УДК 004.023+004.413+004.891

К.В. Попов, П.А. Шелупанова

Информационные системы для анализа угроз национальной безопасности

Информационные технологии играют значительную роль в такой области знаний, которая получила название «вычислительная социальная наука» (computational social science). Разработка онлайн-технологий и цифровых методов исследований в предметных областях общественных наук связана с применением новых междисциплинарных подходов и интеграцией исследователей из различных областей знания. В качестве эффективного и перспективного инструмента для исследователей в области изучения насильственных идеологий, дестабилизирующих общественный строй, в статье рассматриваются примеры баз данных, сервисов и платформ для многомерного анализа. Обозначены возможности и ограничения настройки процесса сбора, обработки и представления данных. Представлен обзор специфических характеристик баз данных, сервисов и платформ, размещенных на веб-сайте исследовательского центра START. Сделаны выводы о том, какие возможности такого рода инструменты открывают для исследователей, и обозначены перспективы реализации наиболее интересных решений для совершенствования процесса автоматизации анализа данных.

Ключевые слова: национальная безопасность, экстремизм, терроризм, радикализация, база данных, Web Mining, big data, исследовательские инструменты, управление.
DOI: 10.21293/1818-0442-2022-25-4-71-79

Возникновение насильственных идеологий связано с различными социальными, политическими и экономическими факторами, которые различаются от страны к стране и меняются со временем [1]. Общими угрозами национальной безопасности, независимо от страны, определяются угрозы терроризма, насильственного экстремизма.

В настоящее время в общественных науках сложилось более-менее единое мнение на природу экстремизма и терроризма, условия их возникновения, методы борьбы с ними и профилактику этих явлений. Значительных расхождений по данным вопросам в публикациях отечественных и зарубежных исследователей не прослеживается.

Поскольку термин «радикализация» используется слишком часто и попадает в различные риторические атрибуты, важно опираться на происхождение этого термина, которое связано со словом «корень», т.е. фундаментальным происхождением идеи или причиной. Таким образом, радикализация в своем эпистемологическом смысле означает привязку себя к своим знаниям, мнениям, ценностям и убеждениям для определения своего поведения [2]. Поэтому мы не будем акцентировать свое внимание на отдельных незначительных расхождениях во взглядах на понятийный аппарат данной проблемы, а сформулируем устоявшееся общественное представление о природе экстремизма и терроризма.

Наиболее полное определение терроризму дал в своей работе Тодд Сандлер: «Терроризм – это преднамеренное применение или угроза применения насилия отдельными лицами или субнациональными группами для достижения политической или социальной цели путем запугивания большой аудитории, помимо непосредственных жертв» [3].

Терроризм, насильственный экстремизм имеют общий предиктор, о котором ученые заговорили еще в начале этого века. Предиктором инцидентов террористической и экстремистской направленности является радикализация. В самом общем виде под радикализацией понимается процесс перехода от ненасильственных форм выражения мнения к насильственным действиям. Насильственный экстремизм рассматривается как «поощрение, оправдание или поддержка совершения насильственного действия для достижения политической цели, идеологических, религиозных, социальных или экономических целей» [4]. Вопрос о том, как и почему одиночки, автономные ячейки, сообщества и целые движения радикализуются, осуществляя переход от насильственного экстремизма, преступлений на почве ненависти к терроризму, является актуальным в сегодняшней повестке обеспечения национальной безопасности как в России, так и во всем мире.

За прошедшие двадцать лет было много исследований, которые применяют эмпирические и теоретические методы изучения насильственных проявлений. Первым способствовала повышенная доступность данных о террористических событиях. В первую очередь, это разработки в области создания инновационных, совместимых между собой исследовательских инструментов и сервисов, которые облегчают поиск, добычу, аналитическое хранилище данных, позволяют заинтересованным субъектам решать сложные исследовательские задачи.

Например, это организации, которые занимались или продолжают заниматься сбором данных, способствуя работе аналитиков, политиков и практиков в понимании тенденций террористической деятельности. Наиболее известные из них – исследовательская организация RAND Corporation [5], Мемориальный институт по предотвращению терроризма (Memorial Institute for the Prevention of Terrorism (MIPT)) [6], Национальный консорциум по изучению терроризма и ответам на террористическую угрозу университета Мэриленд (National

Consortium for the Study of Terrorism and Responses to Terrorism University of Maryland (START)) [7], Национальный контртеррористический центр (The National Counterterrorism Center (NCTC)) [8], Канадская сеть исследований терроризма, безопасности и общества (Canadian network for research on terrorism, security and society (TSAS)) [9], Центр исследований экстремизма (Center for Research on Extremism (C-REX)) [10], Международный центр по борьбе с терроризмом – Гаага (International Centre for Counter-Terrorism, The Netherlands (ICCT)) [11], Международный центр по изучению радикализации (The International Centre for the Study of Radicalisation (ICSR)) [12], Центр изучения терроризма и политического насилия Ханда (The Handa Centre for the Study of Terrorism and Political Violence (CSTPV)) [13], Международный центр исследований терроризма (International Center for Terrorism Studies (ICTS)) [14], Международный центр исследований политического насилия и терроризма (International Centre for Political Violence and Terrorism Research (PSiS)) [15], Центр контртеррористических исследований китайских институтов современных международных отношений (China Institutes of Contemporary International Relations (CICIR)) [16] и др. [17].

Несмотря на то, что каждый центр имеет либо региональную, национальную или конкретную предметную специфику, общим для них является создание и использование баз данных для проведения исследований, касающихся установления рисков и уязвимостей, связанных с насильственным поведением. Собранные данные расширяют возможности исследователей выявлять атрибуты террористических событий, получать доступ к длинному набору данных (за большой временной период). Использование временных рядов и панелей оценки подкрепляет новые эмпирические исследования. Одна вещь остается неизменной в борьбе с терроризмом и насильственным экстремизмом – это информация о радикальных личностях, будь то личная, демографическая, социальная, экономическая или иная информация. Возможность идентифицировать, классифицировать и создавать профили потенциальных радикалов на основе такой информации привлекательна для аналитиков, исследователей, экспертов по всему миру.

