ОБНАРУЖЕНИЕ ПРОТИВОПРАВНОЙ ДЕЯТЕЛЬНОСТИ В КИБЕРПРОСТРАНСТВЕ НА ОСНОВЕ АНАЛИЗА СОЦИАЛЬНЫХ СЕТЕЙ: АЛГОРИТМЫ, МЕТОДЫ И СРЕДСТВА (ОБЗОР)

Басараб М.А.[1], Иванов И.П.[2], Колесников А.В.[3], Матвеев В.А.[4]

Рассмотрены основные методы анализа социальных сетей применительно к задаче выявления подозрительных и преступных сообществ по данным зарубежных источников. Представлен обзор открытых и коммерческих программных средств, предназначенных для обработки и визуализации данных, полученных путем обработки социальных графов. Сделаны выводы об основных характеристиках автоматизированной системы поддержки принятия решений в такого рода задачах.

***Ключевые слова**: социальная сеть, SNA, преступная сеть*

DOI 10.21681/2311-3456-2016-4-11-19

Введение

Анализ отношений в рамках социальных групп и взаимодействие между ними возник, как часть социологии задолго до развития телекоммуникационных технологий. Очевидной трудностью, с которой сталкивались исследователи в данной области являлся недостаток личных данных и сведений о социальной активности различных групп населения [1]. Все изменилось с появлением Интернета и развитием онлайн сервисов социальных сетей, которые предоставили колоссальные объемы данных для анализа взаимодействия и поведения пользователей, выделения сообществ, классификации на основе широкого спектра атрибутов.

Анализ социальных графов (Social Network Analysis - SNA) – это применение теории графов для исследования социальных сетей (социальных графов - СГ) с точки зрения социальных взаимоотношений. В качестве объектов исследований выступают узлы (лица, организации и др.), и связи, характеризующие отношения между ними (дружба, общение, финансовые переводы и др.). Социальные взаимоотношения могут выступать в форме реальных («оффлайновых») социальных сетей либо виртуальных («онлайновых») социальных сетей (Twitter, Facebook и др.). Ранее SNA широко использовался в таких областях как информатика, политология, социальная психология, биология, менеджмент, экономика, разведка. Такие сервисы как Twitter, Facebook и др. имеют развитый инструментарий для разработки стратегий и политик пользователей.

В работе рассмотрены основные методы анализа социальных сетей применительно к задаче выявления подозрительных и преступных сообществ по данным зарубежных источников. Представлен обзор программных средств, предназначенных для обработки и визуализации данных. Сделаны выводы об основных характеристиках автоматизированной системы поддержки принятия решений в такого рода задачах.

Методы SNA

Согласно [2] SNA определяется четырьмя факторами:

1) Мотивация на структурной интуиции, основанной на связях участников сети.

2) Исследование, основанное на систематизированных эмпирических данных.

3) Использование графического представления.

4) Применение математических и/или вычислительных моделях для предсказания будущего поведения.

Сам SNA включает в себя три основных составляющие [3]:

1) Проведение эмпирических исследований структуру СГ с использованием таких средств как интервьюирование, прямое наблюдение, изучение архивных данных и др.

2) Использование математических и статистических методов для получения ответов на вопросы о сообществе.

1 Басараб Михаил Алексеевич, доктор физико-математических наук, профессор, МГТУ им. Н.Э. Баумана, Москва, bmic@mail.ru
2 Иванов Игорь Потапович, доктор технических наук, МГТУ им. Н.Э. Баумана, Москва, ivanov@bmstu.ru
3 Колесников Александр Владимирович, кандидат технических наук, МГТУ им. Н.Э. Баумана, Москва, avkolesnikov90@list.ru
4 Матвеев Валерий Александрович, доктор технических наук, профессор, МГТУ им. Н.Э. Баумана, Москва, a.matveev@bmstu.ru

3) Создание математических и компьютерных моделей для воспроизведения процессов в сетевых системах.

Анализ метрик вершин (узлов) графа (*centrality measures*) эффективно используется для обнаружения связных сообществ. Наиболее распространенными метриками [4] являются следующие.

