DSPA:

Вопросы применения цифровой обработки сигналов

№1

2022

СОДЕРЖАНИЕ

АНАЛИЗ СТОИМОСТИ СОЗДАНИЯ СОВРЕМЕННЫХ БАЗ ДАННЫХ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ С СУБЪЕКТИВНОЙ ОЦЕНКОЙ КАЧЕСТВА

Власюк Игорь Викторович,
Московский Технический Университет Связи и Информатики, Москва, Россия
i.v.vlasiuk@mtuci.ru

Пашковская Алина Рустамовна,
Московский Технический Университет Связи и Информатики, Москва, Россия
alina.pashckovsckaja@yandex.ru

Мясникова Василиса Сергеевна,
Московский Технический Университет Связи и Информатики, Москва, Россия
mvs01@bk.ru

Никольская Дарина Игоревна,
Московский Технический Университет Связи и Информатики, Москва, Россия
darinanikolskaa@icloud.com

Можаева Анастасия Ивановна,
Университет Вайкато, Новая Зеландия;
Московский Технический Университет Связи и Информатики, Москва, Россия
am476@students.waikato.ac.nz

Аннотация

Популярность потокового видео значительно выросла за последние несколько лет. Базы данных видео с субъективными оценками составляют важную основу для обучения метрик качества видео и кодеков, на основе алгоритмов машинного обучения, которые в свою очередь непосредственно влияют на потоковую передачу видеопоследовательностей. В данной работе представлен анализ современных часто используемых баз данных видео с субъективным качеством и сделаны выводы о методах их улучшения. На данном этапе развития субъективные оценки являются наиболее сложной частью создания базы данных видеопоследовательностей, так как эти оценки являются дорогостоящими. Кроме того, субъективные оценки осложняются многими факторами, включая расстояние просмотра, устройство отображения, условия освещения, зрение и настроение испытуемых. В данной работе представлен анализ основных составляющих баз данных видеопоследовательностей с субъективным качеством и их влияние на стоимость создание баз данных. Представленная информация в данной работе позволит исследователям иметь более детальное представление о базах данных видео, а также может помочь в планировании будущих экспериментов.

Ключевые слова: *оценка качества видео, субъективное тестирование, база данных видео, набор данных видео, средний балл мнения (MOS).*

I. Введение

Потоковое видео продолжает занимать все большую долю пропускной способности Интернета, и ожидается, что в 2022 году на потоковое видео будет приходиться 82% интернет-трафика [1]. В условиях стремительного роста видеотрафика совершенствование технологий кодирования видео имеет решающее значение для компаний, занимающихся потоковой передачей видео. На современном этапе развития технологий системы кодирования видео показывают высококачественные и полностью удовлетворительные результаты. Для решения проблем качества потокового видео требуются такие решения, как создание оценок качества видео и кодеков с использованием полного или частичного машинного обучения [2]. Однако создание оценок качества видео и кодеков с использованием полного

или частичного машинного обучения требует наборов видеоданных, которые точно отражают пользовательский опыт человека.

Отсутствие аннотированных баз данных раньше было серьезным препятствием для исследователей, работающих над алгоритмами оценки качества. Даже несжатый видеоконтент было трудно найти [3]. В настоящее время существует более 30 общедоступных баз данных и множество небольших наборов данных, которые используются экспертами для личных исследований. Однако возникает новая проблема, которая отсутствовала ранее и не была подробно рассмотрена научным сообществом в настоящее время. Это проблема выбора наиболее подходящих баз данных для исследования, так как с момента последнего детального анализа баз данных видео прошло более 9 лет, что является критическим аспектом в условиях современного роста информационных технологий. В данной работе мы анализируем 7 наиболее часто используемых баз данных видео для оценки качества кодирования и метрик качества видео на основе машинного обучения, которые точно отражают пользовательский опыт человека. Мы также предлагаем критерии и решение, которые наглядно демонстрируют проблему создания крупномасштабного набора данных видеопоследовательностей.

Сравнение баз данных с использованием одних и тех же критериев полезно для разработчиков моделей, которые могут принять более обоснованное решение о том, какие базы данных могут быть наиболее подходящими для их конкретных сравнительных или других потребностей [3,4]. Кроме того, представленные критерии оценки баз данных видео объясняют проблему с небольшим разнообразием контента, что приводит к ограничениям в разработке и оценке метрик и кодеков, эффективно использующих полное или частичное машинное обучение. Данная работа позволит исследователям получить более детальное представление о базах данных видео, а также может помочь в планировании будущих экспериментов.

Работа организована следующим образом. В разделе II представлен обзор часто используемых современных баз данных видео, с аннотированными субъективными оценками качества. В разделе III предлагаются новый критерий оценки субъективных оценок, которые затем используются для сравнения баз данных. В разделе IV рассматриваются результаты анализа и обсуждаются методы улучшения базы данных и будущей работы.

II. Базы данных видеоматериалов

Здесь мы представляем 7 баз данных видео с субъективной оценкой качества, следуя предыдущему исследованию [5], в котором в 2012 году был представлен всесторонний анализ наборов данных видео на тот момент.

- LIVE Video Quality Database. Сжатый MPEG-2, сжатие H.264, имитация передачи сжатых битовых потоков H.264 по проводным и беспроводным сетям IP, подверженным ошибкам. [6], [7].

- Konstanz Natural Video Database (KoNViD-1k) (2017) 1200 видеороликов с субъективными данными и оценкой атрибутов. [8]

- LIVE YouTube High Frame Rate (LIVE-YT-HFR) Database (2020). Видео обрабатываются с 5 уровнями сжатия при каждой частоте кадров. [9, 10]

- LIVE Wild Compressed Video Quality Database (2020). Видео, снятые на самые разные мобильные камеры, охватывающие широкий диапазон содержания и качества. Большинство из этих видео при захвате искажены с различными достоверными смешанными искажениями. Формат сжатия видео H.264. [11]

- LIVE Netflix Video Quality of Experience Database (2017). 112 видеороликов типичных артефактов адаптивного потокового вещания, оцененных 55+ людьми на мобильном устройстве, Рис. 1. [12,13]

- MCL-JCV Database (2016). Сжатый H.264 / AVC при коэффициентах качества в диапазоне от 1 до 51. [14]

- VideoSet (2017). База данных включает 3520 последовательностей, которые были оценены 800 участниками. [15]

Рис. 1. LIVE Netflix Video Quality of Experience Database

LIVE [6, 7], KoNViD-1k [8], LIVE-YT-HFR [9, 10], LIVE Wild [11] и LIVE-NFLX [12, 13] приняли протокол непрерывной оценки качества одного стимула. Учитывая этот принцип в LIVE Wild [11], видео проигрывались в случайном порядке, при этом каждое видео показывалось только один раз в течение каждой сессии, и по крайней мере 5 видео были разделены между различными искаженными версиями каждого уникального контента. Оригинальные видео были включены в качестве ссылок. Диапазон качества обозначался от низкого до высокого двумя прилагательными: "Плохое" и "Отличное". Субъективные оценки, полученные от испытуемых, были преобразованы в числовые оценки качества в диапазоне [0, 100].

Другой способ используется при создании баз данных MCL-JCV [14] и VideoSet [15], каждая из которых содержит, по утверждению авторов, большое количество обработанных видеорядов с различными уровнями обработки. Однако из 30 клипов, заявленных в [14], только 24 последовательности были выпущены из-за проблемы с интеллектуальной собственностью, а длина клипов составляет всего 5 секунд. Стоит также отметить, что искажения во всех этих базах данных были созданы изолированно синтетическим способом. Испытуемые сравнивали качество двух последовательностей, показанных друг за другом, и определяли, существенно ли они отличаются, выбирая "да" или "нет". То есть в [14] и [15] участник, просмотрев видеопакет, устанавливал пороги изменения сжатия для одной видеопоследовательности, обработав 51 вариант артефактов, понятно, что участник смотрел не все 51 вариант, так как в [12] исключалась вся левая или правая половина интервала, а в [15] отбрасывается только четверть исходного бина в самом удаленном от области интереса месте. В [15] участнику потребовалось 35 минут для оценки 14-15 видеопоследовательностей по 5 секунд каждая, чтобы установить 1-й пороговый уровень (51 уровень) для 14-15 видео. Следовательно, участник мог обрабатывать примерно 4-5 последовательностей по 5 секунд в минуту. Данный подход показывает самые высокие результаты среди существующих баз данных с открытым доступом для оценки количества последовательностей людьми за определенный период времени, однако, тренируясь на выходе, предоставляет очень мало видеоданных (всего 5 часов). Создатели базы данных не указывают, сколько оценок приходится на одну последовательность, однако, даже при минимальном количестве в 10 оценок, хотя в работе MCL-JCV их было 50, отношение времени, затраченного на субъективные тесты (время людей), к итоговому общему времени базы данных значительно уступает первому подходу. Также остается вопрос, точно ли определен порог минимально приемлемого качества для человека из трех точек, найденных в этой базе данных, после которого качество перестает устраивать участника. Не сделает ли такой подход к сбору субъективных оценок эту базу склонной к высокому качеству?

III. Анализ

Существует множество критериев, которые можно использовать для оценки и сравнения баз данных [3]. Оценка баз данных основана на трех компонентах: количественные сравнения исходного контента, условий тестирования и субъективных оценок [16]. Сегодня существует достаточно решений для проблемы исходного контента и условий тестирования [17]. Однако, по субъективным оценкам, все еще нет оптимального решения. Субъективные оценки являются наиболее ценной и, возможно, самой сложной частью при создании наборов данных видеопоследовательностей.

Субъективные оценки дорогостоящие, и требуют значительного времени. Кроме того, субъективные эксперименты осложняются многими факторами, включая расстояние просмотра, дисплей устройство, условия освещения, зрение и настроение испытуемых [4]. Если создание базы происходит в разные дни, или, что еще более проблематично, эксперименты проводятся параллельно в нескольких лабораториях, необходимо строго соблюдать одинаковые условия для всех участников, результаты которых затем будут оцениваться совместно. Кроме того, субъективные испытания, видео тесты кодирования, обычно проводятся очень небольшим количеством экспертов, которых называют золотым стандартом для анализа наихудшего случая [19]. Однако анализ наихудшего случая не отражает фактического качества для визуального контента, а количество и качество воспринимаемых уровней искажений индивидуально. Существующие базы данных имеют очень разные и часто, почти несравнимые подходы к сбору субъективных данных.

Мы предлагаем рассмотреть здесь оценку стоимости одного артефакта в базе данных видео, которая была представлена ранее как вспомогательная оценка, и которой, по нашему мнению, уделили мало внимания, несмотря на ее значимость [19].

$$C = \left(\left(N_s t_e \right) / 60 + t_p \right) S / N_s \right) k \quad (1)$$

где, C – стоимость одного артефакта в базе данных, Ns – количество артефактов, t_e – время одной последовательности (секунды), t_p – время подготовки, k – коэффициент перевода оценок в последовательность, длиной 10 секунд. S – минимальная заработная плата в час. При расчете также необходимо учесть погрешность, вызванную тем, что некоторые базы данных содержат просмотр артефактов и оригинального видео, другие содержат только видео с артефактами.

Данные с учетом этого критерия для рассматриваемых баз данных приведены в Табл. 1. В столбце "Критерий" мы можем увидеть стоимость одного артефакта, значение – российские рубли. Однако, как видно из таблицы 1 есть выбросы, где стоимость артефактов превышает 1300 рублей. Исходя из приведенной выше информации, мы ограничили ценовой диапазон от 0 до 1239 рублей. Количество баз данных, оптимальное по отношению к стоимости артефакта, уменьшилось с 7 до 6 (рис. 2). База данных KoNViD-1k выходит за наш предел. На рисунках 2-7 представлены зависимости основных факторов, влияющих на создание баз данных видеопоследовательностей с субъективным качеством от критерия стоимости.

Таблица 1

СРАВНЕНИЕ БАЗ ДАННЫХ ПО КРИТЕРИЮ СТОИМОСТИ

Базы данных	**Количество последовательностей**	**Количество артефактов**	**Количество экспертов**	**Время одной последовательности (секунды)**	**Количество экспертов для 1 последовательности**	**Критерий**
LIVE VQD	165	150	38	10	38	842
KoNViD-1k	1200	1200	642	8	114	5753
LIVE-YT-HFR	496	480		8	40	1234
LIVE Wild	275	220		10	40	848
LIVE Netflix	126	112	55	10	44	1239
MCL-JCV	96	1124	120	5	50	1152
VideoSet	3520	44880	800	5		

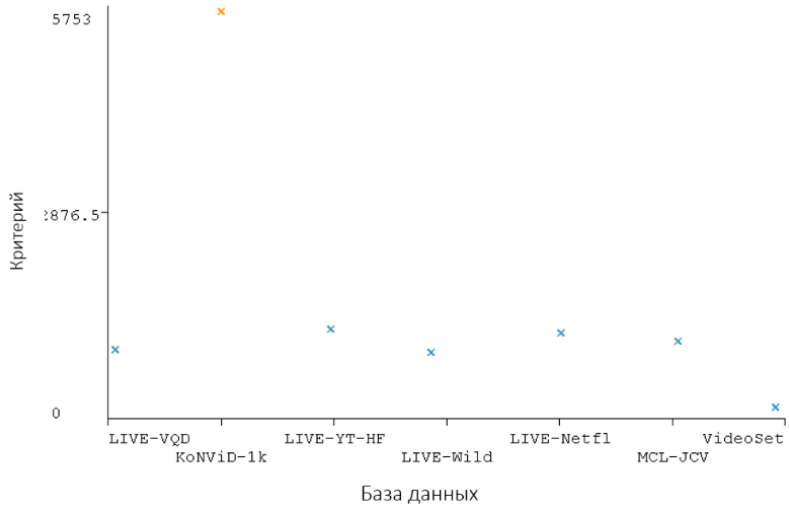

Рис. 2. Распределение критерия для каждой базы данных

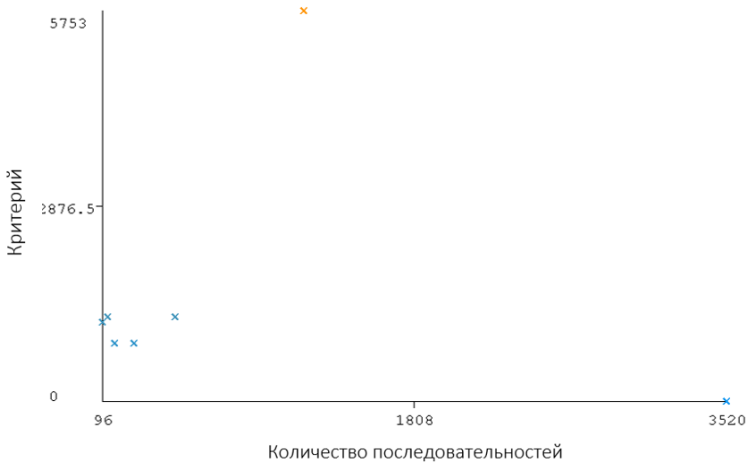

Рис. 3. Зависимость критерия от количества последовательностей

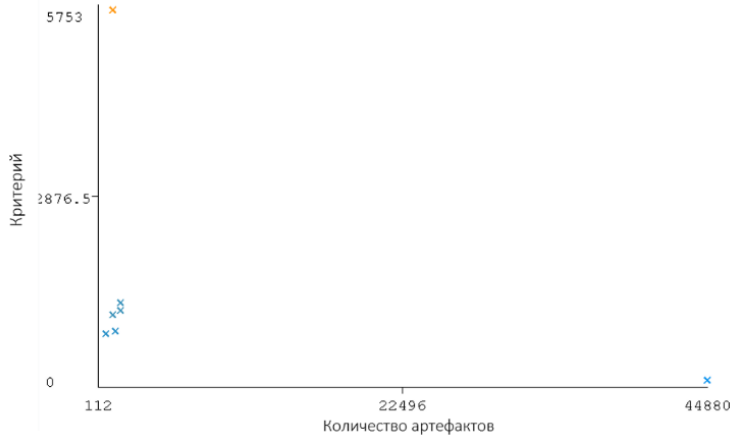

Рис. 4. Зависимость критерия от количества артефактов

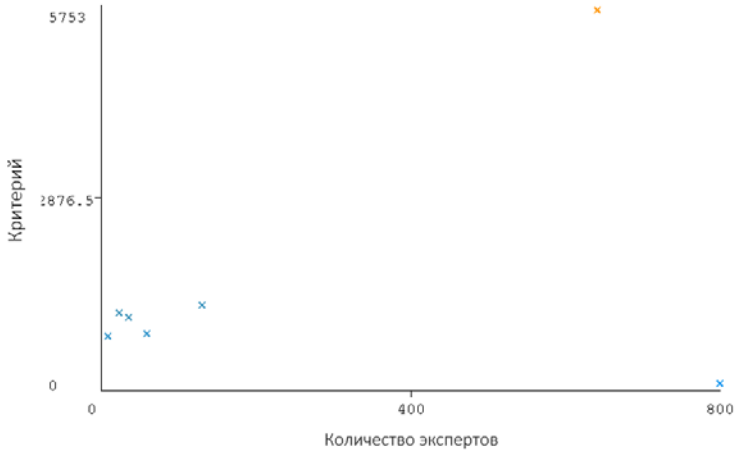

Рис. 5. Зависимость критерия от количества экспертов

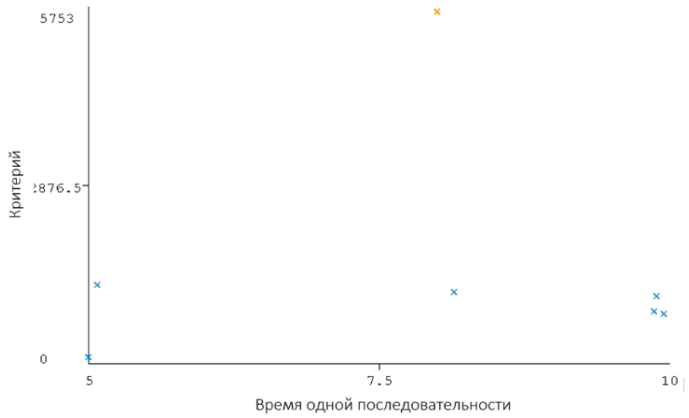

Рис. 6. Зависимость критерия от времени одной последовательности в секундах

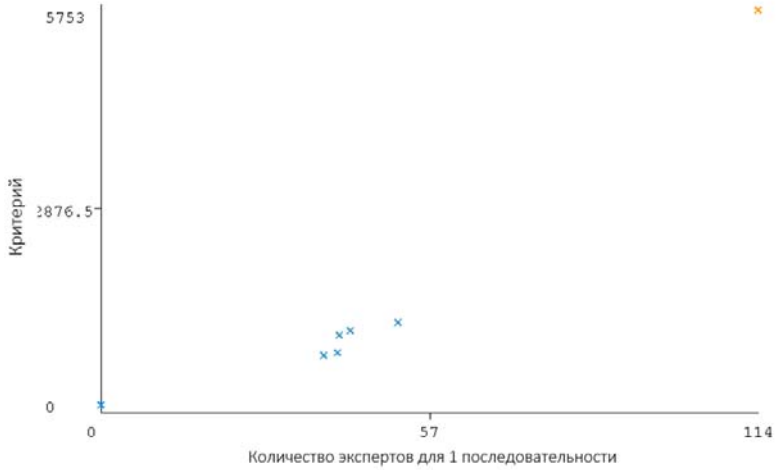

Рис. 7. Зависимость критерия от количествоа экспертов для 1 последовательности

Как видно из представленной выше информации, количество артефактов не сильно влияет на величину критерия. Из Рис. 5. видны 3 исключения с ростом числа экспертов. В остальных случаях критерий увеличивается прямо пропорционально. Так же, время одной последовательности не сказалось на стоимости одного артефакта. С увеличением количества экспертов для 1 последовательности наблюдается рост критерия стоимости, Рис. 7.

IV. Дискуссия

В настоящее время самой большой проблемой при создании метрик и кодеков с помощью машинного обучения является недостаток данных, или, другими словами, крупномасштабности баз данных. Для решения этой проблемы путем привлечения людей к тестированию необходимы новые подходы к сбору субъективных данных, учитывающие максимально возможный объем обработки артефактов при минимальных затратах на субъективные тесты. В этой статье мы предложили анализ влияния основных составляющих баз данных видеопоследовательностей с субъективным качеством на стоимость создание данных баз данных. Как мы видим из таблицы 1 стоимость одного артефакта оказывается весьма значительной с учетом необходимости создание крупномасштабных баз данных.

В нашей предыдущей работе было представлено новое устройство, позволяющее собирать оценки субъективного уровня качества, используя идею нахождения приемлемого минимального уровня качества для участника, или, другими словами, порога восприятия [20]. Данное устройство оптимизирует сбор субъективных оценок для данного этапа развития современных телекоммуникационных технологий и позволяет создавать видеоряд постоянного качества. Иными словами, оно создает условия для создания базы данных с максимальным количеством артефактов и минимумом субъективных экспериментов. Для создания базы данных с помощью данного устройства можно использовать аналогичное время с такими базами данных как LIVE Wild [11], но со

значительно большим количеством артефактов. Также такие базы данных KoNViD-1 [8] будут значительно уступать базам данных, созданным с помощью устройства для сбора субъективных оценок из [20], по затратам времени на субъективные тесты при одинаковом количестве обрабатываемых артефактов. Подход, предложенный в [20], позволит получить видео с постоянной оценкой качества, которое создают сами пользователи и к которому должны стремиться разработчики кодеков как к оптимальному для пользователей.

Учитывая представленный выше анализ баз данных, в настоящий момент в развитии телекоммуникационных технологий [22-30] существуют возможности реализации более оптимальных подходов к сбору субъективных оценок для создания новых баз данных с различными искажениями и хорошо маркированными данными в большем объеме.

V. Заключение

На современном этапе развития технологий существует более трех десятков общедоступных баз данных качества видео, а также большое количество наборов данных для частного тестирования. Это облегчает проверку качества алгоритмов, но все еще недостаточно для создания моделей оценки качества или видеокодеков на основе полного или частичного машинного обучения. В работе представлены и проанализированы современные часто используемые базы данных видео с субъективными оценками качества. Проанализированы основные составляющих баз данных видеопоследовательностей с субъективным качеством и их влияние на стоимость создание баз данных. Также обсуждается метод улучшения создания будущих баз данных видео с субъективными оценками качества.

