Программа VOSviewer для кластеризации и визуализации информации по киберугрозам

Р.И. Горохова, Д.И. Коровин

Аннотация—В исследовании рассматриваются ключевые аспекты обработки и анализа описания информационных угроз, включая кибератаки и утечки данных. Оцениваются возможности VOSviewer в контексте кластеризации И визуализации данных информационных угрозах, что помогает в выявлении взаимосвязей и закономерностей в данных. В тексте также обсуждаются преимущества данного инструмента по сравнению с другими программными продуктами на рынке, а именно его простота использования и поддержка разных форматов данных. Цель исследования заключается в оценке важности методов кластеризации и визуализации для понимания динамики киберугроз, с акцентом на VOSviewer. Метолологический возможности исследования включает использование кластеризации, что позволяет выделять группы схожих угроз и уязвимостей, тем самым способствуя более эффективному анализу и реагированию на потенциальные Рассмотрены библиометрического и наукометрического анализа через призму использования карт и кластеризации в VOSviewer, позволяет одновременно визуализировать структуру исследовательской сети и выделять кластеры. Описаны методы интерактивной настройки визуализации, включая выбор различных методов кластеризации, что способствует более глубокому пониманию групп, связанных с киберугрозами. В процессе анализа выделены четыре основных кластера: типы кибератак, уязвимости и риски, методы защиты и профилактика, а также актуальные тренды и технологии. Каждый кластер включает ключевые слова, отражающие специфические темы, что способствует формированию всестороннего представления о киберугрозах и подходах к их защите. Процесс построения карты ключевых слов включает в себя этапы сбора, импорта и анализа данных, а также кластеризации и визуализации результатов. Дифференциация кластеров по цветам и размеру ключевых слов позволяет лучше интерпретировать связи между терминами. В результате, использование VOSviewer предоставляет исследователю удобные инструменты для анализа структуры и динамики научных тем в области кибербезопасности.

Ключевые слова— киберугрозы, VOSviewer, кластеризация, визуализация данных, библиометрический анализ, информационная безопасность.

I. Введение

В условиях быстро развивающихся технологий и увеличения объемов данных, обработка и анализ информационных угроз становятся важнейшими задачами для обеспечения кибербезопасности [1, 2, 3, 4].

Кибератаки, фишинг, утечки информации — все это представляет серьезные риски для организаций и индивидуальных пользователей [5, 6, 7]. В этом контексте технологии кластеризации и визуализации данных играют ключевую роль в быстром выявлении и анализе таких угроз [8]. VOSviewer был создан в 2009 году исследовательской группой из Университета Твенте в Нидерландах, в рамках проекта, направленного на визуализацию и анализ научных данных. Изначально разрабатывалась для обработки библиометрических данных, таких как ссылки и цитирования в научных публикациях. С течением времени VOSviewer эволюционировала и стала популярной среди исследователей и аналитиков в самых разных областях, включая социальные науки, экономику, медицину и др. [9, 10, 11]. Программа постоянно обновляется и улучшается, добавляя новые функции и алгоритмы, что делает ее мощным инструментом для визуализации данных и анализа взаимосвязей.

Программа VOSviewer, изначально разработанная для визуализации научных данных, демонстрирует широкий потенциал в области анализа текстовой информации и может быть адаптирована для решения задач в сфере информационной безопасности и для обнаружения киберугроз [12].

Для анализа и визуализации данных используются Tableau, Gephi и VOSviewer. VOSviewer выделяется возможностями для обработки больших объемов данных и простотой в использовании. Использование VOSviewer для задач кластеризации и визуализации позволяет выделять угрозы и их взаимосвязи.

В большинстве статей, опубликованных по этой тематике, программа VOSviewer не использовалась для кластеризации и визуализации информации по киберугрозам.

Остальная часть статьи организована следующим образом: Раздел 2 описывает методологию. Раздел 3 представляет эксперименты и результаты. Наконец, Раздел 4 суммирует выводы.

