УДК 004.056
МРНТИ 81.93.29

https://doi.org/10.55452/1998-6688-2025-22-4-107-118

[1]*Самуйлова А.,
магистрант, ORCID ID: 0009-0003-5806-8542,
*e-mail: anastassyiasamuilova@gmail.com

[1]Казахский национальный университет им. аль-Фараби, г. Алматы, Казахстан

СРЕДСТВА ВЫЯВЛЕНИЯ УЯЗВИМОСТЕЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВАНИИ ДАННЫХ ТЕМАТИЧЕСКИХ ИНТЕРНЕТ-РЕСУРСОВ

Аннотация

По мере усложнения киберугроз традиционные методы обнаружения уязвимостей теряют эффективность. Цель данной работы – разработка и апробация подхода к выявлению уязвимостей на основе анализа данных с тематических интернет-ресурсов: форумов, блогов и социальных сетей. Эти источники содержат большое количество неструктурированной информации, что требует применения методов интеллектуального анализа данных. В работе используется интеграция современных технологий: предобученной языковой модели SecBERT (Security Bidirectional Encoder Representations from Transformers), предназначенной для задач в области кибербезопасности, и адаптивной нейро-нечеткой системы вывода DENFIS (Dynamic Evolving Neural-Fuzzy Inference System). Предлагаемая система позволяет фильтровать нерелевантные сообщения, выделять индикаторы компрометации и потенциальные угрозы. Применение нечеткой логики дает возможность эффективно обрабатывать неопределенную и неполную информацию. Проведенные эксперименты подтвердили высокую точность классификации и устойчивость нечеткой кластеризации (FPC = 0,93; PE = 0,28; XB = 0,042). Система продемонстрировала способность к своевременному выявлению признаков киберугроз и обладает потенциалом масштабирования для задач мониторинга и предсказания атак. Результаты свидетельствуют о его потенциале в повышении скорости реагирования на киберугрозы и усилении защиты информационных систем.

Ключевые слова: уязвимости информационной безопасности, тематические интернет-ресурсы, трансформеры, SecBERT, нейро-нечеткая логика.

Введение

В современных условиях стремительного развития цифровых технологий и роста объемов информации увеличивается вероятность кибератак и выявления новых уязвимостей в информационных системах. Традиционные методы обнаружения угроз, такие как сигнатурный анализ, поведенческий анализ, анализ аномалий и статистические подходы, обладают существенными ограничениями, поскольку они ориентированы на известные шаблоны атак или аномалии, выявленные в прошлом [1]. Однако учитывая сложность и динамичность современных угроз, требуется более гибкий и адаптивный подход, способный эффективно анализировать разнородные источники информации, включая специализированные интернет-ресурсы, хакерские форумы, базы данных уязвимостей и новостные агрегаторы [2].

На сегодняшний день ключевыми источниками данных о киберугрозах остаются такие базы, как NVD или CVE, а также отчеты государственных и частных организаций, где сведения структурируются экспертами и исследовательскими центрами [3]. Однако эти источники имеют очевидный недостаток: процесс верификации и внесения данных может занимать значительное время, в результате чего появляется временной лаг между выявлением угрозы и ее официальной регистрацией [4]. В противоположность этому специализированные форумы и анонимные площадки предоставляют информацию о новых уязвимостях и потенциальных

векторах атак практически в режиме реального времени [5]. Их высокая популярность среди специалистов в области информационной безопасности делает такие ресурсы ценным источником для мониторинга и анализа киберугроз, позволяя фиксировать и прогнозировать тенденции еще до их официального признания [6].

Актуальность данного исследования обусловлена необходимостью создания автоматизированного подхода, который позволит эффективно извлекать полезную информацию о киберугрозах из открытых источников [5]. Традиционные методы обнаружения уязвимостей, основанные на статическом анализе данных, не способны оперативно реагировать на появление новых типов атак, в то время как интеграция технологий трансформеров и нейро-нечеткой логики может существенно повысить качество прогнозирования и выявления потенциальных угроз [2].

Объектом исследования выступают информационные системы, функционирующие в условиях активного использования интернет-ресурсов. Предметом – методы автоматического выявления уязвимостей с помощью интеллектуального анализа данных. Основная цель работы – разработать комплексный метод, который позволит повысить точность и оперативность обнаружения киберугроз и потенциальных рисков.

Для реализации данной цели необходимо решить следующие задачи:

- проанализировать существующие методы выявления уязвимостей и выявить их ограничения в условиях постоянно меняющейся киберугрозы;
- разработать методику, сочетающую трансформеры для извлечения семантических признаков и нейро-нечеткую логику для работы с неопределенностью в данных;
- провести эксперименты на реальных данных, собранных с форумов и баз уязвимостей;
- провести сравнительный анализ предложенного метода с традиционными подходами, такими как сигнатурный анализ, статистический анализ и методы машинного обучения.

Методологическая основа исследования включает в себя методы глубокого обучения, применяемые для классификации и категоризации сообщений об уязвимостях, нейро-нечеткую логику, обеспечивающую обработку неопределенных данных и вероятностную оценку угроз, а также статистический анализ, позволяющий выявлять закономерности в распространении уязвимостей и их потенциальное влияние на системы [5].

