УДК 004.81

Когнитивно-алгоритмическая система детекции фишинговых атак в социальных сетях

Волокитина Татьяна Сергеевна

Аспирант кафедры информационной безопасности,
Юго-Западный государственный университет,
305040, Российская Федерация, Курск, ул. 50 лет Октября, 94;
e-mail: tativolokitina@gmail.com

Таныгин Максим Олегович

Доктор технических наук,
доцент кафедры информационной безопасности,
Юго-Западный государственный университет,
305040, Российская Федерация, Курск, ул. 50 лет Октября, 94;
e-mail: tativolokitina@gmail.com

Аннотация

Работа посвящена исследованию эффективности систем защиты от фишинга в социальных сетях на основе когнитивного моделирования. Разработана архитектура когнитивно-алгоритмической системы детекции, способной адаптироваться к новым типам угроз. Проведено эмпирическое исследование на базе социальной сети Одноклассники с анализом 800 тысяч записей. Результаты показали, что традиционные методы защиты обеспечивают лишь 40% эффективности блокировки вредоносных URL. Предложенная система повышает точность обнаружения до 80% за счет интеграции эвристических алгоритмов и когнитивного анализа поведенческих паттернов пользователей. Исследование демонстрирует критическую необходимость внедрения проактивных подходов к защите от фишинга в условиях растущих киберугроз.

Для цитирования в научных исследованиях

Волокитина Т.С., Таныгин М.О. Когнитивно-алгоритмическая система детекции фишинговых атак в социальных сетях // Психология. Историко-критические обзоры и современные исследования. 2025. Т. 14. № 7А. С. 97-104.

Ключевые слова

Фишинг, социальные сети, когнитивное моделирование, машинное обучение, кибербезопасность, детекция атак, эвристические алгоритмы, поведенческий анализ.

Введение

В наши дни цифровая среда переживает настоящий взрыв киберугроз, и фишинговые атаки здесь выделяются своим разрушительным потенциалом и сложностью их своевременного обнаружения. Социальные платформы, которые сегодня связывают миллиарды людей по всему земному шару, превратились в излюбленную охотничью территорию киберпреступников, мастерски владеющих искусством социальной инженерии и нацеленных на кражу личной информации и денежных средств пользователей [Антонов., Сидоров, 2023, с. 45].

Цифры красноречиво говорят о том, насколько серьёзна ситуация: за минувшие три года число фишинговых атак в социальных сетях подскочило на 220%, тогда как привычные методы защиты показывают откровенно слабые результаты [Петров, Козлов, 2024, с. 127]. Для России, где социальные платформы играют центральную роль в ежедневном общении людей, эта проблема становится по-настоящему болезненной. Наши исследования выявили тревожную картину: 78% российских пользователей соцсетей сталкивались с фишинговыми попытками, а каждый четвёртый (23%) действительно попался на удочку мошенников [Иванова, Смирнов, 2023, с. 89].

Что делает современные фишинговые атаки особенно опасными, так это их удивительная способность к адаптации и активное применение искусственного интеллекта для создания более правдоподобных и индивидуально настроенных угроз. Киберпреступники всё чаще обращаются к методам социальной инженерии, тщательно изучая привычки и поведение пользователей, чтобы сделать свои атаки максимально эффективными [Волков, 2024, с. 156]. Всё это заставляет нас искать принципиально новые пути защиты, которые основывались бы на когнитивном моделировании и технологиях машинного обучения.

Наше исследование становится актуальным именно потому, что мы остро нуждаемся в разработке упреждающих систем защиты, которые смогли бы противостоять постоянно эволюционирующим угрозам в режиме реального времени. Привычные методы, которые полагаются на чёрные списки и анализ сигнатур, демонстрируют ограниченную результативность против новых видов атак и постоянно требуют освежения баз данных с информацией об угрозах [Николаев, Федоров, 2023, с. 78].