II. МЕТОДОЛОГИЯ

В этом разделе подробно объясняется предлагаемый метод, который состоит в оценке возможностей программы VOSviewer для кластеризации и визуализации информации по киберугрозам. Она представляет собой инструмент, который используется для визуализации и анализа научных публикаций и

Д.И. Коровин – Финансовый университет при правительстве Российской Федерации (e-mail: dikorovin@fa.ru)

цитат, но его функциональность может быть эффективно применена и для кластеризации и визуализации информационных угроз.

А. Функции и возможности

VOSviewer программа, которая имеет набор функций, позволяющих выполнять анализ и визуализацию данных. Программа может выполнять целый рад задач на основе использования алгоритмов кластеризации взаимного сходства, что позволяет автоматически группировать объекты, основываясь на их характеристиках или взаимных связях [13, 14]. Это особенно полезно при исследовании информационных угроз, где можно выделить группы похожих угроз или уязвимостей. VOSviewer поддерживает множество форматов данных, что обеспечивает гибкость в работе с различными наборами информации. В основном используются библиометрические данные, базы данных, пользовательские наборы данных, к которым относятся полученные в результате парсинга новостных лент. обрабатывать Программа может ланные, смоделированные в CSV, ТХТ и других стандартных форматах, которые содержат информацию публикациях, ссылках и цитированиях [13]. VOSviewer может импортировать данные из популярных научных баз данных, таких как Web of Science и Scopus, что позволяет легко интегрировать уже доступные наборы данных [15]. Пользователи могут загружать собственные данные в различных форматах, что делает VOSviewer универсальным инструментом для анализа различных типов информации.

В. Предлагаемый метод

В программе VOSviewer используются несколько математических и статистических концепций для кластеризации данных [16, 17]. В кластеризации в VOSviewer используются алгоритмы кластеризации k-средних (k-means) и иерархическая кластеризация.

Алгоритм k-средних (k-means) представляет собой очень хорошо разработанный и самый распространенный алгоритм, который основывается на разбиении данных на k кластеров. Начиная с выбора k случайных центров кластеров, каждую точку данных назначают к ближайшему центру кластера и выполняют пересчет центров кластеров на основе средних значений точек, принадлежащих каждому кластеру. Итерируемый процесс продолжается, пока центры кластеров не стабилизируются (точки не будут менять свои кластеры).

Иерархическая кластеризация основана на построении иерархии кластеров, начиная с каждой точки как отдельного кластера и постепенно объединяя их. Сначала каждую точку рассматривают как отдельный кластер, затем на каждом шаге объединяются два ближайших кластера и процесс продолжается до тех пор, пока все точки не останутся в одном кластере или не достигнут предопределенного числа кластеров.

Для определения близости объектов в пространстве используются различные метрики расстояний, что критично для кластеризации. Обычно применяются следующие метрики:

- Евклидово расстояние,
- Манхэттенское расстояние сумма абсолютных

разностей координат,

- косинусное сходство - используется для измерения угол между векторами в многомерном пространстве, что особенно полезно для текстовых данных [18].

VOSviewer часто использует представление данных в виде графов, представленных матрицей смежности или графом. Матрица смежности описывает связи между объектами, где ячейка (A_ij) указывает на наличие/отсутствие или вес связи между объектами і и ј. Элементы матрицы могут быть заполнены на основе сходства, веса и других факторов. Граф визуализирует объекты как узлы, а связи между ними как ребра, что позволяет анализировать структуру данных и их взаимосвязи [19].

Еще один подход к кластеризации, который применяется в контексте обработки текстовых данных - метод локальных плотностей LDA (Latent Dirichlet Allocation). LDA воспроизводит темы в наборе документов и распределяет вероятности тем для каждого документа. Реализация такой модели может быть использована для выявления групп сообщений или угроз, связанных по содержанию.

Результаты работы VOSviewer контексте кластеризации информации киберугрозам представлены в виде наглядных визуализаций [20]. В первую очередь это карта, показывающая как различные типы угроз связаны друг с другом, узлы карты представляют типы угроз, а ребра показывают степени их взаимосвязи. Особое место занимает временной анализ, который позволяет выявить, как менялась структура угроз со временем, в частности показать, как в информационном поле новые типы атак становятся более распространёнными, И какие уязвимости сопровождают.