Современные исследования в области анализа киберугроз демонстрируют эффективность методов машинного обучения и глубокого анализа текста в выявлении потенциальных угроз [2]. В работах [1] и [5] показано, что применение нейросетевых моделей для анализа сообщений на хакерских форумах позволяет выявлять новые уязвимости и векторы атак задолго до их эксплуатации злоумышленниками. Одним из ключевых направлений исследований является анализ индикаторов компрометации (IoC), который используется в платформах AlienVault OTX, VirusTotal и MISP для сопоставления выявленных угроз с реальными атаками [7].

Особенно перспективным направлением считается использование трансформеров, таких как BERT, и его специализированных версий – SecBERT, DarkBERT. Эти модели показали высокую эффективность в обработке естественного языка в контексте задач кибербезопасности [8]. Исследования [9] также подтверждают, что механизмы внимания (attention mechanism) значительно повышают точность извлечения ключевых элементов из текстов, что критически важно при анализе форумов с неструктурированной информацией. В статье [10] представлена модель SecureBERT 2.0, обученная на обширном и тщательно подобранном корпусе данных, включающем более 13,6 млрд текстовых токенов и 53,3 млн токенов кода. Корпус включает публичные отчеты о безопасности, уведомления об уязвимостях, открытые статьи по кибербезопасности, технические книги и рецензируемые научные работы, что позволяет модели получать богатое контекстное представление о кибербезопасности в различных поддоменах, включая сетевую архитектуру, управление доступом, механизмы аутентификации и криптографические протоколы. Модель достигает F1-score 0.945 с высоким recall (0.965, доля

корректно выявленных сущностей) и точностью (0.927, доля правильно классифицированных положительных объектов), что позволяет эффективно выявлять киберугрозы и ключевые сущности при минимуме ложноположительных срабатываний. В статье [11] представлена модель DarkBERT, обученная на текстах Даркнета из датасетов DUTA-10K и CoDA, включающих страницы с onion-доменов, HTML-теги title и body, а также различные активности пользователей Даркнета. Модель использует архитектуру RoBERTa и обучалась как на сырых, так и на предобработанных данных. На датасете DUTA F1-score составляет 80.01 на сырых данных и 79.98 на предобработанных, с precision и recall около 80, а на CoDA достигает 94.25 на сырых и 94.42 на предобработанных данных.

Однако несмотря на успехи глубокого обучения, оно остается недостаточным для полноценной работы с динамически изменяющимися угрозами. В этой связи в ряде исследований предлагается интеграция с другими методами. Так, в обзоре [12] авторы рассматривают применение больших языковых моделей (LLM) в кибербезопасности и отмечают важность решения вопросов этики, объяснимости моделей и их интеграции с другими методами, а для повышения эффективности предлагают сотрудничество с экспертами по когнитивным наукам и психологии. Помимо этого, для работы с неопределенностью и динамичностью угроз в ряде исследований предлагается использование нейро-нечеткой логики [13]. Подходы, основанные на адаптивных нейро-нечетких системах вывода (ANFIS, DENFIS), позволяют повысить точность прогнозирования угроз без необходимости полной переобучаемости модели. Это особенно актуально в условиях постоянной смены тактик атак и появления новых типов уязвимостей [14].

Материалы и методы

В данном исследовании применяется интегрированный подход, основанный на сборе, обработке и анализе данных с тематических интернет-ресурсов с использованием методов вебскрапинга, предобученных моделей и нейро-нечеткой логики, согласно рисунку 1.

Рисунок 1 – Архитектура разрабатываемой модели

На первом этапе осуществляется парсинг данных с хакерских форумов и других специализированных интернет-ресурсов. Используется Python-библиотека BeautifulSoup, Requests, Praw для извлечения постов из Reddit и Sqlite3 для динамического сбора данных с сайтов, требующих взаимодействия пользователя. Собранные данные сохраняются в SQL-базе данных для последующей обработки. Структура базы данных показана на рисунке 2.

На втором этапе осуществляется предобработка текстовой информации, включая очистку данных, удаление дубликатов и фильтрацию нерелевантных сообщений. Применяется лемма-

тизация и удаление стоп-слов с использованием библиотек NLTK и spaCy. Данные подвергаются первичной фильтрации с использованием ключевых слов, связанных с информационной безопасностью, таких как exploit, vulnerability, malware.

