А. О. ИСХАКОВА
ИПУ РАН, Москва
А. В. ЦАРЕГОРОДЦЕВ
МГЛУ, Москва

АНАЛИЗ ПРОЯВЛЕНИЙ КИБЕРУГРОЗ В СОВРЕМЕННОЙ ЦИФРОВОЙ СРЕДЕ И ПОДХОДОВ К ИХ НЕЙТРАЛИЗАЦИИ

Доклад посвящен анализу киберугроз, связанных с воздействием на социальные процессы в Интернете. В результате проведенного анализа были выработаны методологические основы детектирования проявлений киберугроз в цифровой среде. При этом, были учтены не только текстовые формы представления информации, но и различные формы мультимедиа. Доклад является частью исследования по разработке поливекторной методики обнаружения, локализации и нейтрализации мультимодальных проявлений киберугроз в современной цифровой среде.

Введение. Понятие киберугроз в цифровой среде охватывает широкий пласт условий и факторов, представляющих опасность нарушения информационной безопасности виртуального пространства. Такие угрозы могут быть направлены на достижение материального обогащения, а также политических, социальных или иных целей [1]. Стремительное развитие цифровой среды, наблюдаемое обществом в последние 20 лет, повсеместная информатизация всех сфер нашей жизни обусловливает необходимость постоянного совершенствования не только аппаратных средств и программных решений для обеспечения стабильной работы вычислительной инфраструктуры, но и вызывает серьезную озабоченность, связанную с отставанием существующих методов, средств и научно-технических решений в области защиты информации [2].

Предлагаемый доклад посвящен сегменту киберугроз, связанных с воздействием на социальные процессы в Интернете. Была проведена систематизация киберугроз посредством ретроспективного анализа ландшафта за последние 10 лет, изучения основных предпосылок и векторов проявления APT-атак, а также проведен обзор основных трендов рынка кибербезопасности и защиты информации для ближайшей перспективы.

В результате проведенного анализа были выработаны методологические основы детектирования проявлений киберугроз в цифровой среде (интернет-контенте) с помощью современных методов машинного обучения. При этом, были учтены не только текстовые формы представления информации, но и различные формы мультимедиа**.**

Киберугрозы в цифровой среде. Задачей исследования в целом является формирование поливекторной методики обнаружения, локализации и нейтрализации мультимодальных проявлений киберугроз в современной цифровой среде. Задачей настоящего доклада является систематизация и анализ актуальных киберугроз и трендов кибербезопасности для дальнейшего формирования методического и программного комплекса противодействия.

Распределенные атаки типа «Отказ в обслуживании». Мощность DDoS-атак, которой подвергаются российские компании в 2022 году, выросла в несколько раз. Взрывные показатели эксперты связывают с деятельностью так называемых хактивистов, которые вовлекают в атаки пользователей сети в угоду политическим или финансовым интересам. Одним из основных площадок для «хактивизма» является мессенджер Telegram, где с помощью чатов и каналов распространяются инструкции по организации DDoS-атак. С их помощью любой пользователь может стать участником киберинцидента. Например, 26 февраля 2022 г. с мощнейшей атакой в своей истории столкнулся портал «Госуслуги»; 2 и 3 марта из-за DDoS-атак некоторое время часть пользователей не могли загрузить страницы «Газеты.Ru», «Известий», РБК. Эти примеры характеризуют уязвимость крупных информационных систем, однако более незащищенные ресурсы сталкиваются с подобными действиями злоумышленников практически ежедневно.

Атаки на системы искусственного интеллекта. Тематика обеспечения безопасности технологии искусственного интеллекта стала как никогда актуальной из-за первых, официально

подтвержденных киберпреступлений, связанных с технологией Deep Fake. С помощью технологии Deep Fake, позволяющей придать некоему компьютерному «фантому» образ конкретного человека. Имитация человеческих действий может быть использована для усиления атак социальной инженерии [3]. Следующий шаг – создание голосовых и видеоприложений, которые могут пройти биометрический анализ, компрометируя системы аутентификации, таких как голосовые отпечатки или распознавание лиц.

Как известно, небольшие целенаправленные изменения в изображении способны «сломать» систему машинного обучения, так что она ошибочно распознает недостоверные объекты. Такие «троянские» картинки называются «состязательными примерами» (adversarial examples) и представляют собой одно из известных ограничений глубинного обучения.