Традиционные методы анализа, такие как топологический анализ, могут быть более сложными и требующими значительных вычислительных ресурсов, VOSviewer предлагает быстрый и интуитивно понятный способ визуализации данных. Это может сделать его более доступным для исследователей, не обладающих высоким уровнем математической подготовки.

В отличие от более сложных алгоритмов машинного обучения, таких как кластеризация на основе метрики расстояния (например, алгоритмы DBSCAN или Gaussian Mixture Models), VOSviewer ориентирован на легкость анализа и визуализации [21]. Хотя методы машинного обучения могут выдавать более точные результаты, процесс интерпретации их выводов может быть менее интуитивным по сравнению с наглядными визуализациями VOSviewer.

Сравнение со статистическими методами показывает, что некоторые статистические методы могут предложить формальный подход к анализу данных о угрозах, в то время как VOSviewer позволяет представлять результаты в виде интерактивных карт и графов, что существенно упрощает интерпретацию и визуализацию сложной информации.

III. ЭКСПЕРИМЕНТЫ И РЕЗУЛЬТАТЫ

Эффективность и осуществимость предлагаемого

метода были проверены путем проведения нескольких экспериментов. Эксперименты проводились на основе базы данных РБК.ру. Создание карты ключевых слов основано на анализе 48 256 публикаций. В загруженном наборе данных было выявлено 98 000 ключевых слов. Если установить минимальное количество совместно встречающихся ключевых слов на уровне 11, то лишь

3168 из них соответствуют этому критерию и могут быть использованы для создания карты. Для этих 3168 выбранных ключевых слов рассчитывается общая сила связи, и в итоговую визуализацию включаются только 800 ключевых слов (это количество задано по умолчанию, но исследователь может изменить его по своему усмотрению).

Таблица 1: Время обработки

		Таблица 1: Время обработки	
Шаг	Действие 1	Действие 2	Действие 3
1. Сбор данных	Определение источника данных: - выберите подходящую базу данных (РБК.ру), - убедитесь, что база данных содержит статьи по теме «киберугрозы».	Формулирование запросов: - сформулируйте поисковые запросы, используя ключевые слова, такие как «киберугрозы», «кибербезопасность», «вредоносное ПО», «фишинг», и другие релевантные термины уточните фильтры поиска по датам, типам публикаций и другим параметрам в зависимости от ваших потребностей.	Экспорт данных: - экспортируйте результаты поиска в формате, совместимом с VOSviewer (например, в формате CSV, TXT или специфическом формате для вашей базы данных).
2. Импорт данных в VOSviewer	Открытие VOSviewer: - запустите VOSviewer на вашем компьютере.	Выбор типа данных: - на главном экране выберите опцию для создания карты на основе данных, например, «Создание карты на основе данных» (Create a map based on bibliographic data), - загрузите ранее экспортированный файл.	Настройка параметров импорта: - VOSviewer может предложить настройки импорта. Убедитесь, что все необходимые поля (например, заголовки, аннотации, ключевые слова) правильно распознаны программой.
3. Построение карты ключевых слов	Анализ данных: - после импорта данных VOSviewer проанализирует их, извлекая ключевые слова, цитирования и другие метрики, - определите, как вы хотите сформировать карту. Обычно можно выбрать опцию создания карты на основе ключевых слов.	Кластеризация ключевых слов: - VOSviewer автоматически произведет кластеризацию ключевых слов на основе их частоты появления и взаимосвязей. Это может включать использование алгоритмов, таких как К-means или другие методы, как, например, метод ближайших соседей.	Настройка визуализации: - настройка визуализации, изменяя размеры узлов (соответствующих частоте ключевых слов) и цвета. Разные кластеры могут быть выделены различными цветами, что облегчает интерпретацию, - можно также включить/исключить определенные ключевые слова или кластеры по желанию.
4. Визуализа- ция и анализ карты	Отображение карты: - на экране отобразится интерактивная карта ключевых слов. Узлы представляют ключевые слова, а связи между узлами показывают, как они соотносятся друг с другом в контексте собранных данных.	Интерактивное взаимодействие: - пользователи могут взаимодействовать с картой, приближать и удалять, щелкать на узлы для получения дополнительной информации, такой как ссылки на публикации или другие связанные ключевые слова.	Анализ результатов: - анализируйте полученную карту, выявляя ключевые тренды, основные направления исследований и ключевые темы относительно киберугроз, - выявите центральные термины и заголовки, которые наиболее активно фигурируют в научной литературе, и оцените их взаимосвязи.
5. Экспорт и отчет	Сохранение карты: - VOSviewer позволяет экспортировать визуализацию карты в различных форматах (например, графические форматы изображений или PDF), что упрощает	Подготовка отчета: - создается отчет на основе анализа, включая полученные карты и ключевые insights о развитии тематики киберугроз.	Завершение алгоритма