Рисунок 2 – Структура базы данных

Для классификации сообщений и определения релевантности используется предобученная модель трансформеров SecBERT. SecBERT – это языковая модель на основе архитектуры BERT, показанной на рисунке 3, адаптированная для обработки текстов в сфере кибербезопасности. Он использует механизм самовнимания, который позволяет учитывать контекстные зависимости между словами и выявлять значимые элементы в тексте. Модель принимает на вход токенизированный текст и пропускает его через слои трансформеров, после чего формируются эмбеддинги слов. Для представления входных данных используется векторное представление токенов $X = [x_1, x_2, ..., x_i]$, где x_i – это эмбеддинг i-го токена. Эти эмбеддинги передаются через слои самовнимания, определяемые следующим уравнением:

$$A = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) * V \quad (1)$$

где Q, K, V – это матрицы запросов, ключей и значений, а d_k – размерность ключей. Итоговый скрытый слой H формируется после нескольких таких операций, проходя через механизмы нормализации и активации. Для классификации угроз выходное представление H передается в полносвязный слой с функцией активации softmax, который вычисляет вероятность принадлежности сообщения к категории угроз.

Модель была обучена на специализированном корпусе, включающем такие источники, как APTnotes, Stucco-Data, CASIE, а также данные из задачи SemEval-2018 Task 8 (SecureNLP). Для повышения точности работы с текстами данной тематики в модели используется специализированный словарь (secvocab), сформированный на основе терминологии, характерной для кибербезопасности. Помимо стандартной версии SecBERT, существуют также ее модификации, такие как SecRoBERTa, созданные на основе архитектуры RoBERTa и обладающие схожими возможностями. Для решения задачи классификации в данной работе модель была дообучена на размеченном корпусе, содержащем 41 тысячу текстов, относящихся и не относящихся к области информационной безопасности, и продемонстрировала точность 98–99% на данном корпусе.

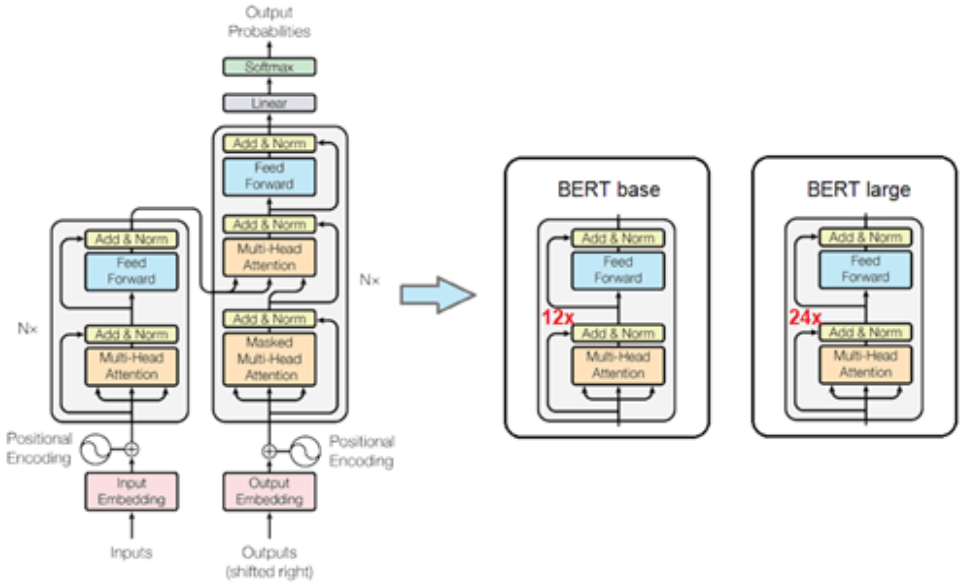


Рисунок 3 – Архитектура BERT модели

Отфильтрованные данные сохраняются в отдельную базу данных как возможные прецеденты и затем передаются в систему нейро-нечеткой логики для дополнительного анализа. Нечеткая логика представляет собой метод обработки информации, позволяющий моделировать неопределенность и приближенные значения, в отличие от классической логики, где переменные принимают строго определенные значения [15]. В данной работе применяется адаптивная нейро-нечеткая система вывода DENFIS, которая сочетает элементы нечеткой логики и нейросетевых алгоритмов, динамически адаптируясь к изменениям входных данных. Архитектура данной системы представлена на рисунке 4 и отражает процесс формирования правил на основе входных данных в реальном времени с использованием механизма эволюционного обучения [15]. Примеры правил показаны в таблице 1.

Таблица 1 – Правила нечеткой логики

1	Если количество комментариев на форуме высокое и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности высокая.
2	Если количество комментариев на форуме высокое и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности средняя.
3	Если количество комментариев на форуме высокое и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.
4	Если количество комментариев на форуме среднее и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности высокая.
5	Если количество комментариев на форуме среднее и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности средняя.
6	Если количество комментариев на форуме среднее и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.
7	Если количество комментариев на форуме низкое и уровень рейтинга авторов высокий, то вероятность возникновения угрозы информационной безопасности средняя.
8	Если количество комментариев на форуме низкое и уровень рейтинга авторов средний, то вероятность возникновения угрозы информационной безопасности низкая.
9	Если количество комментариев на форуме низкое и уровень рейтинга авторов низкий, то вероятность возникновения угрозы информационной безопасности низкая.

