DOI 10.36622/VSTU.2022.25.1.003
УДК 004.056.57

СПОСОБЫ ПОСТРОЕНИЯ И АНАЛИЗА ИНФОРМАЦИОННЫХ КАРТ ЗАЩИЩАЕМОГО КИБЕРПРОСТРАНСТВА

А.Л. Сердечный

В статье рассматриваются типовые комбинации способов построения и анализа информационной карты. В частности, проиллюстрирована комбинация таких способов на примере библиометрических сетей. Рассмотрены особенности реализации способов построения и анализа информационных карт, определяемых их масштабом (крупным, средним и мелким), одним из которых является эффект совмещения несвязанных кластеров, обусловленный ошибками понижения размерности. Также показана специфика реализации способов, связанная с использованием модели данных на примере графа связей. Информационные карты по типу сущностей модели разделяются на объектные и субъектные. В свою очередь типы связей между сущностями определяют возможности отображения структурных и функциональных взаимодействий между субъектами и объектами. Вышеизложенное проиллюстрировано на примере графов терминов и связей публикаций и ключевых слов для учебной дисциплины «Компьютерные преступления». Рассмотренные комбинации способов построения и анализа информационных карт являются типовыми для проведения картографического анализа большинства сетевых структур и могут быть использованы в ходе решения широкого спектра задач в области обеспечения информационной безопасности и защиты информации.

Ключевые слова: способы, информационная карта, научная школа, граф, модель данных.

Одной из важнейших задач при реализации картографического подхода исследования защищаемого пространства [1, 2] является выбор способов построения и анализа информационных карт. Настоящая работа посвящена обобщению опыта, полученного в ходе решения практических задач с использованием информационных карт [3-5]. Способы детализируют методы построения и анализа информационных карт с учётом конкретных особенностей сетевой структуры, лежащей в основе информационной карты. В общем случае таки особенности определяются:

- масштабом информационной карты;
- моделью данных.

Способы построения и анализа информационной карты, определяемые её масштабом

Масштабируемые информационные карты характеризуются возможностью детализации отображаемых с их помощью информационных объектов. В зависимости от масштаба способы построения и анализа информационных карт могут варьироваться. В работе [3] были описаны результаты разработки информационной карты научных школ, сложившихся в области защиты информации и обеспечения информационной безопасности. На данном примере были наглядно показаны особенности масштабирования информационных объектов, представляющих:

- научное сообщество;
- агломерации научных коллективов;
- группы научных коллективов;
- конкретные научные коллективы;
- отдельных авторов.

В качестве модели данных рассматривался граф соавторов научных публикаций в области защиты информации и обеспечения информационной безопасности. Узлом являлся идентификатор автора научной публикации, а связью – общая научная публикация.

Операции построения информационных карт, состав которых показан на рис. 1, были подробно описаны в работе [3]. В результате их выполнения сформирована генеральная карта, на которой можно видеть крупные структуры и их отношения (рис. 2).

