



УНИВЕРСИТЕТ ИТМО

3



Альманах

НАУЧНЫХ РАБОТ
МОЛОДЫХ
УЧЕНЫХ

2020

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

УНИВЕРСИТЕТ ИТМО

**АЛЬМАНАХ НАУЧНЫХ РАБОТ МОЛОДЫХ УЧЁНЫХ
УНИВЕРСИТЕТА ИТМО**

Том 3

 **УНИВЕРСИТЕТ ИТМО**

Санкт-Петербург

2020

Альманах научных работ молодых учёных Университета ИТМО. Том 3. СПб.: Университет ИТМО, 2020. 219 с.

Издание содержит результаты научных работ молодых учёных, доложенные на XLIX научной и учебно-методической конференции Университета ИТМО по тематике: информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Председатель редколлегии:

Бухановский Александр Валерьевич

доктор технических наук, директор мегафакультета трансляционных информационных технологий Университета ИТМО.

Члены редколлегии:

Парфенов Владимир Глебович

доктор технических наук, декан факультета информационных технологий и программирования

Чугунов Андрей Владимирович

кандидат политических наук, доцент институт дизайна и урбанистики

Быстрянцева Наталья Владимировна

кандидат архитектуры, доцент институт дизайна и урбанистики

Матвеев Юрий Николаевич

доктор технических наук, профессор факультета информационных технологий и программирования

Духанов Алексей Валентинович

доктор технических наук, профессор института финансовых кибертехнологий

ISBN 978-5-7577-0636-8

ISBN 978-5-7577-0639-9 (Том 3)



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, получивший в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Миссия Университета ИТМО – открывать возможности для гармоничного развития конкурентоспособной личности и вдохновлять на решение глобальных задач.

© Университет ИТМО, 2020

© Авторы, 2020

ВВЕДЕНИЕ

Издание содержит результаты научных работ молодых ученых, доложенные 29 января – 1 февраля 2020 года на XLIX научной и учебно-методической конференции Университета ИТМО по тематике: информационные технологии и программирование; инфокоммуникационные технологии; дизайн и урбанистика.

Конференция проводится в целях усиления интегрирующей роли университета в области научных исследований по приоритетным направлениям развития науки, технологий и техники и ознакомления научной общественности с результатами исследований, выполненных в рамках государственного задания Министерства образования и науки РФ, программы развития Университета ИТМО на 2009–2018 годы, программы повышения конкурентоспособности Университета ИТМО среди ведущих мировых научно-образовательных центров на 2013–2020 гг., Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы», грантов Президента РФ для поддержки молодых российских ученых и ведущих научных школ, грантов РФФИ, РГНФ, РНФ и Правительства РФ (по постановлению № 220 от 09.04.2010 г.) и по инициативным научно-исследовательским проектам, проводимым учеными, преподавателями, научными сотрудниками, аспирантами, магистрантами и студентами университета, в том числе в содружестве с предприятиями и организациями Санкт-Петербурга, а также с целью повышения эффективности научно-исследовательской деятельности и ее вклада в повышение качества подготовки специалистов.

**НАПРАВЛЕНИЕ
ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
И ПРОГРАММИРОВАНИЕ; ИНФОКОММУНИКАЦИОННЫЕ
ТЕХНОЛОГИИ; ДИЗАЙН И УРБАНИСТИКА**



Азина Любовь Владимировна

Год рождения: 1996

Университет ИТМО,
факультет прикладной оптики,
студент группы № В42801

направление подготовки: 16.04.01 – Университет ИТМО,

e-mail: lu.az.spb@gmail.com



Быстрынцева Наталья Владимировна

Год рождения: 1981

Университет ИТМО,
факультет прикладной оптики,
канд. архитектуры, доцент ФПО,

e-mail: sv.s.lighting@gmail.com

УДК 628.979

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОСВЕЩЕНИЯ ДВОРОВЫХ ПРОСТРАНСТВ МНОГOKВАРТИРНЫХ ДОМОВ НА ПОЛЬЗОВАТЕЛЬСКУЮ АКТИВНОСТЬ НА ПРИМЕРЕ СПБ

Л.В. Азина

Научный руководитель – канд. арх., доцент Н.В. Быстрынцева

Работа выполнена в рамках исследовательской работы «Разработка и прототипирование световых решений в архитектуре и городской среде» на международной образовательной программе «Световой дизайн».

Аннотация

В работе исследовано освещение разных типов дворовых пространств Санкт-Петербурга, проведены светотехнические измерения уровней освещённости, выявлены проблемные моменты в освещении дворов разных типов. Проведен анализ существующей ситуации, оценено влияние освещения на пользовательскую активность. Предложены рекомендации по созданию комфортной световой среды с учетом значимых светотехнических параметров.

Ключевые слова

Световая среда, освещенность, двор, пользовательская активность, дворовое пространство многоквартирных домов, функциональные зоны двора, городская среда.

Требования жителей современных мегаполисов к общественным пространствам изменяются. Потребность в социальной и физической активности стимулирует людей проводить время вне дома. Все больше людей предпочитают проводить свой досуг в общественных местах: как на открытых территориях (улицы, набережные, парки, скверы, площади, дворы, детские и спортивные площадки и т.п.), так и в закрытых пространствах (торговые и развлекательные центры, кафе, культурные учреждения и т.д.). Выделяют три основных критерия привлекательности общественного

пространства [1] – 1) обеспечение физической и психологической безопасности (комфортности), 2) доступность и 3) мультифункциональность. Согласно проведенным в Екатеринбурге исследованиям [1], привлекательность мест отдыха на примере парковых зон в равной степени зависит от эстетической привлекательности и ухоженности территории, культурной ценности места (наличие памятников архитектуры, музеев, арт-объектов), благоустроенности и наличия удобств (скамеек, кафе и пр.), возможности заниматься разными видами деятельности, транспортной доступности, размера территории, наличия развлечений. Одним из факторов, определяющих широту выбора мест отдыха, является фактор наличия детей дошкольного и школьного возраста – респонденты, имеющие детей, более сфокусированы в выборе мест отдыха и чаще попадают в аудиторию «целевых посетителей».

В вечернее время посещаемость мест отдыха открытого типа зависит от освещения территории. Наличие качественного освещения снижает угрозы, связанные с криминальной активностью и повышает безопасность передвижения, связанную с травмируемостью пешеходов. Проведенные исследования показали [2], что для предупреждения преступности и снижения криминальной активности важно соблюдать минимальные значения светотехнических параметров полуцилиндрической освещенности (отвечает за идентификацию лиц) и горизонтальной освещенности, а для безопасного перемещения – значения средней горизонтальной и вертикальной освещенностей (видимость дороги, объектов и фасадов зданий).

Для проведения вечернего досуга как в центральных районах города, так и на периферии люди выбирают места, расположенные поблизости с местом проживания. Таким местом является дворовое пространство – это ближайшая территория, где можно провести досуг, отдохнуть с семьей или посвятить время активному отдыху и спорту [3]. В новых застраиваемых районах на стадии проектирования закладывается необходимая инфраструктура, обеспечивающая жителям возможность вечернего досуга. Следовательно, обеспечение качественного освещения территории является необходимым условием комфортного проведения вечернего времени в общественных пространствах, в том числе, дворового типа. Таким образом становится актуальным запрос на комфортную световую среду. Поэтому создание качественной световой среды, повышающей пользовательскую активность и привлекательность дворовых пространств в вечернее время, является актуальной задачей при проектировании и модернизации освещения на территории жилых районов.

В данной работе исследуется освещение дворовых пространств и оценивается влияние освещения на пользовательскую активность и привлекательность территории для пользователей в вечернее время. Для оценки световой среды проводится измерение светотехнических параметров горизонтальной и полуцилиндрической освещенности.

В ходе сбора первичных данных были определены характерные для большинства дворовых пространств функциональные зоны. В дальнейшем исследовании будут рассмотрены пешеходные проходы и проезды, детские площадки, спортивные площадки и зоны отдыха. В зависимости от комбинации функциональных зон и с учетом освещения предложена классификация дворов Санкт-Петербурга по типам застройки: 1) дворы в исторической застройке, 2) дворы в советской и современной массовой застройке, 3) современные жилые комплексы застройки последних нескольких лет. Для дворов первых двух типов характерно неупорядоченное одновременное использование различных источников освещения (газоразрядные лампы и светодиодные светильники), а для современных жилых комплексов характерно светодиодное освещение территории в единой стилистике. В работе исследованы следующие дворовые пространства Санкт-Петербурга:

- 4 двора в исторической застройке, в т.ч. т.н. дворы-колодцы;

- 5 дворов в массовой советской и современной застройке;
- 2 современных жилых комплекса (в жилых комплексах выделено 4 и 6 дворовых пространств соответственно).

В процессе исследования было установлено, что для всех типов дворов характерны проблемы в освещении:

- наличие неисправных осветительных приборов;
- использование в хаотичных комбинациях различных источников освещения; для светодиодных систем – одновременное использование ламп теплой и холодной цветовой температуры (возможная причина такой ситуации – замена нерабочих ламп на лампы другого типа);
- отсутствие освещения некоторых функциональных зон.

Измерение горизонтальной освещенности проводились только для освещенных функциональных зон, полуцилиндрическая освещенность измерялась вне зависимости от наличия освещения для конкретной зоны. Исследование показало: 1) все исследуемые пешеходные проходы и проезды освещены полностью или частично; 2) доля дворов, в которых проходы и проезды освещены частично, уменьшается от исторической застройки к современным жилым комплексам; 3) усредненные показатели уровня горизонтальной освещенности проходов и проездов удовлетворяют требованиям отечественных норм (ГОСТ 55706-2013) и составляют более 4 лк, однако, имеют широкий разброс от двора ко двору в пределах одного типа пространства; 4) параметр полуцилиндрической освещенности не нормируется для территории жилых районов. В европейском стандарте (CEI Recommendation) нормируется минимальное значение полуцилиндрической освещенности (0,4 лк), а в американском стандарте (American Standard RP-8) – средняя вертикальная освещенность (1 лк). Усредненные показатели параметра средней полуцилиндрической освещенности удовлетворяют требованиям последнего стандарта, однако, минимальное значение полуцилиндрической освещенности в каждом дворовом пространстве составляло 0 лк (в направлении неосвещенных зон). Значение параметра равномерности освещения невысокое для всех зон и особенно низкое для зон проходов и проездов. Данный факт связан с характеристиками установленных светильников, которые выдают высокий уровень освещенности около осветительных приборов и практические нулевые значения между ними. Максимальный уровень освещенности в зонах проходов и проездов, выдаваемый светильниками по направлению излучения, колебался от 30 лк до 140 лк, а минимальное значение составляло ~ 0 лк.

Проведенные исследования позволяют сделать следующие выводы:

- 1) соблюдения минимальных норм уровня средней горизонтальной освещенности и учета минимального значения полуцилиндрической освещенности достаточно для обеспечения безопасности в вечернее время и возможности идентифицировать лица других людей;
- 2) равномерность освещения не оказывает видимого влияния на пользовательскую активность.

Сравнение полученных результатов светотехнических измерений с отечественными нормами (ГОСТ 55706-2013) приведены в таблице, отдельно выделены функциональные зоны с наличием и отсутствием освещения, отсутствующие зоны отмечены «н».

Качество освещения, безусловно, влияет на пользовательскую активность. Среди пользователей были выделены две группы по цели посещения дворового пространства:

- 1) транзитные пешеходы – кратковременное пользование пространством;
- 2) долговременные пользователи – родители с детьми, люди с собаками, пары, компании, одиночные посетители.