Литература

1. Cisco Visual Networking Index: Forecast and Methodology 2017-2022, Feb. 2019, [online] Available:
2. https://www.cisco.com/c/en/us/solutions/collateral/serviceprovider/vis ual-networking-index-vni/white-paper-c11741490.html.
3. *S.-F. Chang and A. Vetro*, "Video adaptation: Concepts technologies and open issues", Proc. of the IEEE, vol. 93, no. 1, pp. 148-158, 2005.
4. *A. Mozhaeva, L. Streeter, I. Vlasuyk and A. Potashnikov*, "Full Reference Video Quality Assessment Metric on Base Human Visual System Consistent with PSNR," 2021 28th Conference of Open Innovations Association (FRUCT), 2021, pp. 309-315.
5. *A. I. Mozhaeva, I. V. Vlasuyk, A. M. Potashnikov, M. J. Cree and L. Streeter*, "The Method and Devices for Research the Parameters of the Human Visual System to Video Quality Assessment," 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, pp. 1-5.
6. *S. Winkler*, "Analysis of Public Image and Video Databases for Quality Assessment," in IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 6, pp. 616-625, Oct. 2012.
7. *K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack*, “LIVE video quality database,” 2010 [Online]. Available: http://live. ece.utexas.edu/research/quality/live_video.html
8. *K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack*, “Study of subjective and objective quality assessment of video,” IEEE Trans. Image Process., vol. 19, no. 6, pp. 1427–1441, Jun. 2010.
9. *V. Hosu* et al., "The Konstanz natural video database (KoNViD-1k)," 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), 2017, pp. 1-6, doi: 10.1109/QoMEX.2017.7965673.
10. *P. C. Madhusudana, X. Yu, N. Birkbeck, Y. Wang, B. Adsumilli and A. C. Bovik*, "Subjective and Objective Quality Assessment of High Frame Rate Videos", submitted to IEEE Transactions on Image Processing, 2020 [paper]
11. *P. C. Madhusudana, N. Birkbeck, Y. Wang, B. Adsumilli and A. C. Bovik*, "Capturing Video Frame Rate Variations through Entropic Differencing", arXiv preprint arXiv:2006.11424, 2020
12. *X. Yu, N. Birkbeck, Y. Wang, C. G. Bampis, B. Adsumilli and A. C. Bovik*, "Predicting the Quality of Compressed Videos with Pre-Existing Distortions", submitted to IEEE Transactions on Circuits and Systems for Video Technology. [paper]
13. *C. G. Bampis, Z. Li, A. K. Moorthy, I. Katsavounidis, A. Aaron, and A. C. Bovik*, “Study of Temporal Effects on Subjective Video Quality of Experience,” IEEE Trans. Image Process., vol. 26, no. 11, pp. 5217–5231, 2017.
14. *C. G. Bampis, Z. Li, A. K. Moorthy, I. Katsavounidis, A. Aaron and A. C. Bovik*, "LIVE Netflix Video Quality of Experience Database," Online: http://live.ece.utexas.edu/research/LIVE_NFLXStudy/index.html, 2016.
15. “MCL-JCV Dataset”, [online] Available at: http://mcl.usc.edu/mcl- jcv-dataset/.
16. *H. Wang* et al., "VideoSet: A large-scale compressed video quality dataset based on JND measurement", J. Vis. Commun. Image Represent, vol. 46, pp. 292-302, Jul. 2017.

17. *S. Winkler*, “Image and video quality resources,” 2012 [Online]. Avail- able: http://stefan.winkler.net/resources.html

18. *Y. Wang, S. Inguva and B. Adsumilli*, "YouTube UGC dataset for video compression research", Proc. IEEE 21st Int. Workshop Multimedia Signal Process. (MMSP), pp. 1-5, Sep. 2019.

19. *P. Mohammadu, A. Ebrahimi-Moghadam, S. Shirani*, "Subjective and Objective Quality Assessment of Image: A Survey", Majlesi Journal of Electrical Engineering, vol.9 (1), Mar 2015, pp.55-83.

20. *H. Wang* et al., "MCL-JCV: A JND-based H.264/AVC video quality assessment dataset", Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 1509-1513, Sep. 2016

21. *A. Mozhaeva, A. Potashnikov, I. Vlasuyk and L. Streeter*, "Constant Subjective Quality Database: The Research and Device of Generating Video Sequences of Constant Quality," 2021 International Conference on Engineering Management of Communication and Technology (EMCTECH), 2021, pp. 1-5, doi: 10.1109/EMCTECH53459.2021.9618977.

22. *Valitskaya N.S., Vlasyuk I.V., Potashnikov A.M.* Video compression method on the basis of discrete wavelet transform for application in video information systems with non-standard parameters // T-Comm. 2020. T. 14. № 3. C. 47-53.

23. *Поташников А.М., Власюк И.В*. Метод построения равноконтрастного цветового пространства для заданной системы отображения информации и условий контроля // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 4. С. 15-22.

24. *Власюк И.В., Любецкая В.Ю.* Анализ методов подавления артефактов звона, возникающих на изображениях в процессе кодирования с wavelet-преобразованием // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 4. С. 53-58.

25. *Валицкая Н.С., Власюк И.В*. Методы синхронизации потоков в видеоинформационных системах // Телекоммуникации и информационные технологии. 2019. Т. 6. № 2. С. 51-57.

26. *Валицкая Н.С., Власюк И.В.* Протоколы и стандарты передачи медиаконтента по IP-сетям // Телекоммуникации и информационные технологии. 2020. Т. 7. № 2. С. 5-12.

27. *Егоров Д.А., Федоров В.Д., Лейман В.В., Власюк И.В*. Методика оценки пространственно-частотной характеристики камер на основе генеративных случайных последовательностей // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 47-53.

28. *Иванчев В.В., Калужских Е.А., Власюк И.В.* Разработка локального метода сжатия динамического диапазона // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 2. С. 31-43.

29. *Можаева А.И., Власюк И.В., Поташников А.М., Стритер Ли.* Эталонная объективная метрика оценки качества видео совместимая с PSNR учитывающая частотные и периферическую характеристики зрения человека // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 2. С. 44-54.

30. *Кремлева Э.А., Власюк И.В*. Оценка эффективности методов визуализации одноканальных изображений в условных цветах // Телекоммуникации и информационные технологии. 2021. Т. 8. № 1. С. 29-37.

МЕТОДИКА ОПЕРАТИВНОГО ОБНАРУЖЕНИЯ КОСМИЧЕСКИХ ОБЪЕКТОВ НА ВЫСОКИХ ОРБИТАХ

Гураль Дмитрий Александрович,
Военно-космическая академия имени А.Ф.Можайского, г. Санкт-Петербург, Россия
vka@mil.ru

Аннотация
Оптические исследования космических объектов, в том числе космического мусора на высотных орбитах проводятся уже более десяти лет. Первоначально эти усилия были сосредоточены главным образом на геостационарной орбите (ГСО). Были разработаны и успешно применены соответствующие стратегии наблюдения, методы обработки и подходы к каталогизации. Например, геосъемка привела к обнаружению значительного скопления мелкого мусора, а затем к обнаружению объектов с высоким отношением площади к массе на геостационарных орбитах. Сравнительно меньший опыт (как с точки зрения практического наблюдения, так и определения стратегии) имеется для эксцентрических орбит, которые (по крайней мере, частично) находятся в области средней околоземной орбиты (СОО).

Ключевые слова: *модель, виртуальный прибор, средство наземного наблюдения, прогнозирование движения, космический объект*

В ходе исследования были разработаны различные стратегии обследования и последующих действий для поиска объектов космического мусора на орбитах с высокой эксцентриситетностью в области средней околоземной орбиты (СОО) и для получения орбит, которые являются достаточно точными для каталогизации таких объектов и поддержания их орбит в течение более длительных периодов времени. Моделирование было проведено для сравнения эффективности различных стратегий обследования и каталогизации. В конце концов, оптические наблюдения были проведены в рамках исследования с использованием телескопа космического мусора ЕКА (ESASDT), 1-метрового телескопа Zeiss, расположенного на Оптической наземной станции (OGS).

За шесть месяцев было проведено тринадцать ночей наблюдений орбит типа «Молния». В конечном итоге за эти тринадцать ночей было проведено 255 обследований, что соответствует примерно 47 часам наблюдений. Всего было обнаружено 30 некоррелированных слабых объектов. В среднем один некоррелированный объект обнаруживался каждые 100 минут наблюдений. Некоторые из этих объектов демонстрируют значительное изменение яркости и имеют высокое отношение площади к массе, как определено в процессе оценки орбиты на рисунке 1 [1-5].

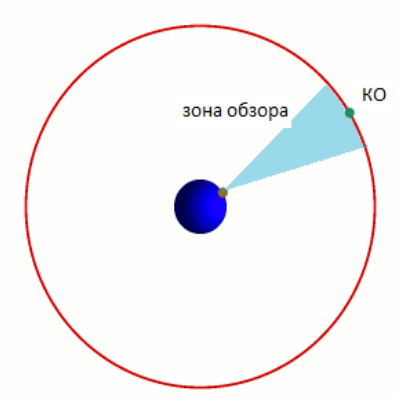

Рис. 1. Орбита КО и конус зоны действа наземного средства наблюдения

Популяция космического мусора в области низкой околоземной орбиты (НОО), которая определяется как область высотой до 2000 километров, была тщательно изучена в течение последних десятилетий, и разумные модели, такие как МАСТЕР-модель ЕКА (Видеманн и др., 2011, Флегель и

др., 2011) и модель НАСА ORDEM (Криско и др., 2015, Сюй и др., 2009), которые охватывают все диапазоны размеров. Признавая первостепенную важность защиты региона геостационарной орбиты (ГСО) от загрязнения космическим мусором, Европейское космическое агентство (ЕКА) в 1999 году инициировало оптический поиск фрагментов на ГСО и ГПО (геосинхронная передаточная орбита), чтобы улучшить знания об их населении мусором и понять будущую эволюцию этих объектов (Шильдкнехт и др., 2004, 2005). Для полноты картины мы должны также упомянуть, что аналогичные геосъемки проводились также другими группами по всему миру, такими как Аберкромби и др. (2009) или Молотов и др. (2008).

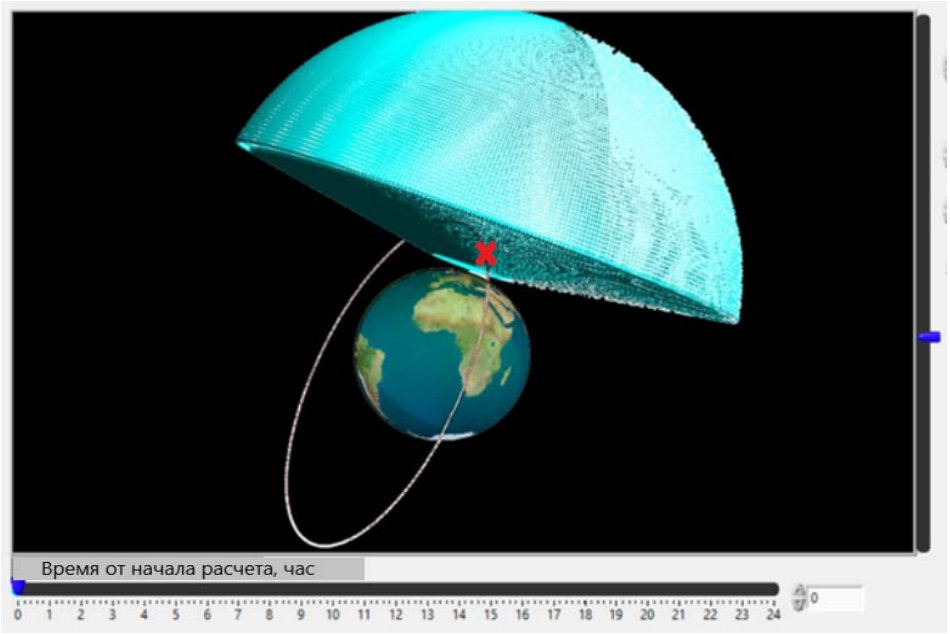

Рис. 2. Визуализация задачи прогнозирования движения КО

Окружающая среда космического мусора в районе средней околоземной орбиты (СОО) до сих пор систематически не исследовалась и, следовательно, в значительной степени неизвестна. Недавно было предпринято расширение продолжающихся исследований космического мусора на новые орбитальные районы, в частности на все более населенный регион СОО (Schildknecht et al., 2012). В последнем исследовании работа была больше сосредоточена на круговых орбитах СОО, включая спутники GPS и ГЛОНАСС. Наблюдение высоко эксцентричных орбит СОО, представленное в этой статье, является дальнейшим продолжением предыдущего исследования. Эти эксцентричные и высокоэллиптические орбиты включают, в частности, орбиты «Молнии» и «Тундры». Орбиты «Молнии» имеют наклон около 63,4°, в основном с аргументом перигея около -90° и периодом обращения в половину звездных суток. С эксцентриситетом около 0,7 и высотой апогея около 40 000 км орбита спроектирована таким образом, что космический аппарат будет проводить значительный период времени вокруг апогея над северным полушарием. Часть населения «Молнии» имеет орбиты с аргументом перигея около 90°. Объекты на этих орбитах проводят большую часть своего времени над южным полушарием. Наклон 63,4 ° выбран таким образом, чтобы свести к минимуму вековое возмущение аргумента перигея, вызванное сжатием Земли (Beutler, 2005). Для сравнения, орбиты «Тундры» имеют тот же наклон, что и орбиты «Молнии», но орбитальный период составляет один звездный день. Основной задачей космических аппаратов на орбитах типа «Молния» в большинстве случаев является обеспечение военной и коммерческой связи.

В работе представлена каталогизированная популяция «Молнии» с ее динамическими характеристиками. Кроме того, эта совокупность была использована для разработки стратегии наблюдения для поиска новых некаталогизированных объектов. Наконец, были представлены результаты тринадцати ночных съемок, проведенных в 2013 году с помощью телескопа ESADT, который является частью Оптической наземной станции (ОНС).

Чтобы выбрать объекты, представляющие интерес при разработке обзора, мы определили объекты «Молния» как объекты с орбитальными элементами, удовлетворяющими следующим критериям. Для

наклона орбиты границы были приняты равными 60° и 67°, для эксцентриситета - 0,5 и 0,8, а для большой полуоси - 20 000 и 30 000 км. Интервал большой полуоси соответствует среднему движению между 3,0 и 1,7 оборотами в день.

В общедоступном каталоге USSTRATCOM (www.space-track.org) был обнаружен в общей сложности 171 неклассифицированный объект «Молния» (далее именуемый выбранной совокупностью TLE). Эта выбранная совокупность TLE включает спутники «Молния» (41), спутники «Меридиан» (4), спутники «Око», корпус ракеты или верхние ступени (73), а также некоторые другие типы мусора (53). Узлы распределены по всему диапазону прямого восхождения с концентрацией от 0° до 210°, а наклоны сосредоточены вокруг номинального значения 63,4° (так называемый критический наклон) (Beutler, 2005). Основываясь на распределении эксцентриситета, можно было выделить подгруппу из выбранной популяции «Молния» с e < 0,65. Эта подгруппа состоит из 35 объектов, включая корпуса ракет (10), обломки (20) и российские спутники Oko (5) с эксцентриситетом менее 0,65. Соотношения между орбитальными элементами выбранной совокупности TLE показаны на рис. 3, где объекты с эксцентриситетом менее 0,65 обозначены крестиками. Объекты с перигеем менее 180° и, следовательно, с апогеем над южным полушарием обозначены пустыми квадратами. Остальная часть населения представлена заполненными кружками. Для наземного датчика такие объекты на сильно эксцентричных орбитах и со средним движением около двух оборотов в сутки показывают в точке своих перигеев относительно высокие угловые скорости относительно звезд. Поэтому было бы трудно наблюдать объекты вокруг этой позиции. Из всех выбранных объектов TLE 55 объектов имеют аргумент перигея от 0° до 180°, что означает, что эти перигеи расположены над северным полушарием [6-13].

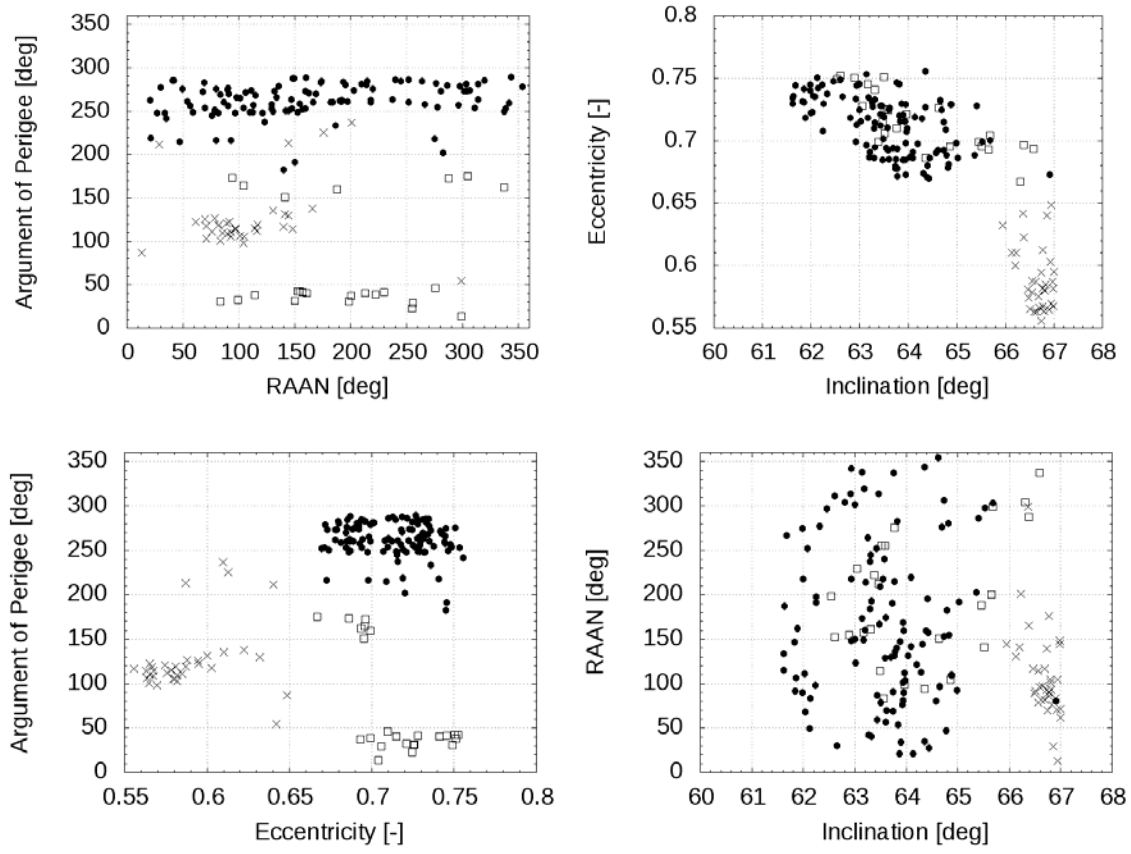

Рис. 3. Соотношение между орбитальными элементами выбранной популяции TLE. Крестики указывают на то, что эти объекты имеют эксцентриситет менее 0,65, а пустые квадраты указывают на то, что эти объекты имеют перигей менее 180°. Заполненные круги представляют остальную часть «Молнии».

На рис. 4 (слева) показаны видимые проходы выбранной популяции TLE в системе восхождения и склонения вправо (RA /DE), как видно из OGS в течение одной ночи. В полосе склонения между 55°

и 65° имеется область с повышенной видимой плотностью. В этой области кульминации большинство объектов, а именно объекты, орбиты которых имеют аргумент перигея около 270° (отмечены на рис. 3 в виде черных точек и крестиков) достигают своего апогея.

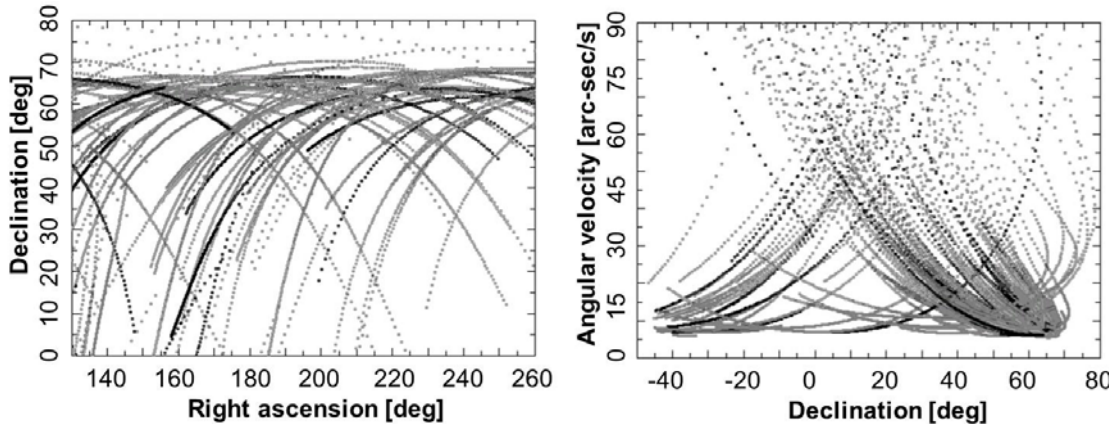

Рис. 4. Слева, видимые проходы выбранной популяции в правильном восхождении и склонении, наблюдаемые с OGS в течение одной ночи. Справа - топоцентрические угловые скорости как функция топоцентрического склонения, видимого с OGS.

Поскольку реализация эффективной стратегии обнаружения требует оптимизации скорости отслеживания и времени интеграции, предпочитают регионы, где изменения кажущейся скорости невелики. На рис. 5 (справа) показана топоцентрическая угловая скорость в системе координат RA/DE в зависимости от склонения, как видно из OGS. Минимальная угловая скорость составляет около 5,5 угловых секунд/с для объектов с наклоном от 60° до 70°. Объекты с перигеем около 90° достигают своей минимальной угловой скорости при склонении от -30° до -40°. К сожалению, это те регионы, которые едва видны из местоположения OGS.

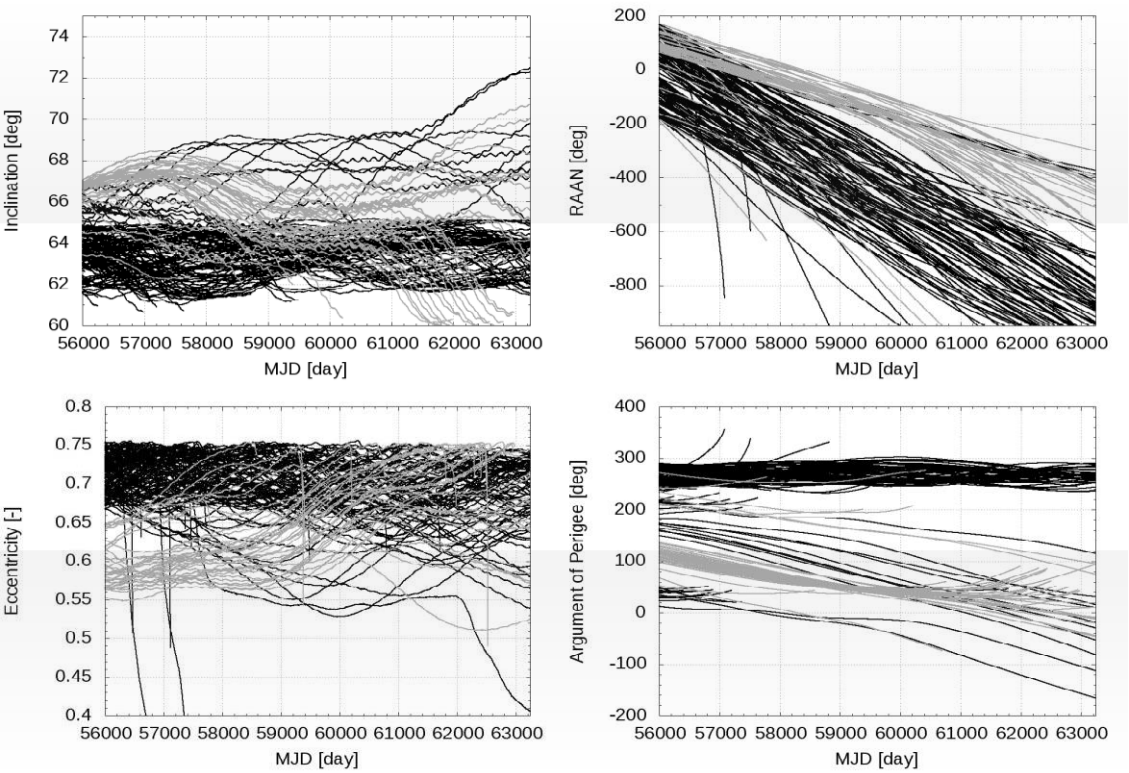

Рис. 5. Эволюция орбитальных элементов в течение 20 лет для выбранной популяции TLE при значении AMR 0,009 кг/м2 для атмосферного сопротивления и радиационного давления. Объекты с эксцентриситетом менее 0,65 выделены серым цветом.