дальнейшую работу с	
данными.	

Построение карты ключевых слов по теме «киберугрозы» в VOSviewer может быть выполнено через несколько этапов. Этот процесс включает сбор данных, их обработку и визуализацию в виде карты [23]. В таблице 1 представлен пошаговый алгоритм построения карты ключевых слов по теме «киберугрозы»

В библиометрическом и наукометрическом анализе использование карт и кластеризации тесно связано и служит одной общей цели — предоставлению информации о структуре исследовательской сети. В VOSviewer реализована возможность одновременного создания карт и кластеризации, что позволяет одновременно визуализировать общую структуру сети и выделять кластеры внутри этой сети.

В VOSviewer пользователи могут интерактивно настраивать параметры визуализации, включая выбор различных методов кластеризации, что помогает им находить наиболее подходящие способы интерпретации выделенных групп, связанных с киберугрозами. В ходе работы, независимо от используемого метода подсчета, при создании карты выделяются четыре четко обозначенных кластера [22].

Разные кластеры отображаются в различных цветах, что помогает пользователям различать их. Важно отметить, что цветовые решения для полного и фракционного методов подсчета могут немного различаться, и некоторые кластеры могут иметь разные цветовые обозначения в зависимости от выбранного метода.

При визуализации сети в VOSviewer размер ключевых слов (объектов на карте) демонстрирует уровень их взаимосвязей, а толщина линий и расстояние между двумя ключевыми словами указывает на силу их связи. Это позволяет более эффективно анализировать и интерпретировать связи между терминами, связанными с киберугрозами, и выявлять ключевые направления в данной области исследования [23].

Для кластеризации киберугроз можно выделить четыре главных кластера, каждый из которых включает восемь ключевых слов, отражающих специфические аспекты темы.

- 1. Типы кибератак представлены словами «фишинг», вредоносное ПО», «DDoS-атака», «эксплойт», «шпионское ПО», «сетевая атака», «социальная инженерия» и т.д.
- 2. Уязвимости и риски представлены словами «уязвимость», «патч», «конфигурация», «уязвимости веб-приложений», «риски», «система управления уязвимостями», «аудит безопасности» и др.
- 3. Методы защиты и профилактика представлены словами «антивирус», «фаервол», «шифрование», «аутентификация», «обучение пользователей», «резервное копирование», «управление инцидентами», «анализ рисков» и т.д.
- 4. Актуальные тренды и технологии представлены словами «киберугрозы», «искусственный интеллект», «машинное обучение», «криптография», «ІоТ (интернет вещей)», «облачные технологии», «блокчейн»,

«цифровая трансформация» и т.д.

Эти кластеры и ключевые слова в информационном поле помогают формировать обширное представление о киберугрозах, их классификации и методах защиты, а также актуальных тенденциях в области кибербезопасности. Эта кластеризация помогает лучше понять различные аспекты киберугроз и организовать их для более эффективного анализа и защиты.

Таким образом, алгоритм построения карты ключевых слов по теме «киберугрозы» в VOSviewer включает сбор, импорт и анализ данных, кластеризацию ключевых слов и визуализацию полученных результатов. Этот процесс позволяет исследователю наглядно увидеть структуру и динамику научной темы, выявить ключевые направления и взаимосвязи между различными аспектами кибербезопасности.