Усредненная степень вершины (*degree centrality*)

$$D_i = \frac{k_i}{N-1} = \frac{\sum_{j \in G} a_{ij}}{N-1}$$

определяется по матрице связности $[a_{ij}]$ СГ G и равна количеству смежных вершин, нормированных на максимально возможное их количество ($N-1$); не зависит от размера сети N и изменяется в интервале от 0 до 1; согласно гипотезе вершина с большим значением параметра D_i имеет высокую степень активности и информационного влияния в своей окрестности;

Степень промежуточности (*betweenness centrality*)

$$B_i = \frac{\sum_{i<k \in G} n_{jk}(i) / n_{jk}}{(N-1)(N-2)},$$

где n_{jk} - количество всех кратчайших путей между узлами j и k; $n_{jk}(i)$ - количество кратчайших путей между узлами j и k, проходящих через узел i.

Узлы с высоким значением показателя B_i («информационные брокеры») играют ключевую роль в распространении информации между двумя или более плотно связанными подмножествами СГ; удаление таких вершин может оказать сильное негативное воздействие на весь СГ.

Степень близости (*closeness centrality*)

$$C_i = (L_i)^{-1} = \frac{N-1}{\sum_{j \in G} d_{ij}},$$

где d_{ij} - расстояние между узлами i и j; L_i - нормированное расстояние между узлом i и другими вершинами СГ.

Через узлы с малой степенью близости можно легче и быстрее достичь остальных вершин СГ.

Метрика собственных векторов (*eigenvector centrality*)

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j = \frac{1}{\lambda} \sum_{j \in G} a_{ij} x_j,$$

где $M(i)$ - набор соседей вершины i; λ - константа (собственное число).

определяет, насколько хорошо узел i связан с другими хорошо связанными узлами.

Индекс PageRank [5] дает возможность сравнить относительную «важность» узлов на основе аналогии между гиперссылками веб-страниц и связями в СГ. Рекурсивно определяется как

$$PR(A) = (1-d) + d \sum_{B \in M(A)} \frac{PA(B)}{L(B)},$$

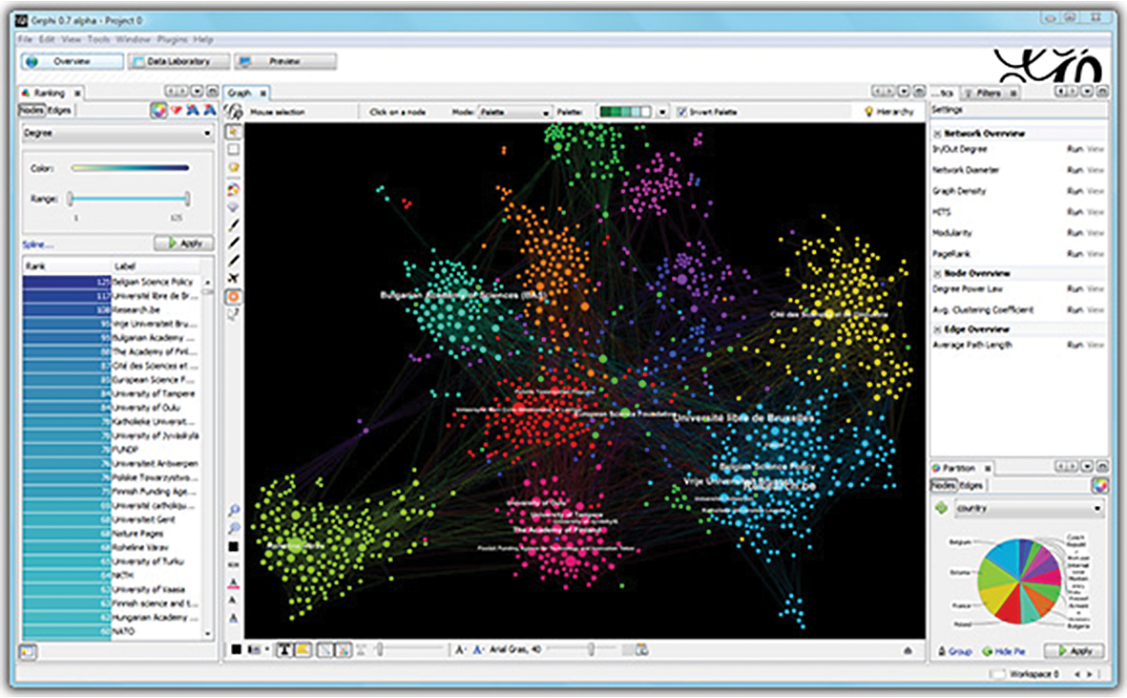

Рис. 1. *Скриншот ПО Gephi*

где $M(A)$ - множество соседей узла A; $L(B)$ - количество исходящих из узла B связей; d – коэффициент затухания.

