

Интеллектуальная оценка QoE с помощью методов машинного обучения

А. В. Ивченко

Московский физико-технический институт
ivchenko.a.v@phystech.edu

Аннотация. Работа посвящена анализу качества передаваемого и воспроизводимого мультимедийного контента. Оценка качества конечным пользователем базируется на субъективном восприятии, однако она определяет, будет ли человек пользоваться тем или иным продуктом или услугой. При этом большинство пользователей предоставляемых мультимедийных услуг находятся не в собственных сетях провайдеров с возможностью предоставления и контроля требуемого качества сервиса QoS (к примеру, IPTV), а в сетях других операторов, в которых информация об ухудшении качества предоставляемой услуги приходит постфактум в виде жалоб пользователей (услуги OTT). До сих пор не существует единых методик оценки качества сервиса (QoS) и качества восприятия (QoE). Предлагаемая методика основывается на методах машинного обучения и нейронных сетей и позволяет в автоматическом режиме оценивать качество восприятия на основе объективных параметров. Также рассматриваются методы объективизации и параметризации оценки.

Ключевые слова: машинное обучение; нейронные сети; QoE; QoS, метрология; искажения; автоматизация

I. ВВЕДЕНИЕ

Анализ качества передаваемого и воспроизводимого мультимедийного контента – необходимое условие эффективного функционирования систем цифровой передачи информации. Оценка качества, базирующаяся на субъективном восприятии человека, называется качеством восприятия – QoE (Quality of Experience). В свою очередь, на восприятие влияют объективные показатели, определяющие оценку, называемую качеством сервиса – QoS (Quality of Service) [1].

При использовании цифровых сетей потери данных складываются из информационных потерь при компрессии исходного сигнала и из случаев нарушения непрерывности потока пакетов с данными. Дegradaция носит характер постепенного снижения четкости картинки, роста вероятности появления мозаичных артефактов и, при больших потерях в тракте передачи, «замерзания» видео в стоп-кадр. При этом зависимость степени деградации от величины потерь носит существенно нелинейный характер, ускоряясь от почти незаметного нарастания при малых потерях до срыва и полного прекращения воспроизведения в худших случаях.

В результате возникают две проблемы, специфические для цифрового мультимедийного вещания. Во-первых, из-за нелинейности соответствий существует большое расхождение между измеренным по сетевым параметрам пакетной передачи QoS объективным качеством любой определенной функциональной характеристики сети и фактическим «воспринятым» абонентами качеством потребляемой услуги QoE. Во-вторых, оказывается множество различных параметров деградации, в связи с чем возникает проблема выведения скалярной оценки качества услуг из векторной величины отклонений.

A. Литературный обзор

Существует ряд работ, посвященных анализу субъективного восприятия. В 2004 году Международный союз электросвязи провёл исследование [3] методов и подходов оценки качества. В результате было доказано, что наиболее прогрессивные объективные методики дают статистически эквивалентные результаты на больших выборках. Основной причиной эквивалентности оказалась зависимость экспертной оценки непосредственно от контента.

Ш. Кэ в 2012 году [4] описал способ уведомления о несоответствии состояния полосы пропускания на устройстве пользователя и текущим требованием к полосе пропускания. Уведомления происходят в полуавтоматическом режиме с помощью SMS, MMS, электронной почты и портала поставщика услуг. Оценка QoE производится анкетированием пользователей. Существенный недостаток данного решения заключается в многократном взаимодействии с конечным пользователем для однократной оценки параметра QoE, что практически исключает возможность построения эффективной автоматической системы мониторинга QoE в крупном масштабе.

Близкой к нашему исследованию является работа З. Кана [7], в которой предлагается способ измерения QoE, основанный на нейронной сети с применением метода оптимизации роя частиц (PSO). Методика рассчитана на применение в мобильных приложениях. В ходе работы приложение опрашивает пользователя о верности оценки и определяет величину ошибки, в случае выхода за допустимый предел применяется алгоритм адаптации весов. В ходе работы анализируется структура кадров по сравнению с эталонными (без искажений), что является

существенным препятствием для массового применения, к примеру, для оценки качества потокового видео. Методика также оперирует фиксированным набором анализируемых параметров.