IV. ЗАКЛЮЧЕНИЕ

Программа VOSviewer в библиометрическом и наукометрическом анализе дает возможности для кластеризации визуализации И информации киберугрозах в различных источниках, например лентах новостей. Анализ терминов, связанных с киберугрозами, позволяет наиболее оперативно исследовать изменения, происходящие в информационном поле. Выделение кластеров и определение их объемов позволит рассматривать направление развития киберугроз и, как следствие, оперативно реагировать на них. Четко обозначенные кластеры, представляющие различные аспекты киберугроз — от типов атак до методов защиты, помогают формировать системное представление о текущем состоянии и тенденциях кибербезопасности.

Методология, изложенная в работе, включает в себя все ключевые этапы: от сбора и импорта данных до анализа и визуализации, что позволяет не только увидеть структуру исследовательской сети, но и выявить ключевые направления, области риска и перспективы развития. Таким образом, VOSviewer является ценным инструментом для анализа информации киберугрозам, позволяющим исследователям более эффективно организовывать И интерпретировать сведения, что в свою очередь способствует принятию обоснованных решений в области информационной безопасности.

БИБЛИОГРАФИЯ

- [1] Басшыкызы, Д. Обеспечение кибербезопасности в современном мире / Д. Басшыкызы // Наука, техника и образование. −2022. № 3(86). С. 36-37. EDN ZULOMV.
- [2] Таков, А. З. Проблемы обеспечения кибербезопасности в современных цифровых системах / А. З. Таков // Пробелы в российском законодательстве. 2023. Т. 16, № 5. С. 232-236. EDN OTHUWF.
- [3] Гылышаев, М. Международно правовые основы обеспечения кибербезопасности / М. Гылышаев // Ео ipso. 2023. № 3. С. 47-48. EDN BGJOCI.
- [4] Исичко, В. Д. Аспекты обеспечения кибербезопасности в условиях цифровой трансформации/В. Д. Исичко, А. В. Линкина // Вестник Воронежского института высоких технологий. – 2021. – № 4(39). – С. 48-51. – EDN IBXQBY.