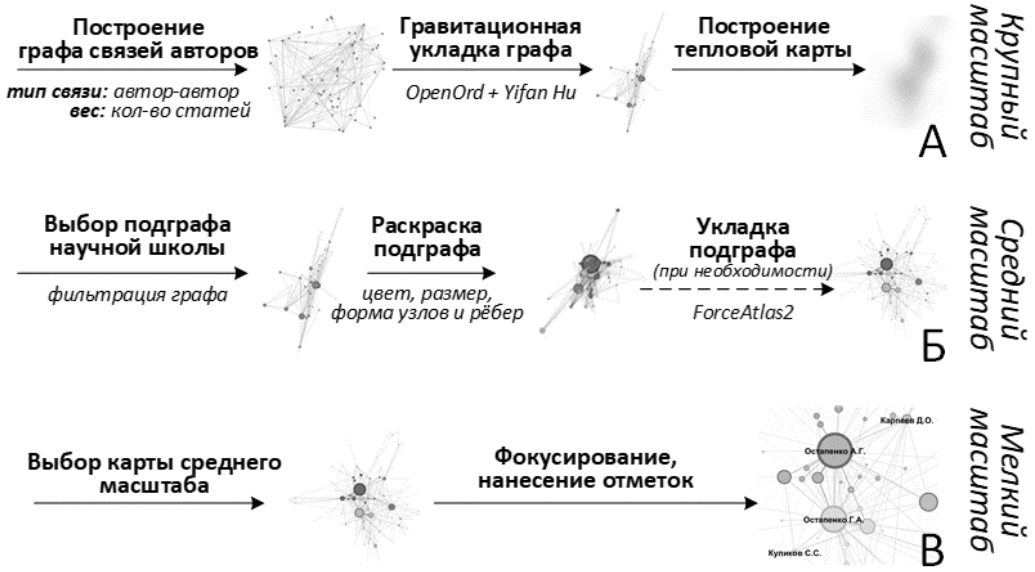

Рис. 1. Состав операций построения информационных карт научных школ, сложившихся в области защиты информации и обеспечения информационной безопасности (А– генеральная карта, крупный масштаб, Б – карта среднего масштаба, где на экране отображается группа научных коллективов, В – карты мелкого масштаба, где отображается конкретный научный коллектив) [3]

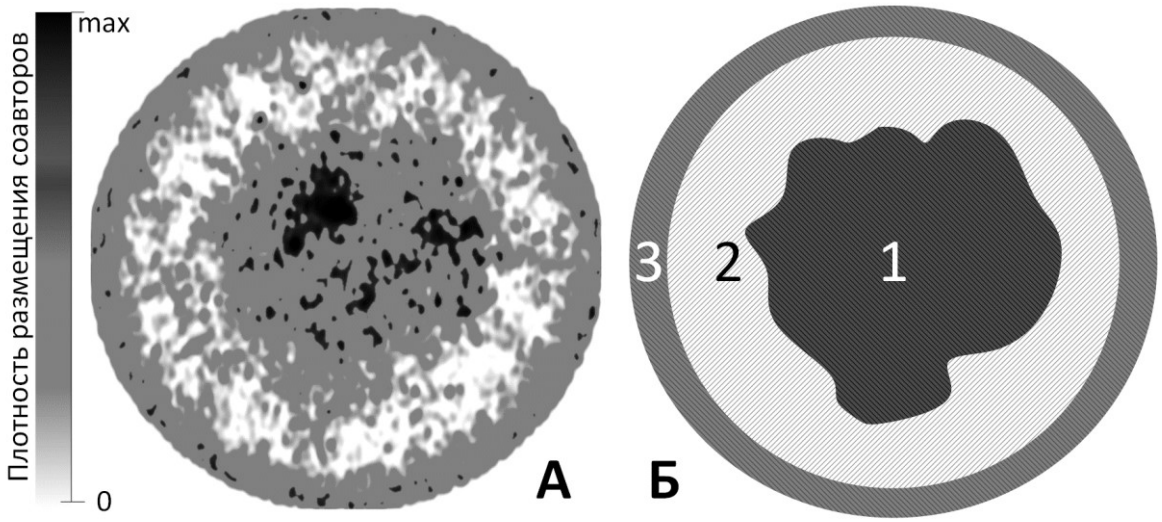

Рис. 2. Информационная карта научных школ (крупный масштаб) и схематическое изображение её основных частей

На рис. 2-Б можно наблюдать три обособленные области: ядро библиометрической сети (1), периферия (2) и изолированные узлы (3). Генеральную карту удобно использовать для демонстрации общих тенденций, а также положения на ней конкретных объектов и их связь с соседними, аналогично тому, как демонстрируются конкретные научные коллективы.

Структуру научных сообществ лучше всего исследовать на информационной карте среднего масштаба (рис. 1-Б).