Гравитационные и негравитационные силы постоянно влияют на динамику каждого объекта, размещенного на геоцентрической орбите. Следовательно, эти силы приводят к тому, что орбитальные элементы объекта могут значительно изменяться с течением времени. Чтобы выяснить, на каких типах орбит сегодня можно найти бывшие объекты «Молния», и лучше оптимизировать стратегию исследования, взяли выбранную популяцию TLE и смоделировали эволюцию орбитальных элементов на двадцать лет вперед. Для этого распространения использовалась модель полной силы, включающая гравитационное притяжение Солнца и Луны, коэффициенты гравитационного потенциала Земли с точностью до степени и порядка 12, возмущения, вызванные земными приливами, поправки, обусловленные общей теорией относительности, и простая модель для прямого радиационного давления. Проведенные исследования актуальны для обеспечения надежной работы инфокоммуникационных сетей связи и управления различного назначения [14 - 37].

Литература

1. *Алдохина В.Н., Куликов С.В., Лиференко В.Д., Чесноков Д.С.* Виртуальный прибор для исследования формы трассы полета КО от значений элементов орбиты // Компоненты и технологии. 2017. № 2. С. 128-130.
2. *Алдохина В.Н., Гудаев Р.А., Смирнов М.С., Шаймухаметов Ш.И.* Модель системы мониторинга и контроля воздушно-космического пространства // Труды Военно-космической академии имени А.Ф. Можайского. 2019. № 668. С. 8-20
3. *Бойкова А.В.* Использование информационных технологий в образовательном процессе военного вуза // Интернет-журнал «Мир науки». 2017. Том 5, № 6.; URL: https://mir-nauki.com/PDF/96PDMN617.pdf (дата обращения 05.01.2021)
4. *Эскобал П.* Методы определения орбит. М.: Мир, 1970. 472 с.
5. *Иванов Н.М., Лысенко Л.Н.* Баллистика и навигация космических аппаратов. М.: МГТУ имени Н.Э.Баумана, 2016. 528 с.
6. *Федоренко Д.С., Легков К.Е.* Моделирование спектра отражения высокоорбитального искусственного спутника Земли // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 11. С. 14-20.
7. *Liferenko V.D., Legkov K.E., Kolesnik D.Y.* Method for recognizing the type of space object in airspace based on the use of radar images // В сборнике: 2020 Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2020. 2020. С. 9166055.
8. *Liferenko V.D., Fedorenko D.S., Legkov K.E.* Verification of the model forming the space object reflection spectrum based on normal-hemispheric reflection coefficients of reflection of materials and coatings // В сборнике: 2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020. 2020. С. 9131551.
9. *Aldokhina V.N., Kolesnik D.Y., Liferenko V.D., Legkov K.E.* Model of recognition of cosmic objects based on informative signs obtained by radar means // В сборнике: 2020 Wave Electronics and its Application in Information and Telecommunication Systems, WECONF 2020. 2020. С. 9131444.
10. *Liferenko V.D., Legkov K.E., Gural D.A*. Organization of effective functioning of the information subsystem of a network of distributed heterogeneous information and computing resources // В сборнике: 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, Conference Proceedings. 2021. С. 9416101.
11. *Aldohina V.N., Fedorenko D.S., Liferenko V.D., Legkov K.E.* Methodology for creating a reference reflection spectra database for space objects monitoring // В сборнике: 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, Conference Proceedings. 2021. С. 9416097.
12. *Гураль Д.А., Легков К.Е.* Программный комплекс моделирования процессов получения информации от средств сопровождения космических объектов. Свидетельство о регистрации программы для ЭВМ 2021680088, 07.12.2021. Заявка № 2021669737 от 29.11.2021.
13. *Гураль Д.А., Легков К.Е.* Программный комплекс моделирования информационных подсистем автоматизированных систем управления объектами специального назначения. Свидетельство о регистрации программы для ЭВМ 2021680877, 15.12.2021. Заявка № 2021669953 от 22.11.2021.
14. *Смирнов Б.П., Зверев А.Б., Легков К.Е.* Методика формирования единого комплекса описания данных в системе информационных технологий единого информационного пространства специального назначения // Наукоемкие технологии в космических исследованиях Земли. 2020. Т. 12. № 5. С. 74-82.
15. *Буренин А.Н., Легков К.Е*. Основные проблемы безопасности подсистем обеспечения единым временем элементов систем управления сложными организационно-техническими объектами // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 1. С. 45-53.
16. *Буренин А.Н., Легков К.Е., Терещенко Г.В.* Управление безопасностью функционирования подсистемы обеспечения единым временем элементов системы управления сложным организационно-техническим объектом // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 2. С. 36-45.
17. *Буренин А.Н., Легков К.Е., Левко И.В.* О моделях информационных структур комплексов обеспечения единым временем системы управления сложным организационно-техническим объектом // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 3. С. 44-51.

18. *Басыров А.Г., Легков К.Е.* Метод эвристико-комбинированного решения трудоемких задач в параллельных вычислительных системах реального времени // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 3. С. 52-56.

19. *Буренин А.Н., Легков К.Е.* Основные подходы к организации оперативного управления комплексами обеспечения единым временем системы управления сложным организационно-техническим объектом // Наукоемкие технологии в космических исследованиях Земли. 2019. Т. 11. № 1. С. 20-32.

20. *Буренин А.Н., Легков К.Е.* Модели состояния современных инфокоммуникационных сетей при организации стохастического управления ими // Наукоемкие технологии в космических исследованиях Земли. 2019. Т. 11. № 2. С. 32-39.

21. *Буренин А.Н., Легков К.Е.* Модели стохастического управления современными инфокоммуникационными сетями // Наукоемкие технологии в космических исследованиях Земли. 2019. Т. 11. № 3. С. 26-31.

22. *Оркин В.В., Легков К.Е.* Алгоритм выбора процедур распределенного программного управления потоками заявок в информационной системе в условиях возмущений // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 10. С. 41-45.

23. *Легков К.Е., Левко И.В., Оркин В.В.* Методика адаптивного управления информационной системой критически важных объектов в условиях массовых возмущений // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 11. С. 51-56.

24. *Буренин А.Н., Голубев В.Е., Легков К.Е.* Организация подсистемы обеспечения единым временем решающих элементов автоматизированной системы управления сложными организационно-техническими объектами специального назначения // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 2. С. 27-34.

25. *Легков К.Е.* Методические основы управления информационными подсистемами автоматизированных систем управления сложными объектами специального назначения // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 5. С. 31-40.

26. *Буренин А.Н., Легков К.Е., Первов М.С.* Вероятностно-временные характеристики функционирования защищенной агрегативной автоматизированной системы управления сложной организационно-технической системой в условиях интенсивных кибератак // Наукоемкие технологии в космических исследованиях Земли. 2018. Т. 10. № 5. С. 56-63.

27. *Буренин А.Н., Легков К.Е., Первов М.С.* Организация процедур по выявлению и локализации нарушений политик безопасности при управлении безопасностью функционирования подсистемы обеспечения единым временем автоматизированной системы управления сложной организационно-технической системой // Наукоемкие технологии в космических исследованиях Земли. 2018. Т. 10. № 6. С. 28-36.

28. *Нестеренко О.Е., Легков К.Е.* Методика формирования информационной структуры параллельных программ вычислительной системы специального назначения // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 3. С. 43-49.

29. *Легков К.Е.* Модели информационных подсистем автоматизированных систем управления сложными объектами // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 5. С. 33-44.

30. *Буренин А.Н., Легков К.Е.* Метод повышения эффективности функционирования на основе процедур оперативного управления структурой информационных подсистем // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 9. С. 48-57.

31. *Буренин А.Н., Легков К.Е., Оркин В.В*. Постановка задачи управления функционированием информационной системы специального назначения // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2017. № 1. С. 92-98.

32. *Легков К.Е., Нестеренко О.Е.* Алгоритм формирования информационной структуры параллельных программ иерархической вычислительной системы // Наукоемкие технологии в космических исследованиях Земли. 2017. Т. 9. № 1. С. 52-59.

33. *Буренин А.Н., Легков К.Е., Левко И.В.* Организация эффективного функционирования информационных подсистем автоматизированных систем управления сложными организационно-техническими объектами на основе методов управления процессами предоставления информационных услуг // Наукоемкие технологии в космических исследованиях Земли. 2017. Т. 9. № 3. С. 45-54.

34. *Буренин А.Н., Легков К.Е.* Основы обеспечения эффективного функционирования информационных подсистем автоматизированных систем управления сложными организационно-техническими объектами в условиях воздействий // Наукоемкие технологии в космических исследованиях Земли. 2017. Т. 9. № 4. С. 79-86.

35. *Легков К.Е., Левко И.В*. Системный подход к организации управления информационными подсистемами автоматизированных систем управления сложными объектами специального назначения // Наукоемкие технологии в космических исследованиях Земли. 2017. Т. 9. № 5. С. 84-91.

36. *Буренин А.Н., Легков К.Е., Оркин В.В*. Алгоритм адаптивного управления информационными системами в условиях массовых возмущений // Наукоемкие технологии в космических исследованиях Земли. 2017. Т. 9. № 6. С. 90-95.

37. *Буренин А.Н., Легков К.Е., Оркин В.В.* Функционирования информационных подсистем автоматизированных систем управления сложными организационно-техническими объектами в условиях внешних воздействий // Системы синхронизации, формирования и обработки сигналов. 2017. Т. 8. № 2. С. 4-9.

АРХИТЕКТУРА АСИМПТОТИЧЕСКОГО УПРАВЛЕНИЯ БЕЗОПАСНОСТЬЮ КРИТИЧЕСКИХ ИНФОРМАЦИОННЫХ ИНФРАСТРУКТУР

Ерохин Сергей Дмитриевич,
Московский технический университет связи и информатики, ректор, к.т.н., Москва, Россия

Петухов Андрей Николаевич,
М*осковский технический университет связи и информатики, начальник отдела, к.т.н., Москва, Россия*
anpetukhov@yandex.ru

Аннотация

В статье описана архитектура формирования и реализации управленческих решений для обеспечения безопасности КИИ. Методология асимптотического управления безопасностью различает компьютерную атаку и инцидент. Компьютерные атаки детектируются параллельно с индикацией инцидента на основании мониторинга событий безопасности. В случае обнаружения атаки привлекаются средства анализа развития инцидента, которые определяют возможность его возникновения в результате обнаруженной атаки. Адаптивный контур корректирует комплекс средств защиты, обеспечивая асимптотический рост уровня защищенности. В архитектуре используются методы машинного обучения и привязки к условиям конкретной КИИ

Ключевые слова: *информационная безопасность, критические информационные инфраструктуры, асимптотическое управление, прогнозирование инцидента, обнаружение компьютерных атак, методы машинного обучения*

Введение

Обзор современной нормативной базы и практики обеспечения информационной безопасности (ИБ) показывает, что к настоящему времени поддерживается широкий спектр подходов и решений управления ИБ. Так, основная парадигма, полагающая целью управления соответствие декларированной политике безопасности в условиях установленной совокупности возможных вредоносных воздействий - угроз (декларативный подход), поддерживается концепцией системы менеджмента ИБ.

Также, среди методологических решений в этой области нашли свое место те, которые поддерживают подход, использующий процедуры идентификации, анализа и управления рисками и устанавливающий в качестве цели управления достижения приемлемого уровня остаточного риска (рискориентированный подход). Еще одним относительно распространенным подходом к управлению ИБ является гармонизация установленных политик, предположений и угроз безопасности в виде системы требований (функциональных и доверия) и достижение выполнения этих требований («Общие критерии»). В рамках этих подходов развитие сферы управления ИБ видится, прежде всего, в сторону детализации процедур и функций, а также в сторону спецификации их применения для конкретных индустрий и технологий. Выделение критических информационных инфраструктур (КИИ) в качестве специфического вида объектов защиты привлекло внимание к особенностям управления ИБ в рамках этого вида объектов защиты.

Принципы асимптотического управления безопасностью КИИ

Все доступные источники определяют КИИ не как самостоятельную сущность, а в составе критического объекта. Эти определения как правило содержат два общих аспекта: перечисление индустрий, где используются эти инфраструктуры, и неприемлемость (недопустимость, чрезмерность, несоотносимость) того ущерба, который возникает при нарушении функционирования самого критического объекта. Первый аспект (перечисление индустрий) указывает на то, что информационная инфраструктура является критической не в силу своих особенных собственных свойств, а в связи с участием в определенной деятельности объекта, на котором она эксплуатируется.

Критерии безопасности для КИИ определяются допустимостью функционирования критического объекта и выражаются в состояниях КИИ, обеспечивающих допустимое функционирование. Следовательно, одной из основных функций управления ИБ КИИ должна быть индикация состояний КИИ, не обеспечивающих допустимое функционирование критического объекта (инцидентов).

Второй аспект (неисчислимость ущерба) указывает о том, что какие бы усилия ни были предприняты в отношении КИИ, ущерб от нарушения безопасности критического объекта будет оставаться неприемлемым, что соответствует недопустимости любого остаточного риска и, как следствие, невозможности использования понятия остаточного риска для описания целевого состояния безопасности КИИ. Другой стороной этого аспекта является осознание принципиальной возможности инцидента независимо от реальной защищенности КИИ. В этих условиях роль модели угроз деформируется, допуская, что включенные в модель идентифицированные угрозы составляют лишь часть реально существующих угроз, наряду с которой есть неопределяемая часть за пределами модели.

Первоначальной постановкой задачи управления ИБ для упомянутых выше подходов является:

- для декларативного подхода - эффективное внедрение нормативного состава решений безопасности, определенного политикой.
- для рискориентированного подхода - достижение объектом защиты приемлемого уровня остаточного риска.
- для «Общих критериев» - обоснование правомерности и достижение уровня доверия, установленного на основании принятой аксиоматики.

Таким образом, в рамках этих подходов целевое состояние объекта защиты изначально задается в явной или опосредованной форме, и предполагается существование способа, с помощью которого можно установить достижение объектом этого состояния. В случае КИИ при условии сделанных выше утверждений такое постулирование не совсем правомерно, т.к. целевое безопасное состояние, строго говоря, недостижимо (нулевой остаточный риск). Это приводит к выводу о том, что любой из этих подходов, нуждается в некотором расширении для использования его в условиях управления безопасностью КИИ.

Утверждения о неопределенности целевого состояния КИИ и о неполноте модели угроз приводят к необходимости расширить смысл процессов управления безопасностью - при невозможности достижения цели ввиду ее неопределенности следует перейти к парадигме неуклонного приближения к цели, и целеполагание достижения рубежа дополнить целеполаганием направления движения к рубежу. Таким образом, вводится в рассмотрение понятие асимптотического управления безопасностью КИИ [1], для которого каждое последовательное решение означает неуклонный рост уровня безопасности КИИ. В этой парадигме целевое состояние безопасности определяется не в терминах набора актуальных и скомпенсированных угроз, а в терминах инцидентов, для предотвращения которых применяются все доступные меры, Уровень защищенности в этом случае определяется знаниями об инциденте как состоянии КИИ и о мерах предотвращения инцидента, а с расширением таких знаний, уровень защищенности растет.

Парадигма асимптотического управления не является альтернативой традиционным подходам, она призвана в необходимых случаях дополнить их и использоваться в сочетании с ними. Архитектурно эти подходы включают этапы и решения, условно объединяемые в составе двух контуров: пассивного – оценка соответствия объекта защиты целевому состоянию и анализ факторов расхождения, и активного – проектирование и реализация защитных мер, т.е. приведение в соответствие. В асимптотическом управлении аналогичную роль играют прогностический и адаптивный контуры.

Прогностический контур

В связи с тем, что в условиях КИИ понятие инцидента имеет определенную специфику, методология асимптотического управления безопасностью различает компьютерную атаку и инцидент, имея в виду, что инцидент – это специфицированное состояние КИИ, которое определяется неудовлетворительностью или невозможностью функционирования соответствующего критического объекта. Индикация инцидента - это проверка выполнения заранее сформулированных условий, специфичных для каждой конкретной КИИ, поэтому стандартизация этих процессов представляется ограниченной, однако типизировать их структуру возможно.

Компьютерные атаки детектируются параллельно с индикацией инцидента на основании мониторинга событий безопасности КИИ. В случае обнаружения атаки привлекаются средства анализа развития инцидента, которые определяют возможность его возникновения в результате обнаруженной атаки, т.е. прогнозируется, достаточен ли потенциал защиты для отражения обнаруженной атаки. Индикация инцидента, детектирование компьютерных атак и анализ развития инцидента составляют функциональность прогностического контура асимптотического управления, именно здесь сконцентрированы средства прогнозирования. Архитектура этого контура допускает вариативность инструментов обнаружения атак и прогноза инцидента, что позволяет комплексировать потенциал отдельных решений и принимать общее решение на основании нескольких частных, а также более полно охватывать пространство возможных атак.

Главная задача прогностического контура состоит в идентификации текущего состояния КИИ с точки зрения безопасности и в том, чтобы на основании знаний о текущем состоянии объекта прогнозировать возникновение инцидента, т.е. оценить возможность такого возникновения в текущий момент времени или в будущем. Прогностический контур использует результаты наблюдения для формирования гипотезы (решения) о том, что инцидент произошел, происходит или может произойти. Результаты прогнозирования установленным порядком сигнализируются и используются в адаптивном контуре для изменения конфигураций механизмов защиты и для адаптации множества сигнатур, признакового пространства аномалий, корпусов обучающих и тестовых данных.

В прогностическом контуре можно выделить функциональные блоки:

- мониторинга,
- индикации инцидента,
- детектирования компьютерных атак
- анализа развития инцидента.

Мониторинг

Целью мониторинга обнаружение событий безопасности на основе наблюдаемых параметров. Структура мониторинга включает сенсоры, схемы агрегации, коллектор и хранилище первичных событий. С точки зрения мониторинга первичное событие состоит в условно одномоментном измерении вектора шкалируемых параметров. В архитектуре асимптотического управления предполагается, что агрегированное событие это логическая переменная, полностью определяемая, значениями параметров первичных событий и схемой агрегирования. Схемы агрегирования поддерживаются специальной базой данных – планом контролируемых событий безопасности, это одна из центральных информационных структур, поскольку она определяет пространство признаков для входного интерфейса детектирования компьютерных атак и подвергается корректирующим воздействиям управления со стороны адаптивного контура. Множество этих признаков оптимизируется в смысле понижения размерности и средства поддержания такой оптимизации [2] входят в состав архитектуры асимптотического управления

Вычислительные процедуры по схемам агрегации осуществляет коллектор. При агрегации событий безопасности проявляются следующие факторы:

- в процессе агрегации информация, содержащаяся в событии, может только уменьшаться;
- в процессе агрегации меняется (уменьшается) как полезная, так и избыточная информация;
- восстановить информацию, утраченную в процессе агрегации, невозможно.

Для снижения негативной роли этих факторов архитектурой предусмотрено хранилище для событий безопасности. Это средство накопления и сохранения исходных данных для реализации схем агрегации работает под управлением коллектора.

Функциональный блок мониторинга допускает привлечение специализированных протоколов мониторинга для анализа потоков, который опирается на сбор метаданных о сетевом трафике. Анализ потоков не заменяет и не отменяет захвата пакетов, он лучше всего подходит для мониторинга внутренней инфраструктуры сети. Специально установленный сенсор компьютерная атака может обойти, а сетевое устройство, поддерживающее протокол мониторинга для анализа потоков обойти нельзя. Специализированные протоколы мониторинга собирают как семплированную (отфильтрованную, частичную) информацию, так и несемплированную (полную). Кроме

мониторинга сетевых протоколов допускается мониторинг узлов сети на основе хостовых сенсоров (многоагентный мониторинг). Аудит журналов может рассматриваться как «расщепленная» во времени индикация двух первичных событий: сначала формируется запись регистрационного журнала, а затем в момент собственно аудита осуществляется агрегация - анализ совокупности таких накопленных записей

Индикация инцидента

В рамках архитектуры асимптотического управления предполагается, что средства мониторинга в состоянии обеспечить индикацию инцидента. При этом не всегда состояние системы несет в себе следы свершившегося инцидента, например, утечка данных никаких следов в самой системе не оставляет. Поэтому среди признаков инцидента должны присутствовать события, свидетельствующие не только о свершившимся инциденте, но и о процессе ему предшествующем и однозначно приводящему к инциденту, (в приведенном примере это процесс несанкционированной передачи по каналу утечки). Кроме того, во многих КИИ инцидент первоначально определяется в категориях, которые не являются атрибутами собственно информационной системы, и безопасное или небезопасное состояние информационной инфраструктуры определяется через правильное или неправильное функционирование самого критического объекта. Чтобы учесть это обстоятельство, необходимо моделировать причинно-следственные связи между событиями в информационной системе и признаками инцидента в критическом объекте. Результаты такого моделирования фиксируются в базе данных условий индикации инцидента. Для объектов информационной инфраструктуры, где атрибуты инцидента выражены в категориях информационных технологий, например, КИИ класса ИТКС, формирование такой базы упрощается. Допускается, что инцидент устанавливается за пределами автоматизированного управления, и сведения о нем попадают в функциональный блок индикации и учета инцидентов из контура администрирования.

Наряду с передачей установленным порядком сообщения об инциденте в инстанции для каждого обнаруженного инцидента формируется отчет, который содержит информацию для проведения анализа и подготовки решений о внесении изменений в базу данных условий индикации инцидента, особенно это важно для инцидентов, установленных вне автоматизированного контура. Такие изменения готовятся и вносятся в конуре администрирования, автоматическое внесение изменений не предусматривается. Сведения об инциденте заносятся в базу данных учета инцидентов и компьютерных атак, состав и содержание сведений, зафиксированных в отчете и базе данных, должны быть достаточными, чтобы обеспечить проведение необходимого расследования.

Детектирование атак

Основу функционального блока детектирования атак составляют анализаторы (системы обнаружения атак), которые обрабатывают поток событий безопасности, предоставляемый мониторингом. Наряду с сетевым трафиком в процессе детектирования компьютерных атак участвуют данные из журналов регистрации, поэтому обязательным компонентом этого блока являются процедуры нормализации и анализа таких журналов. Детектирование атак допускает применение разнообразных методов, среди которых, в частности, могут быть:

- подходы на основе соответствия сигнатурным образцам,
- выявления статистических аномалий
- использования эвристических моделей (искусственного интеллекта).

Суть методов сигнатурного анализа заключается в задании множества сигнатур атак в виде регулярных выражений или правил на основе сопоставления с образцом и проверке соответствия наблюдаемых событий этим выражениям. Сигнатурный анализ позволяет идентифицировать несанкционированные действия, имеющие точное описание в виде сигнатур атак, т.е. совокупности действий, применяя которые к текущему набору признаков можно получить свидетельство несанкционированных действий и установить тип таких действий. Даже несущественные изменения в сценарии атаки могут привести к невозможности ее обнаружения методами сигнатурного анализа, поэтому задаваемые правила должны быть относительно универсальными и описывать все известные модификации атак. Методы сигнатурного анализа являются эффективным инструментом для выявления известных типов атак, но их применимость по отношению к новым атакам, а также к модификациям известных атак является сомнительной.