- [5] Васильев, А. В. Приемы и средства обеспечения кибербезопасности на предприятиях / А. В. Васильев // Инновационная наука. – 2023. – № 9-2. – С. 10-14. – EDN XXXFAQ.
- [6] Лошкарев, А. В. Методы по обеспечению кибербезопасности и расчет киберрисков / А. В. Лошкарев // Эксплуатация морского транспорта. 2022. № 3(104). С. 153-159. DOI 10.34046/aumsuomtl04/25.—EDN MDTSBM.
- [7] Кравец, А. Г. Прогнозирование технологических тенденций на основе анализа разнородных данных / А. Г. Кравец, Т. В. Нгуен // Программные продукты и системы. 2022. № 3. С. 396-412. DOI 10.15827/0236-235X.139.396-412. EDNLCARBO.
- [8] Благинин, В. А. Тематический анализ российских исследований в области цифровизации экономики / В. А. Благинин, Е. В. Соколова // Теория и практика общественного развития. – 2023. – № 12(188). – С. 295-302. – DOI 10.24158/tipor.2023.12.37. – EDN NOYDFH.
- [9] Синь, Н. Переводоведческие исследования в Китае (2001-2020 гг.): визуализация на основе программы VOSviewer / Н. Синь // Вестник Московского университета. Серия 19: Лингвистика и межкультурная коммуникация. 2023. № 2. С. 159-176. DOI 10.55959/MSU-2074-1588-19-2023-2-01-13. EDN RMHHCN.
- [10] Xiyang, H. A visual analysis of the research on the use of mobile phones by college students based on VOSviewer / H. Xiyang // International Journal of Education and Management Engineering – 2020. – Vol. 10, No. 6. – P. 10-16. – DOI 10.5815/ijeme.2020.06.02. – EDN OAWMQT.
- [11] Наукометрический анализ отечественных статей по профессиональному выгоранию с использованием программы VOSviewer / В. И. Евдокимов, Р. К. Назыров, М. С. Плужник, Б. А. Низомутдинов // Вестник психотерапии. 2023. № 88. С. 38-61. DOI 10.25016/2782-652X-2023-0-88-38-61. EDN MBQIZF.
- [12] Чарыев, А. Б. Способы обеспечения кибербезопасности в локальной сети / А. Б. Чарыев, Д. Н. Чарыева, Ш. Г. Мамметгулыева // Символ науки: международный научный журнал. 2023. № 10-2. С. 62-64. EDN BPCDET.
- [13] Maulani, A. Bibliometric analysis: Adoption of big data analytics in financial auditing / A. Maulani, R. Widuri // Business Informatics. – 2024. – Vol. 18, No. 2. – P. 78-89. – DOI 10.17323/2587-814X.2024.2.78.89. – EDN ZQDFOM.
- [14] Моргун, А. Н. Оценка деятельности научных школ при помощи библиометрического картирования / А. Н. Моргун, А. П. Эттингер // Методология и технология непрерывного профессионального образования. 2021. № 1(5). С. 38-53. DOI 10.24075/МТСРЕ.2021.004. EDN FUHKGT.
- [15] Problem Solving and Decision-Making Skills for ESD: A Bibliometric Analysis / M. Ubaidillah, P. Marwoto, W. Wiyanto, S. Bambang // International Journal of Cognitive Research in Science, Engineering and Education. – 2023. – Vol. 11, No. 3. – P. 401-415. – DOI 10.23947/2334-8496-2023-11-3-401-415. – EDN EAOXLD.
- [16] Yerzhanova, A. Trends in the Study of the Development of Social Entrepreneurship: Bibliometric Analysis / A. Yerzhanova, L. Ashirbekova // Труды университета. 2024. No. 1(94). Р. 406-413. DOI 10.52209/1609-1825_2024_1_406. EDN IDBDPN.
- [17] Анализ направлений научных исследований в зарубежных статьях по подводной медицине с использованием программы VOSviewer / В. И. Евдокимов, Д. П. Зверев, И. Г. Мосягин, М. С. Плужник // Морская медицина. 2024. 1.0, № 1. 0.84-98. DOI 10.22328/2413-0.5747-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084-0.084
- [18] Вьет, Н. Т. Новый метод прогнозирования технологических трендов на основе анализа научных статей и патентов / Н. Т. Вьет, А. Г. Кравец // International Journal of Open Information Technologies. 2022. Т. 10, № 10. С. 49-62. EDN RBERSN.
- [19] Семантический анализ ключевых слов в зарубежных статьях по психотерапии (2012-2021 гг.) / В. И. Евдокимов, Р. К. Назыров, М. С. Плужник [и др.] // Вестник психотерапии. 2023. № 87. С. 5-19. DOI 10.25016/2782-652X-2023-0-87-05-19. EDN VMSUFS.
- [20] Lozano S., Calzada-Infante L., Adenso-Díaz B., García S. Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. Scientometrics, 2019, no. 120, pp. 609–629.
- [21] Прокофьева, М. С. Исследование наукометрических данных в области исскуственного интеллекта / М. С. Прокофьева // Системный анализ и логистика. 2021. N = 4(30). С. 57-67. DOI 10.31799/2077-5687-2021-4-57-67. EDNNTCMRL.
- [22] Павлова, И. А. Построение карты соприсутствия ключевых слов по теме «Капитал здоровья» в программе Vosviewer / И. А. Павлова // Векторы благополучия: экономика и социум. 2023. Т. 49, № 2. С. 38-54. DOI 10.18799/26584956/2023/2/1592. EDN MIPGCW.

[23] Wang Jun. Visualized analysis of hotspot issues of table tennis based on the prospective of Vosviewer / Wang Jun, Li Tie, S. I. Kolodeznikova // Human. Sport. Medicine. – 2023. – Vol. 23, No. 2. – P. 140-149. – DOI 10.14529/hsm230217. – EDN HKHVAN.