Обзор других типов метрик и различных типов исследуемых графов приведен в [3].

Визуализация социального графа. Для визуализации социального графа наиболее популярными в настоящее время являются алгоритмы, основанные на аналогиях физических принципов притяжения и отталкивания тел или частиц по закону Гука, Кулона и др. с целью минимизации энергии системы (Force-directed graph drawing) [6]. В частности, одним из широко используемых алгоритмов является Force Atlas 2 [7] и специализированное ПО Gephi [8], находящееся в открытом доступе по ссылке https://gephi.org/ (рис. 1).

Программные средства анализа и визуализации сетевого графа. В работе [9] приведен достаточно полный обзор ПО для SNA и визуализации (табл. 1).

Таблица 1.

Сравнение различных программных средств анализа СГ [9]

ПО	Функции	Платформа	Лицензия	Форматы файлов	Ограничения
UCINET	SNA и визуализация посредством NetDraw, различные метрики центральности и мощности, алгоритмы кластеризации и объединения, импорт и экспорт различных форматов данных, работа с большими объемами данных	Windows	коммерческая и академическая версии	.dl, .net, .vna, .csv, матрицы	платформо-зависимость, немасштабируемость, отсутствие динамического анализа
Gephi	SNA и визуализация, различные метрики центральности, кластеризация и модульность, ранжирование исегментация сети, временная шкала, доступна поддержка плагинов	Windows, Linux	open-source	.dot, .gml, .gdf, .graphml, .net, .dl, .gexf, .csv, databases	не пригоден для очень больших СГ
ORA	SNA и визуализация двухрежимных и мультимодальных сетей, различные средства визуализации, динамический анализ, метрики центральности и мощности, кластеризация, генерация отчета	Windows	коммерческая и академическая версии	.xml, .dynetml, .zip	платформо-зависимость, ограниченное число узлов
NodeXL	шаблон MS Excel, визуализация и SNA, прямой импорт данных из онлайновой социальной сети, различные форматы экспорта данных, макеты визуализации, метрики анализа	Windows	open-source	.dl, .net, .graphml, матрицы	нет возможности управления и визуализации данных

ПО	Функции	Платформа	Лицензия	Форматы файлов	Ограничения
JUNG	Java API и библиотеки для моделирования, анализа и визуализации СГ, различные алгоритмы центральности, кластерного анализа, потоков и др., возможность работы с большими массивами данных	все платформы	open-source	.net, .graphml	отсутствие пользовательского интерфейса, нет динамического анализа
Pajek	SNA и визуализация больших СГ, анализ двухрежимных, мультимодальных и временных СГ, алгоритмы центральности и макеты СГ	Windows, Linux	open-source	.dl, .net	ограниченное число узлов, слабые возможности настройки
NetworkX	библиотека Python для SNA, двухрежимные и мультимодальные сети, некоторые центральные метрики, работа с большими сетями	Windows, Linux	open-source	.gml, .graphml, .net	нет интерфейса, нет прямого импорта графа
Igraph	библиотеки R, Python, C для SNA, поддержка двухрежимных СГ, множество метрик центральности, кластеризации и макетов, работа с очень большими сетями (миллионы связей)	Windows, Linux	open-source	.gml, .graphml, .net	нет интерфейса, нет прямого импорта данных из онлайновой социальной сети

Анализ СГ в задачах выявления преступных сообществ и информационного противоборства. Использование методов анализа социальных сетей для решения задач информационного противоборства и влияния всегда привлекало внимание исследователей [10]. Особую роль SNA приобрел для выявления преступных сообществ после террористических атак 11.09.2001 г. на башни Всемирного Торгового Центра в Нью-Йорке. Одной из первых работ, посвященных вопросам организации террористической сети «Аль-Каида» была статья [11]. В ней V. Krebs сделал попытку сконструировать граф 19 соучастников «Аль-Каиды» по открытым источникам и информации, опубликованной в ведущих изданиях (the New York Times, the Wall Street Journal, the Washington Post и the Los Angeles Times). Самые сильные связи были между узлами, соответствующими лицам, которые жили либо учились вместе; средняя сила связей соответствовала лицам, путешествовавшим вместе либо участвовавших на одних и тех же встречах; наконец, слабые связи были между лицами, имевшими случайные связи либо финансовые взаимоотношения. Анализ СГ показал, что один из пилотов (Mohammed Atta) имел максимальную степень вершины и степень близости, что говорило о том ,что он контактировал практически со всеми угонщиками. Однако его степень связности (betweenness) была низка, в силу чего он не мог являться их руководителем и не мог иметь наибольшего влияния на членов группы [11].