Результаты измерений уровней освещенности во дворах разных типов застройки

Светотехнические параметры	Функциональная зона	Нормы	Тип застройки		
		ГОСТ	Историческая застройка	Советская и современная массовая застройка	Современные жилые комплексы
Горизонтальная освещенность $E_{ср} \pm \Delta E_{ср}, лк$	Проходы и проезды	4	$5,6 \pm 4,2$	$6,9 \pm 4,3$	$17,4 \pm 14,1$
	Детская площадка (освещена)	10	$21,9 \pm 7,3$	$5,9 \pm 7,2$	$12,4 \pm 2,4$
	Детская площадка (не освещена)		0	0	н
	Спортивная площадка		0	0,6	$15,3 \pm 5,7$
	Рекреация (освещена)	2	н	8,6	$7,2 \pm 2,7$
	Рекреация (не освещена)		0	0	н
Равномерность освещенности $U \pm \Delta U, \text{от. ед.}$	Проходы и проезды	0,2	$0,20 \pm 0,15$	$0,06 \pm 0,04$	$0,04 \pm 0,01$
	Детская площадка (освещена)	0,3	$0,22 \pm 0,19$	$0,19 \pm 0,05$	$0,12 \pm 0,08$
	Детская площадка (не освещена)		-	-	н
	Спортивная площадка		-	0,16	$0,10 \pm 0,08$
	Рекреация (освещена)	0,1	н	0,01	$0,11 \pm 0,06$
	Рекреация (не освещена)		-	-	н
Средняя ПЦ освещенность $E_{пц ср} \pm \Delta E_{пц ср}, лк$	Проходы и проезды		$7,2 \pm 2,6$	$5,2 \pm 3,2$	$4,1 \pm 0,2$
	Детская площадка (освещена)		$4,1 \pm 2,6$	$2,2 \pm 0,3$	$4,3 \pm 0,1$
	Детская площадка (не освещена)		$2,4 \pm 0,4$	0	н
	Спортивная площадка		0,9	15	$4,3 \pm 0,3$

продолжение таблицы

Светотехнические параметры	Функциональная зона	Нормы	Тип застройки		
		ГОСТ	Историческая застройка	Советская и современная массовая застройка	Современные жилые комплексы
	Рекреация (освещена)		н	1,5	3,0 ± 1,1
Рекреация (не освещена)		0,6	0,8 ± 0,3	н	

Оценка пользовательской активности производилась с разделением на транзитную активность и активность, связанную с проведением досуга. Динамика пользовательской активности рассчитывалась по наполняемости пространства посетителями за час, и представлена на рисунке 1б. На рисунке 1а приведено сравнительное распределение освещенных и не освещенных функциональных зон.

Наибольшая пользовательская активность, связанная с досугом в вечернее время, выявлена в жилых комплексах, где все рассматриваемые зоны освещены (рисунок). Основными пользователями в данном типе дворовых пространств являются родители с детьми, люди с собаками и небольшие группы отдыхающих людей. Данный фактор можно объяснить как разнообразием функциональных зон со значимой долей зон отдыха и широким выбором видов деятельности, так и хорошими показателями светотехнических параметров. На территории жилых комплексов для разных функциональных зон параметры средней горизонтальной освещенности (кроме зон проходов и проездов) и средней полуцилиндрической освещенности имеют примерно равные значения, что говорит о равномерном освещении территории, без контрастов между функциональными зонами. Результаты светотехнических измерений представлены в шестом столбце таблицы.

Динамика пользовательской активности во дворах исторической застройки схожа с динамикой во дворах советской и современной массовой застройки (рисунок б). Основная пользовательская активность в таких дворовых пространствах – транзит. Общая причина невысокого использования мест отдыха – плохое качество освещения некоторых функциональных зон или его отсутствие и ограниченный выбор видов деятельности (рисунок. а). Стоит отметить, что во дворах обоих типов присутствуют функциональные зоны, где соблюдены нормы освещения, и где нормы не соблюдены.

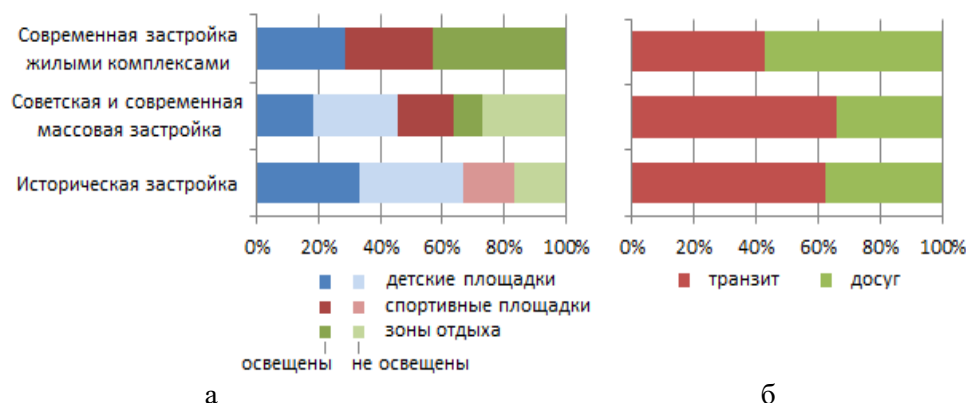


Рисунок. Процентные соотношения для дворовых пространств разных типов: освещенных и не освещенных функциональных зон (а), вечерней пользовательской активности (б)

Сравнительно небольшие размеры дворов в исторической застройке и преимущественно закрытый тип пространства (двор с трех или более сторон ограничен зданиями) накладывают ограничения на разнообразие возможных функциональных зон и видов деятельности (рисунок а), поэтому пользовательская активность в большей степени приходится на транзит (рисунок б). Основной функциональной зоной такого типа дворов становятся пешеходные проходы и проезды, где пользовательская активность зависит от обеспечения хороших условий видимости. В дворовых пространствах исторической застройки выявлены проблемы, связанные с частичным освещением пешеходных проходов и проездов, в т.ч. арок, ведущих во двор. Неосвещенные проходы пользователи предпочитают избегать, выбирая освещенные пути. Результаты светотехнических измерений для дворов в исторической застройке представлены в 4-ом столбце таблицы.

Дворы советской и современной массовой застройки отличаются большими размерами территории. Ограниченный выбор видов деятельности часто связан с неосвещенностью или некачественным освещением представленных функциональных зон (рисунок а). Дворы такого типа застройки наиболее неоднозначны по качеству освещения. Среди данных дворов выявлен наибольший разброс в показателях средней горизонтальной освещенности для одной функциональной зоны, наблюдается ситуация, когда значения горизонтальной освещенности в разы ниже, чем полуцилиндрической (возможная причина – использование светильника класса преимущественно прямого распределения света), показания средней полуцилиндрической освещенности также изменяются в зависимости от функциональной зоны. Результаты светотехнических измерений для дворов в массовой современной застройке представлены в пятом столбце таблицы. Среди пользователей в вечернее время в равной степени выделяются родители с детьми и люди с собаками.

Заключение

1. Дворовые пространства в исторической застройке обладают сравнительно небольшим размером территории, и основной запрос направлен на качественное освещение пешеходных проходов и проездов, являющихся основной функциональной зоной, а также на обеспечение хороших условий видимости. При проектировании освещения необходимо учитывать, что основная пользовательская активность – транзит.

2. В советской и современной массовой застройке существует проблема в качестве освещения, поскольку внутри одного дворового пространства все функциональные зоны могут иметь разные уровни освещенности. Т.к. наличие различных функциональных зон позволяет увеличить виды деятельности в вечернее время, то основная задача – создание качественного освещения мест проведения досуга.

3. Проблемы в освещении современных жилых комплексов незначительны и в основном связаны с низкой равномерностью освещения, обусловленной характеристиками используемых светильников. На данных дворовых пространствах наблюдается наибольшая пользовательская активность, чему способствуют разнообразие функциональных зон, существующий уровень полуцилиндрической освещенности, позволяющий идентифицировать лица, а также соблюдение норм средней горизонтальной освещенности.

4. Вне зависимости от типа дворового пространства пользовательская активность и привлекательность территории наблюдается выше при соблюдении требований норм освещения для средней горизонтальной освещенности, обеспечении возможности идентификации лиц (параметр полуцилиндрической освещенности) и наибольшем разнообразии функциональных зон. Последняя зависимость

подтверждается рядом других исследований [1]. Равномерность освещения - наименее значимый параметр, влияющий на пользовательскую активность. Данные зависимости говорят о значимости и важности соблюдения норм светотехнических параметров полуцилиндрической и средней горизонтальной освещенности.

Полученные результаты позволяют сформировать рекомендации по освещению дворовых пространств. При проектировании освещения любого дворового пространства и создании комфортной световой среды следует:

- соблюдать требования ГОСТа;
- обеспечивать минимальные требования норм средней горизонтальной освещенности;
- обращать внимание на параметр полуцилиндрической освещенности, учитывать его минимальное и среднее значение;
- при выборе источников освещения учитывать характеристики светильников;
- учитывать тип дворового пространства и его индивидуальные особенности.

Соблюдение данных рекомендаций позволит обеспечить безопасность и повысить привлекательность дворового пространства для вечернего отдыха. Однако стоит учитывать, что на пользовательскую активность в вечернее время большое влияние оказывает существующая наполненность пространства функциональными зонами и благоустройство территории.

Литература

1. Нотман О.В., Багирова А.П. Привлекательность общественных пространств мегаполиса в оценках его жителей (на примере ТРЦ и парковых зон Екатеринбурга //Мониторинг общественного мнения: экономические и социальные перемены. 2018. №. 6 (148). С. 347-364.
2. Bommel W. Road lighting: fundamentals, technology and application. 2015. 343 с.
3. Крашенинников А.В. Жилые кварталы: учебное пособие для архитектурно-строительных вузов //Высшая школа. 1988. 86 с.



Аксёнов Александр Александрович

Год рождения: 1995

Санкт-Петербургский институт информатики
и автоматизации РАН, младший научный сотрудник
e-mail: a.aksenov95@mail.ru



Рюмина Елена Витальевна

Год рождения: 1991

Санкт-Петербургский институт информатики
и автоматизации РАН, программист;
Университет ИТМО, факультет информационных
технологий и программирования, студент группы № М41211,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: ryumina_ev@mail.ru



Карпов Алексей Анатольевич

Год рождения: 1978

Санкт-Петербургский институт информатики
и автоматизации РАН,
главный научный сотрудник

УДК 004.852

**АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ
ДЕТЕКТИРОВАНИЯ ЛИЦ**

А.А. Аксёнов, Е.В. Рюмина

Научный руководитель – д.т.н., профессор А.А. Карпов

Данное исследование проведено в рамках Соглашения с Минобрнауки России №075-15-2019-1295 (идентификатор проекта RFMEFI61618X0095).

Аннотация

Работа содержит аналитический обзор современных методов обнаружения лиц: детектор на основе метода Виолы-Джонса, детектор лиц на основе глубокого обучения, детектор лиц на основе функций HoG и SVM, Детектор лица на основе сверточной нейронной сети. Проведен сравнительный анализ всех рассматриваемых методов и проиллюстрированы результаты. Предложено дальнейшее применение системы в задачах биометрии, компьютерном зрении, машинном обучении, автоматических системах распознавания лиц, речи и элементов жестовых языков.

Ключевые слова

Биометрия, распознавание лиц, компьютерное зрение, машинное обучение.

Введение

Задачи автоматизации и безопасности в текущее время играют важную роль, в связи с этим детектирование и распознавание лиц является неотъемлемой частью почти любых современных систем. В частности, технологии обнаружения и распознавания лиц используются в различных системах контроля доступа и платежных системах, определения пола, возраста и эмоций, а также чтении речи по губам [1]. Задача детектирования лиц является первоначальным и особо важным условием в различных исследованиях. Для решения подобного рода задач необходимо использовать максимально доступную информацию с целью достичь приемлемых результатов. Такие системы детектирования лиц должны работать при различных условиях освещения, а также способны распознавать лица под разными углами и частичными перекрытиями областей лица.

Нынешнее исследование направлено на анализ современных методов детектирования лиц. Цель работы заключается в проведении анализа и сравнении современных методов детектирования лиц.