VOSviewer Program for Clustering and Visualizing Cyber Threat Information

Rimma Gorokhova, Dmitriy Korovin

Abstract - This study considers a method for classifying vulnerabilities and cyber threats using data from information and news resources. The authors propose to apply text analysis and machine learning methods to automate the detection and classification of threats in information systems. Various approaches are explored, including latent semantic analysis, topic modeling, n-gram analysis, and vector representation of text, which allows identifying semantic relationships and thematic structures. These methods help to adapt approaches to the requirements for interpretability and accuracy of cyber threat data. The paper considers an approach to analyzing the co-occurrence of words and phrases in texts to identify thematic clusters of cyber threats. The main analysis tool is the calculation of the association measure between terms, which allows for a quantitative assessment of their relationships and contributes to the construction of more effective classification models using machine learning methods. To visualize the results, it is proposed to use strategic diagrams based on two key indicators: centrality and cluster density. Clusters are classified into four categories depending on the values of these indicators, which allows for a more accurate characterization of their essence and relationships with other threats. To analyze the dynamics of changes in clusters, the method of directed graphs is used, which allows tracking the transformation of cyber threat components from one time period to another. Based on the presented approach, it is possible to create more informative cyber threat monitoring systems, which ultimately contributes to increasing the level of protection of information systems from potential attacks. Thus, the study opens up new horizons for automating the analysis and classification of threats, providing more effective solutions in the field of cybersecurity.

Keywords - cyber threats, VOSviewer, clustering, data visualization, bibliometric analysis, information security.

REFERENCES

- Basshykyzy, D. Ensuring cybersecurity in the modem world / D. Basshykyzy // Science, technology and education. - 2022. - No. 3 (86). - P. 36-37. - EDN ZULOMV.
- [2] Takov, A. Z. Problems of ensuring cybersecurity in modem digital systems / A. Z. Takov // Gaps in Russian legislation. - 2023. - Vol. 16, No. 5. - P. 232-236. - EDN OTHUWF.
- [3] Gylyshaev, M. International legal framework for ensuring cybersecurity / M. Gylyshaev // Eo ipso. - 2023. - No. 3. - P. 47-48. -EDN BGJOCI.
- [4] Isichko, V. D. Aspects of Ensuring Cybersecurity in the Context of Digital Transformation / V. D. Isichko, A. V. Linkina// Bulletin of the Voronezh Institute of High Technologies. - 2021. - No. 4 (39). - P. 48-51. - EDN IBXOBY.
- [5] Vasiliev, A. V. Techniques and Means of Ensuring Cybersecurity at Enterprises / A. V. Vasiliev // Innovative Science. - 2023. - No. 9-2. -P. 10-14. - EDN XXXFAQ.
- [6] Loshkarev, A. V. Methods for Ensuring Cybersecurity and Calculating Cyber Risks / A. V. Loshkarev // Operation of Maritime Transport. -2022. - No. 3 (104). - P. 153-159. – DOI 10.34046/aumsuomtl04/25. – EDN MDTSBM.
- [7] Kravets, A. G. Forecasting technological trends based on the analysis of heterogeneous data / A. G. Kravets, T. V. Nguyen // Software products and systems. – 2022. – No. 3. – P. 396-412. – DOI 10.15827/0236-235X.139.396-412. – EDN LCARBO.
- [8] Blaginin, V. A. Thematic analysis of Russian studies in the field of digitalization of the economy / V. A. Blaginin, E. V. Sokolova //