В ходе её анализа необходимо обращать внимание на форму и расположение плотно расположенных. На рис. 3 изображен кластер, форма которого была определена в рамках укладки полного графа соавторов. Видно центральную группу узлов (ядро кластера) и отдельные группы авторов. Авторы ядра относятся к одной научной школе так как часто публикуют совместные работы. Наиболее крупные узлы являются руководителями направлений. Авторы из соседних кластеров также могут входить в соответствующую научную школу, но имеют некоторое количество работ с участием авторов из других коллективов.

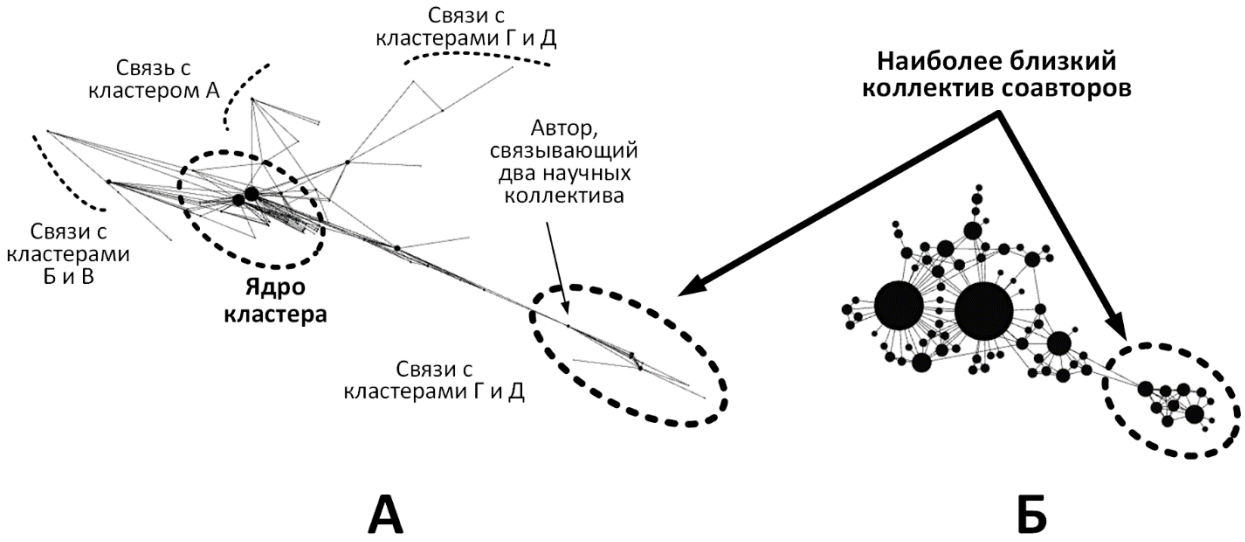

Рис. 3. Кластер научных коллективов, представленный в различных формах (А – форма кластера определяется глобальной укладкой графа, Б – форма кластера определяется его собственной укладкой) [3]

Средний масштаб позволяет типизировать формы различных авторских коллективов. Так, например, для рассмотренного случая наиболее распространёнными являются сообщества со структурами, показанными на рис. 4.

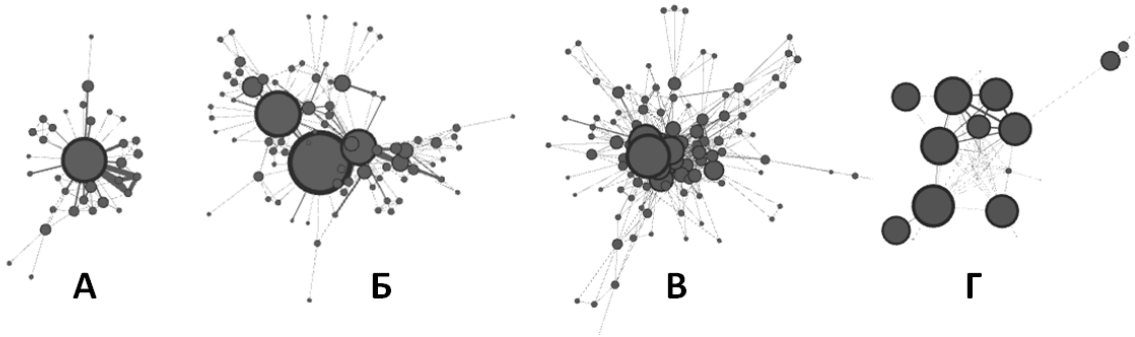

Рис. 4. Типовые формы научных коллективов, выявленные в ходе анализа графа соавторов [3]

Более подробно о результатах типизации форм научных коллективов было сказано в [3].