Набор данных

Для эксперимента использовался набор данных LFW (Labeled Faces in the Wild) [2], характеристики которого представлены далее. На рис. 1 показаны примеры изображений из базы данных. Характеристика набора данных описана в таб. 1.

Таблица 1

Характеристики набора данных LFW

Количество изображений	13 233
Количество человек	5749
Изображения с двумя или более людьми	1680
Размер изображения	250x250 пикселей



Рис. 1. Изображения из набора данных LFW

Современные методы детектирования

В этом разделе проанализированы различные методы обнаружения лиц, разработанные с помощью библиотек OpenCV и Dlib [3]:

- детектор лиц на основе метода Виолы-Джонса [4] (OpenCV);
- детектор лиц на основе глубокого обучения (OpenCV);
- детектор лиц на основе функций HoG и SVM [5] (Dlib);
- детектор лица на основе сверточной нейронной сети (Dlib).

Детектор лиц на основе примитивов Хаара оставался самым современным в задачах обнаружения лиц в течение многих лет, начиная с 2001 года в момент, когда он был представлен Виолой и Джонсом. За последние годы было предложено множество улучшений этого алгоритма, в том числе с использованием библиотеки компьютерного

зрения OpenCV, в которой имеется большое количество моделей на основе примитивов Хаара.

Детектор лиц на основе метода Виолы-Джонса работает посредством интегрального представления изображения, в результате чего находятся необходимые объекты с высокой скоростью, поиск выполняется при помощи признаков Хаара. Выбор наиболее подходящих признаков у искомого объекта на выбранной части изображения выполняется с помощью бустинга. Классификатор на вход принимает признаки и выдает результат «истинна» или «ложь». Для быстрого отбрасывания окон, в которых не найден объект применяются каскады признаков. На рис. 2 (а) изображены примитивы Хаара в стандартном алгоритме Виолы-Джонса, на рис. 2 (б) показаны дополнительные примитивы Хаара. На рис. 3 отображены результаты работы детектора лиц на основе метода Виолы-Джонса.

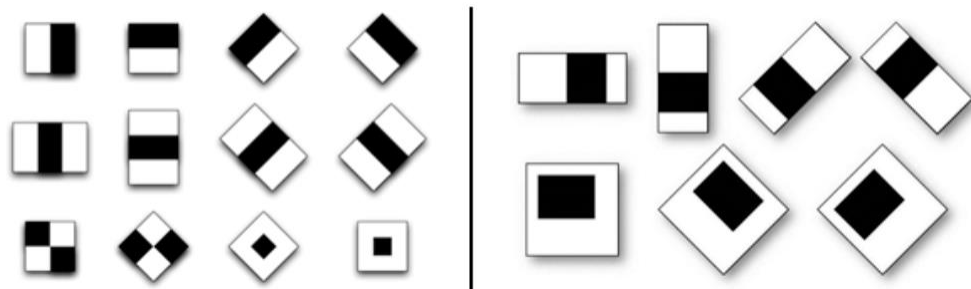


Рис. 2. (а – слева, б – справа). Примитивы Хаара

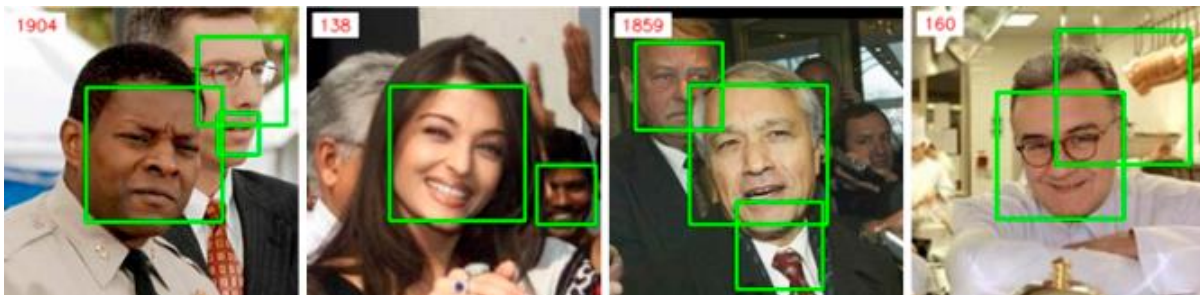


Рис. 3. Результат работы детектора Виолы-Джонса

Детектор лиц на основе глубокого обучения находится в библиотеке компьютерного зрения OpenCV, основанный на Single-Shot-Multibox Detector. В качестве архитектуры используется ResNet-10. Для обучения модели используются изображения, доступные в Интернете, источник которых не разглашается. Пользователю предоставляется две обученные платформы: платформа Caffe с плавающей запятой и платформа с использованием фреймворка TensorFlow. Рис. 4 показывает результаты работы детектора лиц на основе глубокого обучения с использованием платформы Caffe. На рис. 5 показаны результаты работы детектора лиц на основе глубокого обучения с использованием платформы TensorFlow.

Детектор лиц, основанный на функциях HoG и SVM, входит в состав библиотеки Dlib. Особенности данного детектора являются: использование классификатора на основе метода опорных векторов (SVM); для обнаружения необходимой области применяются гистограммы ориентированных градиентов (HoG).

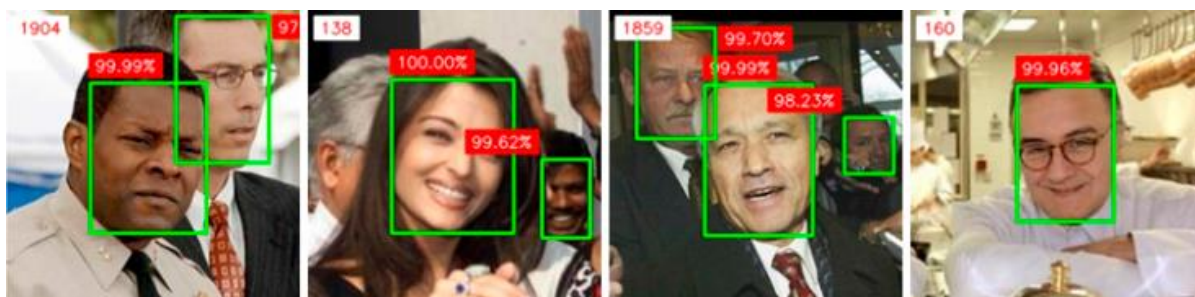


Рис. 4. Результат работы на платформе Caffe

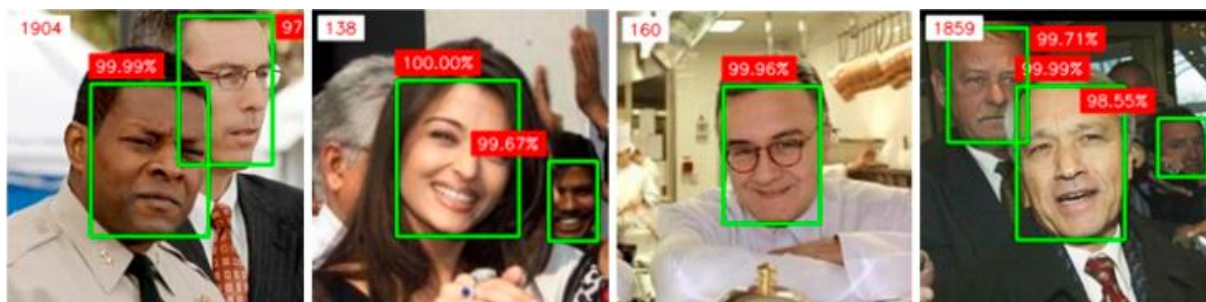


Рис. 5. Результат работы на платформе TensorFlow

Модель детектора состоит из 5 фильтров гистограммы ориентированных элементов:

- вид спереди;
- вид слева;
- вид справа;
- вид спереди, повернутый влево;
- вид спереди, повернутый вправо.

Результат работы детектора на показан на рис. 6.

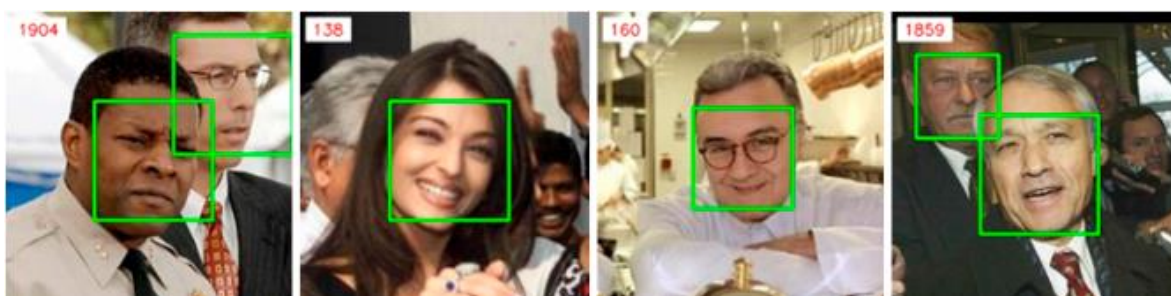


Рис. 6. Результат работы детектора лиц на основе функций HoG и SVM

Следующий детектор, который рассматривается – детектор, основанный на сверточных нейронных сетях (CNN), реализованный в библиотеке Dlib. К особенностям данного детектора можно отнести следующее: для обнаружения объектов используется объемный детектор максимальной границы (MMOD), с использованием функций, на основе CNN. При обучении модели этого детектора использовалось 7220 изображений, размеченных самим автором вручную, из нескольких наборов данных ImageNet, PASCAL, VOC, VGG, WIDER, Face Scrub. На рисунке 7 показаны результаты работы данного детектора.

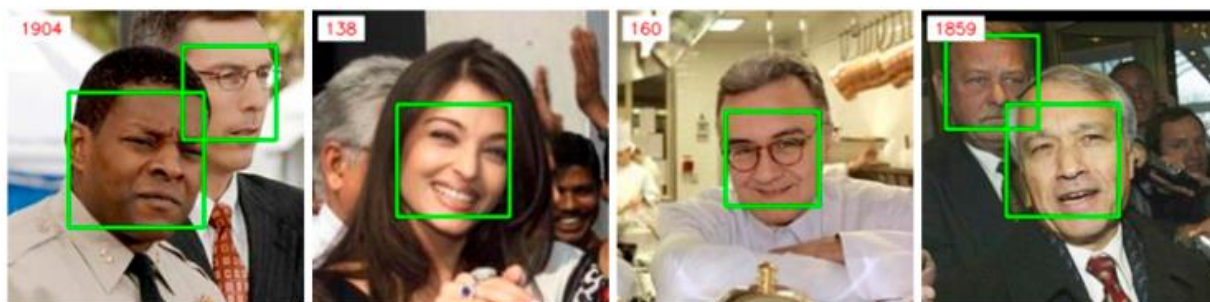


Рис. 7. Результат работы детектора лиц, основанного на CNN

Сравнительный анализ рассматриваемых методов детектирования

В качестве оценки качества использовались следующие метрики: precision (точность), recall (полнота) и F-мера. Иллюстрационное сравнение всех методов представлено на рисунке 8. Эксперименты проводились на машине со следующими характеристиками, указанными в таб. 2. Комплексное сравнение всех рассмотренных методов представлена в таб. 3.