Фундаментальные и прикладные научные исследования в цифровой среде привели к созданию новой области знаний, которая получила название «вычислительная социальная наука» (computational social science), которая основана на применении в социальных исследованиях нового поколения онлайн-технологий и цифровых методов исследований. Это опосредованные компьютерными технологиями методы сбора, анализа и обработки данных, которые обеспечивают большой методологический потенциал для исследований во всех областях общественных наук. Эта область знаний выстраивается на применении междисциплинарных подходов и методов, а также объединении специалистов предметной области, специалистов Data Science и поставщиков профессиональных услуг (разработчиков программного продукта) [18–20].

Для получения представления о существующей схеме работы по выявлению действий насильственной идеологии отобразим ее графически в виде нотаций IDEF0 на рис. 1.

Рис. 1. Контекстная диаграмма процесса выявления насильственного действия

Входящая стрелка – «Информация» (сообщения средств массовой информации, в социальных сетях и мессенджерах). Управляющая – «Указание на выявление насильственных действий». Ограничительная – «Закон (правовая норма)». В роли «механизмов» выступают специалисты (силы) и программное и/или аппаратное обеспечение (средства). Аналитик в ходе своей деятельности получает классифицированный, стандартизированный продукт наполнения базы данных.

Несмотря на рост количества открытых сервисов для исследователей по всему миру в виде баз данных, существует ряд задач, которые до сих пор остаются нерешенными, вследствие чего набор данных по-прежнему дает нам лишь небольшой процент знаний. Условно такого рода задачи можно разделить на несколько уровней, отражающих весь процесс сбора, обработки и представления данных, влияющий на эффективность автоматизации процесса анализа.

Первый уровень. Задача создания архитектуры баз данных и комплектования функций прототипа для набора данных [21–23]. Часто применяются методы, позволяющие исследовать закономерности только по одному измерению за определенный временной промежуток. Например, данные могут быть изучены в хронологическом порядке, чтобы определить скорость изменений (меняющиеся временные тенденции). Данные также могут быть организованы географически путем создания картограммы или карты плотности (например, по конкретной стране/географические вариации) [24–26]. Данные также могут быть обобщены по целевым типам, чтобы понять характеристики инцидентов/измене-

ние характеристик [27, 28]. Например, разработчики глобальной базы данных о терроризме (The Global Terrorism Database™ (GTD)) [29] используют метод разбиения больших данных о террористических инцидентах на типы событий и предложили метод калибровки, который позволил выявлять потенциальные погрешности в GTD в результате недоучета или переучета террористических инцидентов. Эти же авторы применили методы векторной авторегрессии (vector autoregression – VAR) для исследования реакции, вызванной шоком, метод декомпозиции дисперсии (variance decomposition) и тесты на причинно-следственную связь по Грейнджеру (Granger-causality tests), расширив тем самым возможности аналитиков в работе с базой данных GTD [30, 31].

Второй уровень. Задача визуализации данных. Трудно визуализировать данные по нескольким параметрам (включая географию, время и несколько атрибутов) и представить их в доступной для понимания форме. Например, если исследователь хочет изучить пространственные вариации целевых кластеров или какие цели выбирают террористы в конкретной стране. Поэтому, чтобы получить полное представление о данных, необходимо визуализировать их с разных точек зрения и искать разные типы закономерностей.

Возможности для решения этой задачи открываются в применении многомерных пространственных моделей. Особой подзадачей визуализации данных становится многомерное картирование для создания единой среды визуализации, которая может гибко поддерживать визуализацию: пространственно-многомерных, пространственно-временных, временных многомерных и пространственно-временных многомерных моделей [32, 33]. Наиболее перспективными в многомерной визуализации на данный момент времени выделяются пиксельно-ориентированные подходы (pixel-oriented approaches). Это связано с возможностями их широкого использования.

Во-первых, они могут использоваться в качестве автономных инструментов исследования. Во-вторых, могут быть интегрированы в методы интеллектуального анализа данных, объединяя и усиливая существующие алгоритмы и участие человека. В-третьих, они могут быть использованы в кластеризации атрибутов по сходству для улучшения методов многомерной визуализации [34–37].

Помимо способности построить целостное визуальное представление сложных шаблонов базы данных, система многомерной визуализации поддерживает различные взаимодействия с пользователем, чтобы помочь аналитику понять закономерности. Многомерная система визуализации по своей природе относительно сложнее, чем традиционные подходы. В целом обозначенные выше подходы к моделированию основаны на строгих статистических или математических моделях, сформулированы с использованием априорных теоретических гипотез и откалиброваны с помощью данных наблюдений. Поэтому подходы, основанные на применении как визуальных, так и вычислительных методов, могут многое предложить при создании баз данных для изучения насильственных проявлений. Это позволит аналитикам выявлять неизвестные ранее тенденции или закономерности, таким образом побуждая к дальнейшим размышлениям и формулировке новых гипотез.

Третий уровень. Задачи интеллектуального анализа данных социальных медиа. За последнее десятилетие использование социальных сетей для распространения насильственных идеологий, дестабилизирующих общественный строй среди пользователей, резко возросло. Задача сбора информации из социальных сетей, даже при наличии ограничений, устанавливаемых провайдерами социальных сетей, является к настоящему моменту решенной благодаря разработкам и совершенствованию методов в области Web Mining. Проблема в качестве формализации переменных, кодировании исходных данных и точности установления корреляционной зависимости между значениями различных параметров.

Процесс интеллектуального анализа данных представим в виде блок-схемы на рис. 2.