Принцип работы DENFIS основан на том, что входные данные оцениваются через нечеткие функции принадлежности (например, гауссовых или треугольных), которые определяют степень соответствия каждого параметра правилам, аналогичным приведенным в таблице 1. Каждое значение, например рейтинг автора или количество комментариев, проверяется относительно этих функций, и вычисляется степень принадлежности к каждому возможному условию правила. Затем минимальное расстояние между входными данными и центрами кластеров используется для активации наиболее релевантных правил и формирования итоговой оценки уровня угрозы. Входные параметры включают уровень достоверности сообщения, рейтинг автора и уровень обсуждаемости темы, то есть количество комментариев и активность пользователей. Выходной параметр представляет собой оценку уровня угрозы, включающую низкий, средний и высокий уровни.

Система адаптивна: при поступлении новых данных DENFIS корректирует свои правила функции принадлежности и обновляет центры кластеров, автоматически уточняя условия существующих правил и создавая новые комбинации характеристик постов.

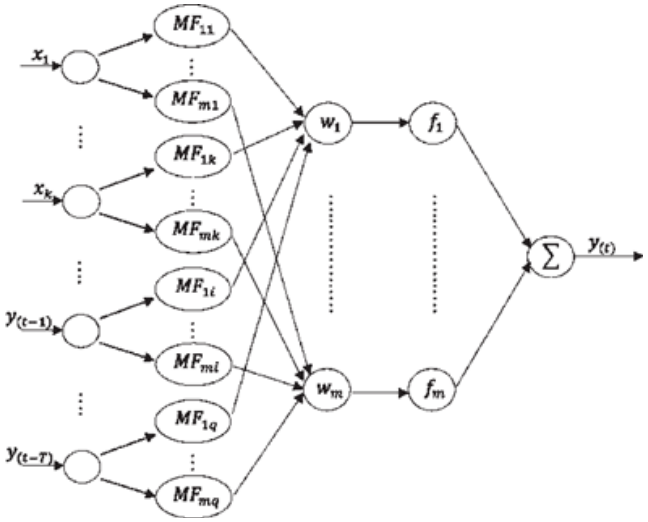


Рисунок 4 – Архитектура DENFIS

Полученные результаты визуализируются и представляются в виде отчетов, содержащих информацию о выявленных угрозах, их классификации и вероятностной оценке значимости. Данные могут быть использованы для автоматизированного мониторинга угроз в реальном времени и принятия решений по реагированию на потенциальные кибератаки. Таким образом, предложенная методика сочетает в себе веб-скрапинг, обработку данных, анализ с использованием предобученной модели и нейро-нечеткую логику, что позволяет эффективно выявлять новые угрозы и обеспечивать их классификацию с учетом вероятностной оценки риска.

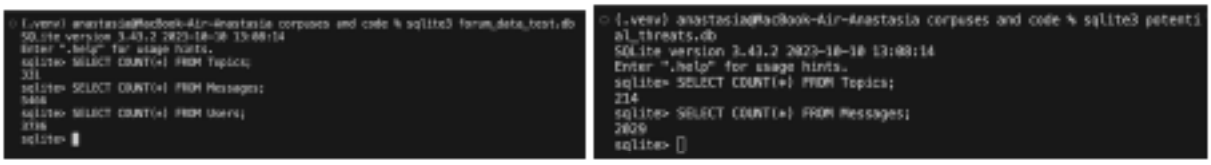
Результаты и обсуждение

Гипотеза данного исследования заключалась в том, что тематические интернет-ресурсы могут служить достоверным источником данных для автоматизированного выявления уязвимостей информационной безопасности при условии применения современных методов анализа неструктурированной информации. В ходе эксперимента был реализован интегрированный подход, сочетающий предобученную трансформерную модель SecBERT и нейро-нечеткую систему DENFIS, что позволило эффективно классифицировать посты и формировать правила на основе кластеризации. Для оценки точности кластеризации использовались метрики FPC, XB и PE: FPC (Fuzzy Partition Coefficient) отражает степень четкости принадлежности объектов к кластерам, XB (Xie-Beni index) оценивает компактность и разделимость кластеров, а PE (Partition Entropy) измеряет неопределенность распределения объектов по кластерам.

Собранный датасет включал более 300 тем, согласно рисунку 5a, и свыше 5000 комментариев с шести доступных форумов, включая Reddit, HackerNews, Antichat и других специализированных ресурсов, посвященных вопросам информационной безопасности, и использовался как тестовый корпус для проверки работы модели. Однако процесс веб-скрапинга сопровождался рядом ограничений, так как некоторые форумы были заблокированы государственными учреждениями, а другие требовали авторизацию для просмотра содержимого, что ограничило доступный объем данных.

Предварительно обученная модель SecBERT успешно выделила 214 релевантных тем, согласно рисунку 5b, соответствующих признакам потенциальных угроз. Пример таких релевантных постов представлен на рисунке 6. Далее с применением алгоритма fuzzy c-means были выделены три устойчивых кластера и рассчитаны их центры для дальнейшей классификации собранных данных.