Главной проблемой при создании любого средства сигнатурного анализа является эффективное проектирование механизма задания сигнатур атак, поэтому в архитектуре асимптотического управления предусматривается возможность внесения изменений в множество учитываемых сигнатур в базе данных правил вычисления сигнатур. Основное преимущество методов сигнатурного анализа заключается в том, что обнаружение известных образцов аномальных событий осуществляется максимально эффективно. Но, в то же время, использование базы сигнатур большого объема отрицательно влияет на производительность процедуры обнаружения.

В сигнатурных методах события безопасности представляются в виде цепочек символов из некоторого алфавита, что дает основание предполагать перспективность использования аппарата формальных языков и грамматик.

Диапазон приемов и подходов для выявления статистических аномалий сетевого трафика включает множество вариаций [3]. Методы выявления аномалий могут использовать, например, многоступенчатые процедуры разложения протяженного во времени «сигнала» (количество байт, количество пакетов, средняя длина пакета и т.п.) на заранее установленный набор базисных функций и сравнение коэффициентов такого разложения с образцовыми значениями (вейвлет-анализ). Другими критериями соответствия нормальному поведению сети может служить установление т.н. самоподобия, когда предельные отклонения «сигнала» не превышают некоторые пороги, например, среднеквадратичное отклонение (фрактальный анализ), или устойчивость структуры распределения элементов трафика по группам, обеспечивающего равномерное распределение этих групп (анализ максимальной энтропии).

Методы, которые основаны на сравнении информации о нормальном поведении системы с параметрами наблюдаемого поведения, ориентированы на построение модели нормального функционирования объекта защиты. Построение шаблона нормального поведения является трудоемкой и не всегда выполнимой задачей. В методах выявления аномалий важную роль играет правильный выбор контролируемых параметров, характеризующих отличия в нормальном и аномальном трафиках. Эти отличия должны быть соотнесены и воздействиями, которые оказывает на трафик конкретная компьютерная атака (тип атаки), и, с другой стороны, быть инвариантными к «собственным» вариациям трафика, чтобы не вызывать ложноположительные срабатывания. В настоящее время наиболее проработанным источником систематизированных сведений о техниках и тактиках компьютерных атак являются базы данных MITRE, в частности, «матрица» ATT&CK [4]. Однако их использование для конфигурирования и настройки анализаторов обнаружения атак требует интерпретации этих данных в условиях инфраструктуры и информационных процессов конкретной КИИ.

Среди множества моделей искусственного интеллекта для детектирования компьютерных атак наибольшее распространение получили процедуры на базе искусственных нейронных сетей, которые представляют собой набор обрабатывающих элементов - нейронов, передающих друг другу сигнал через установленные связи – синапсы, и преобразующих входные векторы значений в вектор желаемых выходных значений. Искусственные нейронные сети обладают способностью обучения по образцу и обобщения из зашумленных и неполных данных, в процессе обучения происходит настройка коэффициентов передачи (синаптических весов). Практика применения многослойных структур с распределенной функциональностью (например, многослойных персептронов), показывает возможность получить необходимую степень детализации при классификации и обнаруживать аномалии на основе данных, взятых из системного журнала аудита и лог-файлов отдельных приложений, а также системной информации, в том числе объемов ресурсов, времени работы, частоты использования наиболее распространенных команд и др. Привлечение в качестве моделей рекуррентных нейронных сетей (в т.ч. LSTM) позволило включить в модель элемент памяти и классифицировать компьютерные атаки, протяженные во времени [5].

Важной особенностью архитектуры асимптотического управления является множественность (вариативность) используемых методов и алгоритмов анализаторов. В состав конкретной реализации архитектуры может быть включено несколько процедур детектирования атак различного типа. Существо такого подхода заключается в объединении нескольких построенных на различных принципах первичных анализаторов, которое позволяет нивелировать недостатки их функционирования по отдельности. Выходные значения первичных анализаторов рассматриваются

как предварительные результаты, которые являются исходными данными при формировании окончательного результата интегрирующего решающего правила.

Анализ развития инцидента

В случае детектирования компьютерной атаки включаются средства анализа развития инцидента, которые определяют возможность возникновения инцидента в результате обнаруженной атаки. Прогнозируется, достаточен ли потенциал защиты для отражения обнаруженной атаки, и если прогноз допускает возможность инцидента, то дальнейшие действия аналогичны случаю индикации уже случившегося инцидента (включая сообщение в ГосСОПКА). В процессе анализа развития инцидента прогнозируется развитие ситуации после обнаружения события безопасности (атаки, сигнатуры, аномалии), оценивается достаточность действующих механизмов защиты, необходимость дополнительного реагирования в виде внесения изменений в эти механизмы или каких-то других действий, например, простой сигнализации. Результаты прогнозирования могут быть использованы не только для изменения конфигураций механизмов защиты, но и для адаптации множества сигнатур, признакового пространства аномалий, корпусов обучающих и тестовых данных.

Методология анализа развития инцидента опирается на три компонента:

- методика прогноза, инструмент анализа развития инцидента
- модель вредоносного воздействия, сценарий атаки
- модель безопасности КИИ, информационных потоков и политики безопасности

Одним из наиболее развитых направлений прогнозирования негативных последствий является методология построения и анализа причинно-следственных траекторий развития как результата возникновения некоторого события. В [6] приводится обширный обзор методов оценки последствий событий в разделах анализа сценариев и функционального анализа. Анализ развития включает оценку диапазона возможных последствий события или оценку значимости и уязвимости компонентов КИИ, но в любом случае определяют характер и тип воздействия, которое может произойти при возникновении конкретного события. Для этого надо формально описать проявление факторов, приводящих к инциденту, начиная с первичного измерения (реакции сенсора) и вплоть до агрегированного события, свидетельствующего о совершении злоумышленником действия (техники), а также методику конструирования из таких техник более сложных комплексов действий (сценариев).

Архитектура асимптотического управления формируется в предположении о том, что инцидент может произойти только вследствие действий в составе компьютерной атаки. С появлением матрицы MITRE ATT&CK возникла возможность систематизировать действия злоумышленника по реализации атаки, как факторы приводящие к инциденту, и попытаться (в реальном времени или апостериорно) проявления этих факторов. Комплекс таких признаков может быть распределенным во времени и в информационном (инфраструктурном) пространстве, для сети следует ожидать их проявления на разных уровнях OSI. Кроме того, сама структура функционального анализа развития инцидента может быть многоуровневой (техники, цепочки техник, тактика, атака).

Моделирование воздействия на трафик установленных техник атак является ключевым моментом анализа развития инцидента, поскольку выявляется близкая к детерминированной связь действий, производимых в составе атак, с характеристиками трафика и, тем самым, очерчивается круг признаков, свидетельствующих о применении конкретной техники.

Для эффективного моделирования сценария атаки должна быть определена информация о КИИ, которая может стать известной нарушителям, и использована ими для эксплуатации уязвимостей. Кроме того, прогноз развития инцидента это прогноз поведения КИИ, и он базируется на знании внутренних информационных связей (модель информационных потоков) и ограничений безопасности, налагаемых на эти связи (политика безопасности). Эти модели предусмотрены в составе информационной базы архитектуры асимптотического управления.

Адаптивный контур

Идеология асимптотического управления предполагает приращение уровня защищенности при каждом получении нового знания, и адаптивный контур использует для этого механизмы обратной связи и каналы взаимодействия с внешними источниками События в прогностическом контуре, свидетельствующие о появлении нового знания, и активизирующие адаптивный контур – это

индикация инцидента, обнаружение атаки и положительный прогноз инцидента в результате обнаруженной атаки.

В архитектуре асимптотического управления предусмотрено три направления адаптации:

- изменение настроек и конфигураций средств защиты КИИ в результате реагирования на индицированный или прогнозируемый инцидент;
- «переучивание» (внесение изменений в настройки) анализаторов атак после индикации инцидента и в результате оценки их роли в определении атаки
- коррекция пространства признаков после индикации инцидента и в результате оценки их роли в определении атаки

Есть три ситуации в прогностическом контуре, когда появляется новое знание и активизируются процедуры адаптивного контура – индикация инцидента, обнаружение атаки и положительный прогноз инцидента в результате атаки. В первой ситуации («инцидент произошел, а атаку пропустили») активизируются все три направления адаптации, причем «переучивание» анализаторов сопровождается коррекцией правил формирования общего решения об атаке. В случае обнаружения атаки оценивается роль каждого анализатора и недостаточно эффективные анализаторы подвергаются «переобучению». В случае положительного прогноза инцидента подвергаются коррекции информационные потоки КИИ путем изменения настроек и конфигураций средств защиты.

Внесения изменений в конфигурацию и настройки средств защиты имеют целью повысить защищенность КИИ, добавляя в политику безопасности новые ограничения, которые могли бы воспрепятствовать инциденту, явившемуся причиной этих изменений. В большинстве моделей безопасности такие ограничения накладываются политикой на информационные потоки, поэтому для принятия решений адаптивному контуру необходимо обращаться к модели информационных потоков КИИ и формальному представлению ограничений – политике безопасности КИИ. Причем для сетевых КИИ эти модели должны носить многоуровневый характер (в смысле уровней OSI), поскольку атрибуты и инцидента, и соответствующего корректирующего воздействия могут располагаться на разных уровнях OSI. Кроме того, необходим инструмент, обеспечивающий трансляцию нотаций политики с более высоких уровней на нижележащие. И, наконец, в составе архитектуры предусматриваются средства интерпретации решений адаптивного контура в терминах языков управления конкретными защитными механизмами [7], например, для межсетевого экрана это могут быть предикаты условий прохождения или отбрасывания пакетов.

Внесение изменений в настройки анализаторов, работающих с сигнатурами атак и выявлением сетевых аномалий, базируется на анализе материалов, характеризующих КИИ в период, непосредственно предшествующий инциденту. Это журналы регистрации и корпуса данных трафика, составляющие реальные потоки событий безопасности. Кроме того, для коррекции сведений о сетевых аномалиях может быть использован эталонный (без признаков вредоносного воздействия) образ трафика. Сравнение этого образа с реальными потоками событий безопасности, в результате которых возник инцидент, дает возможность выявить отклонения, характерные для конкретной КИИ и сформировать необходимые изменения настроек анализаторов. Для анализаторов, работающих с сигнатурами атак можно обеспечить монотонное повышение уровня защищенности при пошаговом внесении изменений в состав используемых сигнатур [8], для анализаторов на базе выявления сетевых аномалий о таком результате неизвестно.

Использование анализаторов на базе эвристических алгоритмов с машинным обучением несколько упрощает технологию внесения корректировок, но снижает предсказуемость результата адаптации. Функционирование этих алгоритмов с точки зрения детектирования компьютерных атак полностью определяется тремя факторами: признаковым пространством, процедурой обучения (включая обучающий датасет), и структурными ограничениями самого алгоритма. Даже не подвергая изменению признаковое пространство, которое определяется мониторингом, для совершенствования процесса детектирования достаточно предъявить в качестве обучающего датасета реальный поток событий безопасности за период, предшествующий инциденту. Поскольку предполагается, что этот датасет несет в себе признаки необнаруженной атаки, в результате которой возник инцидент, после такого обучения анализатор гипотетически приобретет способность обнаруживать атаки такого типа. Гипотетически потому, что это возможно только при условии, если такой классификации не

препятствуют структурные ограничения алгоритма, которые для большинства алгоритмов ясны далеко не полностью.

Управление признаковым пространством (формирование и изменение набора признаков) является одним из наиболее критичных процессов архитектуры, его результаты оказывают существенное воздействие на эффективность асимптотического управления. К настоящему времени исследовано много различных подходов к решению задачи отбора признаков [9] как в смысле качества классификации, так и в смысле понижения размерности признакового пространства [2].

В контексте использования средств машинного обучения представляется перспективным метод оценки информативности и выделения признаков на основе анализа отличий реального фрагмента трафика, подверженного воздействию атаки, от эталонного датасета. Одним из преимуществ такого подхода является возможность статистической оценки корреляционных характеристик признаков, что, в свою очередь позволяет ставить и решать задачи оптимизации набора признаков [10].

Метод анализа отличий дополняет метод моделирование датасета с признаками воздействия заданной атаки. В случае моделирования в эталонный датасет вносятся изменения, обеспечивающие признаки некоторой атаки (например, в привлечением описаний классификатора MITRE ATT&CK), и формируется «зараженный» датасет, который используется для обучения анализаторов. В случае анализа отличий, про фрагмент реального трафика известно, что он был подвержен воздействию атаки («зараженный» датасет), и задача в этом случае – выявить признаки, соответствующие этой атаке. Оба метода – и моделирования, и анализа отличий обладают общим преимуществом – они используют данные, отражающие характерные особенности конкретной КИИ.

Решения по управлению признаковым пространством распространяются только на атрибуты классификации при детектировании компьютерных атак. Признаки инцидента формируются, обрабатываются и хранятся отдельным образом, поскольку они являются отражением не действий злоумышленника, а состояния критического объекта и являются производными от решений функциональной безопасности.

Таким образом, в первом направлении используется модель информационных потоков и политика безопасности КИИ, во втором – корпус данных (датасет) потока событий безопасности, собранный мониторингом за предшествующий период («история» событий безопасности), в третьем – эталонный (без признаков вредоносного воздействия) датасет и «история» трафика.

Взаимодействие с внешними источниками

Концепция ГосСОПКА определяет систему управления КИИ как многоуровневую распределенную систему реального времени, состоящую из некоторого числа взаимодействующих подсистем, которые обмениваются между собой данными. Такими подсистемами являются сами КИИ, структурные звенья ГосСОПКА, а также внешние источники сведений об установленных уязвимостях и компьютерных атаках (базы данных ФСТЭК, разрешенные SERT-центры, другие КИИ). Информационный обмен этих подсистем является одним из главных факторов развития защищенности КИИ. Характерными особенностями этого обмена являются следующие:

- взаимодействие ориентировано на данные, основным «материалом» управления являются сведения об установленных уязвимостях и атаках;
- каждый субъект взаимодействия (КИИ) обладает спецификой, определяющей ценность и адекватность предоставляемых данных;
- взаимодействие осуществляется в реальном времени, задержки и собственно длительности транзакций по предоставлению этих данных являются существенными факторами функционирования КИИ
- постоянное расширение состава взаимодействующих субъектов требует от системы взаимодействия высокой гибкости и масштабируемости

Все источники предоставляют сведения общим массивом, не ориентируясь на специфику КИИ. В то же время, конкретной КИИ необходима адекватная ее проблематике информация, и использование этой информации должно контролироваться. Поэтому в качестве архитектурной платформы взаимодействия может быть привлечена концепция «издателей-подписчиков» (DDS [11]), обеспечивающая фильтрацию и контролируемый селективный доступ к данным. Следует уточнить, что это касается только сведений об установленных уязвимостях и актуальных атаках, в то время как

сведения об инцидентах (индицированных и прогнозируемых) передаются установленным порядком непосредственно через технический интерфейс, обеспечивающий выполнение форматных требований получателя (структурные звенья ГосСОПКА).

Цель спецификации DDS – облегчить эффективное размещение данных и селективный доступ к ним в распределенной системе. Участники, использующие DDS, могут «читать» и «писать» данные эффективно и естественно с помощью унифицированного интерфейса. Программная среда DDS распределяет данные, чтобы каждый участник чтения мог получить доступ к самым актуальным значениям тех данных на которые он подписан.

Ключевая абстракция, лежащая в основе DDS, это распределенное глобальное пространство данных. Спецификация DDS требует распределенной реализации пространства данных, чтобы избежать единой точки отказа или единой точки разногласий. Издатели и подписчики могут присоединиться или покинуть пространство данных, поскольку они обнаруживаются пространством данных динамически. Все участники процесса будут обнаружены автоматически, и данные начнут к ним поступать в соответствии с профилем их потребности. Отказ одного узла не повлечет за собой негативных последствий для доступности, и система в целом будет продолжать работать, даже если участники взаимодействия выйдут из строя, перезапустятся, подключатся, или отключатся.

В составе ядра безопасности архитектуры DDS специфицированы пять расширяемых, взаимозаменяемых компонентов безопасности:

- для аутентификации (алгоритмы односторонней и взаимной аутентификации);
- криптографические функции (интерфейсы для генерации ключей и обмена ими, шифрование, коды аутентификации сообщений, хеширование и цифровые подписи);
- ведение журнала событий;
- прикрепление защитной метки (мандатное управление доступом);
- авторизация при публикации / подписке.

Включение в состав архитектуры управления обмена и анализа сведений отб уязвимостях обусловлено следующими обстоятельствами:

- нормативными актами установлена необходимость для КИИ выполнять функции по анализу уязвимостей;
- органы централизованного управления и иные внешние источники поддерживают информационные потоки сообщений об установленных уязвимостях.

В качестве внешних источников сведений об уязвимостях используются опубликованные данные разработчиков средств защиты, программного обеспечения, различные публичные базы данных уязвимостей, а также информация, поступающая по каналам централизованного управления безопасностью КИИ (ГосСОПКА).

При анализе уязвимостей и обмене сведениями о них ключевым вопросом является язык такого обмена, базирующийся на принятой классификации уязвимостей и сопряженных понятий. Среди всевозможных таксономий лучшим, с точки зрения разработчика КИИ, является реестр MITRE CWE [*12*] (по показателям полноты, всесторонности классификации, наличия подробных описаний с примерами), а с точки зрения администратора, самой эффективной представляется таксономия MITRE CVE [13*]* (по показателям объема записей, оперативности обновления).

Учитывая опыт использования существующих классификаций, можно сформулировать основные общие требования к языку анализа уязвимостей:

- в классификации или свойствах отдельной уязвимости должна содержаться информация об этапе жизненного цикла, на котором возникает дефект и его области возникновения (общая архитектура, код, внутренняя конфигурация, внешнее окружение);
- в классификации уязвимостей или свойствах отдельных типов дефектов должна быть ссылка на виды опасностей или механизмы атак (техники и тактики атак), при которых возможна эксплуатация этой уязвимости (примеры: атаки внедрения данных, атаки подмены идентификатора, атаки физического доступа и т. п.);
- если первопричиной появления уязвимости является связь с конкретными внешними компонентами (СУБД, web-сервер), то в свойствах отдельного дефекта должна быть указана ссылка на наименование соответствующей уязвимости внешнего компонента.

Кроме того, обязательно должна присутствовать оценка критичности уязвимости, например, с использованием спецификаций CVSS, предусматривающих учет до восьми факторов критичности и формирующих интегральную оценку, эти спецификации используются в составе уведомлений формата НКЦКИ.

Методы машинного обучения

Методы машинного обучения в контексте прогностических и адаптивных процедур асимптотического управления безопасностью обладают рядом преимуществ. Прежде всего, практические результаты показывают, что с помощью этих методов можно достичь довольно высоких показателей эффективности обнаружения компьютерных атак (см. например [14, 15]). Причем некоторые исследования показывают, что ограничения этой эффективности связаны в большей степени с типом обучаемого алгоритма, нежели с компонентами машинного обучения. Например, в [5] на примере сравнительного анализа использования в качестве анализатора двух типов искусственных нейросетей показано, что тип алгоритма существенно определяет результат, в частности ограничивает состав обнаруживаемых атак. В то же время, изменения характеристик самих процедур машинного обучения (параметры обучающих наборов данных, вид пороговой функции и др.) лишь незначительно влияют на статистические показатели ошибок обнаружения.

Использование методов машинного обучения в сочетании с эвристическими алгоритмами (решениями искусственного интеллекта) привлекает внимание идеей так называемого «потенциала экстраполяции», которая восходит еще к ранним работам по распознаванию образов. Суть этой идеи состоит в предположении, что в процессе обучения анализатор наряду с запоминанием предъявленных ему конкретных образцов приобретает некоторое неявное обобщающее знание, которое позволяет впоследствии выявлять более широкое множество паттернов, объединенных некоторыми (может быть, заранее неизвестными) свойствами. Применительно к обнаружению компьютерных атак это означает, что анализатор, будучи обученным на некотором наборе событий безопасности, характерных для какой-то компьютерной атаки, приобретает способность диагностировать эту атаку и на других событиях, тоже характерных для нее, но которые анализатору не предъявлялись. Практика показывает, что в ряде случаев эта идея с большей или меньшей эффективностью реализуется, особенно когда можно предполагать выполнение условий компактности на пространстве признаков.

С точки зрения архитектуры асимптотического управления методы машинного обучения привлекательны тем, что предоставляют удобный инструмент для формирования и реализации корректирующих воздействий в процедурах адаптивного контура. При необходимости выявлять новую атаку внести изменения в совокупную функциональность анализаторов можно несколькими способами. Прежде всего, это расширение состава анализаторов, такой способ является радикальным и он должен быть оправдан невозможностью избежать его с помощью других вариантов решения, однако он может быть неизбежен, потому что, как упоминалось выше, алгоритм анализатора это существенный фактор ограничения состава выявляемых атак. Для расширения состава анализаторов необходимо провести глубокий анализ событий безопасности, свидетельствующих о новой атаке, сформировать ее паттерн и оценить его соответствие добавляемому алгоритму, а затем перевести этот паттерн в шаблоны нового алгоритма. Все то же самое надо предпринять, если выбран способ внесения изменений в настройки и конфигурации уже используемых алгоритмов, т.е. дополняется состав их шаблонов. При этом надо быть уверенным, что действующие алгоритмы принципиально допускают обнаружение новой атаки.

Машинное обучение упрощает «внесение изменений в настройки и конфигурации» тем, что предполагает возможность повторного обучения на новых наборах событий безопасности (переобучение). Формирование таких обучающих наборов (корпусов данных, датасетов), включающих паттерны новой атаки, возможно двумя способами, в первом случае можно в качестве обучающего использовать набор событий безопасности, собранный мониторингом за период, в течение которого вероятно проводилась новая атака, например, период, предшествующий индикации инцидента. Этот корпус данных должен постоянно поддерживаться мониторингом в актуальном состоянии, отражая историю событий безопасности в течение определенного времени, предшествующего текущему моменту. Второй способ предполагает наличие эталонного корпуса данных, соответствующего отсутствию каких-либо вредоносных воздействий и инструмента

моделирования на этом наборе изменений, возникающих в результате проявления конкретных атак. Такое моделирование также поддерживает актуализацию состава обнаруживаемых атак при поступлении сведений о них из внешних источников, а также дает возможность учесть публикуемые сведения о тактиках и техника компьютерных атак [3]

Таким образом, в архитектуре асимптотического управления используются три типа корпусов данных (датасетов): эталонный (без признаков атак), текущий (история за определенный период до настоящего времени) и модельные (с признаками заданных атак).

Корпуса данных всех трех типов формируются с использованием реального потока событий безопасности (трафика) конкретной КИИ. Эталонный датасет является основой для формирования модельных датасетов, а также совместно с текущим датасетом используется для оценки эффективности признаков. Кроме того, тестирование анализаторов с помощью эталонного корпуса данных (специальный модельный датасет) позволяет оценить интенсивность ложных срабатываний при обнаружении компьютерных атак. Текущий и модельные датасеты, как уже указывалось, нужны для «переобучения» анализаторов в процессах адаптации, а модельные датасеты, кроме того, являются основой для первичного и регламентного (в соответствии с политикой безопасности) обучения и тестирования. Технологии формирования для этих типов датасетов тоже различаются. Текущий датасет поддерживается непрерывным мониторингом, например, с помощью специализированных протоколов, таких как NetFlow, для этого необходим ресурс производительности и хранения. Эталонный корпус данных формируется аналогично, только это делается однократно и при условии гарантированного отсутствия вредоносных воздействий (атак). Модельные датасеты получаются в результате обработки эталонного корпуса данных специальными средствами моделирования проявлений компьютерных атак.