Карты мелкого масштаба позволяют проводить исследование конкретной научной школы (кластера). Проведение такого анализа представляется нетривиальной задачей, так кластеров может быть достаточно большое количество (в проведённых исследования [3] их число составило более 250 в состав которых входило не менее 7800 учёных). В ходе соответствующих исследований должны проводиться сбор и соотнесение дополнительных сведений о каждом из участников кластера. Например, может потребоваться получение сведений о городе проживания, статусе учёного или его компетенциях.

Применение картографического метода позволяет существенно упростить процесс исследования отдельных кластеров по следующим причинам. В первую очередь, информационные карты позволяют упростить получение дополнительных признаков. В результате визуального анализа можно убедиться, что какой-либо признак локализован в определённых областях карты. Это позволяет распространить значение этого признака на всех членов соответствующих областей, даже если сведения о значении данного признака отсутствуют для некоторых узлов.

Другой возможностью является задействование аналитических возможностей эксперта, связанных с обработкой визуальной информации. Визуальный анализ снижает когнитивную нагрузку на эксперта и

повышает его концентрацию. Например, для того, чтобы выявить центры научных школ достаточно отобразить размер узла с использованием значения метрики центральности. В некоторых задачах значение метрики PageRank для узлов графа соавторов может заменить сведения об учёной степени и научном звании каждого участника.

Также необходимо отметить, что анализ карт мелкого масштаба позволяет сосредоточить внимание исследователя на связях внутри творческих коллективов, поэтому для снятия внешнего влияния требуется выполнение гравитационной укладки соответствующего подграфа.

Таким образом, в зависимости от масштаба методы построения и анализа карт могут существенно отличаться. Комплексный интерактивный анализ масштабируемой информационной карты позволяет существенно сократить время исследований, а в некоторых случаях, получить результаты, которые невозможно добиться традиционными методами.

Способы построения и анализа информационной карты, определяемые спецификой модели данных

В зависимости от особенностей модели исходных данных могут применяться различные комбинации рассмотренных выше методов построения и анализа информационной карты. Специфика модели исходных данных определяется типом сущностей (которые она описывает) и связей между ними.

Информационные карты по типу сущностей модели разделяются на:

а) *объектные карты*:

- *тематические карты* (в основе ландшафта – семантические связи информационного контента и источников информации, а в качестве объектов исследований выступают субъекты и объекты защищаемого киберпространства);

- *ресурсные карты* (в основе ландшафта – объекты защищаемого киберпространства, связанные с информационными и вычислительными ресурсами, объектами исследований выступают процессы информационного противоборства субъектов защищаемого киберпространства);

б) *субъектные карты* (в основе ландшафта – субъекты защищаемого киберпространства, а объектами исследования являются свойства таких субъектов и их объединений).

Типы связей между сущностями определяют следующие информационные карты:

- *структурные карты* (в основе ландшафта – субъекты и (или) объекты защищаемого киберпространства, связанные иерархическими отношениями);

- *функциональные карты* (в основе ландшафта – субъекты и (или) объекты защищаемого киберпространства, которые связаны на основе выполняемых ими функций).