В ходе работы также был рассмотрен альтернативный метод классификации данных на основе онтологий, как показано в исследовании [16]. Однако данный подход обладает рядом существенных ограничений. Онтологическая модель плохо справляется с новыми угрозами, так как требует заранее заданной структуры знаний. Точность классификации напрямую зависит от полноты онтологии и ее соответствия современным угрозам. Онтологии требуют постоянного пополнения и ручного контроля, что делает их менее гибкими в условиях динамически меняющейся киберугрозы. Кроме того, онтологии строятся на фиксированных взаимосвязях между понятиями, что может привести к ошибкам классификации при наличии неоднозначных или неизвестных терминов. Также процесс разработки и поддержки онтологии требует значительных затрат ресурсов, в отличие от машинного обучения, где модели могут автоматически подстраиваться под новые угрозы.



a) До фильтрации b) После фильтрации

Рисунок 5 – Результаты фильтрации при помощи SecBERT

На последнем этапе данные передаются в DENFIS для определения уровня угрозы обсуждений. Сначала происходит подключение к базе данных, содержащей информацию о потенциальных инцидентах информационной безопасности. С помощью SQL-запроса из таблиц с постами и пользователями извлекаются необходимые данные: идентификатор и название темы, рейтинг автора, количество сообщений, время создания поста и имя автора. После извлечения информация преобразуется в удобный для анализа формат – двумерный массив, в котором в качестве признаков используются только числовые значения: рейтинг автора и количество сообщений. Это позволяет эффективно применить алгоритм кластеризации для формирования исходных правил для дальнейшей работы нейро-нечеткой модели.

Алгоритм fuzzy c-means вычисляет центры кластеров и матрицу принадлежности объектов к кластерам. Полученные результаты выводятся в консоль и визуализируются на графике, как показано на рисунке 7: объекты, принадлежащие к каждому кластеру, отображаются разными цветами, а центры кластеров обозначаются крестиками.

В результате работы алгоритма получилось три центра кластеров: (3.98824988e+05, 1.00000024e+00), (1.71489384e+03, 6.88635187e+01) и (1.58592931e+04, 1.75610911e+00). Каждый из этих центров отражает характерные группы объектов, отличающиеся, в частности, по масштабу первого признака и умеренно варьирующиеся по второму.

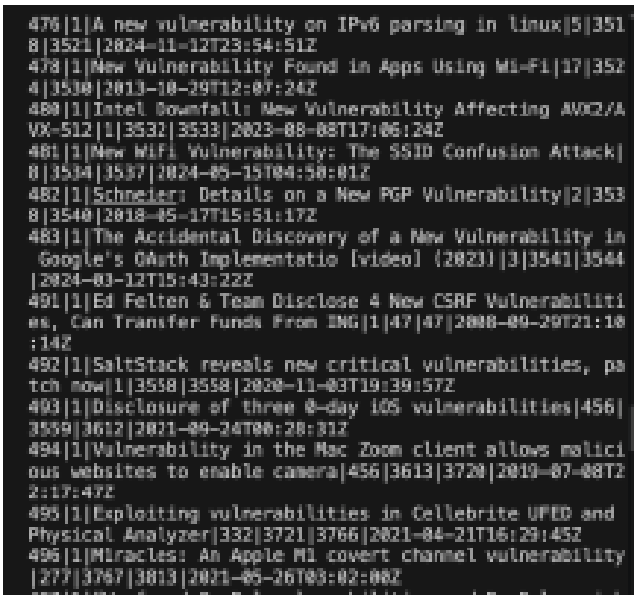


Рисунок 6 – Пример релевантных постов

Суммарно все три кластера, выделенные в ходе анализа, демонстрируют устойчивость и обоснованность: высокое значение FPC (0.93), низкий PE (0.28) и низкий XB (0.042) показывают, что выделенные группы действительно отражают различные типы обсуждений, подтвердив высокую четкость и достоверность кластеризации. Первый кластер при этом характеризуется очень крупным значением первого признака и небольшим вторым, второй кластер включает объекты со значительно меньшим первым признаком и более высокой величиной второго, а третий занимает промежуточное положение по первому признаку с незначительно возросшим вторым.

На основании данной кластеризации были сформированы изначальные правила для последующего применения алгоритма DENFIS, при котором каждому кластеру сопоставляется соответствующая метка, а сами метки затем учитываются в процессе нечеткого неструктурированного моделирования.

После этапа кластеризации данные проходят дополнительную обработку, цель которой – вычисление совокупного балла поста, отражающего его значимость и потенциальный уровень угрозы. Для этого из каждого поста извлекаются два ключевых показателя: рейтинг автора и число сообщений в теме. Чтобы сделать их сопоставимыми, применяется нормализация. Количество сообщений нормализуется линейно, а рейтинг – с помощью логарифмической трансформации, которая позволяет нивелировать влияние экстремально высоких значений и повысить чувствительность к колебаниям в нижнем диапазоне.

После нормализации оба показателя объединяются с равными весами, то есть суммируются и делятся пополам, а затем полученное значение масштабируется в диапазоне от 0 до 100. Этот итоговый балл представляет собой комбинированную меру важности, которая учитывает как активность (количество сообщений), так и авторитетность (рейтинг автора). Таким образом, высокий балл свидетельствует о том, что пост получил высокий рейтинг и/или содержит большое количество сообщений, что может указывать на повышенный уровень обсуждения и потенциальную угрозу.

