб-сервер.

Разработанное приложение, безусловно, уступает по возможностям профессиональными инструментами анализа логов, такими как ARTLAS, LORG, GoAccess, однако, отличается

легковесностью, быстродействием и интуитивно-понятным интерфейсом, оно не требует разворачивания среды разработки или непосредственной установки на сервер.

Данная разработка не предназначена для промышленного использования, так как не отслеживает деятельность на сервере в реальном времени, но может найти применение в процессе расследования инцидентов нарушения информационной безопасности или в качестве вспомогательного инструмента для сортировки данных лог файлов для дальнейшей работе по распознаванию автоматизированных запросов на сервере с классификацией источника, как составной части ботнетов.

4. Анализ результатов

В ходе проводимого исследования с работающего сервера был получен log-файл, содержащий 2,5 мб записей за последние 7 дней.

Вначале авторами был проведён анализ IP адресов, наиболее часто встречающихся в лог-файле. На основе проведенного анализа можно сделать вывод, что 90% из наиболее популярных запросов для автоматической авторизации на сервере приходило из Китая, при этом 70% рассматриваемых адресов уже внесены в список с высоким риском соответствия ботнету, согласно статистике IBM [9].

Таблица 1. Пул наиболее популярных IP адресов и их классификация согласно данным на сайте IBM [9]

IP	Класс	Расположение	Запись WHOIS
31.184.198.71	IP для сканирования (71%)	Россия	Имя регистратора: ORG-PINl1-RIPE e-mail: admin@pinspb.ru
221.131.165.50	IP для сканирования (100%)	Китай	Имя регистратора: APNIC e-mail: abuse@chinamobile.com
221.181.185.94	Не подозрительный	Китай	Имя регистратора: ORG-CM1-AP e-mail: abuse@chinamobile.com
120.197.88.131	Не подозрительный	Китай	Имя регистратора: ORG-CM1-AP e-mail: abuse@chinamobile.com
222.186.42.137	IP для сканирования (86%)	Китай	Имя регистратора: APNIC e-mail: anti-spam@ns.chinanet.cn.net
222.186.42.13	IP для сканирования (86%)	Китай	Имя регистратора: APNIC e-mail: anti-spam@ns.chinanet.cn.net
221.181.185.151	Не подозрительный	Китай	Имя регистратора: ORG-CM1-AP e-mail: abuse@chinamobile.com
222.187.238.58	IP для сканирования (100%) Боты (100%)	Китай	Имя регистратора: APNIC e-mail: anti-spam@ns.chinanet.cn.net
222.186.30.76	IP для сканирования (43%)	Китай	Имя регистратора: APNIC e-mail: anti-spam@ns.chinanet.cn.net
222.186.180.130	IP для сканирования (100%)	Китай	Имя регистратора: APNIC e-mail: anti-spam@ns.chinanet.cn.net

Более того, IP адреса, помеченные как не подозрительные имеют того же регистратора и электронную почту, что и устройства, находящиеся в ботнете, исходя из чего напрашивается вывод, что это - новые устройства, которые уже заражены и благополучно сканируют сеть в поисках возможных уязвимостей на серверах, но ещё не помечены как неблагонадежные.

Результаты, полученные в ходе отображения запросов на карте, не противоречат табличным данным, что свидетельствует, во-первых, о достаточности выборки, а во-вторых о корректности работы алгоритмов в разработанном приложении.

Далее был проведён анализ логинов/паролей, с которыми боты пытаются авторизоваться на сервере. Первое место в рейтинге наиболее используемых логинов занимает root, частота использования которого, более чем в 23 раза превосходит частоту употребления имени пользователя admin, расположившегося на второй строчке рейтинга. Тройку лидеров замыкает имя пользователя user, авторизоваться с которым боты пробовали почти в 2 раза реже, чем с логином admin.

Таблица 2. Топ-10 наиболее используемых логинов

Логин	Число попыток	Логин	Число попыток
root	20309	git	177
admin	850	dev	128
user	540	pi	73
hadoop	421	ubnt	58
test	249	support	42

Рейтинг наиболее часто вводимых паролей при попытке авторизоваться на веб-сервере выглядит следующим образом.

Таблица 3. Топ-10 наиболее используемых паролей

Пароль	Число попыток	Пароль	Число попыток
пустой пароль	20174	root	89
123456	480	password	75
admin	349	test	46
123	211	12345	38
1234	153	guest	12

Согласно статистике [11], отражающей анализ более 18 миллионов паролей в мире, наиболее используемые пароли для авторизации на веб-сервере отличаются от тех, что используют пользователи в повседневной жизни в процесе веб-сёрфинга (рис. 9). Однако алгоритмы ботов стандартизированы, ибо в попытках авторизоваться присутствуют популярные логины и пароли.