Таблица 2

Параметры, используемого оборудования

Процессор	ОЗУ, ГБ	Тип накопителя	Видеоадаптер
Intel Core i7-7700HQ 2,8 ГГц	24	SSD nVME	Nvidia GeForce GTX 1070

Таблица 3

Сравнительный анализ рассмотренных методов

Набор данных	Размер изображений	Скорость распознавания, FPS	Precision	Recall	F-мера
LFW	250x250	метод Виолы-Джонса			
		≈ 31,7	0,9017	0,8229	0,8605
		метод на основе глубокого обучения в OpenCV (Caffe)			
		≈ 13,6	0,9710	0,9732	0,9721
		метод на основе глубокого обучения в OpenCV (Tensorflow)			
		≈ 13,7	0,9710	0,9732	0,9721
		метод на основе HoG и SVM в Dlib			
		≈ 39,3	1	0,7679	0,8687
		метод на основе CNN в Dlib			
≈ 68,2	0,9976	0,9196	0,9570		

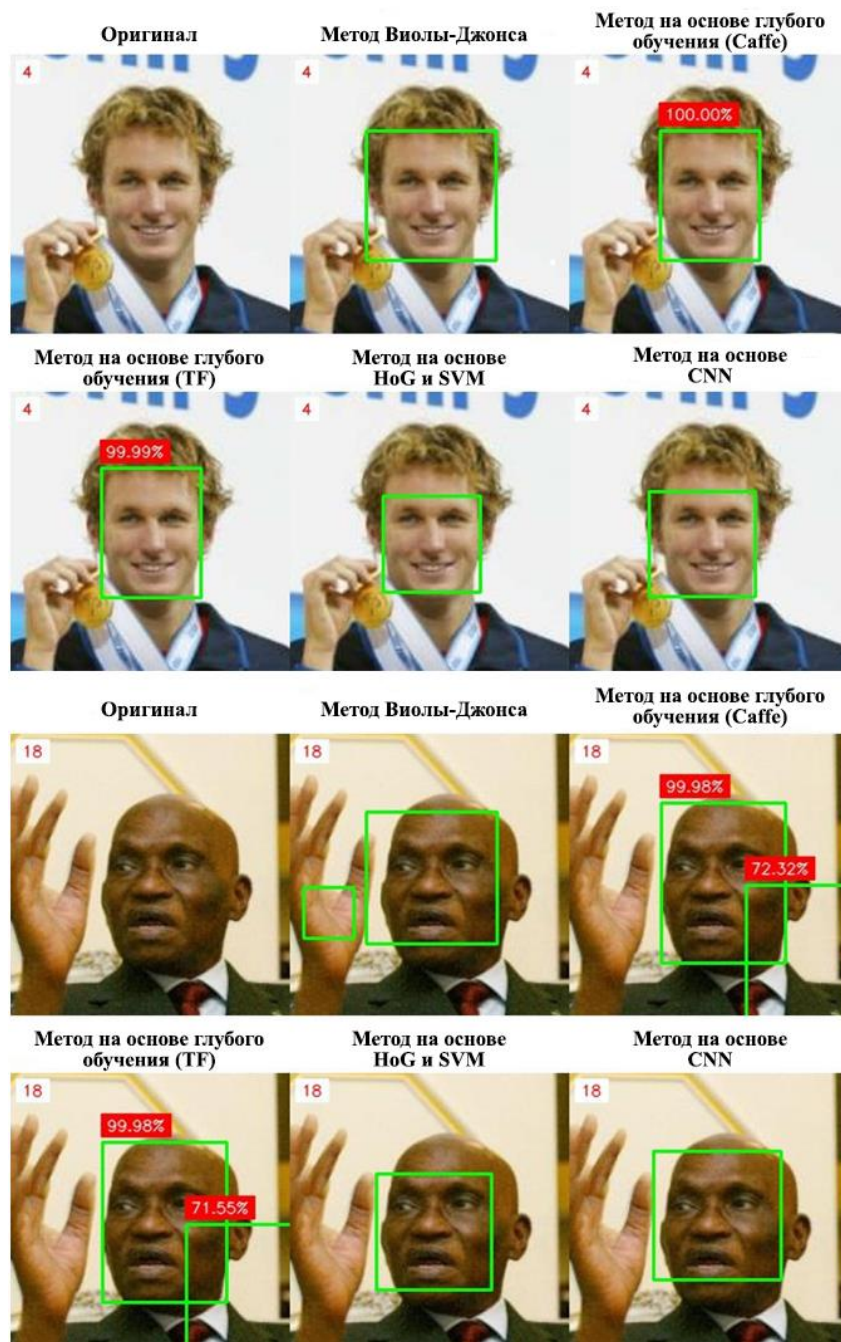


Рис. 8. Сравнительный анализ рассматриваемых методов

Таблица 4

Достоинства и недостатки рассматриваемых методов

Название метода	Достоинства	Недостатки
Детектор лиц на основе метода Виолы-Джонса (OpenCV)	<ul style="list-style-type: none"> — Простота архитектуры; — Обнаружение лица при разных масштабах; — Работает в режиме приближенному к реальному времени на CPU. 	<ul style="list-style-type: none"> — Много ложных срабатываний; — Не определяет лицо в профильном положении; — Не работает при окклюзии.

продолжение таблицы

Название метода	Достоинства	Недостатки
Детектор лиц на основе глубокого обучения (OpenCV)	<ul style="list-style-type: none"> – Точность определения лица; – Обнаружение лица при различной ориентации; – Работа при существенной окклюзии; – Не привязан к масштабу изображений. 	<ul style="list-style-type: none"> – Медленнее всех сравниваемых методов
Детектор лиц на основе функций HoG и SVM (Dlib)	<ul style="list-style-type: none"> – Работа в режиме реального времени на CPU; – Определение лица на фронтальных и слегка не фронтальных положениях; – Облегченная модель по сравнению с остальными; – Работа при незначительной окклюзии. 	<ul style="list-style-type: none"> – Не определяет маленькие лица, поскольку он обучен для минимального размера лица 80×80.
Детектор лиц на основе сверточной нейронной сети	<ul style="list-style-type: none"> – Работает для разных ориентаций лица; – Устойчив к окклюзии; – Очень быстро работает на CPU и GPU. 	<ul style="list-style-type: none"> – Не определяет маленькие лица, поскольку он обучен для минимального размера лица 80×80.

Заключение

В ходе исследования был проведен анализ современных методов детектирования лиц, содержащий комплексное сравнение рассмотренных методов, а также их иллюстрации результатов.

Таким образом, результаты проведенного анализа и экспериментов с различными вариациями детекторов лиц возможно использовать в будущих исследованиях, связанных с задачами по детектированию лиц. С помощью представленных методов можно быстро и довольно точно обнаруживать лица в реальном времени, с наличием разного рода помех, например, окклюзия, которая зачастую встречается в используемой специализированной базе данных. Благодаря своим особенностям каждый детектор может быть использован в задачах биометрии, компьютерном зрении, машинном обучении, автоматических системах распознавания лиц, в зависимости от требуемой задачи.

Литература

1. Аксёнов А.А., Рюмин Д.А., Иванько Д.В. Разработка системы геометрических признаков для автоматического распознавания визуальной русской речи //Информационные технологии в управлении (ИТУ-2018). 2018. С. 526-533.
2. Huang G.B. et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. 2008.
3. King D.E. Dlib-ml: A machine learning toolkit. Journal of Machine Learning Research. V.10. 2009. pp.1755-1758.
4. Viola P., Way O.M., Jones M.J.: Robust real-time face detection. Int. J. Comput. Vision. V.57(2). 2004. pp. 137–154.

5. Déniz O., Bueno G., Salido J., De la Torre F.: Face recognition using histograms of oriented gradients. *Pattern Recognition Letters*. V.32. 2011. pp. 1598–1603.



Акулов Артем Викторович

Год рождения: 1997

Университет ИТМО,

факультет информационных технологий,

и программирования, студент группы № М41212,

направление подготовки: 09.04.02 – Речевые

информационные системы

e-mail: akulov.artem.1997@gmail.com



Шуранов Евгений Витальевич

Год рождения: 1980

Университет ИТМО,

факультет информационных технологий

и программирования,

к.т.н.,

e-mail: evshuranov@itmo.ru

УДК 004

РЕШЕНИЕ ЗАДАЧИ МУЛЬТИЛЕЙБЛИНГА

А.В. Акулов

Научный руководитель – к.т.н. Е.В. Шуранов

Работа выполнена в рамках темы НИР «Нейросетевые подходы к решению задачи классификации акустических событий».

Аннотация

В работе рассмотрены основные подходы в машинном обучении к решению задачи классификации при соответствии нескольких меток одному фрагменту данных, а также проанализирована дальнейшая релевантность исследований на эту тему.

Ключевые слова

Машинное обучение (ml), нейронные сети (nn), классификация.

Традиционно, когда речь идет о классификации каких-либо данных, в машинном обучении задача ставится в виде соответствия одному фрагменту данных лишь одной метки. Однако, такой подход не всегда оправдан ввиду того, что он предполагает взаимоисключающие классы, что случается далеко не всегда.

Обратимся же к теме статьи и опишем, чем является задача мультитейблинга. Задача мультитейблинга в машинном обучении – это задача классификации фрагментов данных (в рассматриваемом случае – акустических данных), в соответствие каждому из которых ставится более чем одна метка класса события [1-4].

Приведем простой пример. Допустим, мы на основании внешних признаков животных (наличие шерсти, зубов и т.д.) пытаемся угадать некоторые их свойства (ядовитый, хищный, травоядный, стадный и т.д.). Составив таблицу, где в качестве данных выступают внешние признаки животных, а целевыми метками являются известные нам их свойства (ядовитый, хищный и т.д.), мы увидим, что мы получаем

задачу мультилейблинга и, если обучиться на таких данных, то можно пытаться предугадать свойства неизвестных животных на основании их внешнего вида.

Подобная задача является применимой во многих областях, в которых применимы методы машинного обучения, к числу которых относятся:

1. Детектирование событий.

Одной из самых крупных баз данных, используемых в задачах по детектированию событий, является AudioSet – база данных Google, в которой содержится порядка 2,000,000 размеченных вручную видеороликов, при общем количестве 550 различных меток.

Более того, метки могут быть как связаны друг с другом иерархически (музыка, духовые инструменты, тромбон), так и быть несвязанными друг с другом совсем (лай, дрель). Пренебрегать таким количеством данных расточительно, поэтому работа с базой предполагает мультилейблинг.

2. Распознавание эмоций.

Задача детектирования эмоций не является тривиальной ввиду того, что спектр эмоций, демонстрируемых человеком, достаточно обширен, чтобы можно было однозначно поставить в соответствие какому-либо речевому фрагменту лишь одну эмоцию, это подтверждают и многие базы со сделанной вручную разметкой аудиофрагментов на предмет наличия эмоций (MOSEI). В такой ситуации, необходимо проверить гипотезу, что каждую эмоцию можно представить в виде сочетания некоторого количества «базовых» эмоций ради потенциального улучшения переносимости результатов разметки одной базы данных на другую, так как оценки становятся более «гибкими».

Подходы к решению задачи

Для подобных задач существует два основных способа решения:

1. Сведение мультилейблинговой базы к «традиционной» базе, где в соответствие одному фрагменту данных ставится лишь одна метка.

Для сведения базы к новому виду приходится прибегнуть к чистке данных, а также выбору метки из совокупности для каждого фрагмента базы.

Исследование было проведено на базе Центра Речевых Технологий (далее ЦРТ). Для эксперимента была рассмотрена выборка из 432 аудиофайлов каждому из которых 7 разметчиков (людей) поставили в соответствие одну эмоцию, характеризующую данный фрагмент. Общий процент однозначно-определяемых ответов составил 17% записей от общего числа. Была выдвинута гипотеза, что некоторые из людей ошибочно определяют эмоции, поэтому необходимо людей, которые часто ошибаются, убрать из рассмотрения при выборе итоговых меток.

Для решения этой задачи помогла метрика оценки совместности распределений Fleiss' карра. Основной плюс данной метрики заключается в том, что она позволяет вычислить степень схожести распределений (в данном случае, оценок аудиофайлов) для любого количества этих распределений (разметчиков), причем результат находится в диапазоне от -1 до 1, где -1 – полное несовпадение оценок, а 1 – их полное совпадение.

Обозначим $(Pe)^-$ – средний процент голосов, отданных одной метке, а P^- – средний процент согласованных разметчиков по фрагменту, тогда:

$$k = \frac{\bar{P} - \bar{Pe}}{1 - \bar{Pe}}$$

Способ выбора разметчика, оценку которого не будем учитывать для трех разметчиков, представлен на рис. 1.