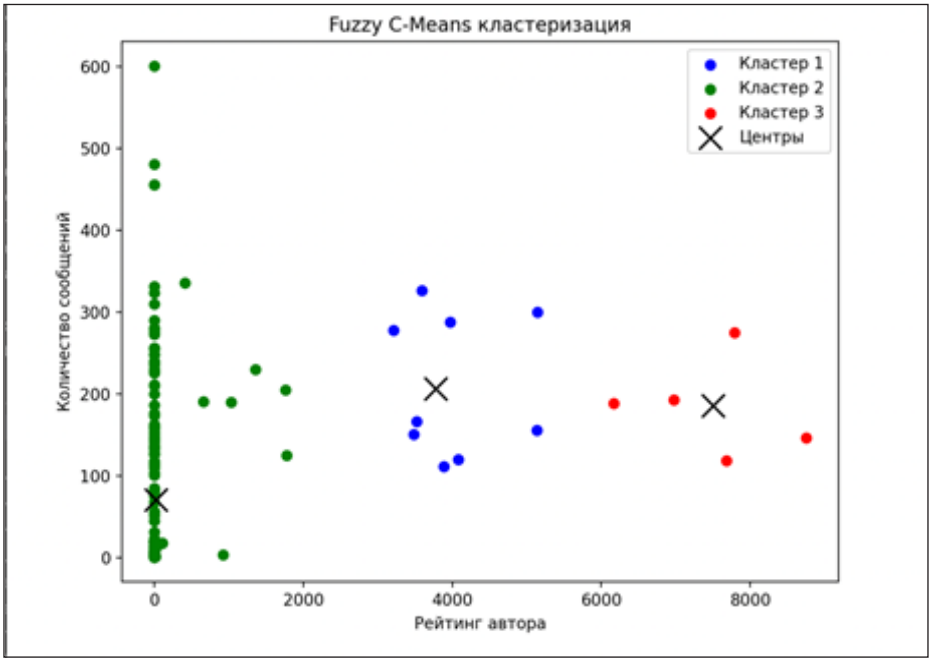


Рисунок 7 – Результат кластеризации

Далее в зависимости от итогового балла посту присваивается метка: если значение балла превышает высокий порог (например, 66), то пост получает метку high; если балл находится в среднем диапазоне (от 33 до 66) – метку medium; а если ниже – метку low. Эта классификация позволяет разделить посты по уровням угрозы, что является важным этапом для последующего анализа. Результаты этой классификации представлены на рисунке 8, где показано распределение количества постов по соответствующим меткам.

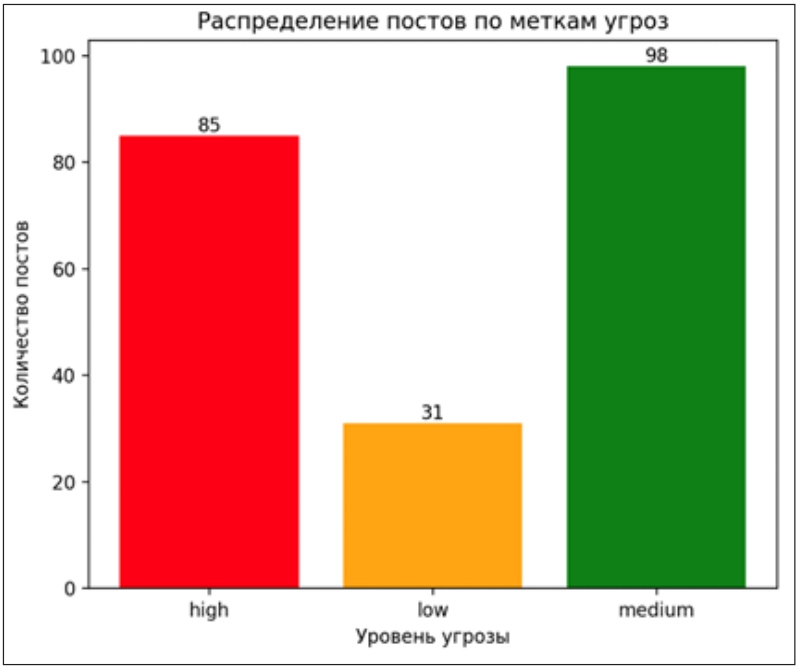


Рисунок 8 – Результат нейро-нечеткого алгоритма

Полученные таким образом результаты подтверждают, что предварительное групповое разбиение данных на три нечетких кластера дает достаточно устойчивые «прототипы», которые впоследствии могут быть эффективно использованы для генерации правил в DENFIS.