Каждый датасет первоначально возникает в «сыром» виде массива первичных событий безопасности, для сетевого трафика это, как правило, фрагменты заголовков пакетов. В то же время анализаторы обнаружения атак в качестве входного потока воспринимают уже агрегированные данные, формирующие пространство признаков. В [14] на примере опубликованного корпуса данных CIC-IDS 2017 [16] показан пошаговый процесс перевода исходного «сырого» массива в рабочий датасет, состоящий из признаков и предназначенный для обучения и тестирования алгоритмов анализаторов. Однако для оценки эффективности признаков при коррекции их состава в контуре адаптации, а также для подготовки модельных датасетов необходим первоначальный необработанный материал. Поэтому архитектурой асимптотического управления предусмотрено два формата поддерживаемых корпусов данных: первоначально собранный мониторингом (для эталонного и текущего датсетов) и переработанный в набор признаков для восприятия анализаторами обнаружения атак (для текущего и модельных датасетов). Причем необязательно постоянно и непрерывно поддерживать данные во втором формате, они могут создаваться и использоваться по мере необходимости.

Функции, связанные с методами машинного обучения реализуются в архитектуре асимптотического управления двумя раздельными блоками: формирования корпусов данных и обучения/тестирования анализаторов. Первый блок непосредственно взаимодействует с мониторингом и базой данных актуальных атак и включает средства преобразования «сырого» трафика в пространство признаков, моделирования проявлений атак и их техник и тактик, а также управления процессами создания и поддержания эталонных и модельных датасетов. Второй блок объединяет стенды для проведения обучения и тестирования анализаторов и средства управления этими стендами в первоначальном и регламентированном режимах обучения/тестирования, а также в случае «переучивания» анализаторов контуром адаптации. Кроме того, этот блок включает средства подготовки датасетов для использования а конкретной среде, эти средства обеспечивают подготовку обучающих и тестовых корпусов непосредственно для предъявления их анализаторам. Взаимодействуют эти функциональные блоки через специальное хранилище корпусов данных.

Информационная база асимптотического управления

Информационное обеспечение архитектуры асимптотического управления включает общие и специальные модели, инструментальные и учетные базы данных, а также обучающие и тестовые корпуса данных (датасеты). В состав общих моделей входят:

- модель информационных потоков КИИ, которая используется для поддержания прогноза развития инцидента после обнаруженной атаки, формирования корректирующих воздействий в адаптивном контуре и создания эталонного набора событий безопасности, т.е. там, где нужно учитывать специфику конкретной КИИ
- модель настраиваемых компонентов комплекса средств защиты КИИ, является управляющим интерфейсом в процессах реагирования на инцидент, обеспечивает корректное внесение изменений в конфигурации и настройки средств защиты
- многоуровневая модель сетевой политики безопасности КИИ, необходима для анализа развития инцидента, интерпретации описаний атак и внесения коррекций в процессе реагирования на инцидент

Специальные модели формируются и поддерживаются в процессе функционирования асимптотического управления и включают:

- правила развития инцидента (библиотека прогнозных анализаторов) как средство проведения анализа последствий обнаруженной атаки и прогнозирования инцидента
- модель интерпретации описаний актуальных атак, техник и тактик матрицы MITRE ATT&CK, необходима для поддержания прогноза развития инцидента после обнаруженной атаки, настройки анализаторов прогностического контура и для подготовки модельных обучающих и тестовых датасетов
- условия индикации инцидента, для контроля возникновения (индикации) инцидента
- план контролируемых событий безопасности, отражает текущее состояние признакового пространства, используется для управления мониторингом, настройки анализаторов прогностического контура, а также платформой управления признаками в адаптивном контуре.

Инструментальные базы данных включают:

- хранилище текущих агрегаций событий безопасности, функционирует под управлением коллектора в составе мониторинга
- журналы состояния безопасности КИИ как один из источников событий безопасности
- базу данных описаний актуальных атак
- базу данных описаний уязвимости
- библиотеку анализаторов детектирования атак
- библиотеку сканеров уязвимостей

Учетные базы данных используются для апостериорного анализа и расследований инцидентов, в их состав входят:

- база данных учета обнаруженных (индицированных) и прогнозированных (установленных) инцидентов
- реестр детектированных атак

Обучающие и тестовые наборы данных используются для настройки анализаторов детектирования атак, а также формирования и оценки признаков атаки на базе анализа отличий реального и эталонного потоков событий безопасности

- эталонный корпус данных потока первичных событий безопасности
- текущий фрагмент потока первичных событий безопасности за период («история»)
- корпус данных признаков текущего фрагмента, согласованный с планом контролируемых событий
- модельные корпуса данных признаков потока первичных событий безопасности, подвергнуто воздействию заданных атак

Заключение

Характерными особенностями архитектуры асимптотического управления информационной безопасностью КИИ являются:

- раздельная и параллельная обработка информации об инциденте и компьютерной атаке
- прогнозирование развития инцидента на основании сведений об обнаруженной атаке.
- многоуровневый комплекс адаптации, включающий внесение коррекций, повышающих защищенность, как в средства защиты, так и в средства собственно управления

- вариативность инструментария анализа и прогнозирования
- использование платформы DDS для организации взаимодействия в рамках управления безопасностью КИИ
- привлечение в качестве обучающих и тестовых корпусов данных, подготовленных на базе реального потока событий безопасности конкретной КИИ

Для развития архитектуры асимптотического управления представляется перспективным решение следующих задач:

- моделирование траекторий развития инцидента с заданными признаками для конкретного состояния КИИ.
- разработка методов и средств формирования шаблона нормального функционирования КИИ.
- интерпретация матрицы MITRE ATT&CK в условиях инфраструктуры и информационных процессов конкретной КИИ, моделирование воздействия на поток событий безопасности установленных техник атак.
- разработка методов и средств формирования и оценки признаков атаки на базе анализа отличий реального и эталонного потоков событий безопасности
- построение многоуровневой модели политики безопасности для сетевых КИИ обеспечивающую трансляцию нотаций политики с более высоких уровней на нижележащие.

Литература

1. *Ерохин С.Д., Петухов А.Н., Пилюгин П.Л.* Управление безопасностью критических информационных инфраструктур. М.: Горячая линия – Телеком, 2021.
2. *Sergey Erokhin, Boris Borisenko, Aleksander Fadeev.* Reducing the Dimension of Input Data for IDS by Using Match Analysis IEEE Xplore Том 28 / С. Баландин, В. Деарт, Т. Туйтина, Материалы 28-й конференции Ассоциации открытых инноваций FRUCT, Москва, Россия. ISSN 2305-7254, 27-29 января 2021 года.
3. *Шелухин О.И.* Сетевые аномалии. Обнаружение, локализация, прогнозирование М.: Горячая линия – Телеком, 2020.
4. https://attack.mitre.org/matrices/
5. *Борисенко Б.Б., Ерохин С.Д., Фадеев А.С., Мартишин И.Д.* Обнаружение компьютерных атак при использовании многослойного персептрона и сетей с долгой краткосрочной памятью // Системы синхронизации, формирования и обработки сигналов. 2021. Т. 12. № 5. С. 4-13.
6. ГОСТ Р ИСО/МЭК 31010-2011 Менеджмент риска. Методы оценки риска.
7. *Котенко И.В., Саенко И.Б., Полубелова О.В., Чечулин А.А.* Технологии управления информацией и событиями безопасности для защиты компьютерных сетей // Проблемы информационной безопасности. Компьютерные системы. 2012 № 2 С. 57-68.
8. *Ерохин С.Д., Петухов А.Н., Пилюгин П.Л.* О сравнении систем защиты информации при асимптотическом управлении информационной безопасностью КИИ Труды учебных заведений связи. 2020. Т. 6. № 3. С. 66-74.
9. *С.Д.Ерохин, А.В.Ванюшина, О.И.Шелухин* Классификация IP-трафика методами машинного обучения М. Горячая линия – Телеком, 2018.
10. *Ерохин С.Д., Петухов А.Н.,* Пилюгин П.Л. Эффективность активного мониторинга событий сетевой безопасности // Электросвязь. 2020. № 2. С. 46-51.
11. OMG Data Distribution Service (DDS) Version 1.4 formal/2015-04-10 http://www.omg.org/spec/DDS/1.4
12. https://cwe.mitre.org.
13. https://cve.mitre.org.
14. *Шелухин О.И., Ерохин С.И.*, *Полковников М.В.* Технологии машинного обучения в сетевой безопасности. М.: Горячая линия – Телеком, 2021.
15. *Ерохина О.В., Борисенко Б.Б., Мартишин И.Д., Фадеев А.С.* Анализ влияния параметров многослойного персептрона на качество идентификации компьютерной атаки // Системы синхронизации, формирования и обработки сигналов. 2021. Т. 12. № 4. С. 19-26.
16. https://www.unb.ca/cic/datasets/ids-2017.html.

О ПРАКТИКЕ НОРМАЛИЗАЦИИ ГРОМКОСТИ ЗВУКА В ТЕЛЕРАДИОВЕЩАНИИ

Илюхина Инна Григорьевна,
МТУСИ, магистрант, Москва, Россия

Рихтер Сергей Георгиевич,
МТУСИ, профессор, к.т.н., Москва, Россия
sergeor@inbox.ru

Аннотация

Дана краткая характеристика измерителей уровня и громкости звука, а также вещательных процессоров, применяющихся в практике работы звукорежиссеров телерадиовещания. Обсуждаются методы обработки звука и используемое оборудование для нормализации громкости. На примере ряда телеканалов российской сети телерадиовещания анализируются методы обработки звука и сведения программ по громкости, а также используемое оборудование.

Ключевые слова: *нормализация громкости звука; вещательные процессоры; измерители уровня и громкости звука; практика сведения программ по громкости*

Введение

Нормативные документы, принятые Международным союзом электросвязи (ITU), Европейским вещательным союзом (EBU) и Минкомсвязи России [1-3] направлены на нормализацию громкости звуковых программ. Однако уровень передачи в звуковом тракте, измеренный приборами со стандартизованными или нормируемыми динамическими параметрами, не может в полной мере соответствовать субъективной оценке громкости.

В соответствии с принципами, разработанными в [2], основа для совпадения субъективного впечатления с объективным измерением уровня громкости формируется путём низкочастотной фильтрации звукового сигнала (ЗС), стробирования контента, а также использования измерителя с высокой частотой дискретизации. Рекомендовано использовать усредненный энергетический параметр ЗС – относительную среднюю мощность RMS (*Root Mean Square*) [4,5].

Все измерения проводятся сертифицированными приборами «*LoudnessMeterControl*» в режиме EBU mode [1,2]. Абсолютный уровень громкости выражается в единицах громкости LUFS (значение взвешенной громкости относительно цифровой полной шкалы), а относительный уровень – в LU (*LoudnessUnit*). Величина – 23 LUFS соответствует значению «правильной громкости для человеческого уха», о которой говорится в методике ФАС РФ.

Ниже, на примере ряда телеканалов российской сети телерадиовещания, обсуждаются методы обработки звука и используемое оборудование для нормализации громкости, а также методы сведения программ по громкости.

Основная часть

Оборудование для измерения и обработки звука. Звуковой сигнал в телерадиовещании практически всегда требует обработки, то есть корректировки его динамического диапазона, удаления шумов, а также создания различных спецэффектов, например, задержанных по времени затухающих копий этого сигнала. Основной же целью обработки, как правило, являются сугубо технические задачи: согласование параметров сигнала с характеристиками электроакустического тракта или, в частности, нормализация уровня громкости в соответствии с требованиями, определяемыми Европейским союзом вещателей (EBU).

В период, когда в России было преимущественно аналоговое телерадиовещание, почти каждый звуковой эффект создавался путем использования отдельного устройства, как правило, весьма дорогого. Была распространена аналоговая запись, а в качестве отображения уровня аудио сигнала измерительные приборы фактически показывали уровень электрического сигнала.

В настоящее время обработка ЗС производится преимущественно в цифровом виде с помощью аудио процессоров, однако часто обработка звука выполняется с помощью звуковых карт различного

назначения с использованием программных звуковых редакторов. В результате большие студии может заменить один хороший компьютер, который по возможностям превосходит их, а по стоимости оказывается кратно дешевле. Цифровая обработка ЗС удешевляет процесс и делает звукозапись доступной как профессионалам, так и широкому кругу любителей. Алгоритмы цифровой обработки звука реализуются как в программном, так и аппаратном исполнении. Именно поэтому цифровая техника уже сегодня почти полностью вытеснила из студий старую аналоговую аппаратуру. Цифровые звуковые рабочие станции – DAW (*digital audio workstation*) – это электронные или компьютерные системы, предназначенные для записи, хранения, редактирования и воспроизведения цифрового звука. Каждая система предусматривает возможность выполнения на ней законченного цикла всех работ: от первичной записи до получения готового результата – продукта.

Сведе́ние (микширование) – стадия создания из отдельных записанных треков конечной записи, следующий после звукозаписи этап создания фонограммы, заключающийся в отборе и редактировании исходных записанных треков, объединении их в единый проект и обработке эффектами. В результате сведе́ния многоканальный проект выводится в монофоническую, стереофоническую или многоканальную фонограмму, которая получает свой окончательный вид в процессе, именуемом мастерингом. Сведе́ние – не чисто технический процесс соединения различных треков, это скорее творческая деятельность, от которой зависят особенности звучания результата.

Звукорежиссёр, формируя треки, оценивает художественные качества звучания, поэтому никакие измерительные приборы не могут заменить его слуха, вкуса и опыта. Однако субъективный контроль необходимо дополнить объективным, поскольку электрические параметры сигналов должны удовлетворять жёстким техническим требованиям. Наибольшие уровни сигнала не должны превышать номинальные значения, при которых нелинейные искажения становятся заметными, наименьшие должны быть значительно выше уровня шумов и помех. Для поддержания примерно одинаковой громкости речи и музыки необходимы определенные соотношения их электрических уровней.

Для записи ЗС, сведе́ния и мастеринга необходим объективный контроль уровня сигнала, без которого невозможно точно настроить регулятор входного уровня, определить момент, когда сигнал начинает попадать в зону клиппинга, а также оценить эффективность динамической обработки и качество сведе́нного для мастеринга микса.

Оценка уровня (величины) сигнала в динамическом режиме осуществляется измерителем уровня (ИУ) в процессе формирования частей программы. Измерение уровней должно помочь поддерживать нужную громкость. Эта задача достаточно сложная, поскольку требования к ИУ для оценки громкости, во многом отличаются от требований к прибору для оценки максимальных уровней.

Несмотря на преимущество цифровых систем, аналоговые устройства обработки ЗС до сих пор используются при обработке аналоговых или архивных записей. С этой целью используются три основных типа измерителей:

– VU-метры (*Volume Units meter*) – пассивные электромеханические устройства с временем интеграции 300 мс, относящиеся к измерителям средних значений уровня.

– QPPM (*quasi-peak program meter*) – измерители квазипиковых значений, которые из-за большого времени восстановления, в некоторых случаях более 1,5 с, не способны адекватно отображать переходные процессы и изменения уровня непрерывного сигнала.

– Измерительные устройства *Dorrough Loudness Meter* – отображают как средние, так и пиковые уровни в одной итерации с чрезвычайно быстрым временем отклика. Важное достоинство этих ИУ – способность достаточно точно отражать транзиент атаки, представляющие собой начальный импульс энергии перкуссионных, т.е. ударных звуков, например, треугольника, барабана, удара молоточка по струне фортепиано, когда сигнал звучит максимально громко [6]. Транзиенты хорошо заметны на звуковой волне, с их помощью подчёркивается динамика звучания, перепады уровня звука, что повышает удовлетворенность от прослушивания музыки [7].

Показания ИУ должны соответствовать субъективной оценке громкости. При этом необходимо учитывать зависимость восприятия громкости как от параметров сигнала, так и от временных и частотных характеристик слуха [8]. Указанные характеристики слуха ввиду их сложности могут быть учтены при создании прибора лишь частично. Например, удар барабана имеет высокое пиковое значение, но устойчивый аккорд с высоким средним значением будет субъективно звучать громче, даже если его пик не будет близок к ударным инструментам. При этом важным является не точная оценка громкости сигнала с помощью прибора, а измерение уровней «по громкости» единой методикой с известной постоянной ошибкой.

Гораздо бóльшие возможности имеют современные цифровые системы. Так, измерение RMS является более точным способом определения громкости по сравнению с традиционными устройствами. RMS-метры, показывая среднеквадратическое значение выходного уровня на временном интервале около 300 мс, предназначены для лучшего отображения воспринимаемой слушателем громкости. Правильное значение RMS зависит от стиля композиции и жанра. Это требует адекватного подхода к миксу для достижения желаемого уровня RMS.

В отличие от человеческого слуха, устаревшие типы ИУ не различают частотные диапазоны. Для устранения этого недостатка, создана новая шкала измерения, в которой абсолютный уровень громкости выражается в единицах громкости LUFS (*Loudness Units Full Scale*) [9]. RMS информативен как точка отсчета для определения громкости трека, но это лишь среднее, а не точное значение. Интегрированный же LUFS показывает громкость аудиофайлов с точным представлением того, как слуховой анализатор человека воспринимает звук.

LUFS или полномасштабная шкала громкости широко используется в индустрии вещательного телевидения. Она также используется в видеоиграх и стриминговых сервисах. Некоторые игровые компании указывают, что музыкальный трек должен находиться в пределах определенных уровней LUFS, чтобы они могли лучше предсказать, как он будет соотноситься с другими звуками игры [9]. Одним из самых сложных моментов при записи музыки является соблюдение определенной меры субъективизма режиссёра при создании конечного продукта. При мастеринге звука громкие треки, цель которых передать интенсивный звук, должны иметь среднеквадратический уровень примерно от -7dBFS до -12dBFS. Для треков, не обладающих такой динамической энергией, лучше подходит уровень от -16dBFS до -18 dBFS.

Для телевизионной трансляции эфирного вещания рекомендуется использовать стандарт -23 LUFS. Также анализ расположения в масштабе LUFS упрощает настройку уровней других треков, чтобы они были согласованными между собой [10]. В настоящее время развитие IP TV привело к тому, что на разных стриминговых доменах существуют различные «корпоративные» вещательные стандарты по шкале LUFS. Домен YouTube, к примеру, требует уровень -13LUFS, а платформа iTunes -16LUFS [11].

Индикаторы для общего контроля уровня громкости используются на всех этапах производства и эфира. Помимо измерения громкости ЗС, устройства позволяют измерять уровень сигнала в каждом аудиоканале и отображать его на индикаторах с динамикой классического пикового/квазипикового студийного измерителя уровня. В государственном российском телепроизводстве нашли широкое применение индикаторы российской фирмы ЗАО «ТРАКТЪ» и TouchMonitor TM9 немецкой фирмы RTW (рис.1). Эти приборы обеспечивают индикацию множества параметров ЗС, легкое управление интерфейсом и быструю настройку [12,13].

Рис. 1. Измеритель уровня и громкости звука TouchMonitor TM9

Одно из основных устройств в современной студии – вещательный процессор, предназначенный для осуществления динамической обработки аудиофайлов при эфирном, цифровом или интернет-вещании. Процессор выравнивает уровни сигналов от разных источников и сводит их к установленным параметрам – пресетам (предустановкам) в процессе трансляции. Вещательные процессоры обеспечивают незаметное для слушателя управление различными пресетами для музыки, для новостных блоков, рекламы и т.п.

Вещательные процессоры бывают программными и аппаратными. Программные решения значительно дешевле аппаратных устройств, поэтому используются значительно чаще, особенно в

интернет-вещании [14]. В упрощенном виде процесс программной обработки сигнала состоит из следующих фаз: компрессирование, суммирование, лимитирование и ограничение амплитуды. Каждая фаза сама по себе сложна и многоэтапна.

Главное преимущество программных вещательных процессоров – относительно низкая цена. Они устанавливаются на любую рабочую станцию и характеризуются параметрами и версией операционной системы. Главный недостаток – обязательное соответствие версии операционной оболочки. К примеру, старые версии Windows непригодны, а с выходом новых версий могут возникать «подвисание» оборудования и самопроизвольные перезагрузки. Для увеличения количества каналов вещания достаточно просто докупить лицензии. Аппаратные устройства для увеличения каналов вещания требуют приобретения полноценного дополнительного оборудования.

Достоинством аппаратных вещательных процессоров является наличие решений, невозможных для осуществления простым программным способом. Например, в аппаратных процессорах есть стереокодер, который формирует для FM-вещания так называемый MPX-сигнал для передачи RDS-данных. RDS (*Radio Data System*) позволяет отображать на приемных устройствах, поддерживающих RDS, дополнительные данные: название, программу, метеоусловия, частоту вещания, телефоны для звонков, текст и прочее [15].

Для окончательной обработки звука в конце звукового канала используются программно-аппаратные системы – финалайзеры. Они представляют собой широкий спектр цифровых алгоритмов компрессии, лимитеров и эквалайзеров, объединенных в общий пакет. Программы визуально отображают изменение сигнала по мере применения обработки. Субъектом, впервые разработавшим аппаратную систему, именуемую «финалайзер», является компания TC Electronics. В нашей стране ныне используют оборудование этой фирмы в пользовательском режиме.

В настоящее время также распространены финалайзеры фирм Junger Audio и Linear Acoustic. Немецкая фирма Junger Audio выпускает различные процессоры и пакеты программного обеспечения для нормализации громкости в ТВ-вещании. Считается, что их анализаторы обладают самой высокой скоростью срабатывания. Компания Linear Acoustic запатентовала собственный алгоритм обработки AERO (рис. 2).

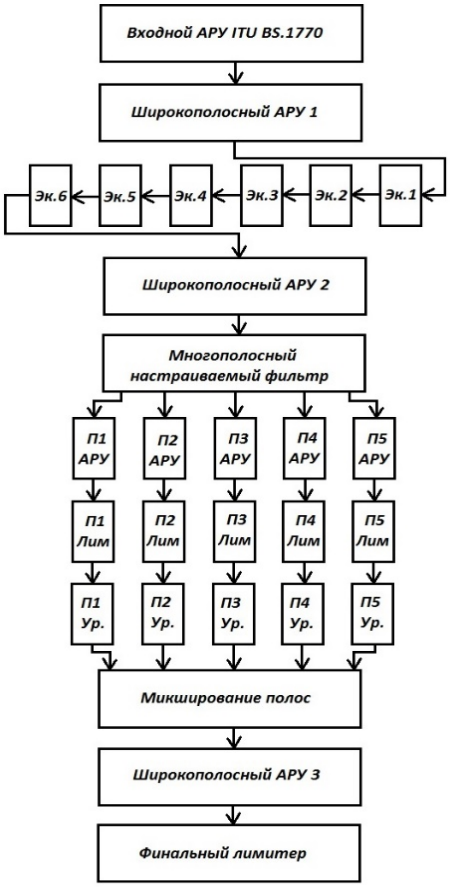

Рис. 2. Упрощенная блок-схема процессоров AERO

В его составе входной широкополосный АРУ, шесть параметрических эквалайзеров, пятиполосная секция обработки с АРУ и лимитерами в каждой полосе, широкополосный АРУ после блока многополосной обработки и выходной финальный лимитер [16]. Процессоры Linear Acoustic обеспечивают отличную разборчивость речи и хорошо справляются с различными по динамическому диапазону сюжетами.

