Оригинальная статья
https://doi.org/10.21683/1729-2646-2020-20-3-53-60

Выявление рисков киберугроз на базе построения событийно-сущностных онтологий по текстам из открытых источников

Михаил К. Ридли, *МАИ (НИУ), Российская Федерация, Москва*
mr@kalabi.ru

Михаил К. Ридли

Резюме. Цель. *Основные результаты в области защиты от киберугроз в настоящее время получены в областях анализа трафика, выявлении зловредного программного обеспечения, блокировании злоумышленников от доступа во внутреннюю сеть, анализе инцидентов и других способах защиты корпоративного периметра. Между тем, эффективность этих методов зависит от своевременности и качества данных об угрозах. Целью работы является исследование способов повышения осведомленности о киберугрозах и возможностях анализа текстов в открытых источниках для задач прогнозирования кибератак, выявления и мониторинга новых угроз, обнаружения уязвимостей нулевого дня до их опубликования и обнаружения утечек.* **Методы.** *Для извлечения общественно доступных знаний о кибербезопасности используется непрерывный сбор данных из сети Интернет (включая фрагменты его неиндексируемой части и специализированные источники) и других сетей общего доступа (включая большое количество профильных ресурсов и площадок в сети TOR). Собранные тексты на разных языках анализируются с помощью методов обработки естественно-языковых текстов для извлечения из них сущностей и событий, которые затем группируются в канонические сущности и события, и вся эта информация используется для непрерывного наполнения событийно-сущностной онтологии, значимой для предметной области: в нее входят общие виды сущностей и событий, необходимые для контекста, и специализированные виды событий и сущностей для задач кибербезопасности (технические идентификаторы, векторы атаки, виды атак, хеши, идентификаторы и так далее). Такая онтология может функционировать как база знаний и использоваться для структурированных запросов от аналитиков в области компьютерной безопасности.* **Результаты.** *Предложенный метод и построенная на его базе система применимы для анализа информации о компьютерной безопасности, мониторинга, обнаружения уязвимостей нулевого дня до их официального опубликования и обнаружения утечек. Извлеченная системой информация может быть использована в качестве признаков с высокой информативностью в статистических моделях: на базе нее построен классификатор, определяющий риск появления эксплойтов для конкретной уязвимости, и балльная система скоринга IP-адресов, которая может использоваться для автоматической блокировки. Кроме того, был разработан метод рискового ранжирования событий и сущностей, связанных с киберугрозами, который позволяет выявить среди всего обилия информации сущности и события, которые требуют особого внимания, а также вовремя принимать соразмерные предупредительные меры.* **Заключение.** *Предложенный метод имеет непосредственную практическую ценность в задачах аналитики, рискового ранжирования киберугроз и мониторинга, а также может использоваться для анализа большого объема текстовой информации и создания информативных признаков для повышения качества работы моделей машинного обучения, используемых в компьютерной безопасности.*

Ключевые слова: *компьютерная безопасность, безопасность объектов железнодорожной инфраструктуры, извлечение знаний, семантическая разметка, онтология, обработка естественно-языковых текстов.*

Для цитирования: *Ридли М.К. Выявление рисков киберугроз на базе построения событийно-сущностных онтологий по текстам из открытых источников // Надежность. 2020. №3. С. 53-60. https://doi.org/10.21683/1729-2646-2020-20-3-53-60*

Поступила *23.06.2020 г.* / **После доработки** *18.07.2020 г.* / **К печати** *21.09.2020 г.*

Введение

За последнее десятилетие компьютерная преступность совершила скачок в развитии и стала большим конкурентным рынком. В 2016 году прямой ущерб мировой экономике составлял 3 миллиарда долларов, а в 2020 году – 6 миллиардов долларов. Эта сумма растет вместе с уровнем цифровизации: чем больше средств автоматизации, тем больше способов нарушить деятельность предприятия. В частности, за шесть недель работы малоизвестного немецкого проекта HoneyTrain, который имитировал системы управления железнодорожной инфраструктурой, было зафиксировано 2,3 миллиона атак [1].

Часто атаки на железнодорожную инфраструктуру приходятся на клиентские сервисы – например, из-за DDoS-атаки (распределенная атака, нацеленная на отказ в обслуживании) в мае 2018 года пассажиры датской железной дороги (DSB) лишились возможность приобретения билетов как через Интернет, так и через стационарные терминалы. Атаки на системы управления случаются реже, но они опаснее: например, в октябре 2017 года транспортная система Швеции лишилась системы мониторинга местоположения железнодорожных подвижных составов и маршрутно-картографических сервисов. Кроме того, случаются атаки в отношении SCADA-систем (системы диспетчерского управления и сбора данных): например, инцидент с заражением вирусом Stuxnet центрифуг на заводе по обогащению урана в Иране или атака на атомную электростанцию в США, расположенную в штате Канзас.