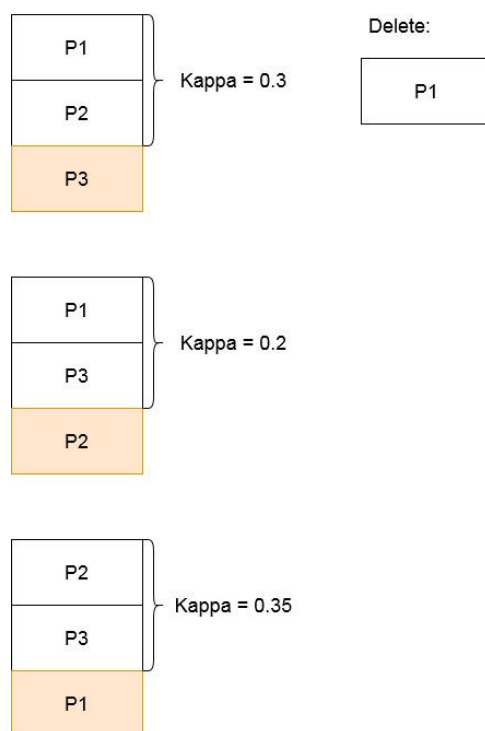


Рис. 1. Способ выбора разметчика

Ввиду того, что нас интересует мнение большинства, а также определенная степень согласованности, удаляем людей до тех пор, пока их количество остается более половины от изначального, а также оценка карра не достигла интересующего нас уровня (в данном случае 0.04), после чего удаляем из рассмотрения каждую запись, согласованность для которой оказалась ниже 0,75.

Также, был протестирован способ удаления разметчиков по общему количеству ошибок относительно самого популярного ответа, результаты однозначно совпавших отметок представлены в таб. 1.

Таблица 1

Результаты чистки базы

Итерации удаления разметчиков	По карра	По непопулярным ответам
1	Злость: 24, Нейтраль: 154	Злость: 24, Нейтраль: 154
2	Злость: 33, Нейтраль: 154	Злость: 27, Нейтраль: 181
3	Злость: 50, Нейтраль: 154, Дефектная эмоция: 2	Злость: 29, Нейтраль: 220
4	Злость: 80, Нейтраль: 154, Дефектная эмоция: 11	Злость: 37, Нейтраль: 253

Как видим, карра стремится сбалансировать выборку, что делает эту метрику достаточно полезной для выбора меток.

2. Использование специализированных нейронных сетей, каждый нейрон выходного слоя которых обрабатывается отдельно.

Традиционно в классификационной нейросети веса выходного слоя пропускаются через функцию, выход которой представляет из себя набор вероятностей того, что данный фрагмент принадлежит той или иной метке (рис. 2).

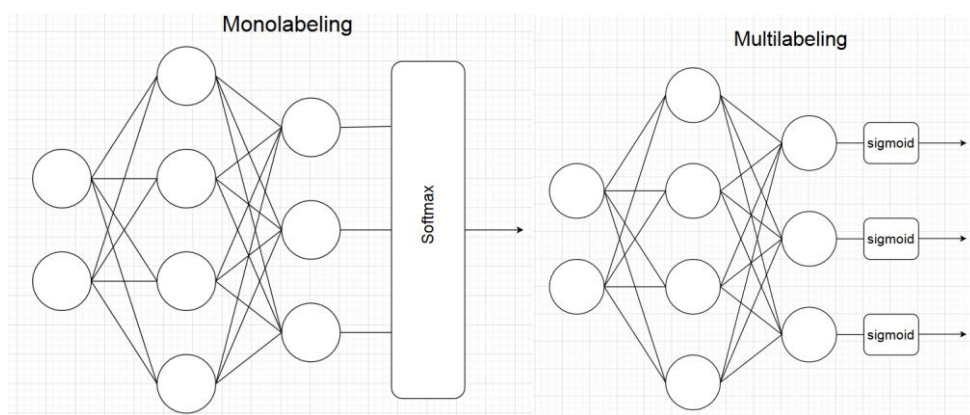


Рис. 2. Нейросетевые подходы к обучению на одной метке и на нескольких

При задаче мультилейблинга, каждый из выходных нейронов обрабатывается отдельно и по итогу вероятности меток выдаются независимо друг от друга (посредством этого несколько меток могут приниматься, как истинные).

Был поставлен эксперимент, где исследовалась возможность нейронной сети обучиться на подобных метках. Гипотеза заключается в том, что при обучении модели без отбрасывания образцов, в оценке класса которых разметчики были неоднозначны, и обучении модели предсказать какое количество разметчиков предсказали каждую из эмоций, повысится значение метрик качества оценки относительно обучения модели только по тем образцам, метки к которым были однозначно определены большинством разметчиков.

При проведении эксперимента использовались следующие наборы образцов:

- IEMOCAP;
- MOSEI;
- RAMAS;
- набор ЦРТ.

В ходе эксперимента модель была обучена на данных набора ЦРТ и протестирована на вышеуказанных выборках (таб. 2, 3).

Таблица 2

Результаты обучения на одной метке

Метрики оценки	Набор тестовой выборки			
	IEMOCAP	MOSEI	RAMAS	набор ЦРТ
Точность	0.412	0.298	0.328	0.346
Полнота	0.549	0.045	0.148	0.702
F1-мера	0.404	0.069	0.172	0.352

Таблица 3

Результаты обучения на нескольких метках

Метрики оценки	Набор тестовой выборки			
	IEMOCAP	MOSEI	RAMAS	набор ЦРТ
Точность	0.339	0.303	0.006	0.243
Полнота	0.470	0.005	0.024	0.684
F1-мера	0.187	0.012	0.012	0.242

По полученным результатам можно сделать вывод о том, что переносимость результатов улучшается на наборах образцов, в которых распределение классов близко к тому, который использовался для обучения модели, такими наборами можно считать RAMAS и IEMOCAP.

Ухудшение результатов связано с тем, что в наборе находится два основных класса эмоций (злость и нейтральность), что могло привести к переобучению.

Литература

1. AudioSet, URL: <https://research.google.com/audioset/> (дата обращения 27.02.2020);
2. IEMOCAP[2] (The Interactive Emotional Dyadic Motion Capture) URL: <https://sail.usc.edu/iemocap/> (дата обращения 27.02.2020).
3. CMU-MOSEI или MOSEI [3] (CMU Multimodal Opinion Sentiment and Emotion Intensity) URL: <https://github.com/A2Zadeh/CMU-MultimodalSDK> (дата обращения 27.02.2020).
4. RAMAS [4] (Russian Multimodal Corpus of Dyadic Interaction for Affective Computing) URL: <https://neurodatalab.com/science/publications/ramas-russian-multimodal-corpus-of-dyadic-interaction-for-affective-computing-for-specem-2018/> (дата обращения 27.02.2020).



Астапов Сергей Сергеевич

Год рождения: 1988
Университет ИТМО,
факультет информационных технологий
и программирования,
ст.н.с., PhD,
e-mail: astapov@speechpro.com



Лаврентьев Александр Валерьевич

Год рождения: 1985
Университет ООО «ЦРТ-Инновации»,
н.с.,
e-mail: lavrentyev@speechpro.com



Кабаров Владимир Иосифович

Год рождения: 1959
Университет ИТМО,
факультет информационных технологий
и программирования, старший преподаватель,
e-mail: kabarov@speechpro.com



Матвеев Юрий Николаевич

Год Рождения: 1955
Университет ИТМО,
факультет информационных технологий
и программирования,
д.т.н., профессор,
e-mail: matveev@speechpro.com

УДК 004.934.2

**ИСПОЛЬЗОВАНИЕ ПРОСТРАНСТВЕННОЙ АКУСТИЧЕСКОЙ
ИНФОРМАЦИИ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ
ИДЕНТИФИКАЦИИ ДИКТОРОВ НА ДАЛЬНОМ МИКРОФОНЕ**

С.С. Астапов, А.В. Лаврентьев, В.И. Кабаров

Научный руководитель – д.т.н., профессор Ю.Н. Матвеев

Аннотация

Рассмотрена возможность использования пространственной акустической информации для повышения точности идентификации дикторов на дальнем микрофоне. Пространственная информация получается с помощью метода акустической локализации SRP-PHAT и используется при диаризации и идентификации дикторов в

сценарии совещания. Результаты экспериментов показывают, что использование пространственных признаков повышает качество идентификации на произнесениях дольше 2 секунд.

Ключевые слова

Идентификация дикторов, микрофонные решетки, SRP-PHAT.

Введение

Качество идентификации дикторов на дальнем микрофоне страдает от неизвестного уделенного расположения диктора и наличия шумов в помещении записи [1]. В данной работе рассматривается возможность применения пространственных акустических признаков для повышения качества идентификации в сценарии совещания. Для акустической локализации мы рассматриваем вариант микрофонной решетки, распределенной по периметру стола для совещаний, где каждый микрофон удален от дикторов как минимум на 1 метр. Локализация производится с использованием метода SRP-PHAT [2], идентификация осуществляется с помощью х-векторной FTDNN [3, 4] речевой модели.

Описание предложенного подхода

Идентификация дикторов рассматривается в рамках сценария совещания, где дикторы расположены за столом для совещаний большой площади, микрофоны расположены по периметру стола на расстоянии 0,6 м от края. Расстояние между микрофонами 0,5–0,8 м, расстояние до диктора 1–1,5 м. Расположение дикторов случайно: рядом с одним микрофоном может находиться больше одного диктора, у других дикторы могут отсутствовать. Расположение микрофонов и вариант возможной рассадки дикторов представлены на рис. 1.

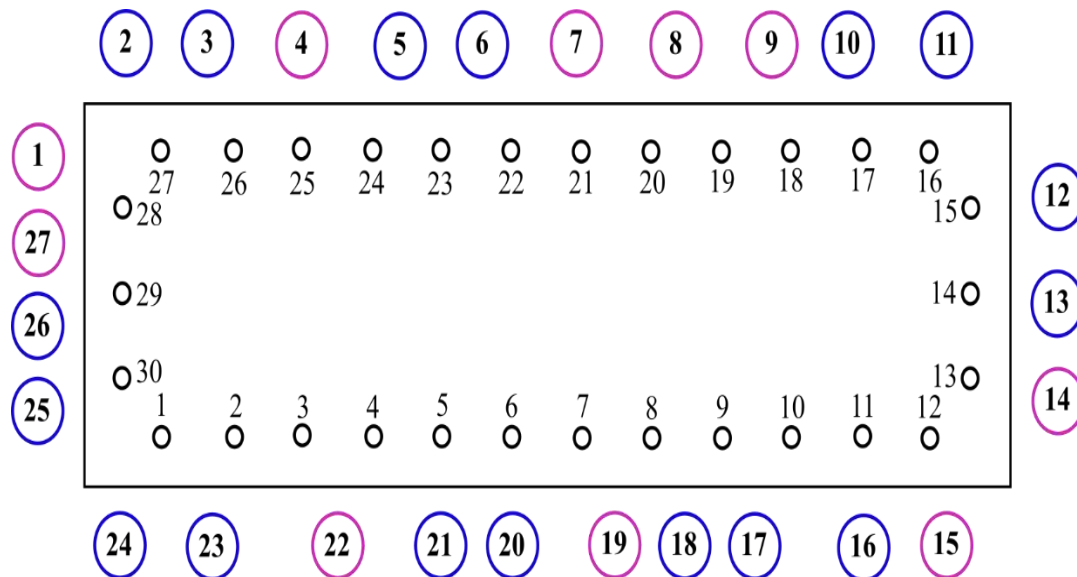


Рис. 1. Схема расположения микрофонов на столе и возможной рассадки дикторов

Схема предложенной системы применения пространственных признаков, полученных с помощью процедуры акустической локализации, для повышения качества идентификации представлена на рис. 2. Многоканальный поток данных сегментируется на участки активности речи с помощью VAD (Voice Activity Detector), затем производится диаризация с применением х-векторной TDNN-LSTM модели [3]. Короткие последовательные сегменты, принадлежащие одному и тому же диктору, склеиваются в более продолжительные произнесения, и по ним затем производится

идентификация с применением x -векторной FTDNN модели [4]. Процессы диаризации и идентификации снабжаются пространственными признаками (x, y, W_{SRP}) для каждого m из всех M каналов, где x, y – оценка координат источника звука, W_{SRP} – т.н. SRP энергия в этой точке координат.