Комбинации рассмотренных сущностей и связей позволяют объединять различные типы информационных карт. Однако слишком большое число разнородных объектов и связей, лежащих в основе ландшафта информационной карты, может привести к потере «физического смысла», который вкладывается в такую карту. Так, например, если для построения ландшафта используются сразу все сущности, включающие «Уязвимость программы», «Источник сведений об уязвимости», «Средство эксплуатации уязвимости», «Уязвимое программное обеспечение», «Объекты, на которых установлено уязвимое программное обеспечение», то с одной стороны, такая карта позволяет провести всесторонний анализ (включающий анализ возможных последствий от эксплуатации уязвимости), а с другой – нагромождение различных объектов и связей приводит к смешиванию влияния различных факторов. В результате полученные ландшафт не будет иметь ярко выраженных особенностей, однозначно характеризующих все объекты, попавшие в определённую область. Это снижает скорость визуального анализа такой карты. Кроме того, становится невозможным использование количественных оценок, полученных в результате измерения расстояний между объектами информационной карты, так как в основе таких оценок будет лежать мера близости,

характеризующая объединение всех факторов без учёта их доли влияния.

Все три типа карт также могут использоваться для анализа полноты, актуальности и достоверности источников исходных данных. Чем больше различных типов карт при этом будет использовано, тем выше точность результатов такого анализа, так как источники будут рассмотрены в разных контекстах.

Для примера рассмотрим особенности построения и анализа тематических информационных карт. Под тематической информационной картой понимается карта, в которой каждая область отражает определённую тематическую направленность содержащихся в ней объектов. К таким картам относятся карты поиска научных публикаций, карты блогов, карты информационных ресурсов, карты исследований организаций и др.

Здесь в качестве модели данных рассматриваются как минимум два типа объектов:

- идентификатор темы (ключевое слово, термин, класс предметной области и др.);
- информационный объект, обладающий содержательным смыслом, который может быть соотнесён с идентификатором темы.

Если идентификаторов тем слишком много (например, в качестве них использованы все термины и именованные сущности, встречающиеся в тексте), то граф связей, на основе которого строится ландшафт информационной карты, будет представлять компоненту с большим значением средней степени, в котором будут присутствовать несколько общих терминов, связанных практически со всеми объектами, а также – большое количество узкоспециализированных сущностей, связанных только с одним информационным объектом (рис. 5).