Отмечалось, что мир вошел в эпоху низкоинтенсивных сетевых войн (netwar), ведомых террористическими, криминальными и экстремистскими организациями, зачастую не имеющими единого центра (лидера), что дает им возможность необычайно быстрого и гибкого реагирования [12]. Там же указывалось на то, что сетевая структура такой организации как «Аль-Каида» напрямую влияет на ее способность воспринимать новые идеи, набирать новых людей и достигать устойчивости». Отмечалось, что неверным является традиционное представление об иерархической структуре «Аль-Каиды» с «эмиром» Бен Ладеном

во главе, окруженным небольшим кругом советников («шура») и многочисленными комитетами в подчинении, отвечающими за военные операции, религиозные вопросы, финансы, производство поддельных документов и др. [13]. Там же отмечалось, что «Аль-Каида» скорее представляет собой необычайно стойкую сеть благодаря наличию большого числа самоорганизованных многоцелевых групп с высокой внутренней дисциплиной и способностью принимать самостоятельные решения, подчиняющихся одной цели, связанной с религиозным фанатизмом.

SNA может быть полезен в исследовании таких различных вопросов функционирования террористических организаций как набор новых людей, эволюция сети, распространение религиозных идей. С ростом задач и уровня сложности SNA специалисты-исследователи стали чаще подразделяться на две группы: специалисты по сбору данных и специалисты по моделированию [12].

Анализ преступных сообществ в СГ позволяет получить ценные знания относительно организации этих структур, в частности, об их размерах и уровне централизации. Организация преступных сообществ осуществляется через скрытые онлайн-сообщества (форумы), позволяющие производить обмен информации и криминальные платежи [14]. Получение более глубоких знаний об особенностях деятельности и коммуникации криминальных сообществ позволяет более эффективно противодействовать достижению их преступных целей. Утечка данных о членах сообществ может быть получена путем анализа их экономической деятельности [15].

Как правило при изучении преступных сообществ наиболее важным является выявление следующих факторов:

- кто из сообщников занимает центральное положение в СГ;
- какие подгруппы и сообщества наличествуют в СГ;
- кто из сообщников выступает в роли передаточных звеньев («брокеров») в осуществлении деятельности и при распространении информации;
- какой ранг имеют сообщники в зависимости от их важности и влиятельности в СГ.

Топологическая структура криминальных сетей (dark networks – темные сети) рассматривалась в работе [16] на примерах сети Global Salafi Jihad, сети метамфетаминового трафика и террористической веб-сети. С использованием аналогичного подхода в пакете Maple 13 была создана модель кругового графа для сети «Аль-Каида» [3] по данным, полученным из статей, книг и правительственных публикаций (рис. 2). Было показано, что рассчитанные параметры средних расстояний между узлами, коэффициентов кластеризации и глобальной эффективности подтвердили тезис о том, что сеть может рассматриваться как «темная» и описывается моделью «малого мира» (small world).

Следует отметить, что несмотря на то, что изучению сетевой структуры «Аль-Каиды» за последние полтора десятка лет было посвящено большое количество работ (кроме упомянутых работ [3, 11-13] можно упомянуть также [17]), интерес вызывали и другие преступные сообщества. В частности, в [18] проводится сетевой анализ пропагандистской кампании в сети YouTube британской террористической организации Al-Muhajiroun. Отмечается, что европейские террористические сообщества могут использовать зарубежные (США) платформы и доменные имена, чтобы под прикрытием Первой поправки к конституции США, гарантирующей свободу речи, уклониться от европейских законов, преследующих подстрекательство и пропаганду ненависти.