Обработка ЗС может осуществляться в ручном или автоматическом режиме. Ручная обработка сигнала используется только на этапе подготовки материала и практически невозможна в прямом эфире. Эта операция сложна для неспециалиста, поскольку требует специальных навыков и хорошего музыкального слуха. Чаще всего регулировка уровня ЗС в ручном режиме применяется на коротких временных отрезках, а также в случае крайней необходимости, если иным способом обработать сигнал невозможно. Ручная регулировка используется также для создания авторских треков.

В настоящее время в профессиональные обязанности звукорежиссера не входит обязанность точного нормирования уровня сигнала. Более того, в учебных заведениях изучаются только граничные значения уровня, за которые не должен выходить сигнал в готовом виде.

Готовые нормалайзеры давно стали стандартным оборудованием теле- и радиостудий, а обработка сигнала происходит в автоматическом режиме. Аппаратная обработка сигнала осуществляется на всех уровнях создания контента, включая прямой эфир. Звукорежиссёр выполняет свою основную работу, на слух придавая окраску программе, а потом программа обрабатывает материал, подгоняя под нужный уровень. Однако, нередко готовый продукт по качеству ниже, чем продукт, созданный вручную, поскольку возможны искажения авторских трактовок.

Методы обработки звука, применяемые на ряде телеканалов. Телерадиовещание в России (как и в мире в целом) характеризуется высокой степенью конкурентности. Поэтому подготовка данного материала была связана с определенными трудностями. Даже такая сугубо утилитарная задача, как методы приведения аудиосигнала в соответствие с определенными параметрами, затрагивает корпоративные интересы со всеми вытекающими обстоятельствами, такими как коммерческая тайна, регламентирующие запреты, договоры о неразглашении. Государственные компании в данном случае не являются исключением. В связи с этим, говоря о практике, можно сделать лишь общий обзор, приведя конкретные примеры.

Каждому телеканалу приходится своими силами подстраиваться под новые параметры, определяемые нормативами EBU. Рассмотрим опыт телеканала «Россия Культура» по формированию программ и сюжетов с предварительной подготовкой материала.

Каждая программа проходит через руки звукорежиссера, который в специальных программах (таких как, например, Pro Tools) редактирует записанную программу и доводит её до совершенства. Звукорежиссеру поступает отснятый продукт в форме сюжета для новостей, интервью или целой программы. Это – сырой материал для работы, т.е. смонтированный видеоряд режиссером со всеми записанными звуковыми дорожками: с петличек, заставок программ, сюжетов и может быть даже запись с «пушки» т.е. микрофона в самой камере. Материал представляет собой нарезки с разных ракурсов, видов, планов и т.д. Звукорежиссёру необходимо этот разрозненный материал превратить в готовую целостную картину. Он вырезает паузы, убирает отдельные несовершенства речи спикеров и междометия говорящих, убирает шипение в паузах и интершумы, записанные, например, с микрофона камеры. И, наоборот, может продлить конечную паузу ведущего перед концом программы и началом титров, а также накладывает музыкальное сопровождение заставок.

Для удобства на этом телеканале все звукорежиссеры работают по старым стандартам, т.е. в дБ - шкалах и стараются держать уровень в -9 дБ. Далее готовый материал прогоняют через программный анализатор и переводят готовый конечный продукт с уровня -9 дБ на шкалу -23 LUFS. Таким образом, готовый продукт нормализуется автоматически, т.е. программным способом.

В основном, звукорежиссеры делают всю работу на слух. Вследствие высокого профессионализма, они сами могут выставлять с высокой точностью значение в -23 LUFS, однако главной задачей звукорежиссёра всё-таки остаётся предание нюансов и акцентов программе. Поэтому используется автоматическая нормализация для подстраховки, со сверкой программным способом готового материала с действующим стандартом.

Обратимся далее к опыту по формированию программ и сюжетов в новостном, прямом эфире без предварительной обработки материала. Когда была введена система измерений по шкале LUFS, возникли трудности, поскольку в цифровых пультах для эфиров не была предусмотрена

переустановка под другие стандарты, а закупка нового оборудования дело дорогостоящее. Поэтому звукорежиссёрам холдинга ВГТРК пришлось вручную, методом подбора на слух и с помощью анализаторов и измерителей уровня выставлять (подгонять) работу пульта к новому нормативному значению уровня. Например, в эфирной аппаратной телеканала Россия Культура для эфира «Новости культуры», эмпирическим методом было выяснено, что если выходной сигнал выставить на уровне -4.5 дБ, то программные измерители будут показывать громкость в пределах нормы.

Однако, это значение индивидуально, поскольку каждый пульт имеет собственные настройки и особенности, выставленные производителем. Кроме того, каждая телекомпания имеет свои внутренние регламенты и принятые нормативы, использует различные шкалы измерения громкости сигнала. Лишь на выходе, в конечной точке, они приводятся к общепринятым значениям.

Перед началом эфира новостей идёт просмотр материала, отобранного продюсером программы, и вручную устанавливается нужный звуковой уровень. Проверяется уровень звука с выходов микрофонов. Под каждого ведущего есть готовые настройки, которые выставляются заранее. Это происходит за полчаса до эфира, когда проходят все технические проверки и тракт-тесты или, как называют их инженеры телевидения, техпробы. За сдачу и приём сигналов звукорежиссёр расписывается в специальном журнале, фиксируя свою ответственность за готовность к эфиру.

Коммутация микрофона ведущего новостного эфира осуществляется с звукорежиссерского пульта. Для удобства звукорежиссера на одном из мониторов есть лист с таймингами программы (рис.3). Помимо отображения чередования тем видеорядов (сюжетов) и ведущих (студий) или видеоряд с закадровым чтением ведущей из студии (синхрон), отображается время до их конца/начала (фиолетовым). Красным для удобства показывают то, что идёт в эфире в данный момент.

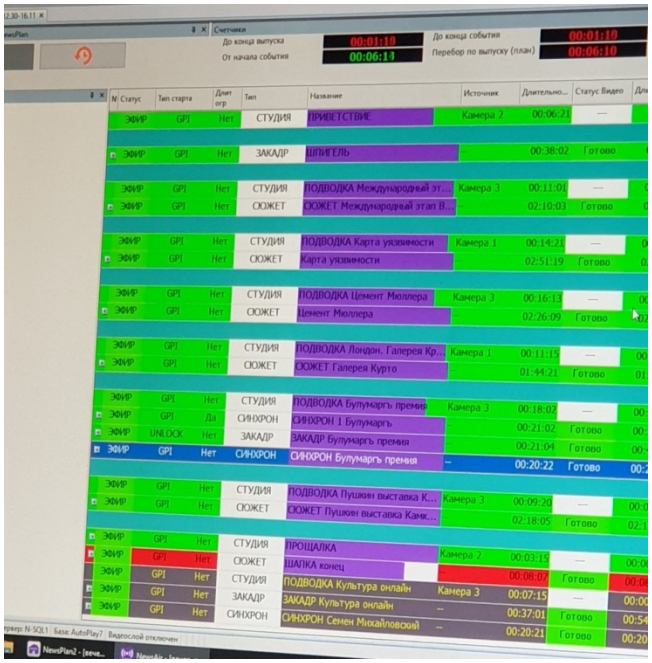

Рис. 3. Лист таймингов (таблица), в которой отображается весь эфир

Телеканалы «Россия 24» и «Москва 24», которые в основном работают в режиме новостного эфира 24/7, удобнее поставить анализатор сигнала на выходе канала. Для этого используют финалайзер. Упомянутые телеканалы используют финалайзер Linear Acoustic AERO.10 DTV Audio Processor.

В новостных студиях прямого эфира звукорежиссеры используют измерители уровня сразу в LUFS, затем сигналы проходят нормализацию и, в принципе, на вход финалайзера приходит уже готовый нормированный сигнал. Таким образом, эти приборы используются для подстраховки и для логирования (создания LOG-файла) сигнала. Логирование законодательно необходимо для отчётности и проверки. Каждый LOG-файл содержит информацию по уровню громкости в каждую секунду эфира, т.е. каждую секунду создаётся LOG-файл. Эти файлы хранятся до трёх месяцев.

Во время подготовки сюжетов, в монтажной аппаратной готовый материал отсматривают и прогоняют через анализаторы и измерители громкости. В случае, если звукорежиссёрам что-то не нравится, то они переходят к регулированию вручную. В итоге, выход студии уже нормализован программным способом. Но человеческий фактор никто не отменял, и поэтому финалайзер в конце канала необходим.

К сожалению, финалайзер Linear Acoustic AERO.10 по скорости срабатывания уступает другим, поэтому пара секунд возможна задержка, но по качеству звука он считается одним из лучших. В этом случае, жертвуя быстродействием, формируется более качественный контент.

На телеканале «Россия 1» на выходе новостной студии стоят программные компрессоры, которые «подтягивают» низкие уровни ЗС, а также лимитеры, которые срезают пики. Так как студия новостей не является студией, формирующей канал, то нормалайзера, как отдельного полноценного устройства, здесь не предусмотрено. Звукорежиссёр вручную подстраивает программу так, чтобы она попала в -23 LUFS. Т.е. на монтаже, в озвучке нормализацию не делают.

На эфире принудительно нормализацию также не делают, но жестко ограничивается лимитером высокий уровень сигнала, и компрессором низкие уровни звука автоматически подтягиваются до уровня, заранее предустановленного в настройках. Звукорежиссёр, наблюдая за интегрированной составляющей за время всего эфира, делает отдельные участки либо чуть громче, либо чуть тише. Нет смысла применять на выходе каждой студии нормалайзер, потому что на выходной канал работает много разных студий и бессмысленно нормализовать звук на каждом выходе студии, когда проще поставить его в конечной точке – в формирующей студии или на выходе канала. Поэтому на выходе канала ставится финалайзер.

На телеканале «РБК» используют регулятор громкости TC Electronic Level Pilot. Компания TC Electronic является одним из ведущих производителей приборов для аппаратной и программной обработки аудиосигнала. Особое место среди них занимает линейка вещательных процессоров на основе алгоритмов TC Electronic, которые позволяют решать различные технические проблемы теле- и радиовещания, включая нормализацию громкости в соответствии с современными стандартами. Серия Pilot считается самой лучшей по звуку, но при этом она наиболее медленная по скорости срабатывания. Особенностью моделей TC Electronic является уникальный индикатор уровня – радар (рис. 4) [17].

Рис. 4. Индикатор уровня – радар

У данного финалайзера такие радары стоят на входе и на выходе, что даёт возможность сравнивать уровни входного и выходного сигнала. Вся информация также записывается в отчёт. Данный финалайзер считается самым продвинутым из представленных на рынке. Они по звуку самые лучшие, но при этом самые медленные.

Заключение

Нормализация громкости звука программ радио и телевидения является задачей государственной важности. Принятые 6 лет назад нормативные акты, привели к необходимости согласования ведомственных нормативов и инструкций по регулированию громкости звука на радио и телевидении с требованиями международных стандартов. Строгое выполнение основных положений этих документов, измерение взвешенной громкости звука относительно цифровой полной шкалы в единицах громкости LUFS, способствуют достижению «правильной громкости для человеческого уха», о чём говорится в методике ФАС РФ.

На момент вёрстки данного материала Российская Федерация вышла из состава Европейского вещательного союза [18], в связи с чем рекомендации EBU на её территории прекратили действие в качестве основания для законодательных норм и приобрели исключительно рекомендательный характер.

Поскольку за последнее десятилетие из состава EBU вышел целый ряд крупных теле- и радиовещательных компаний нескольких государств, в настоящее время встала задача объединения национальных вещателей и, возможно, разработки новых положений для нормирования звука. Возможна также разработка в будущем национального вещательного стандарта, а также, возможно, распространение этого стандарта на вероятные новые межгосударственные союзы вещателей.

Литература

1. ITU-R BS.1770: Algorithms to measure audio programme loudness and true-peak audio level (Методика измерения громкости звуковой программы и пиковых уровней аудиосигналов)
2. EBU Technical Rec. R 128 (2011): Loudness normalisation and permitted maximum level of audio signals (Нормализация громкости и допустимый максимальный уровень громкости звуковых сигналов)
3. Приказ Минкомсвязи России №171 от 21.05.2015 г. «Об утверждении Рекомендаций в области нормирования звуковых сигналов в телерадиовещании»
4. *Попов О.Б., Рихтер С.Г.* К вопросу измерения громкости звука в телерадиовещании / Технологии информационного общества. Сборник трудов XIV Международной отраслевой научно-технической конференции «Технологии информационного общества». М.: ИД Медиа Паблишер, 2020. 580 с.
5. Компандирование сигналов в канале звукового вещания. Учебное пособие для вузов / О.Б.Попов, С.Г.Рихтер, А.Н.Терехов и др.; Под ред. профессора С.Г. Рихтера. М.: Горячая линия – Телеком, 2021. 298 с.
6. Форум специалистов по обработке звука DAStereo URL:https://www.dastereo.ru/t/chto-takoe-bystraya-sistema-i-tranzienty-naskolko-oni-vazhny-razbiraemsya-s-etim-raz-i-navsegda/109446 (Дата обращения 03.12.2021г.)
7. Портал о музыкальной обработке SameSound URL:https://samesound.ru/prod/122932-compression-basics (Дата обращения 10.12.2021)
8. *Цвиккер Э., Фельдкеллер Р.* Ухо как приемник информации. М.: Связь, 1971. 256 с.
9. EBU Tech. Doc. 3343 «Практическое руководство по производству в соответствии с рекомендациями EBU R 128».
10. Сайт производителя вещательного оборудования FMUSER International Group INC. URL: https://ru.fmuser.net/content/?2532.html (Дата обращения 16.12.2021)
11. Сайт поставщика радиовещательного оборудования Тракт-Медиа URL:https://tract.media/lufs/ (Дата обращения 24.12.2021)
12. Паспорт, техническое описание и инструкция по эксплуатации измерителя уровня и громкости звука от производителя «ТРАКТ» ТР-702 URL: https://shop.tract.ru/upload/iblock/09e/09eccaab9241a2738b3a6628895b40cf.pdf (Дата обращения 27.12.2021)
13. Технический паспорт измерителя громкости и уровня звука TouchMonitor TM9 Series от производителя RTW URL:https://www.rtw.com/fileadmin/user_upload/redakteure/07_downloads/produkte/audio_monitore/tm9/RTW_DS_20900_series_20210825.pdf (Дата обращения 04.01.2022)
14. Проект о телевидении, автоматизации и оборудовании AdView URL: https://adview.ru/cat_hardware/monitoring/veshhatelnye-processory-osnovnye-funkcii-i-modeli-omnia/ (Дата обращения 10.01.2022)
15. Проект о телевидении, автоматизации и оборудовании AdView URL: https://adview.ru/cat_hardware/monitoring-audio/programmnye-resheniya-dlya-veshhatelnyx-processorov-za-i-protiv/ (Дата обращения 10.01.2022)
16. *С. Соколов.* Нормализация громкости: цели и средства. Информационно-технический журнал MediaVision URL: http://mediavision-mag.ru/uploads/08-2017/59-72%2008_2017.pdf (Дата обращения 12.01.2022)
17. Сайт производителя оборудования TC Electronic. Индикатор уровня – радар. URL:https://www.tcelectronic.com/product.html?modelCode=HE005 (Дата обращения 13.01.2022)
18. Ресурс ВГТРК «Вести.ру», материал от 26 февраля 2022 г.(Дата обращения 02.03.2022)

ПОМЕХОУСТОЙЧИВОСТЬ И ВЫЧИСЛИТЕЛЬНАЯ СЛОЖНОСТЬ АЛГОРИТМОВ ДЕМОДУЛЯЦИИ ДЛЯ СИСТЕМ MIMO С РАЗНЫМ ЧИСЛОМ АНТЕНН

Комаров Михаил Иванович,
МТУСИ, Москва, Россия
Komarov_mi@mail.ru

Панкратов Денис Юрьевич,
МТУСИ, к.т.н., Москва, Россия
dpankr@mail.ru

Степанова Анастасия Георгиевна,
МТУСИ, ст. преп., Москва, Россия
a.g.stepanova@mtuci.ru

Чуманов Александр Евгеньевич,
МТУСИ, Москва, Россия
vectorchan@yandex.ru

Аннотация

В современных системах часто используется технология MIMO. В данной работе особое внимание уделено расчёту сложности, т.к. информация о необходимом числе вычислительных операций позволяет оценить возможности реализации различных алгоритмов демодуляции. Проведен ряд вычислений и сравнений, позволяющих определить уровень сложности и выделить алгоритм демодуляции, который удовлетворяет запросам по минимизации сложности среди таких алгоритмов, как алгоритм, оптимальный по критерию максимального правдоподобия, алгоритм Zero Forcing, а также алгоритм MMSE, также в работе производится результаты оценки помехоустойчивости этих алгоритмов. Тема, рассматриваемая в данной работе актуальна, так как рассматривается общий случай, который может быть применен для систем MIMO с любым количеством антенн.

Ключевые слова: *MIMO, помехоустойчивость, вычислительная сложность, демодулятор оптимальный по критерию максимального правдоподобия, алгоритм Zero Forcing, MMSE, алгоритмы демодуляции.*

Введение

В настоящее время использование технологии MIMO (Multiple Input Multiple Output) широко распространено во многих системах радиосвязи, например Wi-Fi,Wi-MAX, мобильная связь [1-5]. Технология MIMO, однако, не лишена недостатков – количество антенн увеличивает стоимость и сложность оборудования для цифровой обработки передаваемой и принимаемой информации. Тем не менее, эта технология позволяет благодаря применению нескольких антенн на передающей и приемной стороне существенно повысить качество связи по сравнению с системой с одной передающей и одной приемной антеннами [4-8].

1. Структурная схема и модель системы

Структурная схема системы MIMO представлена на рисунке 1.

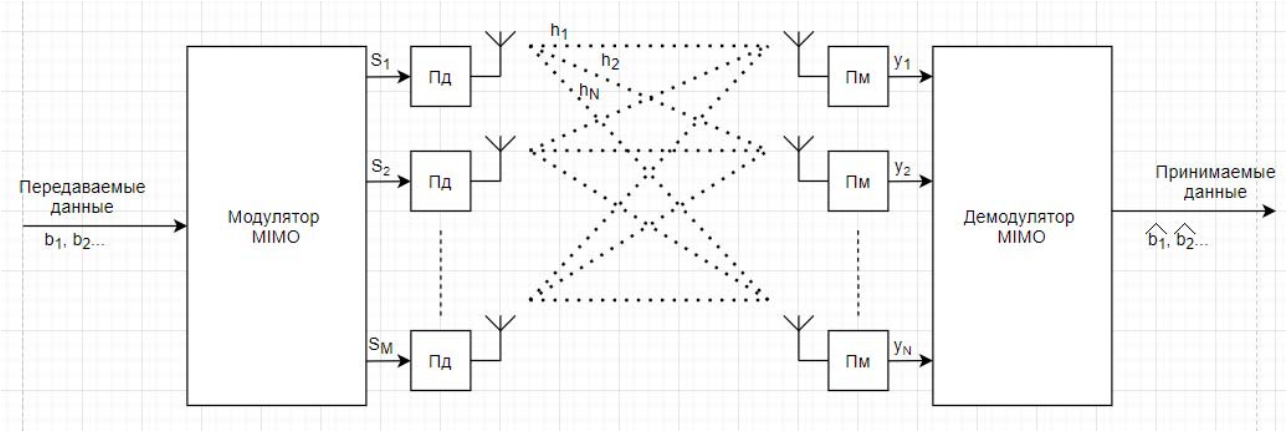

Рис. 1. Структурная схема системы MIMO

В данной систему на вход модулятора MIMO поступают передаваемые данные (биты), которые отображаются в предаваемые сигналы. При использовании в системе MIMO режима пространственного мультиплексирования данные одновременно передаются через M передающих антенн. Переданные сигналы после воздействия релеевских замираний и аддитивного белого гауссовского шума, через N приемных антенн поступают на входы демодулятора MIMO. Здесь происходит демодуляция одним из методов демодуляции, а затем происходит отображение оценок принятых информационных символов обратно в биты в соответствии с используемым видом модуляции.

Благодаря наличию эффекта многолучевости каждый из M переданных сигналов многократно переотражается от различных наземных объектов, таким образом, формируются независимые траектории прохождения сигналов [9,10]. Такой сценарий характерен условиям города, когда отсутствует прямая видимость между передающей и приемной сторонами.

Сигнал, который обрабатывается на приемной стороне, описывается формулой:

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n}, \tag{1}$$

где **y** – вектор принятых комплексных отсчетов;

H – матрица комплексных коэффициентов передачи канала MIMO;

s – вектор переданных информационных символов;

n – шум в канале связи.

Рассмотрим случай, когда на приемной и передающей стороне по 2 антенны. В таком случае модель системы MIMO (1) примет следующий вид:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \tag{2}$$

2. Алгоритмы демодуляции

В данной работе будут рассмотрены несколько алгоритмов демодуляции, а именно: алгоритм, оптимальный по критерию максимального правдоподобия (ML), алгоритм Zero Forcing (ZF), а также алгоритм, оптимальный по критерию минимума среднеквадратической ошибки (MMSE).

Алгоритм демодуляции, оптимальный по критерию максимального правдоподобия

Выражение оценки, оптимальной по критерию максимального правдоподобия (ML) можно записать с помощью формулы [3,5,9]:

$$\mathbf{S}^{\mathbf{ML}} = \underset{\hat{\mathbf{s}} \in \mathbf{S}}{\operatorname{argmin}} \| \mathbf{y} - \mathbf{H}\hat{\mathbf{s}} \|^2 = \underset{\hat{\mathbf{s}} \in \mathbf{S}}{\operatorname{argmin}} (\mathbf{y} - \mathbf{H}\hat{\mathbf{s}})'(\mathbf{y} - \mathbf{H}\hat{\mathbf{s}}) \tag{3}$$

Для нахождения с помощью данного алгоритма требуется провести полный перебор по всем возможным комбинациям вектора **s** из множества **S**. Применение данного алгоритма в реальности

возможно только при низкой кратности модуляции и малом количестве антенн в системе MIMO. Так что такой алгоритм в условиях реальности практически невыполним и слишком затратен, поэтому используются другие, более простые алгоритмы демодуляции, но у которых хуже характеристики помехоустойчивости.

Алгоритм Zero Forcing

Если в (3) поиск минимума выполнять не перебором всех комбинаций вектора **s**, а считая все его компоненты непрерывными комплексными переменными, то этот минимум может быть найден путем дифференцирования выражения $(\mathbf{y}-\mathbf{Hs})^{t}(\mathbf{y}-\mathbf{Hs})$, в результате чего получается следующая

формула [5,10,11]:

$$S^{ZF} = (H^{t}H)^{-1}H^{t}y \tag{4}$$

Выражение определяет линейный алгоритм демодуляции, известный под названием Zero Forcing (ZF) или декоррелятор. Данный алгоритм не учитывает шум наблюдения и его применение приводит к ухудшению точности оценивания с увеличением уровня шума и как следствие существенному ухудшению помехоустойчивости относительно оптимального демодулятора (ML).

Алгоритм демодуляции, оптимальный по критерию минимума среднеквадратической ошибки

В алгоритме, оптимальном по критерию минимума среднеквадратической ошибки (MMSE) оценка S^{MMSE} находится при помощи формулы [3,5,12,13]:

$$S^{MMSE} = [H^{t}H + 2\sigma_{\eta}^{2}\cdot 1]^{-1}H^{t}y \tag{5}$$

где $2\sigma_{\eta}^{2}$ – дисперсия комплексного шума наблюдения, 1 – единичная матрица размера MxM,

где M – число антенн в системе MIMO.