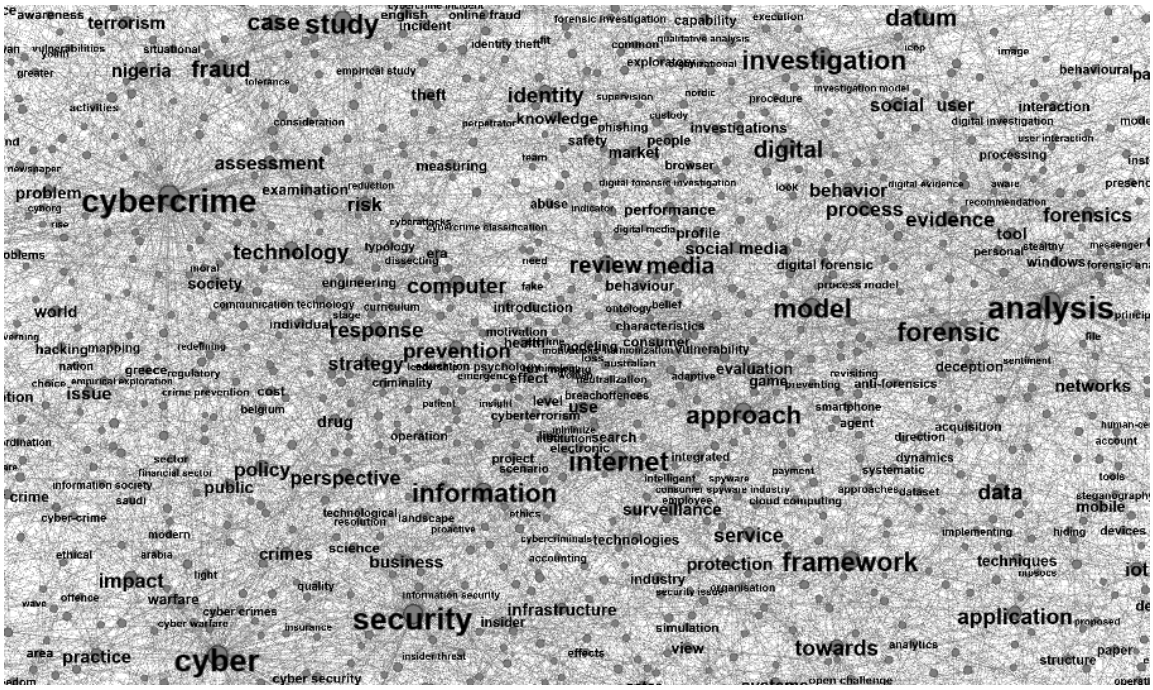

Рис. 5. Фрагмент графа терминов, встречающихся в заголовках англоязычных научных публикаций по теме «Компьютерные преступления»

Ландшафт такой информационной карты не будет обладать стабильностью, так как иметь место большое состояние локальных минимумов ввиду высокой размерности исходного пространства признаков. Иногда использование такого представления оказывается полезным, например, оно может быть использовано вместо «облака тэгов». При этом, если убрать отображение связей, то информационная карта (рис. 6) будет обладать преимуществами по сравнению с «облаком тэгов». Информационная карта, помимо частотности терминов, отражаемой в «облаке тэгов» с помощью размера

соответствующих слов, будет показывать семантику за счёт нахождения слова в определённой точке ландшафта карты.

Для снижения связанности графа могут применяться различные механизмы фильтрации исходных данных, например, ограничение на минимальную частоту встречаемости термина, который может быть включён в состав графа.

Другим подходом построения информационной карты является использование авторских меток в качестве идентификаторов тем. С одной стороны – состав таких меток обычно не превышает 6, что обеспечивает умеренную связанность элементов исходного графа (рис. 6). С другой стороны – каждый автор использует уникальные термины, исходя из собственного понимания их значения, поэтому для формирования общей информационной карты публикаций, размеченных разными авторами, может потребоваться приведение меток к единым терминам.