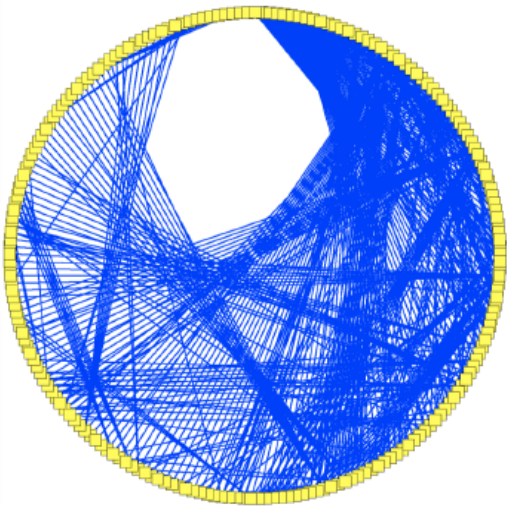

Рис. 2. *Кольцевая модель сетевого графа «Аль-Каиды» [2]*

В работе [19] рассматривался вопрос воссоздания крупного СГ по неполному набору данных, полученных открытым путем и содержащих только электронные адреса членов преступной группы (Nigerian fee fraud scammers social network - нигерийская фродовая мошенническая сеть). В итоге был создан СГ, состоящий из 43 тыс. узлов на основании всего около 1 тыс. исходных адресов, добытых в результате утечки с использованием связанных профилей Facebook, имеющих те же адреса. Далее выполнялся полномасштабный

анализ СГ, включая ручной анализ эффективности алгоритма PageRank [5] идентификации потенциальных криминальных профилей Facebook, для выявления ключевых лиц, крупных преступных подгрупп и особенностей культуры и организации сообщества.

В работе [9] приводится обзор работ по применению SNA для изучения сети террористических организаций, действующих в Индии, в частности, в штате Джамму и Кашмир, а также в г. Мумбаи накануне событий 26.11.2008 г. Указывается на сложности, связанные со сбором информации в СГ (неполнота, недостоверность, высокая степень децентрализации, активная динамика и др.), приводящим к некорректному анализу СГ. Приводится обзор и сравнительных программных средств сбора и анализа данных СГ, даются рекомендации по выбору программных средств и реализации корректного сбора данных.

В качестве резюме следует отметить работу [20], в которой выделены 7 методов, которые должны использоваться в тактике дестабилизации террористических сетей:

1) Выявление ключевых сущностей (сообществ) и связей между ними.

2) Выявление ключевых процессов, с помощью которых добавляются или удаляются сущности и связи, и (в случае последних) изменяется их сила.

3) Сбор данных о системе (скрытой сети).

4) Определение характеристик полученной системы.

5) Определение характеристик возможной оптимальной системы.

6) Определение уязвимостей и выбор стратегий по дестабилизации.

7) Определение характеристик системы в краткосрочной и долгосрочной перспективе после применения выбранных стратегий дестабилизации.

Выводы

Стоит отметить важность изучения различных сетевых структур для решения различных задач информационной безопасности. Например, обнаружение вторжений в компьютерные системы, расследование инцидентов, борьба с ботнетами, изучение динамики распространения вирусов, слежение за телефонными и социальными сетями, информационное влияние и управление. Актуальность исследований подтверждаются тем, что с помощью алгоритмов выделения сообществ можно выделить зараженные узлы в компьютерной сети, кардеров в сети банковских транзакций, ботов и поддельные аккаунты в социальных сетях, недобросовестные СМИ. В сетях доверия PGP многие программы проверяют лишь один уровень доверия, поэтому существует возможность их обмана с помощью нескольких поддельных сертификатов. Подобное мошенничество можно обнаружить с помощью выделения сообществ. Исследуя структуру сети, администратор безопасности может выделить наиболее вероятные точки отказа (например, те узлы, которые обладают наибольшей промежуточностью, наибольшей степенью или наибольшим PageRank и дополнительно их защитить или осуществить резервирование. Злоумышленники, в свою очередь, могут определить таким образом наиболее уязвимые места для атаки, которая может быть направлена на нарушение доступности информации (вызов отказа в обслуживании), конфиденциальности (прослушка, мониторинг трафика) или целостности (подмена/модификация пакетов на маршрутизаторе). Аналогичными способами можно выделить лидеров общественного мнения с целью использования в информационном противоборстве и распространения дезинформации.

Анализ зарубежных литературных источников показал необходимость использования широкого класса программно-математических средств и методов в составе автоматизированной системы поддержки принятия решения при выявлении преступных сообществ в сетях, анализе их деятельности и организации информационного противоборства.

На основании анализа литературных данных и предшествующих работ авторов [21-30] можно перечислить следующие требования к системе поддержки принятия решений.