- Theory and practice of social development. 2023. No. 12(188). P. 295-302. DOI 10.24158/tipor.2023.12.37. EDN NOYDFH.
- [9] Xin, H. Translation Studies in China (2001-2020): Visualization Based on VOSviewer / H. Xin // Bulletin of Moscow University. Series 19: Linguistics and Intercultural Communication. – 2023. – No. 2. – P. 159-176. – DOI 10.55959/MSU-2074-1588-19-2023-2-01-13. – EDN RMHHCN.
- [10] Xiyang, H. A visual analysis of the research on the use of mobile phones by college students based on VOSviewer / H. Xiyang // International Journal of Education and Management Engineering – 2020. – Vol. 10, No. 6. – P. 10-16. – DOI 10.5815/ijeme.2020.06.02. – EDN OAWMOT.
- [11] Scientometric analysis of domestic articles on professional burnout using the VOSviewer program / V. I. Evdokimov, R. K. Nazirov, M. S. Pluzhnik, B. A. Nizomutdinov // Bulletin of Psychotherapy. 2023. No. 88. P. 38-61. DOI 10.25016/2782-652X-2023-0-88-38-61. EDN MBQIZF.
- [12] Charyev, A. B. Methods of Ensuring Cybersecurity in a Local Area Network / A. B. Charyev, D. N. Charyeva, Sh. G. Mammetgulyeva // Symbol of Science: International Scientific Journal. - 2023. - No. 10-2. - P. 62-64. - EDN BPCDET.
- [13] Maulani, A. Bibliometric Analysis: Adoption of Big Data Analytics in Financial Auditing / A. Maulani, R. Widuri // Business Informatics. -2024. - Vol. 18, No. 2. - P. 78-89. - DOI 10.17323/2587-814X.2024.2.78.89. - EDN ZQDFOM.
- [14] Morgun, A. N. Evaluation of the activities of scientific schools using bibliometric mapping / A. N. Morgun, A. P. Ettinger // Methodology and technology of continuous professional education. - 2021. - No. 1 (5). - P. 38-53. - DOI 10.24075 / MTCPE.2021.004. - EDN FUHKGT.
- [15] Problem Solving and Decision-Making Skills for ESD: A Bibliometric Analysis / M. Ubaidillah, P. Marwoto, W. Wiyanto, S. Bambang // International Journal of Cognitive Research in Science, Engineering and Education. – 2023. – Vol. 11, No. 3. – P. 401-415. – DOI 10.23947/2334-8496-2023-11-3-401-415. – EDN EAOXLD.
- [16] Yerzhanova, A. Trends in the Study of the Development of Social Entrepreneurship: Bibliometric Analysis / A. Yerzhanova, L. Ashirbekova // Proceedings of the University. – 2024. – No. 1(94). – P. 406-413. – DOI 10.52209/1609-1825_2024_1_406. – EDN IDBDPN.
- [17] Analysis of scientific research directions in foreign articles on underwater medicine using the VOSviewer program / V. I. Evdokimov, D. P. Zverev, I. G. Mosyagin, M. S. Pluzhnik // Marine Medicine. – 2024. – Vol. 10, No. 1. – Pp. 84-98. – DOI 10.22328/2413-5747-2024-10-1-84-98. – EDNMJBIAA.
- [18] Viet, N. T. A new method for forecasting technological trends based on the analysis of scientific articles and patents / N. T. Viet, A. G. Kravets // International Journal of Open Information Technologies. – 2022. – Vol. 10, No. 10. – P. 49-62. – EDN RBERSN.
- [19] Evdokimov V.I., Nazirov R.K., Pluzhnik M.S. [et al.]. Semantic analysis of keywords in foreign articles on psychotherapy (2012-2021) // Vestn. psychotherapy. 2023. No. 87. P. 5-19. DOI: 10.250162782-652X-2023-0-87-05-19 EDN: VMSUFS.
- [20] Lozano S., Calzada-Infante L., Adenso-Díaz B., García S. Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature. Scientometrics, 2019, no. 120, pp. 609–629.
- [21] Prokofieva, M. S. Issstudy of scientometric data in the field of artificial intelligence / M. S. Prokofieva // Systems analysis and logistics. - 2021. - No. 4 (30). - P. 57-67. - DOI 10.31799 / 2077-5687-2021-4-57-67. -EDN NTCMRL.
- [22] Pavlova, I. A. Construction of a map of the co-presence of keywords on the topic "Health capital" in the Vosviewer program / I. A. Pavlova // Vectors of well-being: economy and society. - 2023. - Vol. 49, No. 2. - P. 38-54. - DOI 10.18799 / 26584956/2023/2/1592. - EDN MIPGCW.
- [23] Wang Jun. Visualized analysis of hotspot issues of table tennis based on the prospective of Vosviewer / Wang Jun, Li Tie, S. I. Kolodeznikova // Human. Sport. Medicine. – 2023. – Vol. 23, No. 2. – P. 140-149. – DOI 10.14529/hsm230217. – EDN HKHVAN.