Литературный обзор

Вопросы защиты от фишинга в социальных сетях привлекают пристальное внимание исследователей во всём мире. Основополагающие работы в сфере обнаружения фишинга можно найти в трудах Zhang и его коллег (2020), которые выдвинули гибридную модель, объединяющую анализ структуры URL с семантическим разбором содержимого [Zhang, Wang, Liu, 2020, с. 234]. Этим исследователям удалось показать рост точности обнаружения до 87% в сравнении с общепринятыми методами.

Отечественные учёные тоже внесли весомый вклад в развитие этой научной области. Исследования Иванова П.С. и Петрова А.В. (2021) сосредоточены на создании адаптивных алгоритмов для выявления фишинга в русскоязычной части интернета [Иванов, Петров, 2021, с. 112]. Эти исследователи обнаружили характерные черты российских фишинговых атак, в том числе использование доменов с кириллическими символами и приспособление под наши местные финансовые платформы.

Когнитивное моделирование в области кибербезопасности получило развитие в работах

Kumar et al. (2022), которые предложили архитектуру самообучающейся системы защиты, способной адаптироваться к новым типам угроз без предварительного обучения [Kumar, Patel., Thompson, 2022, с. 78]. Система демонстрировала 92% точность в обнаружении zero-day атак, что существенно превышает показатели традиционных решений.

Особое внимание в литературе уделяется анализу поведенческих паттернов пользователей. Исследование Chen et al. (2023) показало, что интеграция поведенческого анализа с техническими методами детекции повышает эффективность на 34% [Chen, Li, Anderson, 2023, с. 167]. Авторы выделили ключевые индикаторы подозрительного поведения, включая время реакции на уведомления, частоту кликов и паттерны навигации.

Российские работы в области когнитивного моделирования представлены исследованиями Сидорова М.И. и Козлова В.А. (2023), которые разработали нейросетевую модель для предсказания фишинговых атак на основе анализа социальных графов [Сидоров, Козлов, 2023, с. 134]. Модель показала высокую эффективность в предсказании атак за 72 часа до их реализации.

Несмотря на значительные достижения в области детекции фишинга, большинство существующих решений имеют ограничения в контексте социальных сетей. Традиционные подходы не учитывают специфику социальных платформ, включая многообразие форматов контента, высокую скорость распространения информации и сложность анализа контекста коммуникации.

Материалы и методы

Наше исследование проходило на базе социальной сети Одноклассники в течение десяти дней февраля 2025 года - с 1 по 10 число. Мы остановили свой выбор на этой платформе по двум важным причинам: её огромная популярность в нашей стране (80 миллионов активных пользователей) и возможность проводить исследования через официальные программные интерфейсы.

Чтобы собрать нужные данные, мы создали специальное приложение, которое подключили к программному интерфейсу Одноклассников. Это приложение следило за публикациями и комментариями в реальном времени, выуживая ссылки и всю связанную с ними дополнительную информацию. В итоге нам удалось собрать внушительную выборку из 800 тысяч записей - среди них было 500 тысяч публикаций и 300 тысяч комментариев.

Архитектура когнитивно-алгоритмической системы детекции включала несколько взаимосвязанных компонентов:

– Модуль предварительной обработки данных - осуществлял нормализацию URL, извлечение признаков и первичную фильтрацию контента.
– Эвристический классификатор - использовал машинное обучение для анализа структурных особенностей URL и контента сообщений.
– Когнитивный анализатор поведения - моделировал поведенческие паттерны пользователей для выявления аномальной активности.
– Система принятия решений - интегрировал результаты различных компонентов для окончательной классификации угроз.

Для обучения классификатора использовался метод градиентного бустинга с оптимизацией гиперпараметров через байесовскую оптимизацию. Обучающая выборка включала 50 тысяч размеченных примеров, равномерно распределенных между классами "фишинг" и "легитимный

контент".

Когнитивный анализатор поведения основывался на теории графов и анализе временных рядов. Для каждого пользователя строилась модель обычного поведения, включающая частоту публикаций, время активности, типы контента и социальные связи. Отклонения от нормальных паттернов служили индикаторами потенциальных угроз.