1) Мультиплатформенность (включая Windows и Linux).

2) Возможность сбора (включая прямой импорт данных из социальной сети), хранения и экспорта в различных форматах данных «больших» социальных сетей (~10^6 связей), включая информацию профилей узлов.

3) Развитые возможности визуализации сетевых графов, включая «большие», анализа широкого спектра их характеристик (SNA) с целью выявления потенциальных угроз, осуществления прогноза (предсказание) и моделирования противодействия. На первом этапе возможно использование открытого ПО (Gephi, Tulip и др.) с созданием собственных специализированных программных модулей и библиотек на языках высокого уровня (C++ и др.).

4) Использование широкого спектра средств статического и динамического анализа сетевых графов на основе следующих математических методов и подходов:

- теория вероятности и математическая статистика, включая новые подходы к статистическому анализу социальных графов на основе оценки их спектральных характеристик [21,22];

- методы искусственного интеллекта и машинного обучения на основе графовых нейронных сетей (GNN – Graph Neural Network) и иерархической темпоральной памяти (HTM - Hierarchical Temporal Memory) (предполагается использовать впервые; ранее HTM, в частности, широко использовались в задачах распознавания образов) [23,24];

- обобщение аппарата биматричных и позиционных игр на случай возможности осуществления сторонами информационного противоборства в сети нескольких ходов одновременно, наличия временных ограничений на совершение ходов, изменения ресурсов и целей игроков и др. [24];

- имитационное моделирование, позволяющее сконструировать СГ произвольного уровня сложности с учетом требуемых топологических характеристик: спектральные параметры, фрактальная размерность и др.; это позволит, с одной стороны, получить возможность исследовать процесс образования преступных сообществ, а с другой – получить возможность восстановления большого СГ по неполным и неточным данным о нем.

5) Использование средств лингвистического и мультимодального анализа информации, передаваемой между пользователями социальной сети, с целью выявления потенциальных угроз, уточнения особенностей организации и иерархии сети и др.

6) Наличие развитого интерфейса, а также возможности по автоматической генерации отчета.

***Рецензент**: Цирлов Валентин Леонидович, кандидат технических наук, доцент МГТУ им. Н.Э.Баумана, г. Москва, v.tsirlov@bmstu.ru*

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-29-09517 офи_м.

Литература

1. Wasserman S., Faust K. Social Network Analysis. - Cambridge: Cambridge University Press, 1994.
2. Freeman L.C. The development of social network analysis: A study in the sociology of science. – North Charleston, SC: Book Surge, LLC, 2004.
3. Hopkins A. Graph theory, social networks and counter terrorism. May 19, 2010. Univ. of Massachusetts Dartmouth. URL: https://compmath.files.wordpress.com/2010/05/ ahopkins_freports10.pdf.
4. Freeman L.C. Centrality in social networks conceptual clarification // Social networks. - 1979. – 1(3). - P. 215-239.
5. Page L., Brin S., Motwani R., Winograd T. The pagerank citation ranking: bringing order to the web // January 29, 1998. URL: http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf.
6. Kobourov, Stephen G. Spring Embedders and Force-Directed Graph Drawing Algorithms. 2012. URL: http://arxiv.org/abs/1201.3011.
7. Jacomy M., Heymann S., Venturini T., Bastian M. ForceAtlas2, a graph layout algorithm for handy network visualization // Paris. August 29, 2011. URL: http://webatlas.fr/ tempshare/ForceAtlas2_Paper.pdf
8. Bastian M., Heymann S., Jacomy M., Gephi: an open source software for exploring and manipulating networks // International AAAI Conference on Weblogs and Social Media. Association for the Advancement of Artificial Intelligence, 2009.
9. Choudhary P., Singh U. A survey on social network analysis for counter-terrorism // Int. Journal of Computer Applications. - Feb. 2015. – Vol. 112. – No. 9. – P. 24-29.
10. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. М.: Издательство физико-математической литературы, 2010. 228 С.
11. Krebs V. Mapping networks of terrorist cells // Connections. – 2002. – 24(3). – P.43-52.
12. Ressler S. Social Network Analysis as an Approach to Combat Terrorism: Past, Present, and Future Research // Homeland Security Affairs. – July 2006. – 2. Article 8. URL: https://www.hsaj.org/articles/171.
13. Rothenberg R. From whole cloth: Making up the terrorist network // Connections. – 2002. - 24(3). – P.36-42.
14. Motoyama M., McCoy D., Levchenko K., Savage S., Voelker G.M. An analysis of underground forums // Proc. of the ACM Internet Measurement Conference. Berlin, CA, Nov. 2011.
15. McCoy D., Dharmdasani H., Kreibich C., Voelker G.M., Savage S. Priceless: the role of payments in abuse advertised goods // Proc. of the ACM Conference on Computer and Communications Security. Raleigh, NC, Oct. 2012.
16. Xu J., Chen H. The topology of dark networks // Communications of the ACM. – 2008. -51(10). – P.58-65. New York, NY. URL: http://ai.arizona.edu/intranet/papers/Xu-SNA- 2008.pdf.
17. Wu E., Carleton R., Davies G. Discovering bin-Laden's Replacement in al-Qaeda, using Social Network Analysis: A Methodological Investigation // Perspectives on Terrorism. – 2014. - Vol. 8. - No.1.
18. Klausen J., Tschaen Barbieri E., Reichlin-Melnick A., Zelin A.Y. The YouTube Jihadists: A Social Network Analysis of Al-Muhajiroun's Propaganda Campaign // Perspectives on Terrorism. – 2012. - Vol. 6. - No.1.