Заключение

На основе проведённого анализа логов с помощью разработанного программного продукта, для администраторов информационного ресурса сформулированы следующие рекомендации:

Рис. 9. 30 самых используемых паролей в мире по данным исследования SafetyDetectives

- критически важно использовать менеджер паролей для формирования пароля, используемого для авторизации на сервере;
- необходимо систематически проверять логи SSHD и проводить их анализ активности атакующих на предмет принадлежности к ботнет;
- необходимо систематически обновлять «черный список» IP адресов;
- следует переносить критически важные сервисы на экзотические порты (наименее встречающиеся в логах.

Исходя из того, что действия атакующего могут остаться незаметными для администратора ресурса, так как простой перебор логина и пароля не окажет существенного воздействия на производительность работы среднестатистического сервера, в качестве техники противодействия несанкционированному пентесту, может быть выбран анализ логов сервера [6], а также, как один из способов идетентификации ботнета или его составной части.

Список литературы

1. Косенко, М. Ю. Вопросы обеспечения защиты информационных систем от ботнет атак / М. Ю. Косенко, А. В. Мельников // *Вопросы кибербезопасности*. 2016, № 4(17), с. 20-28. DOI: 10.21681/2311-3456-2016-4-20-28.
2. Раимов, М. Е. Анализ новых систем защиты веб-сервисов / М. Е. Раимов, А. К. Мукашева, Г. Б. Исаева // *Евразийский союз ученых*. 2021, № 3-6(84), с. 18-23.

3. Парфенова, А. С. Компьютерная сеть ботнет / А. С. Парфенова, М. С. Клочек // *Инновационное развитие*. 2019, № 1(28), с. 21-22.

4. Андрианов, В. И. Исследование возможностей сети ботнет и методология защиты от несанкционированного доступа / В. И. Андрианов, А. А. Чекалов // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2017) : Сборник научных статей VI Международной научно-технической и научно-методической конференции. В 4-х томах, Санкт-Петербург, 01–02 марта 2017 года / Под редакцией С.В. Бачевского. – Санкт-Петербург: Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, 2017. С. 41-46.

5. Elisan C. *Malware, Rootkits & Botnets A Beginner's Guide.* O'Reilly Media, Inc., 2012.

6. System Overlord: A blog about security [Электронный ресурс]. Режим доступа: https://systemoverlord.com/2020/09/04/lessons-learned-from-ssh-credential-honeypots.html (дата обращения 01.12.2021).

7. API для получения геоданных по IP [Электронный ресурс]. Режим доступа: https://ip-api.com/ (дата обращения 01.12.2021).

8. Библиотека JavaScript с открытым исходным кодом для интерактивных карт [Электронный ресурс]. Режим доступа: https://leafletjs.com/ (дата обращения 01.12.2021).

9. Информация об угрозах IBM X-Force Exchange [Электронный ресурс]. Режим доступа: https://exchange.xforce.ibmcloud.com/ (дата обращения 01.12.2021).

10. Майнинг криптовалют [Электронный ресурс]. Режим доступа: https://vk.cc/c9C4bB (дата обращения 01.12.2021).

11. Самые взламываемые пароли по версии команды SafetyDetectives [Электронный ресурс]. Режим доступа: https://ru.safetydetectives.com/blog/most-hacked-passwords-in-the-world-ru/ (дата обращения 01.12.2021).

12. Ботнет[Электронный ресурс]: Википедия. Свободная энциклопедия. Режим доступа: https://vk.cc/c9C43H (дата обращения 01.12.2021).

13. Полный перебор [Электронный ресурс]: Википедия. Свободная энциклопедия. Режим доступа: https://vk.cc/c9C45w (дата обращения 01.12.2021).

14. DoS-атака [Электронный ресурс]: Википедия. Свободная энциклопедия. Режим доступа: https://vk.cc/3xW09m (дата обращения 01.12.2021).

Botnet detection via server logs analysis

Sutyrkina, E. A.*[*], *Burmistrov, A. N.

[*]k@jcup.ru
Ulyanovsk state university, Ulyanovsk, Russia

We've worked on botnets detection by analyzing real web-server logs. The special software product has been created to generate a sample of IP addresses, ports, and login/password pairs from the log file, which contains unsuccessful authorizations reports. As the result, a map of potential botnets was compiled, besides the most dangerous passwords, and a blacklist of IP addresses was obtained.

Keywords: *botnet detection, analysis of log files, web server.*