ОАО «РЖД» проводит активную работу в отношении киберзащиты железнодорожной инфраструктуры: в частности, в 2016 году был запущен совместный проект ОАО «РЖД», Positive Technologies и АО НИИАС[1], в котором исследовано устройство сигнализации, централизации и блокировки EBILock 950, которое через объектные контроллеры управляет напольными устройствами типа переездов, рельсовых цепей и стрелочных приводов.

Важно, что технологическое обеспечение ОАО «РЖД» является разнородным: на входящих в состав ОАО «РЖД» шестнадцати железных дорогах используются разные виды оборудования и протоколов. На верхнем уровне используется около 100 автоматизированных систем управления, а на нижнем (локальном) – десятки тысяч микропроцессорных систем управления движением поездов почти 70 различных типов [2].

В настоящее время фокус смещен в сторону защиты и отражения известных атак. Например, описанный выше проект был завершен внедрением системы Positive Technologies Industrial Security Incident Manager. Безусловно, это было верным шагом: за один только 2018 год количество доступных из сети Интернет компонентов систем управления технологическим процессом в России выросло в 1,5 раза, как выросло и число уязвимостей, которые могут использоваться удаленно без привилегированного доступа.

Тем не менее, защиты периметра и анализа трафика недостаточно: нужна возможность оперативного осведомления о новых уязвимостях, о проводимых в мире кибератаках (в том числе запланированных), о действиях хактивистов, об атаках на родственные классы систем и так далее. Подобный мониторинг входит в перечень значимых рекомендаций для объектов железнодорожной инфраструктуры [3]. К примеру, это позволяет отреагировать на уязвимость, которую обнаружили в одном из используемых программных решений, за недели до ее официального опубликования.

Известно, что в открытых источниках информация об уязвимостях и эксплойтах часто появляется до попадания в используемые всеми базы данных CVE (Common Vulnerabilities and Exposures) и NVD (National Vulnerability Database), причем разрыв может составлять месяцы [4]. Часто такая информация появляется в системах отслеживания ошибок программного обеспечения с открытым исходным кодом, в социальной сети Twitter, в тематических блогах, в сервисах вопросов и ответов для программистов типа StackOverflow, в почтовых рассылках, на хакерских форумах и на торговых площадках в анонимных сетях. Для эффективного мониторинга аналитикам в области компьютерной безопасности необходимы методы автоматического извлечения информации из текстов в сетях общего пользования – включая некоторые неиндексируемые сегменты Интернета и анонимные сети типа TOR. Такой мониторинг позволит не только выявлять новые угрозы и идентифицировать риски в области кибербезопасности, но и более полно и системно подходить к анализу и скорингу угроз.

Мониторинг подобных источников – это способ автоматического отслеживания уязвимостей нулевого дня, которые особенно опасны и часто незаметны для анализа сетевого трафика. К таким уязвимостям относятся те, против которых еще не разработаны или не выпущены защитные механизмы, что позволяет злоумышленникам свободно их использовать до момента публикации исправлений, а также мешает защитным средствам обнаруживать попытки их эксплуатации. Это также эффективный инструмент для оперативного поиска утечек информации, подобной случившейся в июне 2019 года краже сотен тысяч документов с внутренних ресурсов ОАО «РЖД»[2]. Поскольку злоумышленники часто ищут способы монетизации украденного, они размещают объявления на специальных площадках, где утечка будет обнаружена через несколько минут после публикации.

Построение событийно-сущностных онтологий для прикладных баз знаний

Один из способов концептуализации информации из текстов – это построение онтологии на базе описанных