19. Sarvari H., Abozinadah E., Mbaziira A., McCoy D. Constructing and analysing criminal networks // IEEE Security and Privacy Workshops, 2014, pp. 84-91. DOI 10.1109/SPW.2014.22.
20. Carley K.M., Reminga J., Kamneva N. Destabilizing terrorist networks // NAACSOS conference proceedings. - 2003. - Pittsburgh, PA.
21. Чесноков В.О., Ключарёв П.Г. Выделение сообществ в социальных графах по множеству признаков с частичной информацией // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. 2015. № 9. С. 188-199. DOI: 10.7463/0915.0811704.
22. Ключарёв П.Г., Чесноков В.О. Исследование спектральных свойств социального графа сети LiveJournal // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. 2013. № 9. С. 391-400. DOI: 10.7463/0913.0603441.
23. Басараб М.А., Вельц С.В. Иерархическое представление компьютерной сети на основе нейронной сети Хопфилда // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. 2013. № 9. С. 335-348. DOI: 10.7463/0913.0630141.
24. Басараб М.А., Вельц С.В. Теоретико-игровой подход к оценке рисков и нахождению уязвимостей в сетях передачи информации // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. 2013. №8. С. 271-280. DOI: 10.7463/0813.0630132.

DETECTION OF ILLEGAL ACTIVITIES IN CYBERSPACE ON THE BASIS OF THE SOCIAL NETWORKS ANALYSIS: ALGORITHMS, METHODS, AND TOOLS (A SURVEY)

Basarab M.[5], Ivanov I.[6], Kolesnikov A.[7], Matveev V.[8]

The basic methods of social network analysis (SNA) as applied to detection of suspicious and criminal groups, are considered. A review of open source and commercial software for processing and visualization of social network data is presented. The conclusions about the main characteristics of the automated system of support of decision–making in this kind of problems are given.

Keywords*: social network, SNA, criminal network*

References

1. Wasserman S., Faust K. Social Network Analysis. - Cambridge: Cambridge University Press, 1994.
2. Freeman L.C. The development of social network analysis: A study in the sociology of science. – North Charleston, SC: Book Surge, LLC, 2004.
3. Hopkins A. Graph theory, social networks and counter terrorism. May 19, 2010. Univ. of Massachusetts Dartmouth. URL: https://compmath.files.wordpress.com/2010/05/ ahopkins_freports10.pdf.
4. Freeman L.C. Centrality in social networks conceptual clarification, Social networks. - 1979. – 1(3). - P. 215-239.
5. Page L., Brin S., Motwani R., Winograd T. The pagerank citation ranking: bringing order to the web // January 29, 1998. URL: http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf.
6. Kobourov, Stephen G. Spring Embedders and Force-Directed Graph Drawing Algorithms. 2012. URL: http://arxiv.org/abs/1201.3011.
7. Jacomy M., Heymann S., Venturini T., Bastian M. ForceAtlas2, a graph layout algorithm for handy network visualization, Paris. August 29, 2011. URL: http://webatlas.fr/ tempshare/ForceAtlas2_Paper.pdf
8. Bastian M., Heymann S., Jacomy M., Gephi: an open source software for exploring and manipulating networks, International AAAI Conference on Weblogs and Social Media. Association for the Advancement of Artificial Intelligence, 2009.
9. Choudhary P., Singh U. A survey on social network analysis for counter-terrorism, Int. Journal of Computer Applications. - Feb. 2015. – Vol. 112. – No. 9. – P. 24-29.
10. Gubanov D.A., Novikov D.A., Chkhartishvili A.G. Sotsial'nye seti: modeli informatsionnogo vliyaniya, upravleniya i protivoborstva. M.: Izdatel'stvo fiziko-matematicheskoy literatury, 2010. 228 P.
11. Krebs V. Mapping networks of terrorist cells, Connections. – 2002. – 24(3). – P.43-52.
12. Ressler S. Social Network Analysis as an Approach to Combat Terrorism: Past, Present, and Future Research, Homeland Security Affairs. – July 2006. – 2. Article 8. URL: https://www.hsaj.org/articles/171.
13. Rothenberg R. From whole cloth: Making up the terrorist network, Connections. – 2002. - 24(3). – P.36-42.