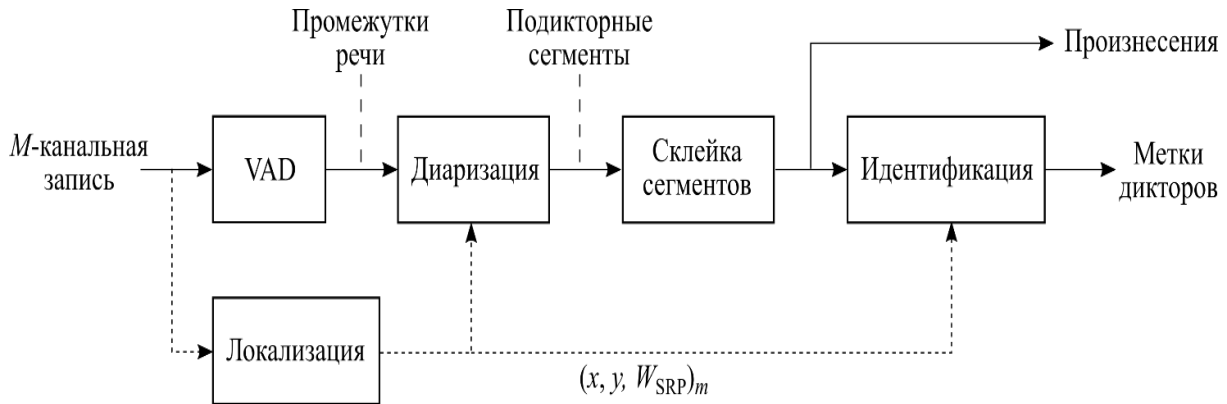


Рис. 2. Схема системы применения акустической локализации для получения пространственных признаков в процессе идентификации дикторов

Акустическая локализация

В качестве метода акустической локализации мы применяем широкополосный метод энергии нацеленного отклика с коэффициентом фазового преобразования SRP-РНАТ. Этот метод является одним из самых робастных и терпимых к реверберации методов акустической локализации [2]. Он основывается на оценке т.н. SRP энергии в каждой точке пространства поиска источника звука.

Допустим, что имеется множество точек \mathbf{a} в системе декартовых координат. SRP энергия на этом множестве $P(\mathbf{a})$ вычисляется как объединенная по всем парам каналов обобщенная кросс-корреляция с фазовыми сдвигами, соответствующими разнице прихода звуковой волны с этих точек. Представим, что имеется пара сигналов $x_k(t), x_l(t)$ с двух микрофонов микрофонной решетки. Время прихода звуковой волны из точки $\mathbf{a} \in \mathbf{a}$ для этих двух микрофонов составляет $\tau(\mathbf{a}, k)$ и $\tau(\mathbf{a}, l)$, соответственно. Временная задержка между сигналами тогда будет равна $\tau_{kl}(\mathbf{a}) = \tau(\mathbf{a}, k) - \tau(\mathbf{a}, l)$. SRP-РНАТ для всех пар микрофонов выражается в следующем виде:

$$P(\mathbf{a}) = \sum_{k=1}^M \sum_{l=k+1}^M \int_{-\infty}^{\infty} \Psi_{kl} X_k(\omega) X_l^*(\omega) e^{j\omega\tau_{kl}(\mathbf{a})} d\omega,$$

где $P(\mathbf{a})$ – значение SRP энергии, $X(\omega)$ – спектр одного окна кратковременного преобразования Фурье (STFT) на угловой частоте ω , $(\cdot)^*$ – операция сопряжения комплексной части, j – мнимая единица, $\Psi_{kl} = (|X(\omega)X^*(\omega)|)^{-1}$ – коэффициент РНАТ. Положение источника звука определяется по координатам точки или множества точек с максимальным значением SRP энергии.

Для снижения объема требуемых вычислительных ресурсов для подсчета кросс-корреляции в SRP-РНАТ мы используем алгоритм поиска единственного источника в ограниченной зоне поиска, именуемый методом стохастического сужения региона поиска SRC (Stochastic Region Contraction) [5]. Для покрытия обширной зоны поиска по площади стола предлагается применить разбиение общей зоны на подзоны поиска, в каждой из которых локализацию проводят 3 самых ближних микрофона. Пределы этих локальных зон определяются как: 0,1 м от микрофонов, на расстоянии 1,5 м; 0,1 м влево от самого левого микрофона тройки, 0,1 м вправо от самого правого микрофона. Исключение составляют микрофоны на углах четырехугольного стола: их левый или

правый пределы продлены до угла стола для обеспечения полного покрытия. Для геометрии, представленной на рис. 1, установленные зоны поиска представлены на рис. 3.

Результаты экспериментов

Эксперимент проводился на записи, сделанной во время совещания, с участием 27 человек – 17 мужчин и 10 женщин. Стол для совещаний имеет размеры 8,8 × 2,8 м, оснащён 30 микрофонами, запись с которых ведется синхронно. Расположение дикторов и микрофонов представлено на рис. 1 и 3, соответственно. Длина записи совещания составила 83 минуты. Длина STFT окна локализации SRP-PHAT и выделения признаков составила 200 мс.

Таблица 1

Результаты измерения ошибки идентификации на тестовых данных

Длина сегмента	Ошибка идентификации IER (%)	
	Без пространственных признаков	С пространственными признаками
< 2 с	42,6	46,7
2–10 с	29,1	22,8
> 10 с	11,6	9,8

Результаты локализации для полной записи представлены на рис. 3. Кластеризация точек оценки положения источников звука выявляет положения дикторов с достаточной точностью для пространственного разделения дикторов. Исключение составляют результаты на микрофонах 28–30, поскольку диаграмма направленности этих микрофонов не предназначена для записи удаленных дикторов. Результаты оценки ошибки идентификации IER представлены в таб. 1. Ошибка идентификации при применении пространственных признаков, полученных с помощью акустической локализации, понижается на произнесениях длиной больше 2 секунд.

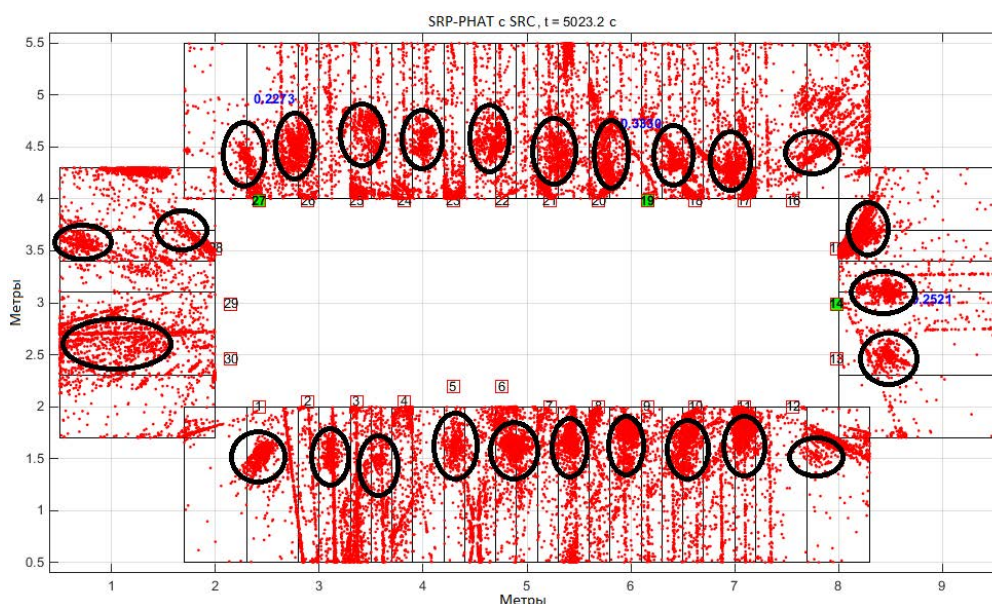


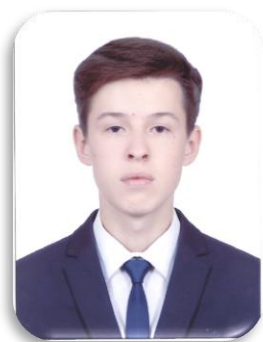
Рис. 3. Результаты локализации дикторов на записи совещания. Квадраты с цифрами – позиции микрофонов; черные четырехугольники – пределы локальных зон поиска источников звука; красные точки – результаты локализации на каждом STFT кадре; черные овалы – кластеры локализации положения дикторов.

Заключение

В работе была проверена возможность применения пространственных признаков для задачи идентификации дикторов в сценарии совещания. Результаты показывают, что применение акустической локализации на распределенной решетке микрофонов способно обеспечивать качественные пространственные признаки, применение которых снижает ошибку идентификации на произнесениях дольше 2 секунд.

Литература

1. Zelenák M., Segura C., Luque J., Hernando J. Javier. Simultaneous Speech Detection with Spatial Features for Speaker Diarization // IEEE Transactions on Audio, Speech & Language Processing. Vol. 20. pp. 436-446, 2012.
2. DiBiase J.H. A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays // PhD thesis, Brown University, Providence, RI, May 2000.
3. Novoselov S. et al. Speaker Diarization with Deep Speaker Embeddings for DIHARD Challenge II // Interspeech 2019, pp. 1003–1007.
4. Novoselov S. et al. STC speaker recognition systems for the VOICES from a distance challenge // Interspeech 2019, pp. 2443–2447.
5. Do H., Silverman H.F., and Y.Yu. A real-time SRP-PHAT source location implementation using stochastic region contraction (SRC) on a large-aperture microphone array // IEEE Int. Conf. Acoustics, 227 Speech and Signal Processing (ICASSP). vol. 1, 2007, pp. 121–124.



Барсуков Алексей Валерьевич
Университет ИТМО,
факультет инфокоммуникационных технологий,
студент группы №К42402,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: alexbarsukov@hotmail.com

УДК 004.921

ИССЛЕДОВАНИЕ ПОПУЛЯРНЫХ АЛГОРИТМОВ ПРОЦЕДУРНОЙ ГЕНЕРАЦИИ ЛАНДШАФТОВ НА ВОЗМОЖНОСТЬ УСОВЕРШЕНСТВОВАНИЯ

А.В. Барсуков

Научный руководитель – к.т.н., доцент А.В. Куркин

Работа выполнена в рамках темы НИР «Исследование алгоритмов процедурной генерации ландшафтов на основе шумов».

Аннотация

В статье приведены результаты исследования алгоритмов diamond-square и шума Перлина, способов повышения реалистичности генерируемых ландшафтов. Выведены способы усовершенствования рассматриваемых алгоритмов. Выведен принцип распределения природных зон (биомов) на сгенерированном ландшафте. Описан способ генерации ландшафта на основе реальных топографических данных и способ их получения. Сделаны выводы об актуальности темы.

Ключевые слова

Процедурная генерация контента, процедурная генерация ландшафтов, карта высот, шум, шум Перлина, diamond-square, биомы, srtm.

Процедурная генерация – это способ создания контента «без участия человека». То есть контент создается автоматически с помощью вычислительных алгоритмов. Эти алгоритмы создают модели домов, ландшафтов или других необходимых объектов на основе заранее заданных параметров [1].