Самый яркий случай Deep Fake – обман государственной системы Китая, которая принимала налоговые документы, подтвержденные фальшивой биометрией. По некоторым данным, злоумышленники подделывали личные данные жертв с 2018 года. Если раньше для реалистичной подмены лица и мимики на видео требовалось использовать большие вычислительные мощности, то сегодня практически на любой смартфон можно установить подобные приложения. Таким образом, развитие этой технологии может стать хорошим подспорьем для злоумышленников.

Критерии методики выявления киберугроз. Для формирования поливекторной методики обнаружения, локализации и нейтрализации мультимодальных проявлений киберугроз в современной цифровой среде были сформированы основные критерии, а также выделены основные используемые подходы на основе текущего состояния научно-технического прогресса. Формализация задачи определения киберугроз в цифровой среде в рамках исследования сводится к изучению контента различных классов, определению признакового пространства для классификации, выбору вычислительного аппарата и разработке необходимого методического и алгоритмического обеспечения для автоматизации процесса.

Для формирования признакового пространства предлагается исследовать следующие свойства материалов (в формате текстового, графического, аудио- и видеоматериалов):

- рубрикация материалов и анализ соответствия рубрики более ранним материалам автора или ресурса;
- анализ наличия ключевых слов, выделенных для искомых материалов, частоты их упоминания, соответствия классу;
- оценка проявления эмоциональных компонент в материале, в том числе резко негативных, резко позитивных, агрессивных, оскорбительных и т. д.;
- выявление массово порожденных текстов, сгенерированных автоматически с применением специальных алгоритмов и программных средств;
- анализ параметров медиаматериалов – звукового ряда, видеоряда, изображений – для определения наличия встроенных шумов, использования неблагоприятных для человека шумов, неприятных для глаза изображений и пр.

Заключение. В настоящем докладе приведены основные дестабилизирующие факторы и угрозы информационных систем с примерами их реализаций. На основании проведенного анализа сформированы критерии планируемой поливекторной методики обнаружения, локализации и нейтрализации мультимодальных проявлений киберугроз в современной цифровой среде. Создание такой методики подразумевает под собой разработку инструмента анализа разнородного медиаконтента с целью выявления различных форм злоумышленных действий. Предполагается, что анализ текстовой информации, графического, а также аудио- и видеоматериалов с применением специальных алгоритмов позволит выявлять и, соответственно, устранять действия злоумышленников в сети Интернет.

Исследование выполнено в рамках гранта Президента Российской Федерации для государственной поддержки молодых российских ученых – кандидатов наук (проект МК-3172.2021.1.6)

ЛИТЕРАТУРА

1. **Шкодинский С.В., Дудин М.Н., Усманов Д.И.** Анализ и оценка киберугроз национальной финансовой системе россии в цифровой экономике. *Финансовый журнал*. 2021. Т. 13. № 3. С. 38-53.
2. **Жукавин В.А.** Анализ средств обеспечения защиты автоматизированных систем управления от киберугроз. Студенческая наука для развития информационного общества. Сборник материалов XI Всероссийской научно-технической конференции в онлайн формате (посвящается светлой памяти профессора Николая Ивановича Червякова). Ставрополь, 2020. С. 452-457.
3. **Фомичева В.А.** Статистический анализ актуальных киберугроз в период мировой пандемии. Решетневские чтения. Материалы XXV Международной научно-практической конференции, посвященной памяти генерального конструктора ракетно-космических систем академика М.Ф. Решетнева. В 2-х частях. Под общей редакцией Ю.Ю. Логинова. Красноярск, 2021. С. 462-463.

A.O.Iskhakova (ICS RAS, Moscow), A.V.Tsaregorodtsev (MSLU, Moscow)

Analysis of manifestations of cyber threats in today's digital environment and approaches to neutralizing them

The report is devoted to the analysis of cyber threats related to the impact on social processes on the Internet. As a result of the analysis we developed a methodological framework for the detection of cyber threats in the digital environment. At the same time, not only textual forms of information presentation, but also various forms of multimedia were taken into account. The report is a part of research on development of polyvector methodology of detection, localization and neutralization of multimodal manifestations of cyberthreats in modern digital environment.