Данный алгоритм учитывает наличие шума в отличие от алгоритма Zero Forcing и поэтому он имеет более высокую по сравнению с ним помехоустойчивость

Моделирование системы MIMO

В результате моделирования построен график зависимости коэффициента ошибок (BER) от отношения сигнал шум (SNR) для различных алгоритмов демодуляции в системе MIMO. Данный график представлен на рисунке 2. Условия моделирования – релеевский канал (условия города), число передающих и приемных антенн M=2, двоичная фазовая модуляция, режим пространственного мультиплексирования.

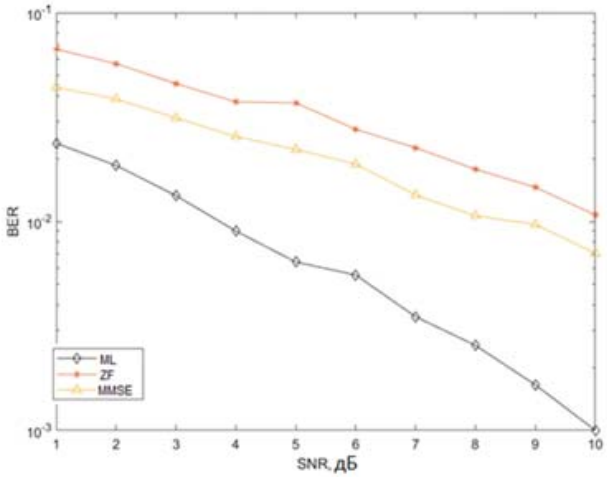

Рис. 2. Зависимость коэффициента (BER) ошибок от отношения сигнал\шум (SNR) для системы MIMO 2x2

По графику видно, что наихудшей помехоустойчивостью обладает алгоритм Zero Forcing. Лучшую помехоустойчивость имеет метод максимального правдоподобия (ML). Выигрыш MMSE по сравнению с ZF составляет порядка 1-2 дБ за счет учета влияния шума в канале связи. Кроме того, приводим результаты моделирования помехоустойчивости различных режимов системы MIMO – пространственное мультиплексирование (SM), пространственно-временное кодирование (STC) по схеме Аламоути и для сравнения режим с одной антенной (SISO) на рисунке 3.

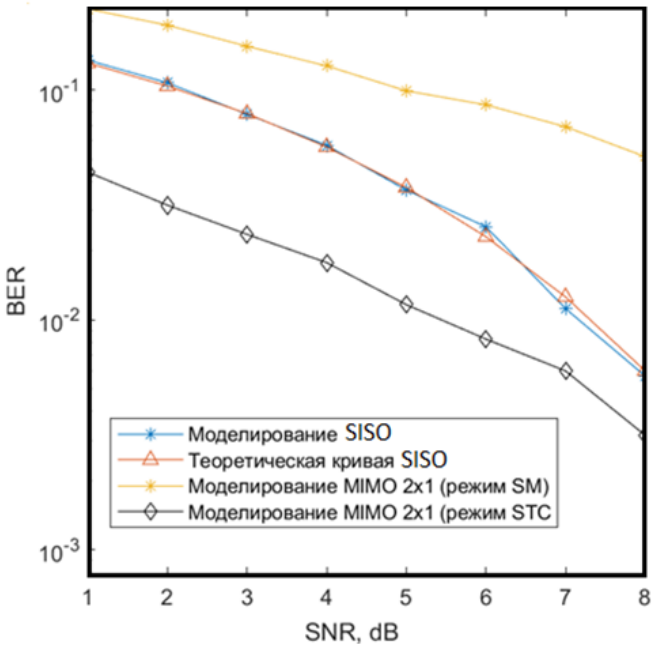

Рис.3 Зависимость коэффициента (BER) ошибок от отношения сигнал\шум (SNR) для различных режимов системы MIMO по сравнению с системой SISO

Из графика видно, что режим STC является наилучшим по помехоустойчивости, его целесообразно применять для абонентов, находящихся вдалеке от базовой станции. Режим SM лучше с точки зрения пропускной способности и его лучше применять для абонентов, находящихся вблизи базовой станции. Далее рассматривается сложность алгоритмов демодуляции для системы MIMO в режиме SM.

3. Расчет вычислительной сложности

Вычислительная сложность алгоритма демодуляции определяется числом операций, необходимых для реализации. При этом порядок сложности важен, так как он определяет возможность реализации метода демодуляции на практике. Рассмотрим расчет вычислительной сложности рассмотренных выше алгоритмов демодуляции для систем MIMO.

Алгоритм, оптимальный по критерию максимального правдоподобия

Алгоритм метода, оптимального по критерию максимального правдоподобия (ML), определяется формулой (3). Введем условие, что количество антенн на приемной стороне равно количеству антенн на передающей стороне, где M – количество антенн.

Обозначим $\mathbf{norm^2 = (y - H\hat{s})'(y - H\hat{s}) = r^2}$, где **r** - это невязка. Распишем вычисления для общего случая с M передающими и M приемными антеннами. Вектор **y** будет иметь размерность Mx1. Матрица **H** будет иметь размерность MxM. Вектор $\hat{\mathbf{s}}$ будет иметь размерность Mx1. Для вычисления вектора $\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{s}}$ потребуется умножить матрицу размерностью MxM на вектор размерностью Mx1, тогда необходимое количество операций, для выполнения: $M^2 - M$ сложений и M^2 операций умножения [5,14,15]. Для вычитания векторов необходимо сделать только же операций вычитания (сложений) сколько и элементов в векторе. В общем случае происходит

вычитание вектора размерностью Mx1 из вектора размерностью Mx1, следовательно необходимо сделать M сложений. Возведение вектора в степень 2 происходит при помощи умножения транспонированного вектора на этот вектор. В данном случае транспонированный вектор размерностью 1xM умножается на вектор Mx1, для данной операции необходимо сделать M умножений и $M-1$ операций сложения.

Всего M^2+M-1 операций сложения без учета перебора. Всего M^2+M операций умножения без учета перебора. Всего операций без учета перебора $2M^2+2M-1$. Данный алгоритм демодуляции использует перебор, поэтому общее количество операций увеличивается в 2^M раз (для двоичной модуляции). Таким образом, число операций необходимое для вычисления оценки по методу ML в общем случае можно представить как $2^M(2M^2+2M-1)$.

Для наглядности приведем данные в виде таблицы 1.

Таблица 1

Результаты анализа сложности демодулятора, оптимального по критерию максимального правдоподобия

Количество антенн	Этапы работы демодулятора, оптимального по критерию максимального правдоподобия (общее число операций для каждого этапа)				
	$\hat{\mathbf{y}}=\mathbf{H}\hat{\mathbf{s}}$	$\mathbf{y}-\hat{\mathbf{y}}$	$\|\mathbf{y}-\hat{\mathbf{y}}\|^2$	$\arg\min_{\hat{\mathbf{s}}\in\mathbf{S}}$	Общее число операций
M	$2M^2-M$	M	$2M-1$	2^M	$2^M(2M^2+2M-1)$

Алгоритм Zero Forcing

Алгоритм Zero Forcing определяется формулой (4). Рассмотрим случай, когда M=N=M. Вектор **y** будет размерностью Mx1. Матрица **H** будет размерностью MxM. Операция $\mathbf{H}'\mathbf{H}$ представляет собой перемножение матрицы размерностью MxM на матрицу размерностью MxM. Для того чтобы выполнить операцию $\mathbf{H}'\mathbf{H}$ необходимо согласно таблице 10 книги [5] сделать M^3-M^2 сложений и M^3 умножений, всего $2M^3-M^2$ операций. Операция $\mathbf{T}=(\mathbf{H}'\mathbf{H})^{-1}$ согласно таблице 12 книги [5] для обращения матрицы размера MxM требуется умножений $\frac{M^3}{2}+\frac{M^2}{2}-M$, сложений $\frac{M^3}{2}-\frac{M^2}{2}$, делений M. Всего требуется M^3 операций. Рассмотрим операцию $\mathbf{Y}=\mathbf{H}'\mathbf{y}$ по перемножению матрицы размерностью MxM на вектор размерностью Mx1. Для данной операции необходимо сделать согласно таблице 10 книги [5] M^2-M сложений и M^2 умножений.

Последняя операция умножения **T** на **Y**, представляет собой умножение матрицы размерностью MxM на вектор размерностью Mx1. Данная операция осуществляется за M^2 умножений, сложений M^2-M.

Подведем итог по каждому виду операций и общему количеству операций для данного метода в случае с M передающими и M приемными антеннами.

Всего $3M^3+3M^2-2M$ операций.

Для лучшего восприятия представим данные в виде таблицы 2.

Таблица 2

Результаты анализа сложности алгоритма ZF

Количество антенн	Этапы обработки алгоритма ZF (общее число операций для каждого этапа)				
	$\mathbf{Y}=\mathbf{H}'\mathbf{y}$	$\mathbf{H}'\mathbf{H}$	$\mathbf{T}=(\mathbf{H}'\mathbf{H})^{-1}$	$\mathbf{TY}$	Общее число операций
M	$2M^2-M$	$2M^3-M^2$	M^3	$2M^2-M$	$3M^3+3M^2-2M$

Алгоритм демодуляции, оптимальный по критерию минимума среднеквадратической ошибки

Алгоритм демодуляции, оптимальный по критерию минимума среднеквадратической ошибки (MMSE) определяется формулой (5):

Рассчитаем сложность данного алгоритма. Рассмотрим общий случай M передающих и M приемных антенн. Вектор **y** будет иметь размерность Mx1. Матрица **H** будет иметь размерность MxM.

Операция $\mathbf{H}'\mathbf{H}$ представляет собой перемножение матрицы размерностью MxM на матрицу размерностью MxM. Для того что бы это выполнить, согласно таблице 10 книги [5] необходимо сделать сложений $M^3 - M^2$ и умножений M^2. Всего: $2M^3 - M^2$. Операция $\mathbf{H}'\mathbf{y}$ представляет собой умножение вектора MxM на вектор Mx1. Для вычисления, необходимо сложений $M^2 - M$,умножений M. Всего: $2M^2 - M$. Операция $\mathbf{H}'\mathbf{H} + 2\sigma_\eta^2 \cdot \mathbf{1}$ представляет собой поэлементное сложение элементов, находящихся на главной диагонали матрицы размерностью MxM, и элементов главной диагонали матрицы размерностью MxM. Данная операция осуществляется в M сложений. Операция $\mathbf{T}^{\mathbf{MMSE}} = \left[\mathbf{H}'\mathbf{H} + 2\sigma_\eta^2 \cdot \mathbf{1}\right]^{-1}$, согласно таблице 12 книги [5] обращения матрицы размера MxM требует умножений $\frac{M^3}{2} + \frac{M^2}{2} - M$, $\frac{M^3}{2} - \frac{M^2}{2}$ сложений, M делений, итого M^3 операций. Операция $\mathbf{T}^{\mathbf{MMSE}}\mathbf{Y} = \left[\mathbf{H}'\mathbf{H} + \mathbf{2\sigma_\eta^2 \cdot 1}\right]^{-1} \mathbf{H'y}$, представляет собой умножение матрицы размерностью MxM на вектор размерностью Mx1. Для расчёта, данной операции необходимо сделать $M^2 - M$ операций сложения и M^2 операций умножения, итого $2M^3 - M^2$ операций. Для наглядности введем данные в таблицу 3.

Таблица 3

Результаты анализа сложности алгоритма MMSE

Количество антенн	Этапы обработки алгоритма MMSE (общее число операция для каждого этапа)					
	$\mathbf{H}'\mathbf{H}$	$\mathbf{Y} = \mathbf{H'y}$	$\mathbf{H}'\mathbf{H} + \mathbf{2\sigma_\eta^2 \cdot 1}$	$\mathbf{T}^{\mathbf{MMMSE}}$	$\mathbf{T}^{\mathbf{MMSE}}\mathbf{Y}$	Общее число операций
M	$2M^3 - M^2$	$2M^2 - M$	M	M^3	$2M^2 - M$	$3\mathrm{M}^3 + 3\mathrm{M}^2 - \mathrm{M}$

Для того чтобы получить общую картину для всех рассмотренных алгоритмов демодуляции обобщим значения в сводную таблицу 4.

Таблица 4

Результаты анализа сложности алгоритмов демодуляции

Алгоритм демодуляции	Общее число операций
ML	$2^M(2M^2 + 2M - 1)$
ZF	$3\mathrm{M}^3 + 3\mathrm{M}^2 - 2\mathrm{M}$
MMSE	$3\mathrm{M}^3 + 3\mathrm{M}^2 - \mathrm{M}$

Для визуализации полученных результатов построим графики зависимости сложности алгоритмов демодуляции от количества антенн для случаев максимального числа антенн 6 и 40.

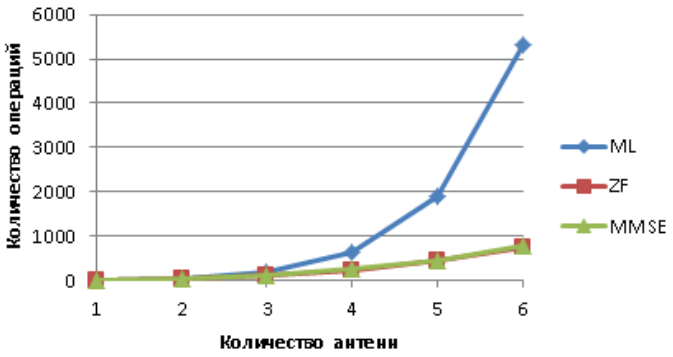

Рис.4 Зависимость сложности при малом количестве антенн

Чтобы лучше оценить, приведем некоторые конкретные числовые значения, так, например, вычислительная сложность для 5 антенн на приемной стороне и на передающей стороне составляет ML – 1888, ZF – 440, MMSE – 445.

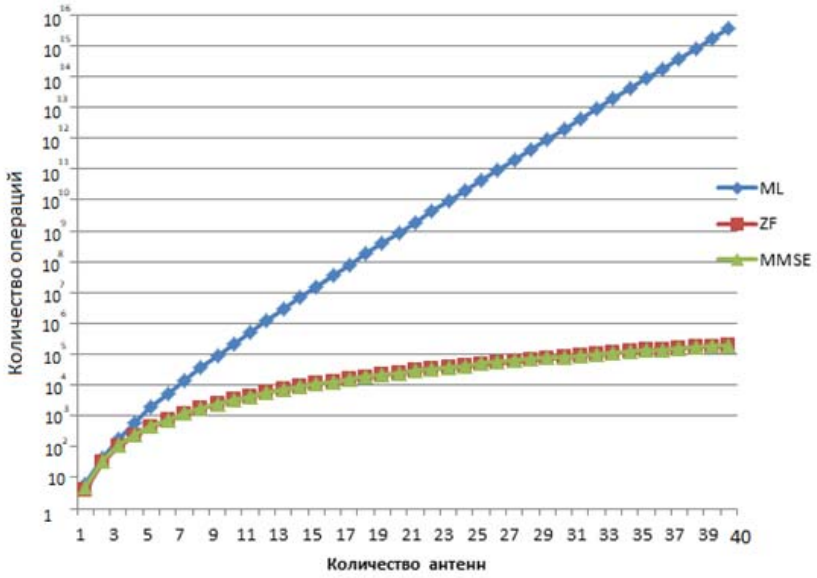

Рис. 5. Зависимость сложности при большом количестве антенн

Для случая 40 антенн на приемной стороне и на передающей стороне, количество операций для алгоритма ML - $3{,}6 \cdot 10^{15}$, ZF – 196 720, MMSE – 198 320.

Выводы

Лучшей помехоустойчивостью обладает алгоритм демодуляции, оптимальный по критерию максимального правдоподобия (ML), но он обладает наиболее высоким, из рассмотренных алгоритмов порядком сложности, в связи с тем, что в данном алгоритме осуществляется перебор по всем возможным комбинациям вектора информационных символов [14-27]. Число таких комбинаций экспоненциально зависит от размера вектора $\hat{\mathbf{s}}$, который зависит от числа передающих антенн. Таким образом, алгоритм ML имеет экспоненциальную сложность порядка 2^M, где M – число антенн. Применение этого алгоритма, как показывает проведенный расчет, возможно только при небольшом числе антенн в системе MIMO.

Было проведено моделирование помехоустойчивости алгоритма демодуляции ML наряду с алгоритмами ZF и MMSE, из которых наглядно видно, что наиболее устойчивым является алгоритм ML, а наименее помехоустойчивым – ZF. Алгоритмы демодуляции ZF и MMSE обладают примерно схожей сложностью. Как видно из расчета эти алгоритмы демодуляции имеют полиномиальную

сложность порядка M^3, где M – число антенн. Однако алгоритм MMSE сложнее алгоритма ZF на M операций. При этом при моделировании было показано, что алгоритм MMSE обладает лучшей помехоустойчивостью по сравнению с алгоритмом ZF (выигрыш порядка 1- 2 дБ). Таким образом, для систем MIMO с большим числом антенн приоритетно использование алгоритмов с полиномиальной вычислительной сложностью, например MMSE.

Литература

1. *Abu-Rgheff, Mosa Ali.* 5G physical layer technologies. UK: JohnWiley & Sons, 2020, 579 p.
2. *Harri Holma,Antti Toskala,Takehiro Nakamura* 5G Technology: 3GPP New Radio ISBN: 978-1-119-23629-0 2019. 536 p.
3. *Hanzo Lajos* and others, MIMO-OFDM for LTE, WiFi and WiMAX: Coherent versus Non-coherent and Cooperative Turbo-Tranceivers / Lajos Hanzo, Yosef (Jos) Akhtman, Li Wang, Ming Jiang //John Wiley & Sons, 2010, 692 p.
4. *G. J. Foschini and M. J. Gans*, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas, Wireless Personal Communication, Vol. 6, № 3, Mar. 1998, p. 311.
5. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.,* Технологии в системах радиосвязи на пути к 5G. М.: Горячая линия – Телеком, 2018. 280 с.
6. *Ngo H.Q.* Massive MIMO: Fundamentals and System Designs. Linköping University Electronic Press, 2015. 301 p.
7. *Manish Mandloi, Devendra Gurjar, Prabina Pattanayak, Ha Nguyen*. 5G and Beyond Wireless Systems. PHY Layer Perspective. Singapure: Springer, 2021, 425 p.
8. *Burg A, Borgmann M, Wenk* M et al (2005) VLSI implementation of MIMO detection using the sphere decoding algorithm. IEEE J Solid-State Circuits, № 40(7), pp. 1566-1577.
9. *Rusek F, Persson D, Lau BK* et al (2012) Scaling up MIMO: opportunities and challenges with very large arrays. Sig Process Mag IEEE, № 30(1), pp. 40-60.
10. *Ермолаев В.Т., Флаксман А.Г.* Теоретические основы обработки сигналов в беспроводных системах связи. Нижний Новгород: Изд-во ННГУ им. Н.И. Лобачевского, 2011. 368 с.
11. *Панкратов Д.Ю., Степанова А.Г.* Компьютерное моделирование технологии MIMO для систем радиосвязи // T-Comm: Телекоммуникации и транспорт. 2018. Том12. №12.С.33-37.
12. *Bjornson E., Larsson E.G., Martezza T.L.* Massive MIMO: ten myths and one critical question // IEEE Communications Magazine. 2016. Vol. 54, issue: 2, pp. 114-123.
13. *Pankratov D.Yu., Stepanova A.G.* Linear iterative demodulation algorithm for mimo systems with large number of antennas B сборнике: Proceedings of the International Conference Technology & Entrepreneurship in Digital Society (TEDS). Proceedings of the International Conference. 2019. С. 61-64. DOI: https://doi.org/10.17747/TEDS-2018-61-64
14. *Бакулин М.Г., Варукина Л.А., Крейнделин В.Б.,* Технология MIMO: принципы и алгоритмы. М.: Горячая линия – Телеком, 2014, 244 с.
15. *Damen M. O., El Gamal H., Caire G.* On Maximum-Likelihood Detection and the Search for the Closest Lattice Point // IEEE Trans. Info. Theory, Vol. 49, №10, Oct. 2003, pp. 2389-2402.
16. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Анализ пропускной способности канала MIMO в условиях замираний // Системы синхронизации, формирования и обработки сигналов. 2018. Т. 9. № 2. С. 13-20.
17. *Крейнделин В.Б., Старовойтов М.Ю.* Повышение помехоустойчивости системы связи MIMO с пространственным мультиплексированием методом додетекторного сложения // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 4. С. 4-13.
18. *Бакулин М.Г., Крейнделин В.Б.* Проблема повышения спектральной эффективности и емкости в перспективных системах связи 6G // T-Comm: Телекоммуникации и транспорт. 2020. Т. 14. № 2. С. 25-31.
19. *Крейнделин В.Б., Резнёв А.А.* Матрица пространственно-временного кода высокой размерности типа "Голден" // T-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 6. С. 34-40.
20. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Исследование вероятностных моделей радиоканала MIMO с учетом взаимной корреляции передающей и приемной сторон с помощью компьютерного моделирования // REDS: Телекоммуникационные устройства и системы. 2017. Т. 7. № 1. С. 64-68.
21. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю*. Алгоритмы нелинейной фильтрации двоичной ЛРП со случайной задержкой и случайной начальной фазой // Системы синхронизации, формирования и обработки сигналов. 2019. Т. 10. № 2. С. 45-51.
22. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Методы приема псевдослучайных последовательностей в системах радиосвязи // REDS: Телекоммуникационные устройства и системы. 2018. Т. 8. № 1. С. 108-112.

23. *Крейнделин В.Б., Григорьева Е.Д.* Анализ быстрого алгоритма умножения матриц и векторов для банка цифровых фильтров // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 1. С. 4-10.

24. *Бакулин М.Г., Бен Режеб Т.Б.К., Крейнделин В.Б., Смирнов А.Э.* Способы минимизации объёма передаваемой информации в обратном канале многоантенных систем MIMO // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 3. С. 17-24.

25. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Применение технологии MIMO в современных системах беспроводной связи разных поколений // T-Comm: Телекоммуникации и транспорт. 2021. Т. 15. № 4. С. 4-12.

26. *Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю.* Исследование вероятностных моделей радиоканала MIMO с учетом взаимной корреляции передающей и приемной сторон с помощью компьютерного моделирования // REDS: Телекоммуникационные устройства и системы. 2017. Т. 7. № 1. С. 64-68.

27. *Крейнделин В.Б., Григорьева Е.Д.* Реализация банка цифровых фильтров с пониженной вычислительной сложностью // T-Comm: Телекоммуникации и транспорт. 2019. Т. 13. № 7. С. 48-53.

СФЕРЫ ПРИМЕНЕНИЯ МАШИННОГО ЗРЕНИЯ В ПРОМЫШЛЕННОСТИ

Мамрега Валерий Викторович,
Частный исследователь машинного зрения, генеральный директор ОсОО «Дарбаза-Автоматик», г. Бишкек, Кыргызстан
mamregavv@gmail.com

Аннотация

В данной статье рассматривается перспективы и преимущества внедрения компьютерного зрения на основе нейронных сетей на производствах. Приводится статистика несчастных случаев на производстве, высокие показатели которой обусловлены большими объемами производства и устаревшей системой мониторинга соблюдения правил техники безопасности и наличия средств индивидуальной защиты сотрудников. Рассмотрена схема взаимодействия компонентов системы компьютерного зрения, которая позволит производить мониторинг событий, происходящих на производстве в период эксплуатации, отслеживать ситуацию на предприятии на предмет возникновения потенциально опасной ситуации для персонала и оборудования, и, соответственно, данная система будет способна предотвратить аварийную ситуацию, а также избежать получения травм персоналом, реагируя даже при незначительном отклонении от рабочих параметров. Исследования проведены на основе изучения и анализа материалов, опубликованных в открытых информационных источниках.