Проблемой является большая доля кропотливого ручного труда при генерации ландшафтов и другого контента. Поэтому возникает необходимость увеличить долю процедурной генерации контента.

Целью настоящей исследовательской работы является исследование и изучение алгоритмов процедурной генерации ландшафтов, а также способов программной реализации этих алгоритмов визуализации и выявление сильных и слабых сторон этих алгоритмов.

В исследованиях алгоритмов генерации ландшафтов особенный интерес представляет проблема процедурного создания карт высот, которые лежат в основе всех задач. Решение данной проблемы обладает практической значимостью, поскольку процедурная генерация карт высот позволит существенно разнообразить ландшафт за счет динамического изменения координат. Кроме того, такие системы генерации могут

быть полезны исследователям алгоритмов искусственного интеллекта в сфере универсальных игровых программ, а также игровым дизайнерам для быстрого создания и тестирования различных игровых прототипов.

Процедурная генерация ландшафтов и другого контента становится очень популярной. В настоящее время это применяется в кинематографе при генерации сцен второго плана, а также в игровой индустрии. Именно процедурная генерация позволяет создавать изображения со специальными характеристиками динамическим образом. Таким образом можно получить большое количество моделей в короткие сроки без кропотливого ручного труда.

Так как весь контент генерируется процедурно, то такой подход позволяет во много раз сократить объем требуемого пространства на накопителе.

Раньше художникам приходилось либо использовать реальные фотографии и на их основе делать модель, либо создавать текстуру с нуля. У этих подходов есть существенные недостатки, например такие как отсутствие быстрого редактирования этих моделей, а также невозможность вносить дополнительные параметры по ходу. Эти два подхода являются представителями рутинного труда.

Преимуществами процедурной генерации моделей являются:

- внесение изменений в модель на любом этапе работы;
- изменение размеров модели без потерь качества;
- скорость генерации контента;
- Небольшой размер данных. Так как этот размер не зависит от разрешения изображения. В файле хранится только алгоритм;

Создание контента содержит в себе большое количество проблем. Например, проблема большого количества необходимых работ для создания какой-либо модели, проблема статичности моделей (они тяжело поддаются изменениям), а ведь в процессе разработки требования к модели могут изменяться.

Карта высот

Чтобы понять базовый принцип процедурной генерации ландшафтов надо иметь представления об основах. В основе любого ландшафта лежит карта высот. Карта высот – двумерный массив, каждый элемент которого является координатой точки, то есть каждый элемент массива является точкой высоты (рис. 1).

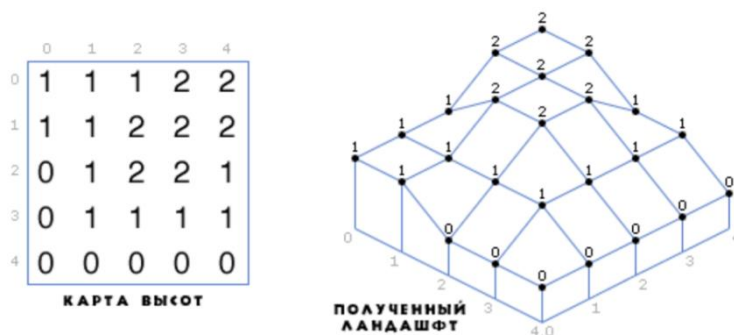


Рис. 1. Карта высот

Карта высот лежит в основе всех базовых алгоритмов процедурной генерации ландшафтов. Для дальнейшей работы с картой высот её нужно визуализировать, самый распространенный способ визуализации карт высот заключается в раскраске пикселей полутонами, где самый низкий пиксель – черный цвет, а самый высокий пиксель – белый цвет [2].

Такой результат визуализации позволяет наглядно работать с результатом и вносить изменения по необходимости.

Шумы

Шум – набор случайных чисел в заданном диапазоне без правил и корреляции между двумя соседними числами. Карта высот сгенерированная произвольным способом является идеальным примером шума.

Процедурный Шум – это шум, который реализуется программными алгоритмами. Наиболее известным из таких алгоритмов является Шум Перлина (Perlin's Noise).

Алгоритмы шума имеют широкое применение в области компьютерной графики и чаще всего используются при генерации текстур и создании сложных графических объектов, таких как огонь, облака, туман. Также, шум Перлина применяется при генерации ландшафта [7].

Шум Перлина

Кен Перлин придумал свой шум в 1983 году, в те времена, когда все пытались достичь реализм в картинке. По своей сути, шум является генератор случайных чисел, которые в последствии визуализируются по определенным правилам. Шум Перлина – «разносторонний» алгоритм, используемый для генерации ландшафтов, облаков, огня. Воды и так далее.

Шум Перлина – градиентный шум, из псевдослучайных векторов, расположенных в определенных точках. (По необходимости применяются функции сглаживания между точками и интерполяция.)

Как было сказано выше, рассматриваемый в данной главе алгоритм шума основан на генерации псевдослучайных чисел и все видимые элементы текстуры обладают одинаковым размером, что позволяет легко ими управлять. Так же наложение различных функций Шума Перлина друг на друга образуют октавы, которые позволяют создавать самые разнообразные текстуры.

Для повышения реалистичности ландшафта, можно взять готовую карту высот (сетку), провести зашумление по необходимому правилу и тем самым процедурно сгенерировать новый ландшафт используя шум Перлина. Таким образом создается шум на исходной сетке, где каждая точка находится в диапазоне от 0 до 1. Выбранный масштаб будет влиять на сложность шума. Выражаясь простым языком, большее число квадратов на сетке – плотный шум, меньшее число квадратов – шум похожий на облака [7].

В каждой точке на сетке строится случайный вектор нормали. Общепринятый способ создания таких векторов – организация справочной таблицы из 256 векторов, которые охватывают полный круг, и последующий случайный выбор одного из них для каждой точки на сетке. Это обеспечивает распределение векторов, которые могут с одинаковой вероятностью указывать любое направление. Следующий шаг – создать четыре диагональных вектора, соединяющих углы ячейки с текущим пикселем (рис. 2) [4-5].



Рис. 1. Карта высот созданная с помощью шума Перлина

Модифицированный алгоритм diamond-square

Модификация рассматриваемого алгоритма заключается в следующем: помимо разбиения сетки на квадраты и «вытягивания» высот, происходит разбиение на ромбы, одна из диагоналей которого должна являться частью границы сетки. Тем самым происходит вытягивание крайних точек высот и как следствие, повышается реалистичность сгенерированного ландшафта (рис. 3).

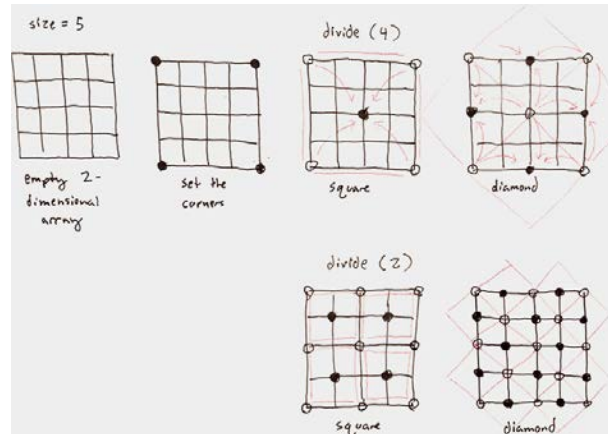


Рис. 2. Схематичное описание алгоритма diamond-square

Определение природных зон на сгенерированной карте высот

Полученная карта высот нуждается в визуализации и доведении ее до вида ландшафта. Для этого необходимо вывести закономерности и правила, по которым будут располагаться природные зоны (биомы).

Для начала необходимо определиться с точкой уровня воды, тогда из этого следует, что все что ниже «точки воды», будет «затапливаться» водой [6].

Аналогично воде, задавая высоту, можем определить зону откуда начинается снег, таким образом получают заснеженные вершины.

Так же мы можем сгенерировать с помощью шума – карту влажности и наложить ее на карту высот. И тогда нужно определиться какая ось будет «осью влажности», а какая ось «осью высоты»

Малые высоты — это океаны и побережья. Большие высоты каменистые и заснеженные. В промежутке между ними мы получаем широкий диапазон биомов.

Еще одна из идей – учет климата. Нужно условиться и обозначить север и юг. В предыдущем этапе климат зависел от высоты, чем выше, тем холоднее и заснеженнее. Из этого следует, что в алгоритм необходимо добавить параметр, отвечающий за стороны света с помощью которого, можно будет распределять температурные зоны [2-3].

Так выглядит вышеописанная идея распределения природных зон в псевдокоде:

```
function Биомы(a, d) {  
  if (a < 0.1) return Океан;  
  if (a < 0.11) return Прибрежная зона;  
  if (a > 0.7) {  
    if (d < 0.3) return Степь;  
    if (d < 0.6) return Тундра;  
    return Снег;  
  }  
  if (a > 0.6) {  
    if (d < 0.33) return Пустыня;  
    if (d < 0.66) return Низкорослые леса;  
    return TAIGA;  
  }  
  if (a > 0.3) {
```

```
if (d < 0.50) return Луга;  
if (d < 0.83) return Леса;  
return Леса;  
}  
if (d < 0.19) return Субтропическая пустыня;  
if (d < 0.44) return Луга;  
if (d < 0.77) return Тропический лес;  
return Тропический влажный лес;  
}
```

Этот алгоритм способен определять, на каком участке будет определенная природная зона и это позволит наполнять карту различными географическими признаками.

Повышение реалистичности генерируемого ландшафта

Иногда на практике необходимо сгенерировать ландшафт очень похожий, на какой-нибудь участок местности нашей планеты, но со своими особенностями и отличиями. Самый действенный и результативный способ решения подобной задачи – зашумление реальной карты высот.

В настоящее время наиболее используемым открытым источником пространственных данных является радарная съемка поверхности Земли, проведенная NASA в 2000 г. с борта космического шаттла, отсюда миссия получила следующее название – Shuttle radar topographic mission (далее SRTM).

При решении задач определения точности данных SRTM, необходимо понимать, что данные представляют непрерывные значения осредненных высот отражающей поверхности (снега, деревья, застройка и т.д.)

Идея данного способа заключается в том, чтобы получить реальную карту высот, соответствующую реальному участку планеты и подавать ее на вход в любой из алгоритмов для последующего зашумления (изменения). Тем самым на выходе получается реалистичный ландшафт, соответствующий выбранному участку местности.

SRTM предоставляют данные в формате geoTiff – открытый формат данных, содержащий в себе информацию о географической привязке, информацию о системе географических координат, модель геоида и другую различную информацию.

Подводя итог, получаемый файл содержит данные по высотам точек земной поверхности.

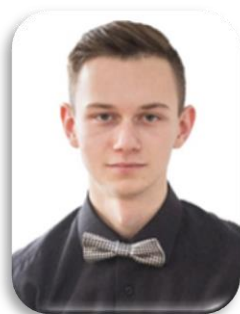
Вывод

В данной статье содержится краткая информация о шуме Перлина, алгоритме «diamond-square», способах их модернизации, а также изложены идеи по повышению реалистичности генерируемых ландшафтов. На основе данной статьи можно сделать вывод, что рассматриваемая тема обладает большим потенциалом развития и почти неограниченна в исследовании и совершенствовании.

Литература

1. Unreal Engine Features [Электронный ресурс] // Epic Games : Возможности движка Unreal Engine 4. Режим доступа: <https://www.unrealengine.com/unreal-engine-4> – яз. англ. (Дата обращения 01.02.2020).
2. Generate Random Cave Levels Using Cellular Automata [Электронный ресурс] / Michael Cook // EnvatoTuts+. Режим доступа: <http://gamedevelopment.tutsplus.com/tutorials/generate-random-cave-levels-using-cellular-automata--gamedev-9664> – яз. англ. (Дата обращения: 01.02.2020).