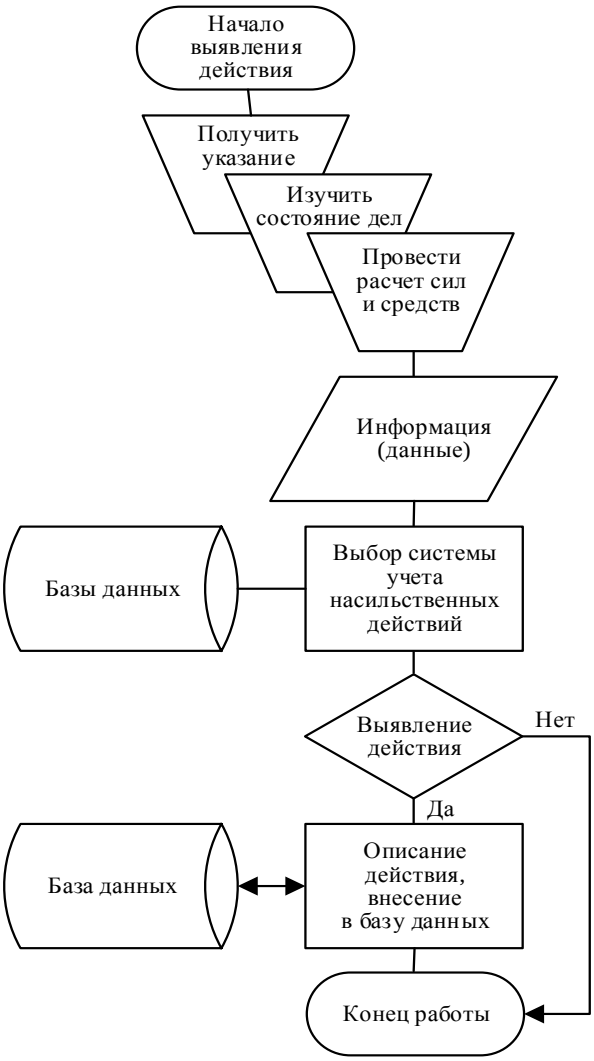

Рис. 2. Блок-схема выявления насильственного действия

Новые методы сбора данных предлагают огромные возможности, которые приводят к обработке все большего и большего количества геопространственных данных, а также связанной с ней увеличивающейся вычислительной нагрузкой. Анализ этих данных становится сложной задачей из-за размера наборов данных, их сложности, проблем с масштабированием и скрытых закономерностей. Применяемые на сегодняшний день статистические методы, искусственный интеллект (с использованием систем на основе правил и деревьев решений), искусственные нейронные сети представляют собой новое решение для анализа данных и распознавания образов. Среди моделей нейронных сетей самоорганизующаяся карта (SOM) часто рассматривается как многообещающий метод исследовательского анализа данных [24].

В качестве примеров рассмотрим несколько наиболее известных открытых баз данных, доступных исследователям со всего мира, обозначим их отличительные характеристики, а также аффилированные с базами данных инструменты и сервисы для анализа.

Одна из первых открытых баз данных – Всемирная система отслеживания инцидентов (Worldwide Incidents Tracking System (WITS)) [38] – начала формироваться в 2003 г. на базе Национального контртеррористического центра (The National Counterterrorism Center (NCTC)) [39]. Содержит информацию о глобальных насильственных экстремистских и террористических инцидентах с 2005 г.

База данных построена на основе переупорядочиваемой матрицы (reorderable matrices), т.е. организована как таблица данных с переменными в виде столбцов и элементами данных в виде строк. Каждый элемент данных (строка) имеет значение для каждой переменной (столбца). Переупорядочиваемая матрица может отсортировывать столбцы и строки так, чтобы похожие столбцы или похожие строки располагались рядом друг с другом.

Таким образом, появляются шаблоны для нескольких переменных и нескольких элементов данных. Значительные ограничения для аналитиков в работе с этой базой заключались в том, что в ней содержалась выборочная информация, не по всем странам. Кроме того, было ограниченное число переменных для анализа, соответственно, это ограничивало возможности прогнозировать тенденции в течение определенного периода времени. С 2012 г. WITS интегрирована в GTD.

GTD создана в 2005 г. на базе Национального консорциума по изучению терроризма и ответов на террористическую угрозу (National Consortium for the Study of Terrorism and Responses to Terrorism (START)) [40]. Представляет собой базу данных с открытым исходным кодом, включая информацию о террористических инцидентах по всему миру с 1970 по 2021 г. (с ежегодными обновлениями, запланированными на будущее). В отличие от многих других баз данных инцидентов GTD включает систематические данные о внутренних, а также международных террористических инцидентах, которые произошли в течение этого периода времени.

Содержит информацию по меньшей мере о 45 переменных для каждого инцидента, включая информацию о более чем 120 переменных. Более 100 структурированных переменных характеризуют местоположение, тактику, оружие, жертвы, пострадавших, а также общую информацию, такую как критерии определения и связи между скоординированными атаками. Это самая полная неклассифицированная база данных о террористических событиях в мире с открытым доступом. Но аналитикам, заинтересованным в работе с базой, открытый доступ предоставляется только по семи переменным (дата, регион, страна, преступные группы, оружие, тип атаки, цель атаки).

По остальным переменным можно получить данные только по запросу с обоснованием. Для того чтобы скачать данные и применить собственные параметры выборки, также необходимо прислать запрос с обоснованием и получить разрешение на использование данных. Файлы могут быть загружены непосредственно с веб-сайта START. Неструктурированные переменные включают краткое описание атак и более подробную информацию об используемом оружии, конкретных мотивах атакующих, имущественном ущербе и требованиях о выкупе (если применялся) [41].

Ключевые характеристики: информация из открытых источников, применение группы методов Web Mining (извлечение и уточнение знаний о людях, сообществах, структурах, системах, организациях, событиях, их взаимосвязях и взаимном влиянии), автоматический и полуавтоматический сбор данных, модели машинного обучения (ML), установленные правила кодирования (для оценки вкл./не вкл. инцидента в базу), критерии для переменных разработаны социологами, аналитика проводится рабочими группами.

Профили индивидуальной радикализации в США (Profiles of Individual Radicalization in the United States – PIRUS (Keshif)) [42]. PIRUS появилась в открытом доступе в 2014 г. Профили индивидуальной радикализации в наборе данных PIRUS содержат идентифицированную информацию по отдельным лицам: предыстории (предикторы процесса радикализации – формализованные переменные), атрибуты (неизменные признаки), фоновая аналитика (формализованные переменные) более 3000 насильственных и ненасильственных экстремистов. Самая большая открытая база данных в США и в мире. Она построена на изучении по четырем категориям идеологических платформ: крайне правые, крайне левые, исламисты, «одинокие волки».

Информация собирается только по экстремистам, радикализовавшимся в США, за период с 1948 г. по настоящее время (постоянно обновляется). В базе зашита обработанная информация, собранная из общедоступных источников информации. Персо-

нальные данные радикалов недоступны, им присвоен идентификационный номер. Создатели базы данных заявляют о 147 переменных, которые доступны для анализа. Но в открытом доступе для аналитиков представлена информация только по 49 переменным. Остальные доступны только по запросу с обоснованием. Инструмент (алгоритм) визуализации данных – Keshif – основан на использовании методов многомерной визуализации, таких как временная визуализация (temporal visualization), пиксельные методы (pixel-based techniques) и анимированные карты (animated maps). Агрегация и преобразование данных во всем наборе геокодирования трансформируются «на лету».