[1] Источник: https://bit.ly/2YkAQ4N

[2] Источник: https://www.kommersant.ru/doc/4252728

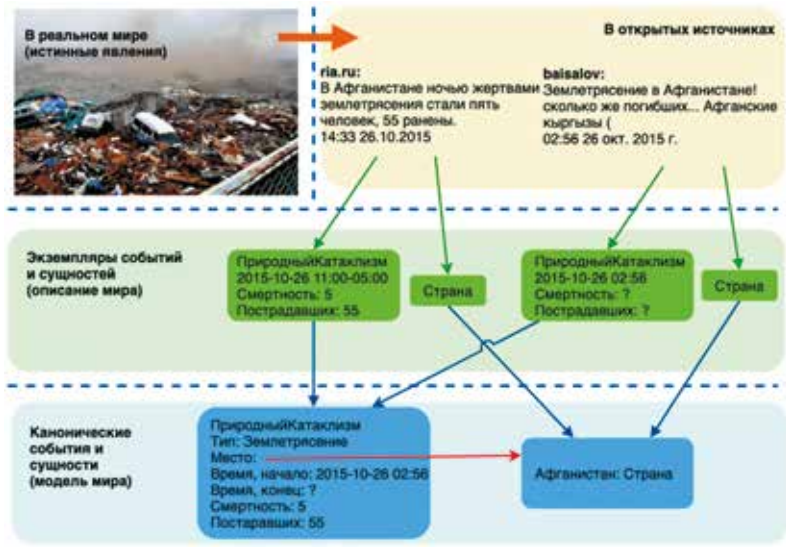

Рис. 1. Пример описания реальных явлений через канонические события и их упоминания

в них фактов. В таких онтологиях представляются понятия предметной области, отношения между ними и их атрибуты. Для их извлечения в соответствии с мета-онтологией (онтологией верхнего уровня, описывающей конкретную онтологию) применяются инструменты компьютерной лингвистики, правила сопоставления типа шаблонов Херста и регулярных выражений, статистические модели.

Автоматическое построение онтологий чаще всего используется для создания универсальных онтологий, основанных на лингвистических категориях типа гипоним/гипероним и мероним/холоним, отношениях IS-A, INSTANCE-OF и других. Это актуально для многих задач искусственного интеллекта, но не для задач по представлению и накоплению прикладных знаний.

Для построения баз знаний практичнее использовать автоматически пополняемые событийно-сущностные онтологии. Автором ранее была разработана информационно-аналитическая система, которая извлекает из источников тексты, проводит их анализ, строит онтологию, а затем предоставляет ее пользователю как базу знаний [5]. Система разработана для новостных и политических приложений, а также для гражданской авиации [6]. Одна из задач настоящей работы – адаптация системы для анализа и мониторинга киберугроз.

Подход событийно-сущностных онтологий подразумевает, что мир моделируется путем отделения документов от того, о чем в них говорится – о канонических сущностях и событиях, которые соответствуют реальным людям, технологиях, компаниям, встречам, бизнес-транзакциям, атакам и политическим событиям. Каждой канонической сущности и каждому каноническому событию могут соответствовать несколько экземпляров, которые соотносят упоминания канонических объектов в текстах со временем, местом и прочей контекстной информацией, как это проиллюстрировано на рис. 1. Подобная онтология позволяет осуществлять структурированные запросы к базе