Заключение

В ходе проведенного исследования удалось успешно достичь поставленной цели – разработать интегрированный подход к обнаружению киберугроз на основе анализа тематических интернет-ресурсов с использованием методов кластеризации и нейро-нечеткой логики. Созданная система позволяет объединять все ключевые этапы – от извлечения и предварительной обработки данных до кластеризации с помощью алгоритма fuzzy c-means и формирования начальных правил для последующей работы нейро-нечеткой модели DENFIS. Полученные результаты демонстрируют высокую эффективность модели, что подтверждается следующими краткими выводами:

Извлечение и предварительная обработка данных позволили сформировать набор признаков (рейтинг автора и количество сообщений), обеспечивающих корректное представление исходной информации.

Применение алгоритма нечеткой кластеризации привело к выделению устойчивых прототипов, характеризуемых высокими значениями коэффициента нечеткого разбиения (FPC ≈ 0.93), что указывает на достаточно четкое распределение объектов.

Дополнительные метрики качества кластеризации – Partition Entropy (0.28) и индекс Xie-Beni (0.042) – подтвердили компактность и разнесенность полученных кластеров, что свидетельствует о высокой определенности группировки данных.

На основании полученных кластерных центров были сформированы первоначальные правила для нейро-нечеткой модели DENFIS, позволяющие присваивать постам метки high, medium или low в зависимости от совокупного балла, рассчитанного на основе нормализованных значений рейтинга и количества сообщений.

Итоговая интеграция предварительной кластеризации с динамической адаптацией правил в системе DENFIS обеспечивает оперативное и точное определение уровня угрозы обсуждений, что является критически важным для своевременного реагирования на потенциальные кибератаки.

В дальнейшем планируется расширение функциональности предлагаемой системы за счет интеграции дополнительных источников данных, таких как специализированные социальные сети, новые тематические интернет-ресурсы и форумы. Также будет вестись работа по улучшению этапа предварительной обработки, чтобы повысить качество извлечения информации из неструктурированных текстов. Особое внимание планируется уделить разработке адаптивных механизмов обновления правил DENFIS с применением онлайн-обучения, что позволит системе быстрее реагировать на появление новых угроз. Также рассматривается возможность внедрения современных трансформерных моделей для более глубокой семантической обработки текстов, что повысит точность классификации постов и обнаружения угроз. Эти направления будущей работы направлены на дальнейшее повышение оперативности, точности и надежности системы мониторинга, что имеет решающее значение для обеспечения информационной безопасности в условиях динамично меняющейся цифровой среды.

ЛИТЕРАТУРА

1 Sommer, R., and Paxson, V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. Proceedings of the IEEE Symposium on Security and Privacy, 305–316 (2010). https://doi.org/10.1109/SP.2010.25.

2 Bilge, L., and Dumitras, T. Before We Knew It: An Empirical Study of Zero-Day Attacks in the Real World. Proceedings of the ACM Conference on Computer and Communications Security (CCS), 83–94 (2012). https://doi.org/10.1145/2382196.2382284.

3 National Institute of Standards and Technology (NIST). National Vulnerability Database (NVD) (2010). https://nvd.nist.gov/.

4 Zhao, X., Wang, X., and Li, X. Vulnerability Disclosure and Information Verification Delays in Cybersecurity. Journal of Cybersecurity Research, 3 (2), 45–60 (2015).

5 Wang, W., and Lu, Y. Mining Cyber Threat Intelligence from the Dark Web. IEEE Transactions on Information Forensics and Security, 13 (2), 275–286 (2018). https://doi.org/10.1109/TIFS.2017.2761918.

6 O'Connor, N., and Torabi, A. Cyber Threat Intelligence: An Introduction. IEEE Security & Privacy, 13 (3), 19–27 (2015).

7 Cadar, C., Dunbar, D., and Engler, D. KLEE: Unassisted and Automatic Generation of High-Coverage Tests for Complex Systems Programs. Proceedings of the USENIX Symposium on Operating Systems Design and Implementation (OSDI), 209–224 (2008).

8 Yadav, S., Sharma, A., and Gupta, M. Indicators of Compromise Analysis Using Threat Intelligence Platforms. Journal of Cybersecurity, 5 (4), 210–223 (2019). https://doi.org/10.5555/1855741.1855756.

9 Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL HLT, 4171–4186 (2019). https://doi.org/10.18653/v1/N19-1423.

10 Aghaei, E., Jain, S., Arun, P., and Sambamoorthy, A. SecureBERT 2.0: Advanced Language Model for Cybersecurity Intelligence. Cisco AI, San Jose, CA, USA (2025). {eaghaei, sjain2, parun, asambamo}@cisco.com.

11 Jin, Y., Jang, E., Cui, J., Chung, J.-W., Lee, Y., and Shin, S. DarkBERT: A Language Model for the Dark Side of the Internet. KAIST, Daejeon, South Korea; S2W Inc., Seongnam, South Korea (2023). {ijinjin, claude}@kaist.ac.kr; {genesith, geeoon19, jwchung, lee}@s2w.inc. https://aclanthology.org/2023.acl-long.415.pdf.

12 Güven, M. A Comprehensive Review of Large Language Models in Cyber Security. International Journal of Computational and Experimental Science and Engineering, 10 (3), 507–516 (2024). https://doi.org/10.22399/ijcesen.469.