Матрица знаний I-VEO (I-VEO Knowledge Matrix) [43] (IVEO – Influencing Violent Extremist Organizations (влияние на воинствующие экстремистские организации)) – это пример сервиса, который адаптирован под кроссплатформенную интеграцию нескольких баз данных для решения исследовательских задач. Матрица появилась в 2011 г. как первая попытка синтезировать теоретические и эмпирические знания, накопленные в START об операциях влияния на воинствующие экстремистские организации.

Междисциплинарный проект, в который интегрированы специалисты из различных областей: политологии, криминологии, государственной политики, психологии, международных отношений и computer science. В качестве эмпирических данных в матрице I-VEO используются данные GTD и PIRUS. Проект по созданию матрицы направлен на то, чтобы объединить теоретические знания с эмпирически проверенными фактами. Матрица содержит 183 гипотезы о том, как влиять на VEO. Каждая гипотеза опирается на анализ качественных и/или количественных исследований. Для функциональности все гипотезы и связанные с ними обзоры литературы отсортированы по темам и могут быть заданы исследователями самостоятельно по нескольким схемам в соответствии с конкретными интересами.

Еще более широкие возможности для исследователей, аналитиков, практиков и заинтересованных лиц открываются на платформе базы данных о терроризме и экстремистском насилии в США (Terrorism and Extremist Violence in the United States (TEVUS) Database and Portal (TEVUS Portal)) [44, 45]. Этот новый инструмент для анализа терроризма и насильственного экстремизма базируется на интеграции информации из четырех баз данных: база данных американского исследования терроризма (American Terrorism Study (ATS)), GTD, база экстремистских преступлений (U.S. Extremist Crime Database (ECDB)), база профилей виновных в терроризме в Соединенных Штатах (Profiles of Perpetrators of Terrorism in the United States (PPT-US)).

Каждая из этих четырех баз данных имеет уникальные особенности как с точки зрения организации и интеграции в единое хранилище эмпирических данных, так и в возможностях многоуровневого представления данных, зашитых в интерфейс. На платформе TEVUS аккумулированы ключевые переменные (поведенческие, географические и временные характеристики) экстремистского насилия в Соединенных Штатах начиная с 1970 г. Через портал пользователи могут создавать поисковые запросы на основе четырех типов данных, включая конкретные события, исполнителей террористического акта или экстремистские преступления, группы и/или судебные дела, связанные с терроризмом и экстремистской преступностью в Соединенных Штатах. За счет интеграции такого массива данных из разных баз данных и разработки удобного веб-интерфейса для доступа к ним платформа TEVUS позволяет пользователям проводить многомерный анализ. Это самая полная база данных о терроризме и экстремистском насилии в Соединенных Штатах.

Все представленные примеры относятся только к одному исследовательскому центру START и имеют национальную специфику. Но именно эти образцы являются наиболее показательными в части организации сбора, обработки, хранения и анализа данных. На основе этих образцов в дальнейшем появились и, безусловно, будут появляться базы данных в разных регионах мира.

Примером обобщенного видения формирования базы данных является блок-схема выявления признаков насильственных действий, представленная на рис. 3.

Представленный обзор информационных систем страдает определенной неполнотой. В нем отсутствуют разработанные и используемые отечественные аналитические информационные системы, доступные в сегменте интернета.

Все поиски открытых баз данных в интересующей нас тематике не привели к положительному результату, хотя силовые структуры Российской Федерации, безусловно, не могут не использовать в своей деятельности по профилактике насильственных идеологий IT-разработки, связанные с формированием, созданием, построением подобных баз данных. Соответственно, не представляется возможным провести оценку ни функционала, ни содержания, ни алгоритмов, используемых в их работе.

Поэтому основное внимание в данной статье обращено на общедоступные базы данных, которые представляют исключительно зарубежные разработки.

Справедливости ради стоит отметить, что попытки создания подобных баз данных, размещенных в свободном доступе, предпринимались в нашей стране.

Информационно-аналитический центр «СОВА» (ИАЦ «СОВА») основан в октябре 2002 г. [46]. Сфера центра – проблемы национализма и ксенофобии, взаимоотношения религии и общества, политический радикализм. Статистику, собираемую по преступлениям ненависти и по антиэкстремистскому правоприменению, можно изучить в открытой базе данных. Данные о совершенных по мотиву ненависти насильственных преступлениях и актах ванда-

лизма (но не о пропагандистской деятельности), а также данные о приговорах по всем статьям УК, относящимся к «экстремистским», накапливаются в трех соответствующих базах данных. Данные систематически вводились с начала 2007 г. и доступны по месяц, закончившийся за 3 месяца до даты запроса. 30 декабря 2016 г. организация внесена в реестр некоммерческих организаций, выполняющих функции иностранного агента.

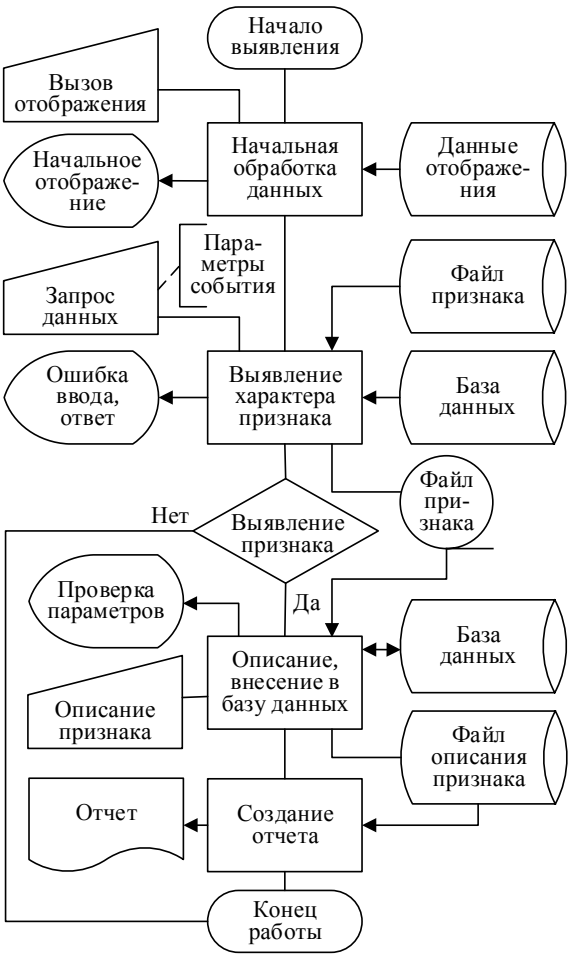

Рис. 3. Блок-схема выявления признаков насильственных действий

Основной недостаток ИАЦ «СОВА», не позволяющий ставить ее в один ряд с обозреваемыми зарубежными, – это отсутствие возможности формирования профиля нарушителя. Следовательно, о выявлении предпосылок возникновения угрозы насильственных актов, а также возможности предупреждения и профилактики этих явлений говорить не приходится.