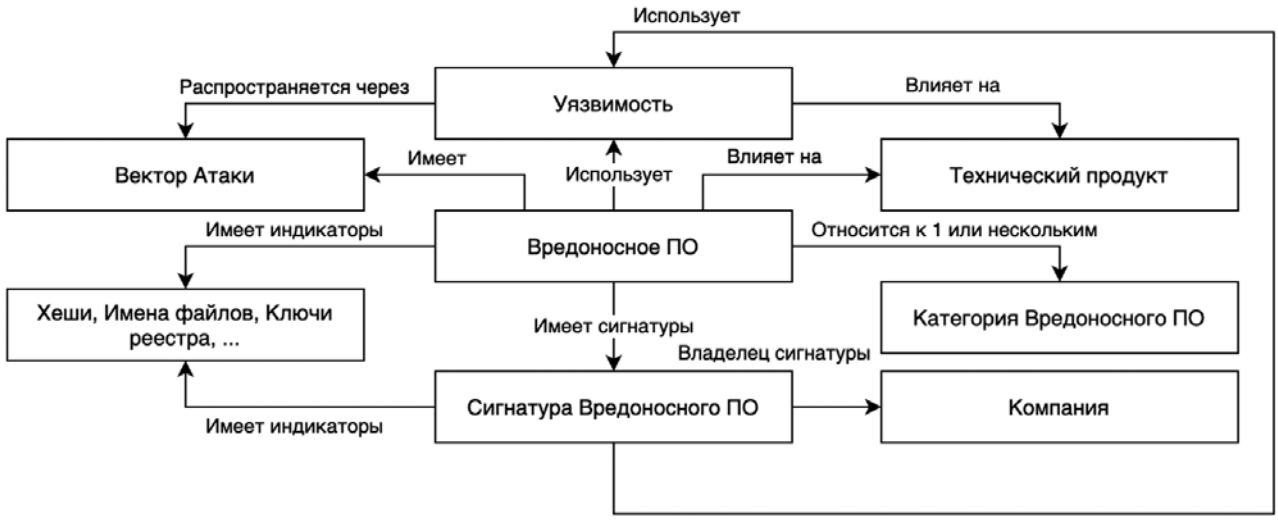

Рис. 2. Фрагмент мета-онтологии сущностей в области кибербезопасности

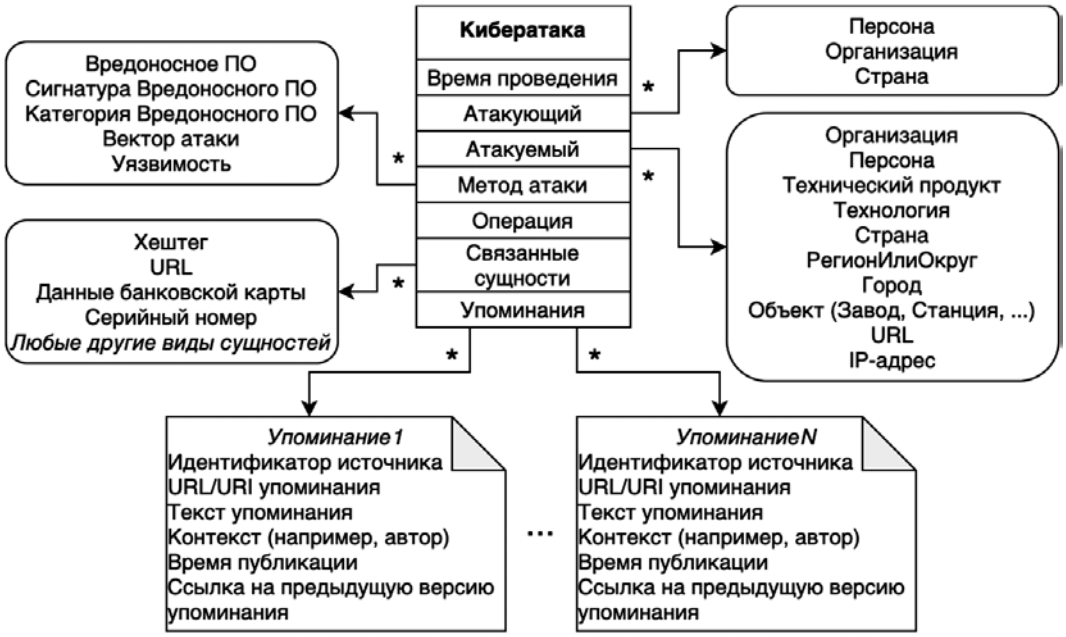

Рис. 3. Схема событий типа «Кибератака»

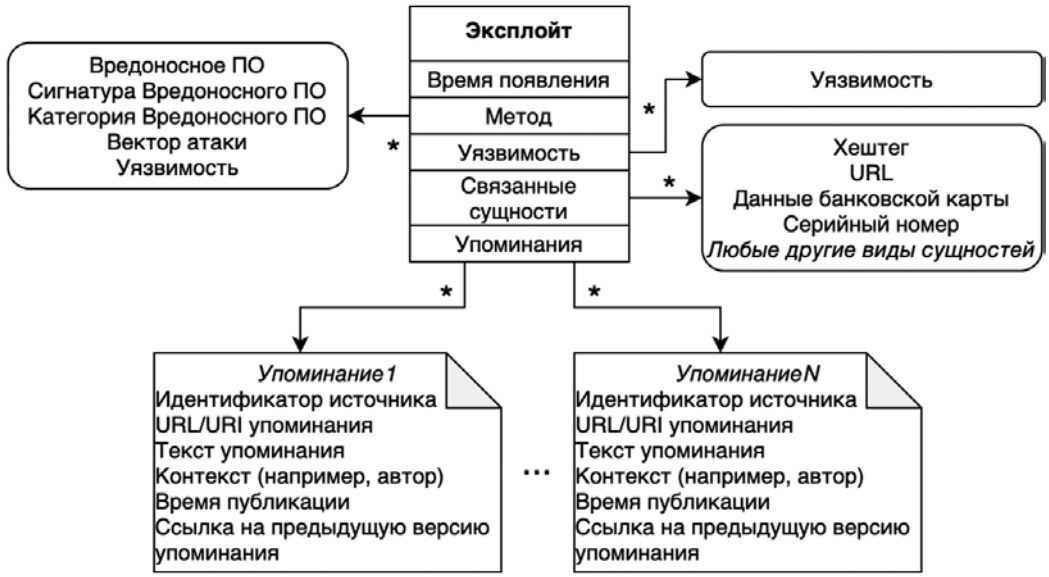

Рис. 4. Схема событий типа «Эксплойт»

знаний: можно узнать технические идентификаторы, связанные с атакой, получить список источников, публиковавших информацию об атаке, узнать, какие типы системы могут быть атакованы с помощью конкретного эксплойта, кто и где продает конкретный набор эксплойтов и другие вопросы, сводящиеся к фильтрации по атрибутам.