Близкой работой является методика U-vMOS [8]. На основе групп параметров «Контент» (разрешение, частота кадров и цветовая гамма), «Загрузка» (время поиска контента, время начала загрузки и просмотра, время переключения каналов) и «Воспроизведение» (уровень блочных искажений картинки, частота и длительность прерываний воспроизведения) выведена формула, по которой рассчитывается значение U-vMOS – отдельно для широкоэкранный трансляции и для видео по запросу (1).

$$U - vMOS = 1 + (sQuality - 1) \cdot \left(\frac{\alpha(sUnteraction - 1) + \beta(sView - 1)}{4(\alpha + \beta)} \right) \quad (1)$$

Методика группирует контент на классы, такие как «активный спорт», «комедии» и пр., определяет ряд сценариев взаимодействия пользователя с сервисами и пороги значений характеристик. Коэффициенты α и β зависят от класса контента и сценария.

Данная методика обладает рядом недостатков. U-vMOS оперирует фиксированным набором параметров для анализа. Если какой-то из них отсутствует или измеряется с недопустимой точностью, то методика не работает. Разбивка контента на классы и состояния пользователей на сценарии является грубой и вносит, строго говоря, ошибку произвольной величины – к примеру, в футбольном матче есть перерывы, в которых мультимедийный контент качественно меняется. Методика может оценивать только те искажения, которые возникли в силу деградации после прохождения канала связи. Если искажения возникли до этого момента – они не будут учтены. Это касается точечных бликов, эффекта «Кругов Ньютона», искажений постобработки и пр.

В. Задача

Задача, которую мы поставили, состоит в создании автоматической методики анализа субъективного восприятия на основе объективных параметров. При этом решение должно быть:

- модульным (добавление или удаление параметров из анализа);
- универсальном (применение в системах с различной архитектурой и оборудованием);
- применимо на стационарных и мобильных устройствах;
- возможность применения в масштабе национальной сети IP-вещания.

Также решение не должно требовать наличия эталона – неискаженного контента для анализа.

II. МЕТОДИКА

А. Нахождение функции QoE

Общий вид функции QoE от некоторого объективного показателя V представлен на рис. 1 и описывается формулой (2), где b_1 , b_2 , b_3 – масштабирующие коэффициенты.

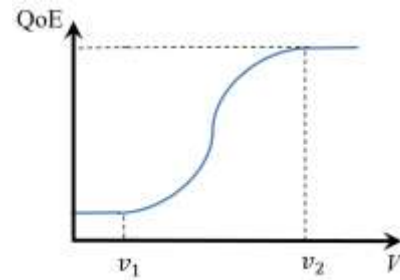


Рис. 1. Зависимость QoE от интегрального показателя V

$$QoE = b_1 / \left[1 + \exp(b_2(V - b_3)) \right] \quad (2)$$

При значениях менее v_1 ухудшения качества восприятия не происходит, аналогично при уровне показателей более v_2 пользователь не замечает улучшения качества. Показатель V можно рассматривать как

$$V = \sum_{i=1}^n w_i(W, P) \psi_i, \quad W = \{w_i\}, \quad P = \{p_j\}, \quad w_i -$$

вычисляемые весовые коэффициенты, ψ_i – значения объективных показателей, p_j – характерные особенности контента. Функция QoE и весовые коэффициенты нелинейны, таким образом мы имеем задачу многомерной нелинейной регрессии.

Для решения данной задачи была написана многослойная нейронная сеть на основе открытой программной библиотеки для машинного обучения Google TensorFlow [19]. Для сбора данных о работе сети был применен анализатор трафика Wireshark [5]. Ряд экспериментов с оболочками для ядра привел к выбору библиотеки Keras [20]. Предварительные эксперименты на тестовых данных показывают до 82% точности.

В. Объективизация параметров

Анализируемые показатели делятся на группы: сеть и оборудование; характерные особенности контента; оценки экспертов. Среди первой группы предлагается отслеживать [11–15]: потери пакетов в сети; скорость трафика; задержку; джиттер; скорость ошибочных звуковых пакетов; разрешение экрана; фрактальность аудио- и видеотрафика.



Рис. 2. Кадры сериалов «Castle» и «Каменская»

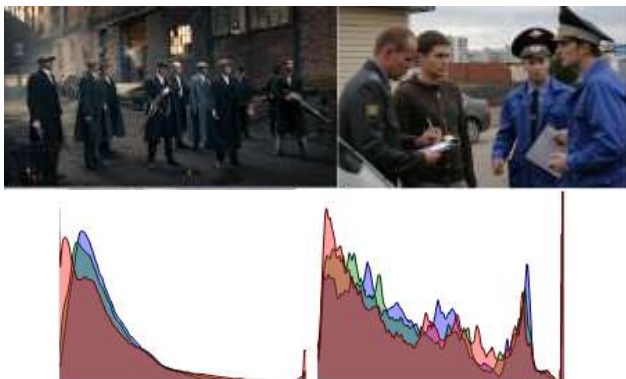


Рис. 3. Кадры и соответствующие им гистограммы фильмов «Peaky blinders» и «Глухарь»

Среди характерных особенностей видеотрафика предлагается принимать во внимание [16–18]:

- активность движения;
- частоту смены ракурсов съёмки и смены сюжета;
- искажения различного толка;
- глубину резкости;
- насыщенность.