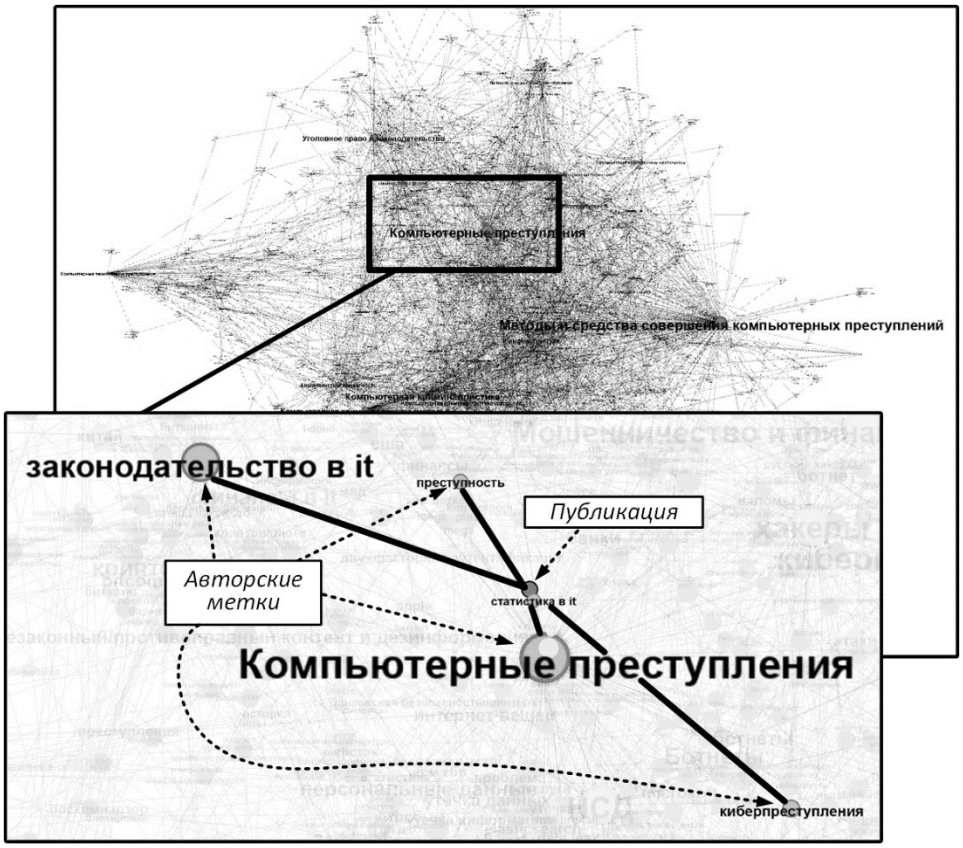

Рис. 6. Фрагмент графа связей публикаций и ключевых слов (меток или тэгов), указанных автором

В результате указанных приёмов может быть сформирован ландшафт, который может успешно использоваться для решения различных задач, включающих тематическое моделирование предметной области, сравнение информационных источников, исследование авторской активности, мониторинг публикаций и др. На рис. 7-А показан пример такого ландшафта (построенного на базе рассмотренного графа). Имея возможность фильтрации по публикациям конкретных авторов, при помощи данной карты можно за доли секунды определить тематическую направленность. Так, например, на рис. 7-Б можно видеть, что автор в основном размещал статьи, связанные с защитой от несанкционированного доступа (именно эта область подсвечена на рис. 7).

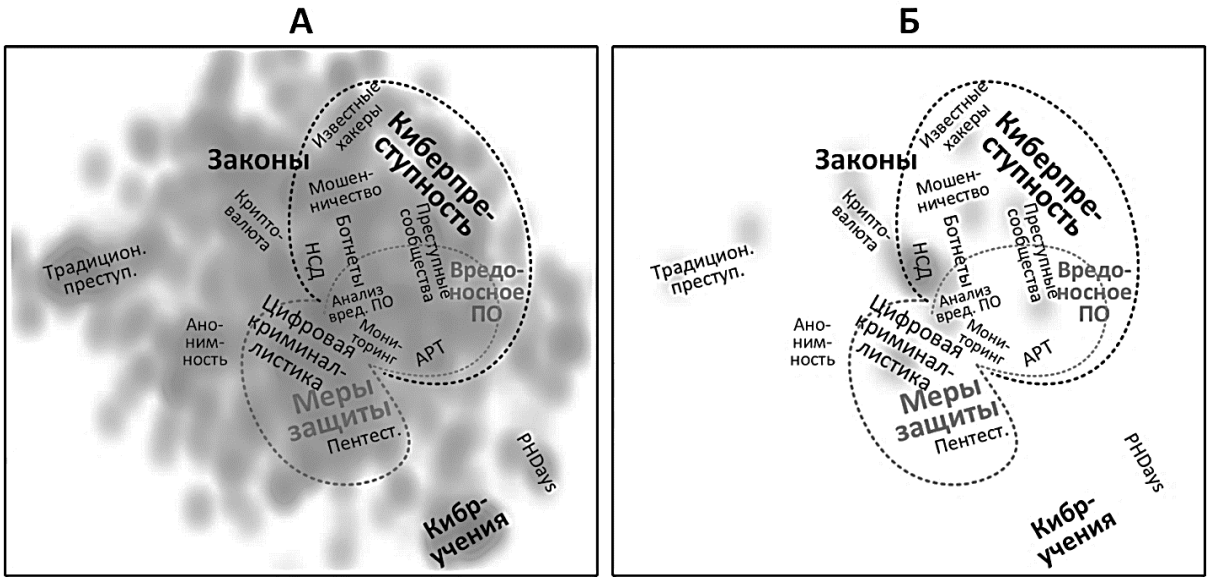

Рис. 7. Тематическая карта публикаций информационного ресурса Хабр [4] по теме «Компьютерные преступления» (А – карта всех публикаций, Б – карта публикаций, размещённых конкретным автором)

Для одних и тех же тематических областей могут быть построены различные проекции, которые раскрывают свои аспекты темы. Для задач обеспечения защиты киберпространства требуется разработка как общих тематических карт, так и тематических карт, детализирующих определённые области. Это позволит сформировать систему вложенных слоёв семантического (когнитивного) уровня киберпространства, что является одним из важных направлений разработки генеральной карты защищаемого киберпространства. Для этого может быть использована информационная карта поиска научных публикаций, так как научные исследования предлагают согласованный и чётко определённый понятийный аппарат по всем областям защищаемого киберпространства.