Именно информационные технологии должны, на наш взгляд, кратно повысить эффективность работы по определению, выявлению и профилактике подобных угроз.

Выводы

В сегодняшней повестке ведется широкое обсуждение практических аспектов сбора данных в интернете как эффективном инструменте для исследований, позволяющем осуществлять быстрый и экономичный сбор данных, доступ к большим выборкам и различным группам населения. Преимущества, которые такой инструмент, как открытые базы данных, представляет для исследователей, заключаются в возможности интегрировать в единый репозиторий накопленный массив уже имеющихся данных, рассредоточенный по разным хранилищам. Это сократит время на поиск исследователями тематической и эмпирической информации.

Базы данных с заданными переменными позволяют исследователям практически мгновенно составить «картину» по заданным параметрам. Например, по конкретному лицу можно увидеть весь профиль с сопутствующими многомерными характеристиками. Кроме того, на материалах базы данных можно создать методологию исследования процесса, конкретного феномена, а также проводить сравнительный анализ сходств и различий проявлений и динамики процесса/феномена на региональном, федеральном, международном уровне. К тому же мы до конца не понимаем, что скрывается за «черным ящиком» алгоритмов описанных зарубежных баз данных в силу их недоступности.

В статье дан обзор существующего порядка вещей, обозначены проблемы и вызовы в стремлении систематизировать, структурировать, а далее автоматизировать систему анализа данных в нашей стране. Технически алгоритмы уже не представляют сложности, но должны быть адаптированы под конкретные задачи, диагностики, прогнозирования, предупреждения угроз и рисков национальной безопасности в Российской Федерации.

Статья подготовлена в рамках реализации программы ЛИЦ «Доверенные сенсорные системы» (договор № 009/20 от 10.04.2020) при финансовой поддержке Минкомсвязи России и АО «РВК». Идентификатор соглашения о предоставлении субсидии – 0000000007119P190002.

Литература

1. Попов К.В. Новые вызовы: стохастические угрозы национальной безопасности / К.В. Попов, А.А. Шелупанов // Доклады ТУСУР. – 2020. – Т. 23, № 4. – С. 23–29. DOI: 10.21293/1818-0442-2020-23-4-23-29.

2. Alava S. Youth and Violent Extremism on Social Media: Mapping the research / S. Alava, D. Frau-Meigs, G. Hassan // United Nations Educational, Scientific and Cultural Organization. – 2017. – URL: https://unesdoc.unesco.org/ark:/48223/pf0000260382 (дата обращения: 28.09.2022).

3. Sandler T. New frontiers of terrorism research: An introduction // Journal of Peace Research. – 2011. – Vol. 48, No. 3. – P. 279–286. DOI: 10.1177/0022343311399131.

4. Borum R. Radicalization into Violent Extremism I: A Review of Social Science Theories // Journal of Strategic Security. – 2011. – Vol. 4, No. 4. – P. 7–36.

5. RAND Corporation. – URL: https://www.rand.org/ (дата обращения: 28.09.2022).

6. Memorial Institute for the Prevention of Terrorism (MIPT). – URL: https://web.archive.org/web/20130620070250/https://www.mipt.org/Home.aspx (дата обращения: 28.09.2022).

7. National Consortium for the Study of Terrorism and Responses to Terrorism (START). – URL: https://www.start.umd.edu/ (дата обращения: 28.09.2022).

8. The National Counterterrorism Center (NCTC). – URL: https://www.dni.gov/index.php/nctc-newsroom/nctc-resources (дата обращения: 28.09.2022).

9. Canadian network for research on terrorism, security and society (TSAS). – URL: https://www.tsas.ca/ (дата обращения: 28.09.2022).

10. Center for Research on Extremism (C-REX). – URL: https://www.sv.uio.no/c-rex/english/ (дата обращения: 28.09.2022).

11. International Centre for Counter-Terrorism, The Netherlands (ICCT). – URL: https://icct.nl/ (дата обращения: 28.09.2022).

12. The International Centre for the Study of Radicalisation (ICSR). – URL: https://icsr.info/ (дата обращения: 28.09.2022).

13. The Handa Centre for the Study of Terrorism and Political Violence (CSTPV). – URL: https://cstpv.wp.st-andrews.ac.uk/ (дата обращения: 28.09.2022).

14. International Center for Terrorism Studies (ICTS). – URL: https://www.potomacinstitute.org/academic-centers/international-center-for-terrorism-studies-icts (дата обращения: 28.09.2022).

15. International Centre for Political Violence and Terrorism Research (PSiS). – URL: https://www.rsis.edu.sg/research/icpvtr/ (дата обращения: 28.09.2022).

16. China Institutes of Contemporary International Relations (CICIR). – URL: http://www.cicir.ac.cn/NEW/en-us/Institution.html?subtype=America&&type=region (дата обращения: 28.09.2022).

17. Freedman B. Terrorism Research Centres: 100 Institutes, Programs and Organisations in the Field of Terrorism, Counter-Terrorism, Radicalisation and Asymmetric Warfare Studies // Perspectives on Terrorism. – 2010. – Vol. 4, No. 5. – P. 48–56.

18. Ультраправая радикализация: методика автоматизированного выявления угроз методами web mining / А.Ю. Карпова, А.О. Савельев, А.Д. Вильнин, А.Ю. Кайда, С.А. Кузнецов, Н.Г. Максимова, Д.В. Чайковский // Вестник Российского фонда фундаментальных исследований. Гуманитарные и общественные науки. – 2020. – № 5 (102). – С. 30–43.