Активность движения оценивается на основе количества и значения нормы длины векторов движения, доступных с декодера. Анализ смены ракурсов, обусловлен снижением QoE длительными статическими сценами. Частота смены ракурса зависит от количества движения в кадре. Для отслеживания данного параметра используется алгоритм анализа смены сюжета, предложенный А.В. Дворковичем и В.П. Дворковичем, который заключается в анализе изменения гистограммы распределения яркости в соседних кадрах [10].

Глубина резкости определяет область кадра, на котором сконцентрировано внимание зрителя (рис. 2). На кадре справа наблюдается огромное количество достаточно чётких неинформативных деталей, тогда как на кадре слева внимание сконцентрировано исключительно на актёре. В качестве числового показателя используется доля контуров объектов. Для анализа насыщенности исследуются гистограммы кадров (рис. 3).

С. Анализ искажений

Для обучения модуля анализа искажений была выбрана база TID2013 [21], содержащая 24 типа искажений. Для первоначального анализа были выбраны:

- гауссовский шум;
- пространственно-коррелированный шум;
- импульсные помехи;
- локальные блочные искажения.

Была написана сверточная нейронная сеть на основе ядра TensorFlow и оболочки Keras. Обработка массивов данных реализована на основе библиотеки Numpy [22].

Работа непосредственно с изображениями ведётся с помощью библиотеки OpenCV [23]. Для повышения точности работы нейронной сети и уменьшения риска переобучения применен алгоритм Dropout [6]. При разрешении изображений 512x384 пикселя удалось добиться 91% точности и 94% полноты при классификации. Использовался полноцветный анализ, так как попытка анализа в градациях серого снизила точность до 40%.

Д. Методика применения

В работе вводится естественное предположение, что внутри одного сегмента сети, территориально ограниченного географией распределения абонентов, используется единый набор технологий и оборудования. Это позволяет применять функцию оценки QoE, вычисленную единожды, для всего сегмента сети.

Первый этап – этап настройки. Оператор выбирает множество параметров, которые возможно анализировать с помощью доступных ему средств. В рассматриваемой сети запускается утилита, собирающая статистику работы сети и анализирующая характерные особенности контента. В этой же точке производится анализ контента по методике MOS [9].

На втором этапе в лабораторных условиях производится анализ полученных данных, включающий анализ использованного множества показателей. В результате оператор также получает значение точности оценки QoE, которую можно получить с помощью данного



Рис. 4. Пример искажений из TID2013 слева направо: гауссовский шум, пространственно-коррелированный шум, импульсные помехи, блочные искажения

множества параметров и его подмножеств. Уточняется есть ли зависимости между показателями, есть ли возможность исключить часть данных из анализа.

На втором этапе методика может заключить, что данных недостаточно для желаемой точности оценки. В таком случае требуется вернуться к первому этапу и включить в анализ дополнительные данные, опираясь на рекомендации работы.

Таким образом, после второго этапа имеется достаточный набор параметров и функции весовых коэффициентов. На третьем этапе – этапе применения, достаточно собирать статистику и характерные особенности контента (в контрольной точке или на конечном устройстве пользователя), чтобы получить оценку QoE по формуле с весовыми коэффициентами, полученными на предыдущем этапе (рис. 2).



Рис. 5. Методика применения

III. РЕЗУЛЬТАТЫ

Создана методика анализа QoE, производится её реализация и предварительные лабораторные испытания. Восстановление функции QoE показывает 82% точности, однако требуется сбор большого количества экспертных оценок для настройки весов нейронной сети и дальнейшей работы.

Для обучения анализатора использовалось от 100 до 600 изображений, равномерно распределённых по 4 классам искажений. В качестве обучающей выборки было использовано 80% собранного материала и, соответственно, 20% составляла тестовая выборка. При этом особенностью базы данных TID2013 является 5 степеней искажения одного и того же кадра. При такой небольшой выборке при подаче одних и тех же кадров классификация искажений падала до 40% точности. Также ведутся работы над анализом степени искажения, а также случаев нескольких типов искажений на одном кадре.