Информационные карты позволяют координировать взаимодействие различных исследователей, так как, по сути, являются универсальным языком междисциплинарного общения между различными людьми. За счёт этого свойства можно организовывать совместную работу по построению информационных карт. В первую очередь, это нанесение на карту объектов и субъектов защищаемого киберпространства, указание их характеристик и связей с другими субъектами.

Таким образом, рассмотренные комбинации способов построения и анализа информационных карт являются типовыми для проведения картографического анализа большинства сетевых структур и могут быть использованы в ходе решения широкого спектра задач в области обеспечения информационной безопасности и защиты информации.

Список литературы

1. Сердечный А.Л. Концептуальные основы картографии защищаемого киберпространства. Часть 1. / Сердечный А.Л. // Информация и безопасность. 2021. Т. 24. Вып. 3. С. 373-386.

2. Сердечный А.Л. Концептуальные основы картографии защищаемого киберпространства. Часть 2. / Сердечный А.Л. // Информация и безопасность. 2021. Т. 24. Вып. 3. С. 387-400.

3. Калашников А.О. Картографический подход в библиометрическом исследовании отечественных научных школ, сложившихся в области защиты информации и обеспечения информационной безопасности. / А.О. Калашников, А.Л. Сердечный, А.Г. Остапенко. // Информация и безопасность. 2019. Т. 22. Вып. 4. С. 455-484.

4. Сердечный А.Л. Технология построения и использования поисковых карт в образовательном процессе на примере поисковой карты по учебной дисциплине «Компьютерные преступления» / А.Л. Сердечный, А.А. Гончаров, А.Г. Остапенко // Интеллектуальные информационные

системы. Тр. Междунар. науч.-практ. конф. : в 2 ч.. Воронеж, 2021. С. 94-98.

5. Сердечный А.Л. Mapping retrieval method for academic publications in the field of aerospace technology safety / А.Л. Сердечный, А.А. Гончаров, А.Г. Остапенко, И.Л. Батаронов // IOP Conference Series: Materials Science And Engineering. Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations. 2020. С. 52028.

Воронежский государственный технический университет
Voronezh State Technical University

Поступила в редакцию 12.09.2021

Информация об авторе

Сердечный Алексей Леонидович – канд. техн. наук, старший преподаватель, Воронежский государственный технический университет, e-mail: alex-voronezh@mail.ru

TECHNIQUES FOR CONSTRUCTING AND ANALYSIS OF PROTECTED CYBERSPACE INFORMATION MAPS

A.L. Serdechnyy

The article discusses typical combinations of techniques for building and analyzing an information card. In particular, a combination of such techniques is illustrated by the example of bibliometric networks. The features of the technique implementation for constructing and analyzing information maps determined by its scale (large, medium and small). One of the feature is an effect of combining unbound clusters, due to errors of a dimension lowering. Also shows the specifics of the implementation of techniques associated with using the data model, on the example of the bond graph. Information maps by type of entity of the model are divided into object and subject maps. In turn, the types of ties between entities determine the possibilities of displaying structural and functional interactions between the subjects and objects. The above is illustrated by the example of the graphs of terms and links of publications and keywords for the educational discipline "Computer crimes". The considered combinations of methods for building and analyzing information cards are typical for making a cartographic analysis of most network structures and can be used during solving a wide range of tasks in the field of information security and information protection

Keywords: techniques, information map, scientific school, graph, data model.

Submitted 09.09.2021

Information about the authors

Alexey L. Serdechnyy – Cand. Sc. (Technical), Senior Lecturer, Voronezh State Technical University, e-mail: alex-voronezh@mail.ru