Ключевые слова: *производство, компьютерное зрение, соблюдение техники безопасности, мониторинг, процесс производства, промышленная безопасность*.

Введение

Рост механизации во всех экономических отраслях – начиная от сельского хозяйства, строительной и энергетической индустрии, заканчивая торговлей и банковским делом, оказывает решающее влияние на повышение роли промышленности в современном мире.

Например, роль промышленности в секторе сельского хозяйства определяется обеспечением высокого качества и объемов производимых комбикормов, которые можно достичь только при помощи высокоточного оборудования.

В энергетическом секторе основным оборудованием являются динамические типы устройств, такое как компрессоры, газовые турбины, насосы, что, соответственно влечет высокие требования по мониторингу рабочих параметров и безопасную эксплуатацию машин, что обеспечивается, опять же, за счет высокоточного оборудования.

То же самое касается и большого количества финансовых операций, которые, как в сфере торговли, так и в банковском деле, требуют колоссальных вычислительных мощностей, что достигается за счет использований вычислительных систем и серверов, способных обрабатывать и сохранять цифровую информацию.

Все это говорит о востребованности развития механизации не только в каждой отрасли материального производства, но и в каждой сфере современной жизни общества.

В настоящий момент, одни из самых революционных открытий в развитии механизации принадлежат технологии компьютерного зрения, позволяющего автоматизировать и упростить работу многих отраслей производства, медицины, образования и организации безопасности.

Существует несколько значимых и на данный момент до конца не решенных задач, которые ставит перед собой технология компьютерного зрения. В контексте темы материальной промышленности можно выделить конкретные из них:

В сфере промышленности выделяется необходимость перехода от цифрового управления отдельной установки к управлению всем производственным процессом, а в целях улучшения безопасности мониторинга оборудования и сбора данных. Одной из задач является внедрение для упрощения и автоматизации процесса производства виртуальных сенсоров, материального баланса и системы управления снабжением. Помимо прочего, технологии компьютерного зрения

так же способствуют решению таких задач, как предсказание спроса, автоматическое управление и беспилотное оборудование.

Основной практической особенностью компьютерного зрения является его способность к автораспознаванию, выбору зоны интереса, возможность коррекции и дообучения, интеграция.

Вклад данной работы в направление исследования компьютерного зрения – это системный и продуктивный подход внедрения технологии на производство для решения многообразных задач с возможностью снижения стоимости единичных внедрений, который позволит не только фиксировать отклонения от рабочих параметров, но и производить мониторинг событий, происходящих на производстве в период эксплуатации, отслеживать различные опасные для жизни ситуации и соответственно, предотвращать аварии, получения травм персоналом.

Цель исследования данной статьи – обзор видения темы в рамках промышленной сферы, рассмотрение перспектив и преимуществ внедрения компьютерного зрения на основе нейронных сетей на производство, формирование основных задач технологии и актуальных методов их решения. **Метод исследования** – изучение и анализ статистики и материалов, опубликованных в открытых информационных источниках и предшествующих научных трудах.

Первые научные труды, раскрывающие наиболее подробно задачи и технические особенности компьютерного зрения, появились еще в 80-х годах прошлого века. Один из них – книга «Цифровая обработка изображений» У. Прэтта [1, 2], которая освещает вопросы математического представления непрерывных и дискретных изображений, алгоритмы улучшения изображений, методы количественного описания, реставрацию и анализа изображений за счет способов цифрового кодирования.

Основы теории о сферах применения технологии алгоритмов цифровой обработки изображения изложил Т.С. Хуанг [3]. Современные исследования не ограничиваются научно-исследовательскими вузовскими работами и монографиями узконаправленной темы, помимо этого можно выделить книги Р. Гонсалеса – «Цифровая обработка изображений» [4], Д.А. Форсайта «Компьютерное зрение. Современный подход» [5] и многие другие работы, иллюстрирующие основные определения и понятия дисциплины компьютерного зрения, методологии цифровой обработки видеоинформации.

Таким образом, **новизна** данной работы отражается отсутствием узконаправленных исследований в совокупности с необходимостью понимания современных мировых тенденций и их практического применения в развитии технологии промышленности и непрерывностью повышения качества и изучения устройств с учетом перспективных задач их развития в будущем.

Рост интереса в научно-техническом сообществе к теме компьютерного зрения можно подтвердить статистикой по количеству участников конференции международной ICCV. Известно, что до 2009 года количество исследовательских работ данной темы не превышало 1500, а в 2017 году возросло почти в два раза. На сегодняшний день, начиная с 2019 года, заинтересованность отразилась в количестве участников конференции до 8000, и до сих пор значение исследователей и научных работ сильно увеличивается. Из этого следует факт востребованности темы компьютерного зрения, который отлично иллюстрирует **актуальность** представленной работы.

Необходимость непрерывного повышения качества технологий машинного зрения и ее адаптация для процесса материального производства не только способствуют решению поставленных задач, но и ведут за собой прогресс в создании более новых и современных технологий.

Материалы, исследование и результаты

Роль промышленности в строительной индустрии определяется объемами производимых конструкций и качеством строительного сырья, которое можно достичь, опять же, только с помощью высокоточного оборудования, позволяющего осуществлять качественную сборку элементов конструкций, поддержание заданных соотношений и требований качества компонентов при подготовке строительных смесей.

Таким образом, промышленность характеризует собой основную сферу материального производства, в области которой создается большая доля ВВП и национального дохода. Именно поэтому такая область как промышленное производство на сегодняшний день действительно

нуждается в автоматизации рабочего процесса с огромным количеством механических действий персонала – это крайне важно даже для распознавания дефектов на конвейере, не говоря уже о контроле безопасности сотрудников предприятия.

При такой важной роли промышленности в современном мире и скорости ее прогресса, все больше предъявляется требований по объемам производимой продукции, ее качеству и безопасности. Стандартных ресурсов, таких как систем видеонаблюдения, компьютерных программ и человеческого мониторинга уже недостаточно, а осуществление найма дополнительного персонала за наблюдением определенных параметров производства экономически нецелесообразно для предприятий.

Поэтому в настоящий момент ведется разработка и оптимизация системы, которая позволит производить мониторинг событий, происходящих на производстве в период эксплуатации, отслеживать ситуацию на предприятии на предмет возникновения потенциально опасной ситуации для персонала и оборудования, и, соответственно, данная система будет способна предотвратить аварийную ситуацию, а также избежать получения травм персоналом, реагируя даже при незначительном отклонении от рабочих параметров. Данная система называется системой компьютерного (машинного) зрения.

Машинное или компьютерное зрение (Computer Vision или CV) – это автоматическая фиксация и обработка как неподвижных, так и объектов, которые находятся в движении, с помощью компьютерных систем и комплекса специального оборудования. Компьютерное зрение и машинное зрение – термины немного разные, поскольку технология машинного зрения больше относится к технике, к видеокамерам для видеонаблюдения, веб-камерам для коммуникации и т.д.

Машинное зрение на настоящий момент активно внедряется и совершенствуется во многих отраслях, где необходим постоянный мониторинг большого количества информации для обеспечения высокого качества производства и уровня безопасности персонала, см. рис.1.

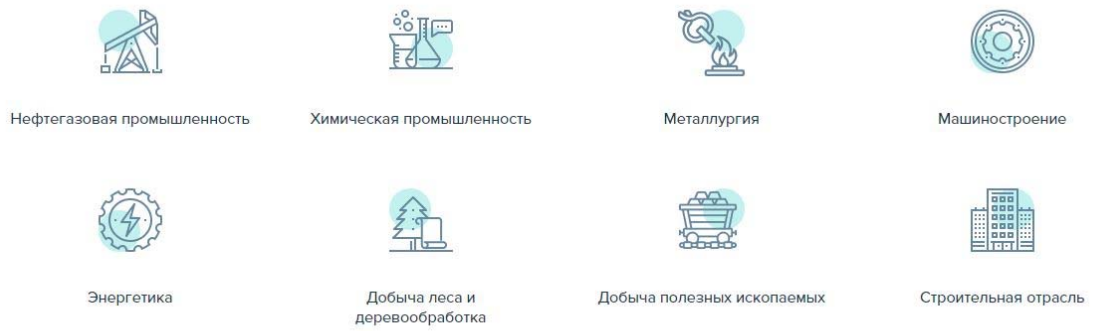

Рис. 1. Основные сферы применения систем компьютерного зрения

В рамках производственной сферы, машинное зрение может быть применимо для решения множества актуальных вопросов – безопасность, контроль качества и процесса работы и т.д. Среди главных задач машинного зрения отметим следующие, применимые в сфере промышленности:

Контроль качества и оценка брака на производстве. Ручной отбор поврежденных, неправильно сделанных или не соответствующих техническим установкам деталей порой предприятию обходится в крупную сумму. К тому же, оценщику брака сложнее обнаружить дефект и быстро принять правильное решение в отличие от автоматизированной системы. Технологии нейросети с большим процентным соотношением справятся с этой задачей – они настроены под автоматическое обнаружение дефектов: например, несоответствие физического размера изделия указанным, наличие брака, отсутствие необходимых конкретных деталей и наклеек, несоответствие цвета и чистоты изделия.

Помимо этого, автоматическая сортировка позволяет убрать с конвейера несоответствующие бракованные детали, не подпуская их к последующему этапу производства – откуда неисправное изделие уже может попасть к потребителю.

К решению задач контроля качества в контексте дисциплины существует два подхода. Классический, выделяемый ранее многими исследователями, заключается в основных способностях машинного зрения – это выделение по цвету, адаптивная бинаризация, морфологические преобразования, повышение контраста изображения т.д. Работа над алгоритмами изображения позволяет камерам видеонаблюдения автоматически фиксировать установленные данные – цвет, качество, размер изделия и наличие у него дефектов.

Второй способ, находящийся сейчас в прогрессирующем развитии – это нейронные сети. Этот способ в области контроле качества выделяется своей уникальной адаптивностью, и позволяет достичь высоких результатов в решении множества задач широкого спектра, касающегося и производства. Появляется возможность настройки алгоритма обучение нейросети, установки для системы конкретных задач. Более того, абсолютно разные задачи можно решить одним инструментом.

Единственное, что может тормозить или вывести в неэкономичный расход такую технологию машинного зрения - это недостаток данных, информации. Поскольку, чтобы выявлять дефекты, нейронных сети требуется видеопоток, который снимает готовую продукцию и демонстрирует возможные дефекты. Такой инновационной системе на сегодняшний день для должной функциональности и эффективности необходимо наличие множества примеров и генерация рабочих процессов.

Безопасность и контроль рабочего процесса. Еще одна популярная область задач, под которую заточена технология компьютерного зрения – это контроль соблюдения установленных норм и, соответственно, безопасности сотрудников на производстве. Характеризуется область контролем соблюдения зон, контролем периметра ношения касок, задачами видеоаналитики, детекцией касок и защитных костюмов.

В целях безопасности, на предприятиях требуется от сотрудников точное следование регламенту и инструкциям как, в целом, при исполнении должностных обязанностей, так и при выполнении определенного процесса работы. Нарушения подобных правил могут кому-то показаться мелочными и безобидными, но в ином случае и в иных обстоятельствах – это опасное для жизни действие.

Технологии машинного зрения на основе нейронных сетей позволяют отследить, были ли совершены подобные нарушения сотрудниками предприятия, были ли использованы необходимые средства защиты, соблюдалась ли установленная дистанция до опасных объектов и нужные рабочие режимы, не осуществлялась ли неправильная обработка на производстве. Кроме того, у большинства предприятий регламент предусматривает и необходимость выполнения конкретных повторяющихся действий, чтобы те, при должном выполнении, привели к большей экономии на издержках.

Следовательно, одной из самых распространенных функций систем компьютерного зрения на производствах является идентификация наличия средств индивидуальной защиты сотрудников, осуществляющих свою рабочую деятельность на потенциально опасных участках. Согласно оперативным данным министерства труда РФ в 2019 году, произошло 4078 случаев с тяжелыми и смертельными исходами на производственных объектах в результате несоблюдения техники безопасности.

Внедрение же систем компьютерного зрения позволит осуществлять постоянный мониторинг соблюдения требований безопасности, а именно определение, где на изображении, полученном в любой момент времени, у человека находится голова, руки и ноги и, самое главное, выявить замечания на предмет наличия средств защиты. Данная система производит анализ, надеты ли перчатки, использует ли сотрудник защитную каску и очки, есть у него с собой портативный газоанализатор, застегнута ли спецодежда и т.д., предупредив о нарушении работника и его руководство. По предварительным расчетам, внедрение систем компьютерного зрения на производстве позволит снизить количество несчастных случаев на 30-35%.

Для предприятия данная система позволяет обеспечить следующие преимущества:

– снизить уровень инцидентов на производстве, связанных с нарушением требований техники безопасности;

– моментально идентифицировать нарушение и сразу же информировать о данном нарушении работника и его руководство;

– охватить большую территорию для мониторинга сравнительно малыми ресурсами.

Так же, компьютерное зрение в значительной степени способно упростить работу специалистам по охране труда – борьба с хищением имущества предприятия, подложными

деталями на конвейере, неправильным перемещением товара по складу, что позволяет им сосредоточиться непосредственно на важных моментах. Для компаний наличие данной системы позволяет сократить затраты и повысить уровень безопасности на производстве.

Специалист по охране труда физически не может находиться на нескольких производственных объектах или в помещениях для осуществления проверки соблюдения выполнения правил по технике безопасности и охране труда. Разумеется, в случае отсутствия на предприятии систем машинного зрения приводит к тому, что специалисту необходимо большее количество времени для обнаружения нарушителей, что, соответственно, влечет возникновение потенциально опасной ситуации для работников.

А отсутствие на производстве данной системы предполагает, что специалисты будут тратить свое время на постоянный мониторинг ситуации на производстве на постоянный просмотр видео, получаемых с камер видеонаблюдения, но для просмотра видео с каждой установки, с каждого помещения требуется большое количество специалистов, что является не выгодным для предприятий с финансовой точки зрения. Схема взаимодействия компонентов системы машинного зрения представлена на рисунке 2.

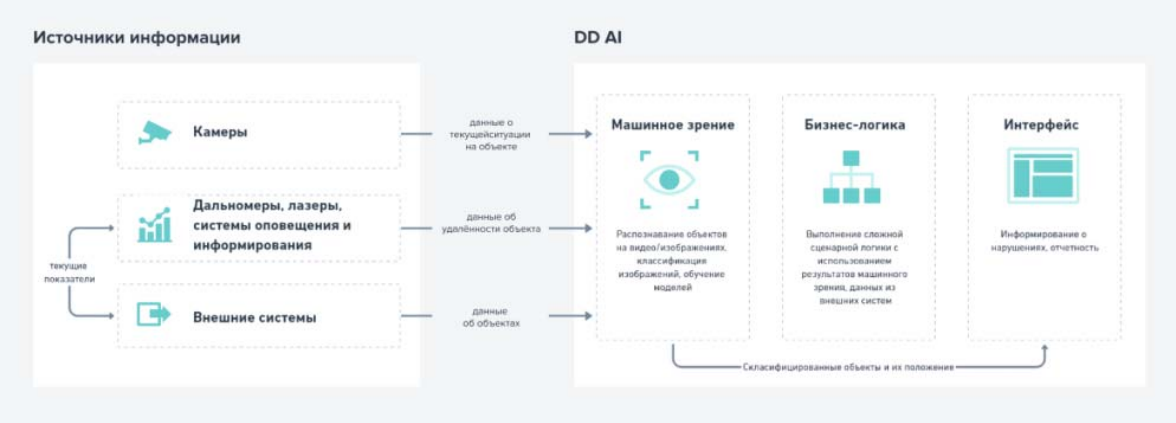

Рис. 2. Схема взаимодействия компонентов системы компьютерного рения

Кроме области промышленной безопасности, компьютерное зрение способно осуществлять мониторинг и контроль производственных параметров, отслеживая их изменения и регулируя протекание процессов на производстве. Задачи склада, учета и логистики – контроль склада, считывания штрих-кодов и контроля движения товара заметно облегчают работу кладовщиков. Аналогична и задача цифровизации и подключение разного оборудование без наличия цифровых интерфейсов – критичная необходимость регистрации процесса дорогостоящей и опасной работы, которая без технологии машинного зрения не была бы осуществлена должным образом.

Проблема использования машинного зрения в качестве системы видеонаблюдения на производстве – это необходимость приобретения неэкономного оборудования с хорошим разрешением. А проблема использования машинного зрения на основе нейронной сети – это обязательное требование соблюдения множества шагов настройки системы, таких как создание разметки и обучения, выбор модели и тюнинг, тестирование, сглаживание результатов.

Заключение

В данной статье были приведены характеристики, особенности и функции технологий машинного зрения, проанализированы известные в исследовательской среде научные работы, проиллюстрированы и приведены данные о сферах применения и статистике, конкретизированы преимущества использования машинного зрения в промышленной сфере, что позволило ответить на поставленные изначально вопросы.

На данный момент развитие технологий искусственного зрения стремительно развивается во многих областях жизни человека – начиная от оптимизации работы мобильных устройств и заканчивая обеспечением безаварийной работы производств, а также прорывами в области медицины. Современный мир ставит множество вопросов и задач перед учеными в плане реализации проектов, основанных на применении компьютерного зрения, которые, разумеется, нужно решать, учитывая потенциал применения данной технологии.

В заключении, можно сделать вывод, что технологии компьютерного зрения на основе нейронной сети все еще успешно развиваются, эволюционируют и совершенствуются в общем контексте развития технических и системных средств, возможностей научного прогресса, и пребывать в подобном состоянии будут еще долгое время. Что касается предпосылок и практического применения технологии в будущем, абстрагировавшись от понятий технологического прогресса или технической революции, можно сказать, что не всегда открытия такого масштаба происходили последовательно в результате длительной эволюции. На это могли влиять многие факторы: от востребованности в определённых сферах применения до соотношения с запросами потребителя в определённый момент прогресса.

А при должной информационной освещенности данной темы, владельцам предприятий в сфере материального производства станет понятно, что технологии компьютерного зрения на основе нейронной сети – это инновационный, устойчивый метод контроля безопасности и процесса производства, использующий системы, предполагающие более высокий контроль качества работы предприятия, предотвращение аварийных ситуаций и массовой дефектности результатов производства. Это эффективный инструмент, который может поспособствовать новому и прогрессивному развитию отрасли производства в будущем.

Литература

1. *Прэтт У.* Цифровая обработка изображений: Пер. с англ. М.: Мир, 1982. Кн.1. 312 с.
2. *Прэтт У.* Цифровая обработка изображений: Пер. с англ. М.: Мир, 1982. Кн. 2. 480 с.
3. *Хуанг Т. С., Эклунд Дж.-О., Нуссбаумер Г. Дж.* и др. Быстрые алгоритмы в цифровой обработке изображений; Под ред. Т. С. Хуанга: Пер. с англ. М.: Радио и связь, 1984. 224 с.
4. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений, Москва: Техносфера, 2005. 1072 с. ISBN 5-94836-028-8.
5. *Форсайт, Дэвид А., Понс, Жан* Компьютерное зрение. Современный подход. : Пер. с англ. М.: Издательский дом "вильямс", 2004. 928 с. ISBN 5-845-0542-7 (рус.)
6. Нейронные сети в промышленности и информационных технологиях [Электронный ресурс] - URL: https://izron.ru/articles/razvitie-tekhnicheskikh-nauk-v-sovremennom-mire-sbornik-nauchnykh-trudov-po-itogam-mezhdunarodnoy-na/sektsiya-2-informatika-vychislitelnaya-tekhnika-i-upravlenie-spetsialnost-05-13-00/neyronnye-seti-v-promyshlennosti-i-informatsionnykh-tekhnologiyakh/
7. Нейронные сети для задач промышленности и безопасности [Электронный ресурс] - URL: http://lib.secuteck.ru/articles2/all-over-ip/neyronnye-seti-dlya-zadach-promyshlennosti-i-bezopasnosti-vstraivaemye-sistemy-mashinnogo-zreniya-novogo-pokoleniya
8. Шлагбаум с постом охраны [Электронный ресурс] - URL: https://remstroytrest.ru/shl.html
9. Нейросети в промышленности: как это работает? [Электронный ресурс] - URL: https://telecomdaily.ru/news/2020/06/05/neyroseti-v-promyshlennosti-kak-eto-rabotaet
10. Нечеловеческие способности [Электронный ресурс] - URL: https://www.gazprom-neft.ru/press-center/sibneft-online/archive/2018-september-projects/1863686/
11. Нейронные сети [Электронный ресурс] - URL: https://bigenc.ru/technology_and_technique/text/4114009
12. Машинное зрение: востребованность и перспективы [Электронный ресурс] - URL: https://www.tbforum.ru/blog/mashinnoe-zrenie-vostrebovannost-i-perspektivy
13. Машинное зрение [Электронный ресурс] - URL: https://www.iksmedia.ru/articles/5685849-Mashinnoe-zrenie-kak-nauchit-lokomo.html
14. Нейронные сети или как обучить искусственный интеллект? [Электронный ресурс] - URL: http://internetinside.ru/neyronnye-seti-ili-kak-obuchit-iskuss/
15. Машинное зрение. Что это и как им пользоваться? Обработка изображений оптического источника [Электронный ресурс] - URL: https://habr.com/ru/post/350918/
16. Как нейросети помогают производствам [Электронный ресурс] - URL: https://pahomov.pro/blog/kak-nejroseti-pomogayut-proizvodstvam-realnye-primery.html
17. Компьютерное зрение [Электронный ресурс] - URL: https://exponenta.ru/comp-vision
18. Анализ промышленного применения алгоритмов нейросетевого моделирования в нефтяной и газовой промышленности [Электронный ресурс] - URL: https://nedraconsult.ru/news/analiz-promyshlennogo-primeneniya-algoritmov-neyrosetevogo-modelirovaniya-v-neftyanoy-i-gazovoy-prom/
19. Какому бизнесу нужен искусственный интеллект [Электронный ресурс] - URL: https://incrussia.ru/understand/nejronki-kakomu-biznesu-nuzhen-iskusstvennyj-intellekt-i-lajfhaki-kak-ego-vnedrit/
20. Компьютерное зрение [Электронный ресурс] - URL: https://yandex.ru/company/technologies/computer_vision/

21. Нейронные сети: общие технологические характеристики [Электронный ресурс] - URL: https://science-engineering.ru/ru/article/view?id=1236

22. Какую роль системы компьютерного зрения играют в четвертой промышленной революции? [Электронный ресурс] - URL: https://www.baslerweb.com/ru/vision-campus/otrasli-i-zadachi/rol-kompyuternogo-zreniya-v-ehpohu-industriya-4-0/

23. Компьютерное зрение. Задачи, области применения, перспективы [Электронный ресурс] - URL: https://vc.ru/ml/166105-kompyuternoe-zrenie-zadachi-oblasti-primeneniya-perspektivy

24. Выгодная имитация: мировая промышленность активно строит нейросети [Электронный ресурс] - URL: http://digital-russia.rbc.ru/articles/vygodnaya-imitatsiya-mirovaya-promyshlennost-aktivno-stroit-neyroseti/