ВЫВОДЫ

Полученные результаты демонстрируют возможность применения предложенного подхода автоматической оценки качества восприятия QoE с помощью свёрточных нейронных сетей и методов машинного обучения. Дальнейшие работы будут направлены на увеличение базы статистических и экспертных оценок, а также базы данных искажений, оптимизацию работы сети с учётом роста количества изображений.

Дополнительные работы будут вестись по обработке и аудиоинформации.

ВЫРАЖЕНИЕ ПРИЗНАТЕЛЬНОСТИ

Выражаем благодарность Дворковичу Александру Викторовичу за консультации и предоставленные материалы.

СПИСОК ЛИТЕРАТУРЫ

- [1] Recommendation ITU-T G.1080. Quality of experience requirements for IPTV services. ITU-T, 12/2008, 44 p.
- [2] Ивченко А. В. Интеллектуальная методика оценки качества восприятия / Дворкович А.В., Москва, Сборник трудов конференции (доклады); международное (ISBN 978-5-905278-29-7); всероссийское (УДК 621.391.004(082); ББК: 32.511)
- [3] ITU-T Tutorial. Objective perceptual assessment of video quality: Full reference television. ITU-T, 2004, 218 p.
- [4] Пат. РФ № 2582573/ III. Кз, Способ уведомления о полосе пропускания пользователя; Оpubл. 21.12.12.
- [5] WireShark.Free and Open Source packet analyzer // Version 2.6.0. URL: <https://wireshark.org>
- [6] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R.. Dropout: A Simple Way to Prevent Neural Networks from Overfitting // Journal of Machine Learning Research. 2014. №15. P. 1929-1958.
- [7] Pat. CN 103179592A / Z. Z. Kan; H. Yuan, W. Wenbo, Method for measuring QoE. 2015.
- [8] J. Kunstler, A. Dobardziev, Wireline Video QoE Using U-vMOS, URL: <https://ovum.informa.com/resources/product-content/2016/11/11/15/09/analyst-white-paper-wireline-video-qoe-using-u-vmos>
- [9] Recommendation ITU-T P.10/G.100. Vocabulary for performance and quality of service. – ITU-T, 07/2006, 42 p.
- [10] В.П. Дворкович, А.В. Дворкович. Цифровые видео-информационные системы (теория и практика). М.: Техносфера, 2012. 1009 с.
- [11] P. Abry, P. Flandrin, M.S. Taqqu, D. Veitch. Wavelets for the analysis, estimation and synthesis of scaling data // Self-Similar Traffic and Performance Evaluation, edited by K. Park and W. Willinger. John Wiley & Sons, 2000. P. 39-88.
- [12] О.И. Шелухин, А.В. Осин, Р.П. Ахметшин. Оценка самоподобности речевого трафика вейвлет-методом с автоматическим определением границ масштабирования // Электротехнические и информационные комплексы и системы. 2007. №1. С. 15-20.
- [13] В.П. Дворкович, А.В. Дворкович. Цифровые видео-информационные системы (теория и практика). М.: Техносфера, 2012. 1009 с.
- [14] Recommendation ITU-R BT.601-7. Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios. ITU-R, 03/2011, 20 p.
- [15] U. Reiter, K. Brunnström, K. De Moor, M.-Ch. Larabi, M. Pereira, A. Pinheiro, J. You, A. Zgank. Factors Influencing Quality of Experience // Quality of Experience. Advanced Concepts, Applications and Methods, edited by S. Möller and A. Raake. Springer, T-Labs Series in Telecommunication Services, 2014. P. 55-72.
- [16] Mariner xVu, Режим доступа <http://www.marinerxvu.com/>
- [17] Datametrie GX, Режим доступа <http://www.ip-label.co.uk/products/datametrie-global-experience/>
- [18] Hughes Sysyique, Режим доступа <https://hsc.com/campaigns/better-QoE-for-effective-OTT-streaming-services>
- [19] TensorFlow. An open source machine learning framework for everyone. // Google. Version 1.6.0. URL: <https://tensorflow.org/>
- [20] Chollet F., Google, Microsoft, others. Keras. // GitHub repository. 2015. URL: <https://keras.io>
- [21] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, Image database TID2013: Peculiarities, results and perspectives, Signal Processing: Image Communication, vol. 30, Jan. 2015, pp. 57-77.
- [22] NumPy. Package for scientific computing with Python. // Version 1.14.2. URL: <http://numpy.org/>
- [23] OpenCV. Open Source Computer Vision Library // Version 3.4.0. URL: <https://opencv.org/>