19. Прищеп С.В. Подходы и критерии оценки рисков информационной безопасности / С.В. Прищеп, С.В. Тимченко, А.А. Шелупанов // Безопасность информационных технологий. – 2007. – № 4. – С. 15–21.

20. Миронова В.Г. Методология формирования угроз безопасности конфиденциальной информации в неопределенных условиях их возникновения / В.Г. Миронова, А.А. Шелупанов // Изв. Южного федерального ун-та. Технические науки. – 2012. – Т. 137, № 12. – С. 39–45.

21. The high-level overview of social media content search engine / A.O. Savelev, A.Yu. Karpova, D.V. Chaykovskiy, A.D. Vilnin, A.Yu. Kaida, S.A. Kuznetsov, L.O. Igumnov, N.G. Maksimova // Proceedings of the 14th International Forum on Strategic Technology (IFOST 2019). – Tomsk: Polytechnic University, 2019. – P. 306–309.

22. Лопарев С.А. Анализ инструментальных средств оценки рисков утечки информации в компьютерной сети предприятия / С.А. Лопарев, А.А. Шелупанов // Вопросы защиты информации. – 2003. – № 4(63). – С. 2–5.

23. Миронова В.Г. Реализация модели Take-Grant как представление систем разграничения прав доступа в помещениях / В.Г. Миронова, А.А. Шелупанов, Н.Т. Югов // Доклады ТУСУР. – 2011. – № 2 (24), ч. 3. – С. 206–210.

24. Koua E.L. Using self - organizing maps for information visualization and knowledge discovery in complex geospatial datasets // In Proceedings of the 21st International Cartographic Conference (ICC) 10–16 Aug 2003. Durban. South Africa. – URL: https://webapps.itc.utwente.nl/library www/papers_2003/art_proc/koua.pdf (дата обращения: 28.09.2022).

25. Цифровизация финансово-кредитной сферы в современной России / Е.В. Агеева, М.А. Афанасова, А.С. Баландина [и др.]; под общ. ред. М.Г. Жигас, А.А. Шелупанова. – Москва; Берлин: Директ-Медиа, 2019. – 408 с. – URL: https://biblioclub.ru/index.php?page=book&id=565080 (дата обращения: 30.09.2022). DOI: 10.23681/565080.

26. Евсютин О.О. Приложения клеточных автоматов в области информационной безопасности и обработки данных / О.О. Евсютин, А.А. Шелупанов // Доклады ТУСУР. – 2012. – № 1 (25), ч. 2. – С. 119–125.

27. Chase L. Internet Research / L. Chase, J. Alvarez // Library & Information Science Research. – 2000. – Vol. 22, No. 4. – P. 357–369. DOI: 10.1016/s0740-8188(00)00050-5.

28. Hewson C. Internet Research Methods / C. Hewson, C. Vogel, D. Laurent // Internet Research Methods (2nd ed.). – London: Sage. – 2016. DOI: 10.4135/9781473920804.

29. The Global Terrorism Database™ (GTD). – URL: https://www.start.umd.edu/gtd/ (дата обращения: 28.09.2022).

30. Enders W. Domestic versus transnational terrorism: Data, decomposition, and dynamics / W. Enders, T. Sandler, K. Gaibulloev // Journal of Peace Research. – 2011. – Vol. 48, No. 3. – P. 319–338. – DOI: 10.1177/0022343311398926.

31. Sandler T. New frontiers of terrorism research: An introduction // Journal of Peace Research. – 2011. – Vol. 48, No. 3. – P. 279–286. DOI: 10.1177/0022343311399131.

32. Guo D. Visualizing patterns in a global terrorism incident database / D. Guo, K. Liao, M. Morgan // Environment and Planning B: Planning and Design. – 2007. – Vol. 34. – P. 767–784.

33. Модель угроз безопасности информации и ее носителей / А.К. Новохрестов, А.А. Конев, А.А. Шелупанов, Н.С. Егошин // Вестник Иркутского государственного технического университета. – 2017. – Т. 21, № 12 (131). – С. 93–104. DOI: 10.21285/1814-3520-2017-12-93-104.

34. Ankers M. Visual Data Mining with Pixel-oriented Visualization Techniques // The Boeing Company. P.O. Box 3707 MC 7L-70, Seattle, WA 98124. – URL: https://www.ics.uci.edu/~kobsa/courses/ICS280/notes/papers/ankerst-kdd2001.pdf (дата обращения: 28.09.2022).

35. Xu D. Comprehensive Survey of Clustering Algorithms / D. Xu, Y. Tian // Annals of Data Science. – 2015. – Vol. 2. – P. 165–193.

36. Wegmann M. A review of systematic selection of clustering algorithms and their evaluation / M. Wegmann, D. Zipperling, J. Hillenbrand, J. Fleischer. – 2021. – URL: https://arxiv.org/ftp/arxiv/papers/2106/2106.12792.pdf (дата обращения: 28.09.2022).

37. Shelupanov A. Information Security Methods-Modern Research Directions / A. Shelupanov, O. Evsyutin, A. Konev, E. Kostyuchenko, D. Kruchinin, D. Nikiforov // Symmetry. – 2019. – Vol. 11, No. 2. – P. 150. DOI: 10.3390/sym11020150.

38. Worldwide Incidents Tracking System (WITS). – URL: https://knoema.ru/tyytrod/violence-statistics-from-worldwide-incidents-tracking-system-wits (дата обращения: 28.09.2022).

39. The National Counterterrorism Center (NCTC). – URL: https://www.dni.gov/index.php/nctc-newsroom/nctc-resources (дата обращения: 28.09.2022).

40. National Consortium for the Study of Terrorism and Responses to Terrorism (START). – URL: https://www.start.umd.edu/ (дата обращения: 28.09.2022).

41. National Consortium for the Study of Terrorism and Responses to Terrorism (START). – URL: http://www.start-dev.umd.edu/gtd/using-gtd/ (дата обращения: 28.09.2022).

42. Profiles of Individual Radicalization in the United States – PIRUS (Keshif). – URL: https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif (дата обращения: 28.09.2022).

43. I-VEO Knowledge Matrix. – URL: http://start.foxtrotdev.com/ (дата обращения: 28.09.2022).

44. Terrorism and Extremist Violence in the United States (TEVUS) Database and Portal (TEVUS Portal). – URL: https://tap.cast.uark.edu/ (дата обращения: 28.09.2022).