Валидация результатов проводилась с использованием методов кросс-валидации и независимого тестирования на отложенной выборке. Для оценки эффективности применялись стандартные метрики: точность (precision), полнота (recall), F-мера и AUC-ROC.

Результаты

Эмпирические испытания нашей когнитивно-алгоритмической системы детекции показали её явное превосходство над традиционными способами защиты. Проанализировав 800 тысяч записей, мы смогли выявить 5 тысяч вредоносных ссылок, что составляет 0,625% от всего проверенного контента.

Сравнительный анализ того, насколько хорошо работают разные подходы, дал нам следующую картину:

– Традиционные системы защиты Одноклассников:
– Общая эффективность блокировки: 40%
– Заблокировано на этапе публикации: 2 тысячи URL
– Пропущено угроз: 3 тысячи URL (60%)
– Предложенная когнитивно-алгоритмическая система:
– Общая эффективность детекции: 80%
– Обнаружено угроз: 4 тысячи URL
– Ложноположительные срабатывания: 2%

Детальный анализ по типам угроз выявил различную эффективность системы в зависимости от характера атак:

– Новые угрозы (не включенные в черные списки на момент публикации) составили 60% от общего числа вредоносных URL. Традиционные методы обнаруживали лишь 20% таких угроз после обновления баз данных (задержка 12-14 часов), в то время как когнитивная система идентифицировала 78% новых атак в режиме реального времени.
– Локальные атаки на российские сервисы (Госуслуги, банки) составили 15% от общего числа угроз. Эффективность обнаружения составила 85% благодаря специализированным алгоритмам анализа доменных имен и контента на русском языке.
– Цепочки редиректов через системы сокращения URL показали 73% эффективности детекции. Система успешно анализировала многоуровневые перенаправления, что традиционные методы не поддерживали.
– Временной анализ показал критическую важность оперативного реагирования. Исследование показало, что 70% пользователей переходят по ссылкам в первые 12 часов после публикации. Задержки традиционных систем защиты (12-14 часов) позволяли злоумышленникам компрометировать значительное количество пользователей до блокировки угрозы.

Когнитивный анализ поведения пользователей выявил характерные паттерны, связанные с фишинговыми атаками:

– Аномальная активность аккаунтов - внезапное увеличение частоты публикаций в 3-5 раз
– Изменение стиля коммуникации - использование нехарактерной лексики и грамматических конструкций
– Временные аномалии - активность в нетипичное для пользователя время
– Сетевые аномалии - резкое изменение круга контактов и получателей сообщений
– Интеграция поведенческого анализа с техническими методами детекции позволила снизить количество ложноположительных срабатываний с 8% до 2%, что критически важно для пользовательского опыта.

Обсуждение

Полученные нами результаты убедительно подтверждают нашу изначальную гипотезу о том, что когнитивно-алгоритмические системы действительно значительно превосходят привычные методы защиты от фишинга в социальных сетях. Двукратный рост эффективности детекции (с 40% до 80%) ясно показывает, насколько перспективным является этот подход для реального использования.

Особенно важной мы считаем способность нашей системы распознавать новые виды угроз, даже не имея заранее подготовленных примеров для обучения. Такая возможность появляется благодаря анализу структурных особенностей URL-адресов, смысловой нагрузки контента и поведенческих шаблонов, что даёт нам шанс замечать аномалии даже при столкновении с совершенно новыми типами атак [Морозов., Лебедев, 2024, с. 89].

Интеграция когнитивного моделирования с традиционными методами машинного обучения создает синергетический эффект, обеспечивая как высокую точность детекции, так и адаптивность к эволюционирующим угрозам. Это особенно важно в контексте современных тенденций развития киберугроз, где злоумышленники активно используют искусственный интеллект для создания более сложных атак [Григорьев., Попов, 2023, с. 134].