Мета-онтология, на базе которой производится непрерывное обновление событийно-сущностной онтологии по текстовым упоминаниям, была дополнена под задачи кибербезопасности. За основу взята онтология кибербезопасности для критической инфраструктуры, которая была переработана под события и сущности [7]. Мета-онтология в части сущностей проиллюстрирована на рис. 2, а в части событий – на рис. 3 и 4.

Реализация анализа текстов для построения событийно-сущностных онтологий

Из текстов на разных естественных языках необходимо извлекать сведения о сущностях (персоны, организации, номера карт, почтовые адреса, хеши, IP-адреса, сигнатуры программного обеспечения, имена файлов и т.д.) и событиях (поглощения компаний, политические протесты, кибератаки, банкротства и другие).

В системе используются пять готовых инструментов анализа естественно-языковых текстов (табл. 1) и набор собственных средств для экзотических сущностей (хеши, серийные номера, коды уязвимостей, фрагменты кода и т.д.). Разработанные средства

Таблица 1. Используемые готовые инструменты компьютерной лингвистики

StanfordNLP	Tomita-парсер	OpenCalais	OpenNLP	Rosette EX
Поддерживаемые языки				
Английский, немецкий, испанский, китайский	Любой, задается словарями и грамматиками	Английский	Европейские языки	55 языков (включая русский, арабский и китайский)
Основные интерфейсы				
API, библиотеки JAVA и Python, web-интерфейс	Консольное приложение, API	API, web-интерфейс	Библиотека JAVA	API, web-интерфейс
Наиболее развитая ветвь функциональности (используется в системе)				
Выделение сущностей с применением статистических алгоритмов и нейронных сетей	Построение грамматик и выделение сущностей при помощи словарей	Выделение сущностей и событий на основании новостной онтологии	Выделение сущностей с применением статистических алгоритмов и нейронных сетей	Выделение сущностей, фактов, разрешение кореферентности
Рекомендуемые разработчиками способы вывода				
JSON, XML, CoNLL, графический	Формат вывода задается грамматикой	RDF, XML, графический	XML	XML, графический

извлечения используют правила на регулярных выражениях или метод условных случайных полей (CRF), чья особенность заключается в отсутствии необходимости в моделировании вероятностных зависимостей между наблюдаемыми переменными и проблемы смещения метки как у марковской модели максимальной энтропии.

Извлеченные сущности и факты сопоставляются и разрешаются согласно онтологии с целью уточнения их значения и разрешения кореферентности. Онтологии