13 Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. Attention Is All You Need. Advances in Neural Information Processing Systems (NeurIPS), 5998–6008 (2017). https://doi.org/10.48550/arXiv.1706.03762.

14 Jang, J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics, 23 (3), 665–685 (1993). https://doi.org/10.1109/21.256541.

15 Kasabov, N., and Song, Q. DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time-Series Prediction. IEEE Transactions on Fuzzy Systems, 10 (2), 144–154 (2002). https://doi.org/10.1109/91.995117.

16 Poletaev, V.S. Informatsionno-analiticheskaya sistema prognozirovaniya ugroz i uyazvimostey informatsionnoy bezopasnosti na osnove analiza dannykh tematicheskikh internet-resursov [Information-Analytical System for Predicting Cybersecurity Threats and Vulnerabilities Based on Analysis of Thematic Internet Resources] (Ulyanovsk: UlSU, 2024), 172 p. (in Russian)

[1]*Самуйлова А.,
магистрант, ORCID ID: 0009-0003-5806-8542,
*e-mail: anastassyiasamuilova@gmail.com

[1]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

ИНТЕРНЕТ РЕСУРСТАРЫ ДЕРЕКТЕРІ НЕГІЗІНДЕ АҚПАРАТТЫҚ ҚАУІПСІЗДІКТІҢ ОСАЛ ТҰСТАРЫН АНЫҚТАУ ҚҰРАЛДАРЫ

Аңдатпа

Киберқауіптердің күрделенуіне байланысты дәстүрлі осалдықтарды анықтау әдістері өз тиімділігін жоғалта бастады. Бұл зерттеудің мақсаты – форумдар, блогтар және әлеуметтік желілер сияқты тақырыптық интернет-ресурстар деректерін талдау негізінде осалдықтарды анықтауға арналған тәсілді әзірлеу және сынақтан өткізу. Аталған дереккөздер құрылымдалмаған ақпараттың үлкен көлемін қамтиды, сондықтан

зияткерлік деректерді талдау әдістерін қолдану қажет. Жұмыста заманауи технологиялар біріктірілген: киберқауіпсіздікке бейімделген алдын ала оқытылған тілдік модель – Security Bidirectional Encoder Representations from Transformers (SecBERT) және динамикалық дамитын нейро-анық емес шығару жүйесі – Dynamic Evolving Neural-Fuzzy Inference System (DENFIS). Ұсынылған жүйе маңызсыз хабарламаларды сүзгіден өткізуге, қауіп туралы индикаторларды және ықтимал осалдықтарды анықтауға мүмкіндік береді. Анық емес логиканы қолдану белгісіз және толық емес ақпаратты тиімді өңдеуге жағдай жасайды. Жүргізілген эксперименттер жоғары дәлдікпен жіктеу және тұрақты шынжырлы кластерлеу нәтижелерін көрсетті (FPC = 0.93; PE = 0.28; XB = 0.042). Жүйе киберқауіптердің белгілерін уақтылы анықтай алатынын көрсетті және мониторинг пен шабуылдарды болжау міндеттері үшін ауқымдауға мүмкіндік бар екенін дәлелдеді. Зерттеу нәтижелері бұл тәсілдің киберқауіптерді уақытылы анықтауға және ақпараттық жүйелердің қауіпсіздігін арттыруға ықпал ете алатынын көрсетеді.

Тірек сөздер: ақпараттық қауіпсіздіктің осал тұстары, тақырыптық интернет-ресурстар, трансформерлер, SecBERT, нейро-анық емес логика.

[1]*Samuilova A.,
Master's student, ORCID ID: 0009-0003-5806-8542,
*e-mail: anastassyiasamuilova@gmail.com

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan

TOOLS FOR IDENTIFYING INFORMATION SECURITY VULNERABILITIES BASED ON DATA FROM INTERNET RESOURCES

Abstract

As cyber threats become more complex, traditional vulnerability detection methods lose their effectiveness. The purpose of this work is to develop and test an approach to identifying vulnerabilities based on the analysis of data from thematic Internet resources: forums, blogs and social networks. These sources contain a large amount of unstructured information, which requires the use of data mining methods. The work uses the integration of modern technologies: the pre-trained SecBERT language model (Security Bidirectional Encoder Representations from Transformers), designed for cybersecurity tasks, and the adaptive neuro-fuzzy inference system DENFIS (Dynamic Evolving Neural-Fuzzy Inference System). The proposed system allows you to filter irrelevant messages, highlight indicators of compromise and potential threats. The use of fuzzy logic makes it possible to efficiently process vague and incomplete information. Experiments confirmed high classification accuracy and stable fuzzy clustering performance (FPC = 0.93; PE = 0.28; XB = 0.042). The system demonstrated the ability to promptly detect signs of cyber threats and has scalability potential for monitoring and attack prediction tasks. The results indicate its potential in increasing the speed of response to cyber threats and strengthening the protection of information systems.

Keywords: information security vulnerabilities, thematic Internet resources, transformers, SecBERT, neuro-fuzzy logic.

Дата поступления статьи в редакцию: 19.05.2025