45. Terrorism and Extremist Violence in the United States (TEVUS) Database and Portal (TEVUS Portal). – URL: https://www.start.umd.edu/tevus-portal (дата обращения: 28.09.2022).

46. Информационно-аналитический центр «СОВА» (ИАЦ «СОВА»). – URL: https://www.sova-center.ru/database/ (дата обращения: 28.09.2022).

Попов Константин Васильевич
Аспирант каф. комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС)
Томского государственного университета
систем управления и радиоэлектроники (ТУСУР)
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7-913-850-98-17
Эл. почта: pokkos@mail.ru

Шелупанова Полина Александровна
Канд. экон. наук, доцент,
зав. каф. экономической безопасности ТУСУРа
Ленина пр-т, 40, г. Томск, Россия, 634050
Тел.: +7 (382-2) 90-71-55
Эл. почта: polina.a.shelupanova@fb.tusur.ru

Popov K.V., Shelupanova P.A.

Information systems for national security threat analysis

Information technology plays a significant role in what has become known as «computational social science». The development of online technologies and digital research methods in the subject areas of social science is associated with the application of new interdisciplinary approaches and the integration of researchers from different fields of knowledge. As an effective and promising tool for researchers in the study of violent ideologies that destabilize the social order, the article discusses examples of databases, services and platforms for multivariate analysis. Opportunities and limitations of customizing data collection, processing, and presentation are outlined. An overview of the specific characteristics of the databases, services and platforms on the START Research Center website is presented. Conclusions are made about what opportunities such tools open for researchers and the prospects of realization of the most interesting solutions for improving the process of automation of data analysis are outlined.

Keywords*:* national security, extremism, terrorism, radicalization, database, web mining, big data, research tools.
DOI: 10.21293/1818-0442-2022-25-4-71-79

References

1. Popov K.V., Shelupanov A.A. [New challenges: stochastic threats to national security]. *Proceedings of TUSUR University*, 2020, vol. 23, no. 4, pp. 23–29. DOI: 10.21293/1818-0442-2020-23-4-23-29 (in Russ.).

2. Alava S., Frau-Meigs D., Hassan G. Youth and Violent Extremism on Social Media: Mapping the research, United Nations Educational, Scientific and Cultural Organization, 2017. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000260382, free (Accessed: September 28, 2022).

3. Sandler T. New frontiers of terrorism research: An introduction. *Journal of Peace Research*, 2011, vol. 48(3), pp. 279–286. DOI: 10.1177/0022343311399131.

4. Borum R. Radicalization into Violent Extremism I: A Review of Social Science Theories. *Journal of Strategic Security*, 2011, vol. 4(4), pp. 7–36.

5. RAND Corporation. Available at: https://www.rand.org/, free (Accessed: September 28, 2022).

6. Memorial Institute for the Prevention of Terrorism (MIPT). Available at: https://web.archive.org/web/20130620070250/https://www.mipt.org/Home.aspx, free (Accessed: September 28, 2022).

7. National Consortium for the Study of Terrorism and Responses to Terrorism (START). Available at: https://www.start.umd.edu/, free (Accessed: September 28, 2022).

8. The National Counterterrorism Center (NCTC). Available at: https://www.dni.gov/index.php/nctc-newsroom/nctc-resources, free (Accessed: September 28, 2022).

9. Canadian network for research on terrorism, security and society (TSAS). Available at: https://www.tsas.ca/, free (Accessed: September 28, 2022).

10. Center for Research on Extremism (C-REX). Available at: https://www.sv.uio.no/c-rex/english/, free (Accessed: September 28, 2022).

11. International Centre for Counter-Terrorism, The Netherlands (ICCT). Available at: https://icct.nl/, free (Accessed: September 28, 2022).

12. The International Centre for the Study of Radicalisation (ICSR). Available at: https://icsr.info/, free (Accessed: September 28, 2022).

13. The Handa Centre for the Study of Terrorism and Political Violence (CSTPV). Available at: https://cstpv.wp.st-andrews.ac.uk/, free (Accessed: September 28, 2022).

14. International Center for Terrorism Studies (ICTS). Available at: https://www.potomacinstitute.org/academic-centers/international-center-for-terrorism-studies-icts, free (Accessed: September 28, 2022).

15. International Centre for Political Violence and Terrorism Research (PSiS). Available at: https://www.rsis.edu.sg/research/icpvtr/, free (Accessed: September 28, 2022).

16. China Institutes of Contemporary International Relations (CICIR). Available at: http://www.cicir.ac.cn/NEW/en-us/Institution.html?subtype=America&&type=region, free (Accessed: September 28, 2022).

17. Freedman B. Terrorism Research Centres: 100 Institutes, Programs and Organisations in the Field of Terrorism, Counter-Terrorism, Radicalisation and Asymmetric Warfare Studies. Perspectives on Terrorism, 2010, vol. 4(5), pp. 48–56.

18. Karpova A.Yu., Savelev A.O., Vilnin A.D., Kaida A.Yu., Kuznetsov S.A., Maksimova N.G., Chaykovskiy D.V. [Ultra-right-wing radicalization: a methodology for automated threat detection using web mining methods]. *Vestnik Rossijskogo fonda fundamental'nyh issledovanij*. Gumanitarnye i obshchestvennye nauki, 2020, vol. 5(102), pp. 30–43 (in Russ.).

19. Prishchep S.V., Timchenko S.V., Shelupanov A.A. [Approaches and criteria for assessing information security

risks]. Bezopasnost' informatsionnykh tekhnologiy, 2007, no. 4, pp. 15–21 (in Russ.).

20. Mironova V.G., Shelupanov A.A. [Methodology for the formation of threats to the security of confidential information in uncertain conditions of their occurrence]. Izvestiya Yuzhnogo federalnogo universiteta. Tekhnicheskiye nauki, 2012, vol. 137, no. 12, pp. 39–45 (in Russ.).

21. Savelev A.O., Karpova A.Yu., Chaykovskiy D.V., Vilnin A.D., Kaida A.Yu., Kuznetsov S.A., Igumnov L.O., Maksimova N.G. [The high-level overview of social media content search engine]. *Proceedings of the 14th International Forum on Strategic Technology (IFOST 2019)*. Tomsk, Tomsk Polytechnic University, 2019, pp. 306–309 (in Russ.).