Результаты исследования имеют важные практические импликации для разработчиков систем безопасности социальных сетей. Внедрение когнитивно-алгоритмических подходов может значительно повысить уровень защиты пользователей и снизить экономические потери от фишинговых атак.

Однако необходимо учитывать ограничения предложенного подхода. Высокая вычислительная сложность когнитивного анализа может создавать проблемы масштабирования для крупных социальных платформ. Требуется дальнейшая оптимизация алгоритмов и разработка эффективных методов распределенных вычислений [Семенов, Кузнецов, 2024, с. 78].

Заключение

Проведённое нами исследование наглядно показывает, насколько эффективными могут быть когнитивно-алгоритмические системы в борьбе с фишинговыми атаками в социальных сетях. Созданная нами архитектура продемонстрировала 80% эффективность в обнаружении угроз, что в два раза лучше показателей привычных методов защиты.

Главные достижения нашей работы можно сформулировать так:

– Мы создали адаптивную систему детекции, которая умеет распознавать новые типы угроз прямо в процессе работы.

– Нам удалось объединить поведенческий анализ с техническими методами детекции.

– Мы сумели снизить количество ложных срабатываний до всего лишь 2%.

– Мы доказали, как критически важно быстро реагировать на появляющиеся угрозы.

Результаты нашего исследования открывают новые горизонты для развития систем кибербезопасности в социальных сетях. В дальнейшем мы планируем оптимизировать алгоритмы для работы с крупномасштабными платформами, изучить возможность применения нашего подхода к другим типам киберугроз и разработать методы коллективной защиты, основанные на обмене информацией между различными социальными платформами.

Библиография

1. Антонов В.В., Сидоров П.А. Современные тенденции развития фишинговых атак в социальных сетях // Вопросы кибербезопасности. 2023. № 2. С. 43-52.
2. Петров И.И., Козлов А.В. Статистический анализ киберугроз в российском сегменте интернета // Информационная безопасность. 2024. Т. 15. № 3. С. 125-138.
3. Иванова М.С., Смирнов Д.В. Социальная инженерия в контексте российских социальных сетей // Безопасность информационных технологий. 2023. № 4. С. 87-95.
4. Волков Е.А. Искусственный интеллект в современных киберугрозах // Защита информации. 2024. № 1. С. 154-162.
5. Николаев К.П., Федоров В.С. Ограничения традиционных методов защиты от фишинга // Информационная безопасность и защита персональных данных. 2023. № 5. С. 76-84.
6. Zhang L., Wang H., Liu M. Hybrid phishing detection model combining URL analysis and content semantics // Computer Security Journal. 2020. Vol. 28. No. 4. P. 232-245.
7. Иванов П.С., Петров А.В. Адаптивные алгоритмы детекции фишинга в русскоязычном интернете // Прикладная информатика. 2021. Т. 16. № 6. С. 110-125.
8. Kumar S., Patel R., Thompson J. Cognitive modeling for adaptive cybersecurity systems // International Journal of Cyber Security. 2022. Vol. 45. No. 2. P. 76-89.
9. Chen Y., Li X., Anderson M. Behavioral pattern analysis for phishing detection // ACM Computing Surveys. 2023. Vol. 55. No. 8. P. 165-178.
10. Сидоров М.И., Козлов В.А. Нейросетевые модели предсказания фишинговых атак // Нейрокомпьютеры: разработка и применение. 2023. № 7. С. 132-147.
11. Морозов А.Б., Лебедев С.К. Структурный анализ URL для детекции фишинга // Программные системы и вычислительные методы. 2024. № 2. С. 87-96.
12. Григорьев Д.М., Попов Н.В. Применение искусственного интеллекта в кибератаках // Информационная безопасность регионов. 2023. № 3. С. 132-140.
13. Семенов О.Л., Кузнецов Р.И. Масштабируемость систем защиты от фишинга // Вестник компьютерных и информационных технологий. 2024. № 1. С. 76-85.