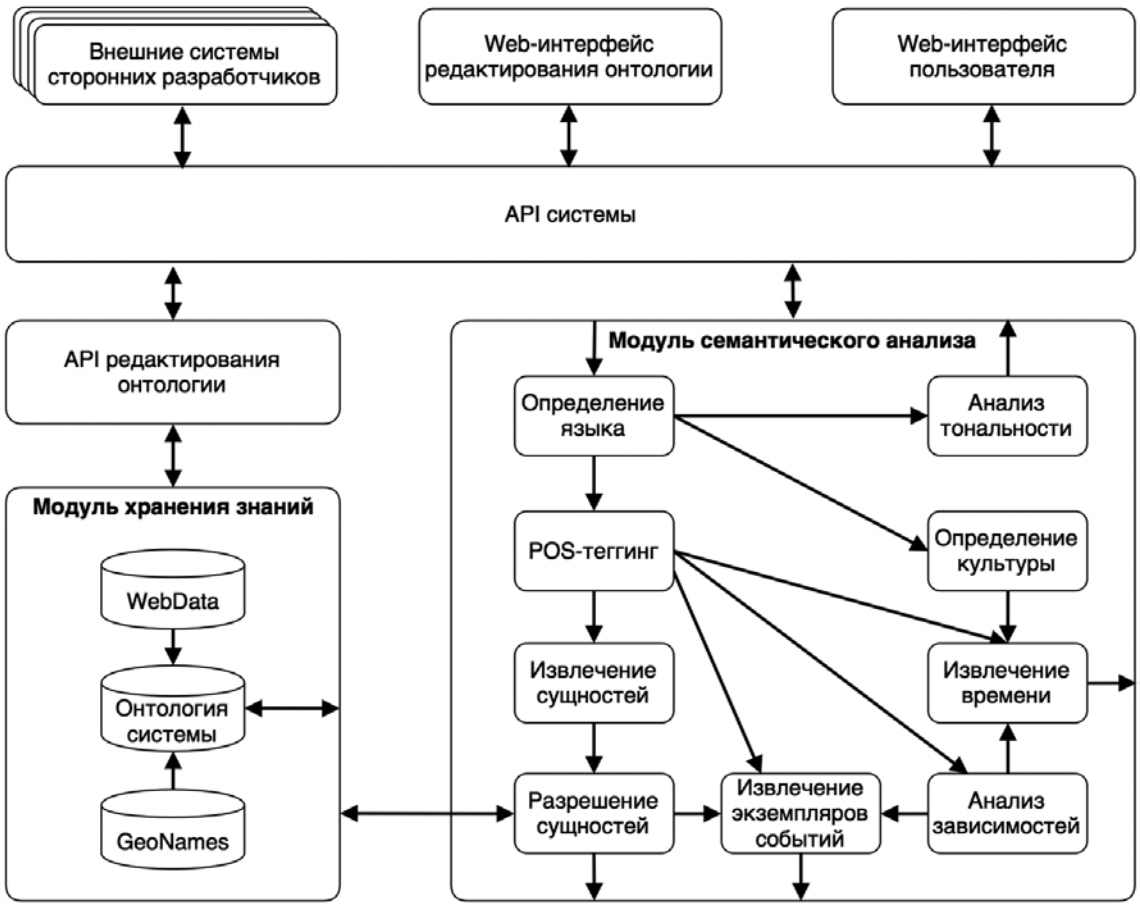

Рис. 5. Верхнеуровневая архитектура подсистемы лингвистического анализа

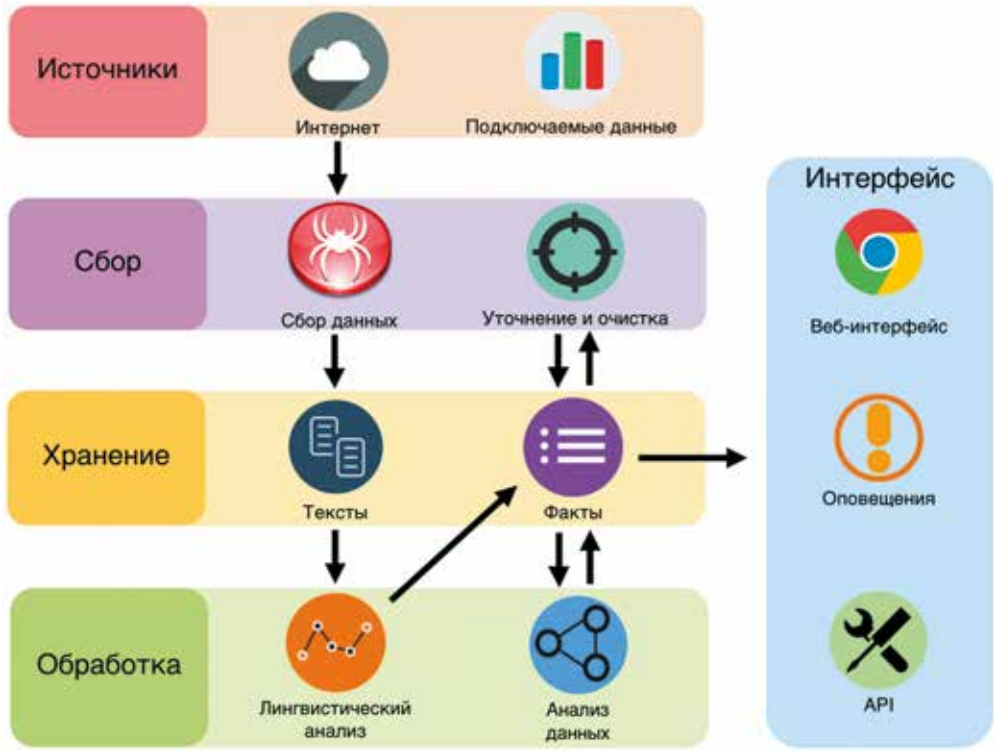

Рис. 6. Верхнеуровневая функциональная схема системы

структурированных отношений между сущностями используются для управления процессом фильтрации и в качестве газеттира для улучшения процесса извлечения.

Получив набор отфильтрованных сущностей, модуль, отвечающий за извлечение высказываний, связывает эти сущности с событиями, упоминаемыми в документе. Событиям присваивается фрагмент текста из документа, наилучшим образом кратко его представляющий – при этом решается задача автореферирования. Для каждого фрагмента анализируется эмоциональная тональность.

Далее факты подвергаются временному анализу. Культурные и региональные категории извлекаются из документа для учета полушария, первого дня недели и формата дат. События могут быть как случившимися, так и ожидаемыми. Будущие события либо запланированы (дни выборов), либо спекулятивны (предположения о том, когда правильнее выйти на митинг). В частности, это позволяет заранее реагировать на действия хактивистов типа группировки Anonymous и выпуск исправлений от производителей программного обеспечения.