5 Mikhail Basarab, Dr.Sc., Professor, Bauman Moscow State Technical University, Moscow, bmic@mail.ru
6 Igor Ivanov, Dr.Sc., Bauman Moscow State Technical University, Moscow, ivanov@bmstu.ru
7 Aleksandr Kolesnikov, Ph.D., Bauman Moscow State Technical University, Moscow, vkolesnikov90@list.ru
8 Valeriy Matveev, Dr.Sc., Professor, Bauman Moscow State Technical University, Moscow, v.a.matveev@bmstu.ru

14. Motoyama M., McCoy D., Levchenko K., Savage S., Voelker G.M. An analysis of underground forums, Proc. of the ACM Internet Measurement Conference. Berlin, CA, Nov. 2011.
15. McCoy D., Dharmdasani H., Kreibich C., Voelker G.M., Savage S. Priceless: the role of payments in abuse advertised goods, Proc. of the ACM Conference on Computer and Communications Security. Raleigh, NC, Oct. 2012.
16. Xu J., Chen H. The topology of dark networks, Communications of the ACM. – 2008. -51(10). – P.58-65. New York, NY. URL: http://ai.arizona.edu/intranet/papers/Xu-SNA- 2008.pdf.
17. Wu E., Carleton R., Davies G. Discovering bin-Laden's Replacement in al-Qaeda, using Social Network Analysis: A Methodological Investigation, Perspectives on Terrorism. – 2014. - Vol. 8. - No.1.
18. Klausen J., Tschaen Barbieri E., Reichlin-Melnick A., Zelin A.Y. The YouTube Jihadists: A Social Network Analysis of Al-Muhajiroun's Propaganda Campaign, Perspectives on Terrorism. – 2012. - Vol. 6. - No.1.
19. Sarvari H., Abozinadah E., Mbaziira A., McCoy D. Constructing and analysing criminal networks, IEEE Security and Privacy Workshops, 2014, pp. 84-91. DOI 10.1109/SPW.2014.22.
20. Carley K.M., Reminga J., Kamneva N. Destabilizing terrorist networks // NAACSOS conference proceedings. - 2003. - Pittsburgh, PA.
21. Chesnokov V.O., Klyucharev P.G. Vydelenie soobshchestv v sotsial'nykh grafakh po mnozhestvu priznakov s chastichnoy informatsiey, Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana. 2015. № 9. S. 188-199. DOI: 10.7463/0915.0811704.
22. Klyucharev P.G., Chesnokov V.O. Issledovanie spektral'nykh svoystv sotsial'nogo grafa seti LiveJournal, Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana. 2013. No 9, pp. 391-400. DOI: 10.7463/0913.0603441.
23. Basarab M.A., Vel'ts S.V. Ierarkhicheskoe predstavlenie komp'yuternoy seti na osnove neyronnoy seti Khopfilda, Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana. 2013. No 9, pp. 335-348. DOI: 10.7463/0913.0630141.
24. Basarab M.A., Vel'ts S.V. Teoretiko-igrovoy podkhod k otsenke riskov i nakhozhdeniyu uyazvimostey v setyakh peredachi informatsii, Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana. 2013. No 8, pp. 271-280. DOI: 10.7463/0813.0630132.