Cognitive-Algorithmic System for Phishing Attack Detection in Social Networks

Tat'yana S. Volokitina

Graduate Student,
Department of Information Security,
Southwest State University,
305040, 94 50 Let Oktyabrya str., Kursk, Russian Federation;
e-mail: tativolokitina@gmail.com

Maksim O. Tanygin

Doctor of Technical Sciences,
Associate Professor,
Department of Information Security,
Southwest State University,
305040, 94 50 Let Oktyabrya str., Kursk, Russian Federation;
e-mail: tativolokitina@gmail.com

Abstract

The study investigates the effectiveness of phishing protection systems in social networks based on cognitive modeling. An architecture of a cognitive-algorithmic detection system capable of adapting to new types of threats has been developed. An empirical study was conducted on the Odnoklassniki social network, analyzing 800,000 entries. The results showed that traditional protection methods provide only 40% effectiveness in blocking malicious URLs. The proposed system increases detection accuracy to 80% through the integration of heuristic algorithms and cognitive analysis of user behavioral patterns. The research demonstrates the critical need for implementing proactive approaches to phishing protection in the context of growing cyber threats.

For citation

Volokitina T.S., Tanygin M.O. (2025) Kognitivno-algoritmicheskaya sistema detektsii fishingovykh atak v sotsial'nykh setyakh [Cognitive-Algorithmic System for Phishing Attack Detection in Social Networks]. *Psikhologiya. Istoriko-kriticheskie obzory i sovremennye issledovaniya* [Psychology. Historical-critical Reviews and Current Researches], 14 (7A), pp. 97-104.

Keywords

Phishing, social networks, cognitive modeling, machine learning, cybersecurity, attack detection, heuristic algorithms, behavioral analysis.

Referents

1. Antonov V.V., Sidorov P.A. Modern trends in phishing attacks development in social networks // Issues of Cybersecurity. 2023. No. 2. P. 43-52.
2. Petrov I.I., Kozlov A.V. Statistical analysis of cyber threats in Russian internet segment // Information Security. 2024. Vol. 15. No. 3. P. 125-138.
3. Ivanova M.S., Smirnov D.V. Social engineering in the context of Russian social networks // Information Technology Security. 2023. No. 4. P. 87-95.
4. Volkov E.A. Artificial intelligence in modern cyber threats // Information Protection. 2024. No. 1. P. 154-162.
5. Nikolaev K.P., Fedorov V.S. Limitations of traditional phishing protection methods // Information Security and Personal Data Protection. 2023. No. 5. P. 76-84.
6. Zhang L., Wang H., Liu M. Hybrid phishing detection model combining URL analysis and content semantics // Computer Security Journal. 2020. Vol. 28. No. 4. P. 232-245.
7. Ivanov P.S., Petrov A.V. Adaptive phishing detection algorithms in Russian-language internet // Applied Informatics. 2021. Vol. 16. No. 6. P. 110-125.
8. Kumar S., Patel R., Thompson J. Cognitive modeling for adaptive cybersecurity systems // International Journal of Cyber Security. 2022. Vol. 45. No. 2. P. 76-89.
9. Chen Y., Li X., Anderson M. Behavioral pattern analysis for phishing detection // ACM Computing Surveys. 2023. Vol. 55. No. 8. P. 165-178.
10. Sidorov M.I., Kozlov V.A. Neural network models for phishing attacks prediction // Neurocomputers: Development

and Application. 2023. No. 7. P. 132-147.
11. Morozov A.B., Lebedev S.K. Structural URL analysis for phishing detection // Software Systems and Computational Methods. 2024. No. 2. P. 87-96.
12. Grigoriev D.M., Popov N.V. Application of artificial intelligence in cyber attacks // Regional Information Security. 2023. No. 3. P. 132-140.
13. Semenov O.L., Kuznetsov R.I. Scalability of phishing protection systems // Bulletin of Computer and Information Technologies. 2024. No. 1. P. 76-85.