В итоге события помечаются типом, временным интервалом, вовлеченными сущностями, их ролями (каким атрибутам события они присваиваются), эмоциональной тональностью, и источником. Схематично архитектура подсистемы лингвистической обработки изображена на рис. 5. Она представляет собой независимую систему,

Рис. 7. Техническая архитектура системы с точки зрения потока данных

интегрированную с основной системой, рассматриваемой в следующем разделе работы.

Архитектура и реализация системы сбора и анализа данных на базе онтологий

На рис. 6 изображена Верхнеуровневая архитектура системы.

Подсистема сбора отвечает за получение текстовых данных из ресурсов сети Интернет. На вход она получает список заранее заданных и настроенных ресурсов и выполняет регулярное извлечение данных с них. Выходом системы является поток неструктурированных текстов с указанием времени сбора, контекста и ресурса-источника.

На вход *подсистемы лингвистической обработки* поступают очищенные тексты из подсистемы сбора. Выходом подсистемы является список сниппетов (текстовых фрагментов), размеченных на языке XML с выделением в тексте сущностей, событий и временных и географических меток. Это единственное место в системе, зависящее от языка: поддержка нового языка требует создания нового модуля в этой подсистеме, а вся остальная система оперирует либо «сырыми» текстами, либо независящими от языка фактами.

Подсистема хранения принимает на вход размеченные сниппеты и разбирает их, после чего извлеченные факты помещаются в хранилище, а также принимает запросы от внешних систем. Доступ к хранилищу фактов предоставляется через интерфейс программирования приложений (API) в формате JSON в соответствии с принципами REST API. Выходом подсистемы являются срезы онтологии (наборы фактов, событий и сущностей, а также их метрики типа важности), соответствующие API-запросу.

Хранилище фактов просматривают и изменяют модули *подсистемы интеграции данных*, занимающиеся обогащением и уточнением данных. Эта подсистема ослабляет неправдопобные события и усиливает резонансные, попутно обогащая их дополнительной структурой и формируя канонические события и сущности.

Система является распределенной и имеет микросервисную архитектуру (рис. 7). Компоненты коммуницируют через очереди сообщений RabbitMQ, порождая порядка 4000 сообщений в секунду. Для хранения базы знаний используется NoSQL-хранилище MongoDB (9 шардов с ролями чтения и записи), а для сниппетов используется система полнотекстового поиска ElasticSearch (7 шардов). Метаданные для ста документов в среднем занимают 1 Мб.

Результаты применения метода в прикладных задачах

Мониторинг и аналитика в области компьютерной безопасности и прогнозирование.

Конструируемая событийно-сущностная онтология напрямую используется для прогнозирования атак хактивистов, обнаружения уязвимостей нулевого дня и поиска утечек. Другое прямое применение – анализ информации о компьютерной безопасности: какие сущности и через что связаны с набором эксплойтов, какие методы использует конкретная группировка, какие секторы экономики в настоящее время под угрозой и так далее.

Раннее обнаружение уязвимостей нулевого дня.

Обнаружено, что 77% уязвимостей из списка CVE (Common Vulnerabilities and Exposures) в отношении Linux были известны до их опубликования в качестве уязвимостей нулевого дня, а средняя задержка между первым упоминанием и датой официального обнародования – 19 дней. При этом все опубликованные в CVE уязвимости можно было найти в социальной сети Twitter [8].

Создание признаков с высокой информативностью для машинного обучения.

Получаемая в итоге и постоянно пополняемая база знаний фактов о мире кибербезопасности может использоваться для создания признаков с высокой информативностью в моделях машинного обучения, как это показано в следующей задаче.

Определение риска появления эксплойта для уязвимости.

Используя обучение с учителем на базе метода опорных векторов, был получен классификатор, который для каждой конкретной уязвимости предсказывает, будет ли для нее создан эксплойт, с точностью 0,79 и полнотой 0,80. Сбалансированная обучающая выборка состояла из 7000 примеров уязвимостей. Классификатор предсказывает риск эксплуатации конкретной уязвимости и подходит для приоритизации работ по разработке мер противодействия, экстренной изоляции уязвимых систем при высоком риске и так далее.

Балльный скоринг IP-адресов.

Скоринг IP-адресов позволяет аналитикам в области кибербезопасности принять решение о дальнейшем анализе, а в ситуации повышенной готовности осуществлять автоматическое блокирование IP-адресов для усложнения доступа для атакующих. Важно отметить, что рынок компьютерной преступности коммодитизирован и предоставляет возможности аренды ботнет-сетей и специализированной инфраструктуры для проведения атак. Поэтому меры по блокировке рискованных IP-адресов на практике весьма эффективны, особенно против DoS-атак, направленных на отказ в обслуживании систем.