22. Loparev S.A., Shelupanov A.A. [Analysis of tools for assessing the risks of information leakage in the computer network of an enterprise]. *Voprosy zashchity informatsii*, 2003, no. 4(63), pp. 2–5 (in Russ.).

23. Mironova V.G., Shelupanov A.A., Yugov N.T. [Implementation of Take-Grant model as a representation of user access rights differentiation system in the building]. *Proceedings of TUSUR University*, 2011, no. 2(24), part 3, pp. 206–210 (in Russ.).

24. Koua E.L. Using self – organizing maps for information visualization and knowledge discovery in complex geospatial datasets, Proceedings of the 21st International Cartographic Conference (ICC), 10-16 Aug 2003. Durban. South Africa. Available at: https://webapps.itc.utwente.nl/library www/papers_2003/art_proc/koua.pdf, free (Accessed: September 28, 2022).

25. Ageeva E.V., Afanasova M.A., Balandina A.S. [and others]; under total ed. Zhigas M.G., Shelupanov A.A. Tsifrovizatsiya finansovo-kreditnoy sfery v sovremennoy Rossii [Digitalization of the financial and credit sphere in modern Russia]. Moscow; Berlin, Direct-Media, 2019. 408 p. Available at: https://biblioclub.ru/index.php?page=book&id=565080, free (Accessed: September 28, 2022). DOI 10.23681/565080 (in Russ.).

26. Evsutin O.O., Shelupanov A.A. [Applications of cellular automata in the field of information security and data processing]. *Proceedings of TUSUR University*, 2012, no. 1(25), part 2, pp. 119–125 (in Russ.).

27. Chase L., Alvarez J. Internet Research. Library & Information Science Research, 2000, vol. 22(4), pp. 357–369. DOI: 10.1016/s0740-8188(00)00050-5.

28. Hewson C., Vogel C., Laurent D. Internet Research Methods. Internet Research Methods (2nd ed). London: Sage, 2016. DOI: 10.4135/9781473920804.

29. The Global Terrorism Database™ (GTD). Available at: https://www.start.umd.edu/gtd/, free (Accessed: September 28, 2022).

30. Enders W., Sandler T., Gaibulloev K. Domestic versus transnational terrorism: Data, decomposition, and dynamics. *Journal of Peace Research*, 2011, vol. 48(3), pp. 319–338. DOI: 10.1177/0022343311398926.

31. Sandler T. New frontiers of terrorism research: An introduction. *Journal of Peace Research*, 2011, vol. 48(3), pp. 279–286. DOI: 10.1177/0022343311399131.

32. Guo D., Liao K., Morgan M. Visualizing patterns in a global terrorism incident database. *Environment and Planning B: Planning and Design*, 2007, vol. 34, pp. 767–784.

33. Novokhrestov A.K., Konev A.A., Shelupanov A.A., Egoshin N.S. [Information and information carrier security threat model]. *Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta*, 2017, vol. 21, no. 12, pp. 93–104. DOI: 10.21285/1814-3520-2017-12-93-104 (in Russ.).

34. Ankers M. Visual Data Mining with Pixel-oriented Visualization Techniques, The Boeing Company. P.O. Box 3707 MC 7L-70, Seattle, WA 98124. Available at: https://www.ics.uci.edu/~kobsa/courses/ICS280/notes/papers/ankerst-kdd2001.pdf, free (Accessed: September 28, 2022).

35. Xu Dongkuan, Ying-jie Tian. Comprehensive Survey of Clustering Algorithms. *Annals of Data Science*, 2015, vol. 2, pp. 165–193.

36. Wegmann M., Zipperling D., Hillenbrand J., Fleischer J. A review of systematic selection of clustering algorithms and their evaluation, 2021. Available at: https://arxiv.org/ftp/arxiv/papers/2106/2106.12792.pdf, free (Accessed: September 28, 2022).

37. Shelupanov A., Evsyutin O., Konev A., Kostyuchenko E., Kruchinin D., Nikiforov D. Information Security Methods-Modern Research Directions. Symmetry, 2019, vol. 11 (2), pp. 150. DOI: 10.3390/sym11020150.

38. Worldwide Incidents Tracking System (WITS). Available at: https://knoema.ru/tyytrod/violence-statistics-from-worldwide-incidents-tracking-system-wits, free (Accessed: September 28, 2022).

39. The National Counterterrorism Center (NCTC). Available at: https://www.dni.gov/index.php/nctc-newsroom/nctc-resources, free (Accessed: September 28, 2022).

40. National Consortium for the Study of Terrorism and Responses to Terrorism (START). Available at: https://www.start.umd.edu/, free (Accessed: September 28, 2022).

41. National Consortium for the Study of Terrorism and Responses to Terrorism (START). Available at: http://www.start-dev.umd.edu/gtd/using-gtd/, free (Accessed: September 28, 2022).

42. Profiles of Individual Radicalization in the United States – PIRUS (Keshif). Available at: https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif, free (Accessed: September 28, 2022).

43. I-VEO Knowledge Matrix. Available at: http://start.foxtrotdev.com/, free (Accessed: September 28, 2022).

44. Terrorism and Extremist Violence in the United States (TEVUS) Database and Portal (TEVUS Portal). Available at: https://tap.cast.uark.edu/, free (Accessed: September 28, 2022).

45. Terrorism and Extremist Violence in the United States (TEVUS) Database and Portal (TEVUS Portal). Available at: https://www.start.umd.edu/tevus-portal, free (Accessed: September 28, 2022).

46. SOVA Center for Information and Analysis. Available at: https://www.sova-center.ru/database/, free (Accessed: September 28, 2022) (in Russ.).

Konstantin V. Popov
Postgraduate Student, Department of Integrated Information Security of Electronic Computing Systems
Tomsk State University of Control Systems and Radioelectronics (TUSUR)
40, Lenin pr., Tomsk, Russia, 634050
Phone: +7-913-850-98-17
Email: pokkos@mail.ru

Polina A. Shelupanova
Candidate of Economic Sciences, Associate Professor,
Head of the Department of Economic Security TUSUR
40, Lenin pr., Tomsk, Russia, 634050
Phone: +7 (382-2) 90-71-55
Email: polina.a.shelupanova@fb.tusur.ru