Рисковое ранжирование событий и сущностей в области кибербезопасности.

Риск сущности или события вычисляется на базе динамики упоминания, наличия значимых атакуемых целей и разнообразия языка в упоминаниях. Все признаки вычисляются по скользящему среднему для оцениваемой сущности, поскольку упоминания часто носят характер «всплесков» (день-ночь, рабочие дни-выходные и т.д). Уровень критичности для сущностей

основан на количестве упоминаний событий-кибератак/эксплойтов, происходящих сегодня или в течение следующего месяца, и включающих сущность. Общий объем упоминаний в отношении какой-либо сущности не влияет на изменение уровня критичности: маленькие всплески и аномалии влияют сильнее, поскольку в кибербезопасности важнее не известное положение дел, а отклонения от него.

Разнообразие языка в упоминаниях оценивается по повторяемости описательной лексики. При этом описания на разных языках считаются различными. Подобная метрика позволяет отличать события, вызывающие реальные обсуждения (признак того, что событие лично затрагивает чьи-то интересы), и позволяет избавиться от завышения оценки, которое во многих системах мониторинга социальных сетей случается из-за повторений.

Заключение

В работе показано, что использование событийно-сущностных онтологий в решении задач кибербезопасности имеет собственную прикладную ценность, а кроме того – может быть значимой составной частью или вспомогательным механизмом в других методах.

Другим результатом является высокая информативность признаков на базе извлеченной информации при их применении в моделях машинного обучения, что позволяет повысить качество моделей, используемых в области кибербезопасности для задач выявления аномалий, экстраполяции и прогнозирования, классификации и кластеризации, поиска закономерностей и ассоциаций.

Теоретический результат заключается в возможности предобработки корпуса текстов, которая позволяет использовать классические количественные и категориальные методы в отношении текстов путем выделения из них информации.

В дальнейшем целесообразно исследовать применимость метода для анализа журналов (в том числе внутри периметра), переписки (поиск утечек), мониторинга социальных сетей, анализа документов в области кибербезопасности.

Библиографический список

1. Kühner H., Seider D. Security Engineering für den Schienenverkehr // Eisenbahn Ingenieur Kompendium. 2018. P. 245-264.

2. Макаров Б.А. Актуальность кибербезопасности на железнодорожном транспорте // Вестник Института проблем естественных монополий: Техника железных дорог. 2015. № 3 (31). С. 10-15.

3. Киселева Е.М. Железная дорога как объект киберзащиты // Международный студенческий научный вестник [Электронный ресурс]. URL: http://www.eduherald.ru/ru/article/view?id=19179 (дата обращения: 15.06.2020).

4. McNeil N., Bridges R.A., Iannacone M.D. et al. Pace: Pattern accurate computationally efficient bootstrapping for timely discovery of cyber-security concepts // 12th International Conference on Machine Learning and Applications. 2013. Vol. 2. P. 60-65.

5. Кузьмина Н.М., Ридли М.К. Об автоматическом построении в информационных системах гражданской авиации онтологий предметной области по корпусу текстов // Научный вестник ГОСНИИ ГА. 2018. № 21. С. 122-131.

6. Кузьмина Н.М., Ридли М.К. Архитектура системы построения онтологий предметной области и смыслового поиска в задачах совершенствования функционирования авиатранспортной системы // Научный вестник ГОСНИИ ГА. 2019. № 28. С. 103-113.

7. Bergner S., Lechner U. Cybersecurity Ontology for Critical Infrastructures // Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. 2017. Volume 2: KEOD. P. 80-85.

8. Trabelsi S., Plate H., Abida A. et al. Mining social networks for software vulnerabilities monitoring // 7th International Conference on New Technologies, Mobility and Security (NTMS). 2015. P. 1-7. DOI: 10.1109/NTMS.2015.7266506

Сведения об авторе

Ридли Михаил Кристофорович – аспирант Московского Авиационного Института (Национальный Исследовательский Университет), Российская Федерация, Москва, e-mail: mr@kalabi.ru

Вклад автора в статью

Автором Ридли М.К. выполнен анализ предметной области, предложен метод извлечения информации из открытых источников на базе событийно-сущностных онтологий, доработана ранее разработанная автором информационно-аналитическая система для извлечения и хранения знаний в форме событийно-сущностных онтологий под предметную область кибербезопасности, решены пять прикладных задач (мониторинг и аналитика в области кибербезопасности, раннее обнаружение уязвимостей нулевого дня, определение риска создания эксплойта на базе уязвимости, балльный скоринг IP-адресов, рисковое ранжирование событий и сущностей в области кибербезопасности).

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.