3. Random level generation in Wasteland Kings [Электронный ресурс] / Jan Willem Nijman // Vlambeer . Режим доступа: <http://www.vlambeer.com/2013/04/02/random-level-generation-in-wasteland-kings/>. яз. англ. (Дата обращения 01.02.2020).
4. Map representations [Электронный ресурс] // Amit's Thoughts on Pathfinding. – Режим доступа: <http://theory.stanford.edu/~amitp/GameProgramming/MapRepresentations.html>. яз. англ. (Дата обращения 01.02.2020).
5. More than 1.2 billion people are playing games [Электронный ресурс] / Dean Takahashi // GamesBeat. Режим доступа: <http://venturebeat.com/2013/11/25/more-than-1-2-billion-people-are-playing-games/>. яз. англ. (Дата обращения 01.02.2020).
6. Supreme Court sees video games as art [Электронный ресурс] / John D. Sutter // CNN.com. Режим доступа: <http://edition.cnn.com/2011/TECH/gaming.gadgets/06/27/supreme.court.video.game.art/>. яз. англ. (Дата обращения 01.02.2020).
7. Static Meshes | Unreal Engine [Электронный ресурс] // Epic Games. Режим доступа: <https://docs.unrealengine.com/latest/INT/Engine/Content/Types/StaticMeshes/>. яз. англ. (Дата обращения 01.02.2020).



Беген Петр Николаевич
Университет ИТМО,
Институт дизайна и урбанистики,
студент группы № С42551,
направление подготовки: 09.04.03 – Прикладная информатика,
e-mail: peetabegen@yandex.ru



Чугунов Андрей Владимирович
Университет ИТМО,
Институт дизайна и урбанистики,
к. полит. н., доцент,
директор Центра технологий электронного правительства,
e-mail: chugunov@itmo.ru

УДК 004.89

АНАЛИЗ РЕЗУЛЬТАТИВНОСТИ РАБОТЫ ИНТЕЛЛЕКТУАЛЬНОГО КЛАССИФИКАТОРА СООБЩЕНИЙ ГРАЖДАН НА ПОРТАЛЕ «НАШ САНКТ-ПЕТЕРБУРГ»

П.Н. Беген

Научный руководитель – к.полит.н., доцент А.В. Чугунов

Работа выполнена при поддержке РФФ, проект №18-18-00360 «Электронное участие как фактор динамики политического процесса и процесса принятия государственных решений» в рамках магистерской НИР «Разработка интеллектуального классификатора сообщений граждан на портале «Наш Санкт-Петербург».

Аннотация

В работе представлены основные результаты разработки интеллектуального классификатора, являющимся одним из подходов оптимизации процесса подачи сообщения на портал «Наш Санкт-Петербург» и его дальнейшей проверки специальными службами. Представлены метрики результативности и релевантности разработанного решения, приведены результаты и анализ тестирования на реальных данных, определены перспективы дальнейших работ.

Ключевые слова

Классификация, машинное обучение, искусственный интеллект, обработка естественного языка, электронное участие граждан, портал электронного участия.

С момента запуска в 2014 году портал «Наш Санкт-Петербург» стал действенным инструментом для решения общегородских проблем. Основной целью создания портала являлось обеспечение эффективного и продуктивного взаимодействия между жителями региона и сотрудниками исполнительной государственной власти, государственных учреждений, управляющих компаний [1]. Перед создателями портала было поставлено выполнение следующих задач:

- обеспечить оперативное решение проблем жителей региона;
- повысить прозрачность деятельности органов исполнительной власти региона;
- привлечь жителей в управление регионом (т. е. обеспечить эффективное электронное участие граждан);
- повысить уровень удовлетворенности жителей региона работой органов исполнительной власти;
- обеспечить внедрение принципов Открытого правительства и Электронной демократии.

Портал на сегодняшний день уверенно развивается и активно используется: по состоянию на январь 2020 г. жителями региона подано более 2.1 млн сообщений о городских проблемах и решено более 96% от общего числа сообщений. Число зарегистрированных пользователей постоянно растет и на данный момент составляет около 160 тыс. человек. Ежедневно жители отправляют на портал более 2.5 тыс. сообщений по различным категориям проблем, зафиксированных в Классификаторе портала. В совокупности это приводит к возрастающей нагрузке на систему портала и работу исполнительных органов, что соответствующим образом сказывается на снижении продуктивности деятельности различного рода служб. Постоянно возрастающая нагрузка на портал и ограниченный кадровый потенциал Городского мониторингового центра (который представляет службу модерации на портале) стимулируют постановку задач по оптимизации имеющегося функционала портала в целях своевременного исполнения строго зафиксированных сроков отработки сообщений в соответствии с регламентом за счет более быстрой и качественной обработки поступающих сообщений граждан для их дальнейшей отработки органами исполнительной власти, подведомственными организациями и управляющими компаниями.

Одной из проблем портала «Наш Санкт-Петербург» является сложность в действиях для пользователя в процессе подачи сообщения на портал, которая заключается в самостоятельном выборе жителем одной из 200 доступных категорий проблем. В результате пользователь часто допускает ошибку в выборе правильной категории. Согласно статистике, размещенной на сайте Администрации Санкт-Петербурга, по причине неправильно выбранной категории проблемы отклоняется от 20 до 25% всех сообщений, поступающих от пользователей. В результате отклонение почти 1/4 сообщений могут негативно влиять на степень удовлетворенности жителями работой и удобством портала в целом, а также деятельностью исполнительных органов.

Другой проблемой, с которой уже сталкиваются модерационные службы портала, является комплексность бизнес-процесса проверки и маршрутизации сообщения, которая приводит к несоблюдению сроков отработки сообщений согласно регламенту из-за возрастающей нагрузки на портал (увеличение количества пользователей, сообщений, категорий проблем) и рост ежедневного количества сообщений в пиковые дни (например, наличие сезонности), в следствие чего служба модерации вынуждена увеличивать данный срок отработки (что не соответствует установленному регламенту) и оперативно сообщать об этом на портале для информирования граждан.

В качестве одного из решений данной проблемы было предложено разработать и внедрить автоматическую классификацию сообщений граждан в систему портала. В целях минимизации риска ошибочного определения категории пользователем и повышения продуктивности деятельности службы модерации по проверке и отработке поступающих сообщений, были сформулированы следующие подходы:

в процессе подачи сообщения о проблеме исключить обязательство для пользователя самостоятельно выбирать категорию проблемы из Классификатора или

вводить ключевые слова в поисковую форму: для этого пользователю достаточно сразу описать проблему в виде текста (аннотации). Последующий порядок действий в процессе подачи, таких как указание местоположения существующей проблемы на карте и загрузка подтверждающих проблему фотографий, сохранить.

Для службы модерации разработать модуль автоматической классификации текста сообщения, который представит результат работы в виде ранжированного списка из трех определенных алгоритмом категорий с соответствующим процентом точности классификации для последующего выбора модератором.

На рис. 1 представлена схема в нотации UML 2.0, фиксирующая предлагаемые изменения в процедуре подачи гражданином сообщения на портал и последующей проверки сообщения службами модерации.

Для реализации подходов было предложено разработать интеллектуальный классификатор сообщений граждан, представляющий веб-модуль, главным компонентом которого является алгоритм автоматической классификации сообщений, основанный на методах машинного обучения и обработки естественного языка.

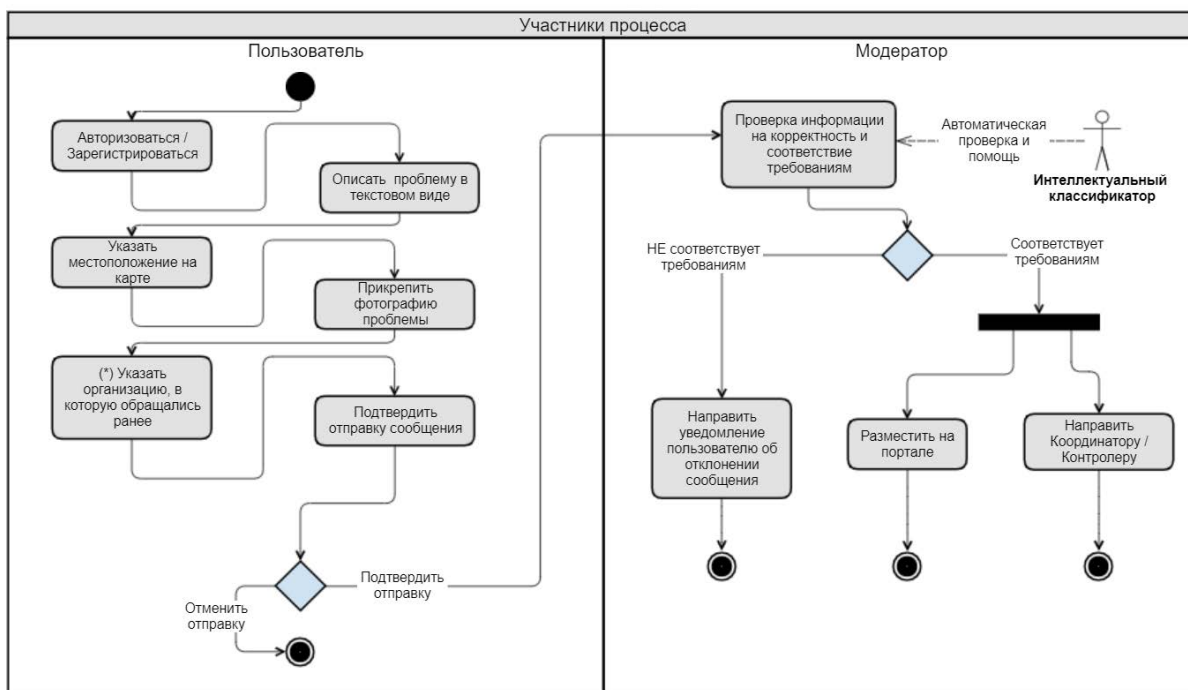


Рис. 1. Процедура подачи сообщения на портал и его модерации с использованием интеллектуального классификатора

Работа алгоритма классификации заключена в том, чтобы в автоматическом режиме по тексту сообщения, поданного гражданином на портале «Наш Санкт-Петербург», определить категорию проблемы, зафиксированной в Классификаторе, с некоторой долей вероятности, выраженной в процентах. При последующей проверке сообщения на основе данной доли служба модерации готовит окончательное решение о правильности и корректности выбора категории пользователем и производит дальнейшую работу с сообщением.

Для технической реализации алгоритма, которая велась на языке программирования Python 3.6.5, использованы основные методы машинного обучения, известные и используемые при работе с текстовой классификацией: наивный байесовский классификатор [2], дерево решений [3], метод опорных векторов (SVM) [4] и искусственные нейронные сети [5]. В качестве нейронных сетей были предложены три сети с разной архитектурой и конфигурацией: сеть прямого распространения

(FFNN), сверточная сеть (CNN) и рекуррентная сеть (RNN) с LSTM-блоком. В результате разработки по каждому методу были получены метрики, показывающие степень результативности произведенного обучения, выраженные в процентах.

В качестве методов обработки естественного языка использовались:

- предварительная очистка данных;
- токенизация;
- лемматизация / стемминг;
- биграммы (т. е. фразы из 2-ух слов);
- работа со списком стоп-слов, удаление их из текста;
- вычисление меры TF-IDF, сокращение векторной размерности;
- формирование модели «мешка слов»;
- формирование модели «связный мешок слов» (Continuous bag-of-words);
- формирование модели Word2Vec;
- разделение на обучающую и тестовую выборки в соотношении 80% на 20%.

Разработка интеллектуального классификатора сообщений граждан подразумевала проектирование веб-модуля с предоставлением API-доступа к методам алгоритма автоматической классификации сообщений и непосредственно программную реализацию данного модуля. Схема основной структуры и логики работы интеллектуального классификатора сообщений на концептуальном уровне в нотации UML 2.0 представлена на рис. 2.

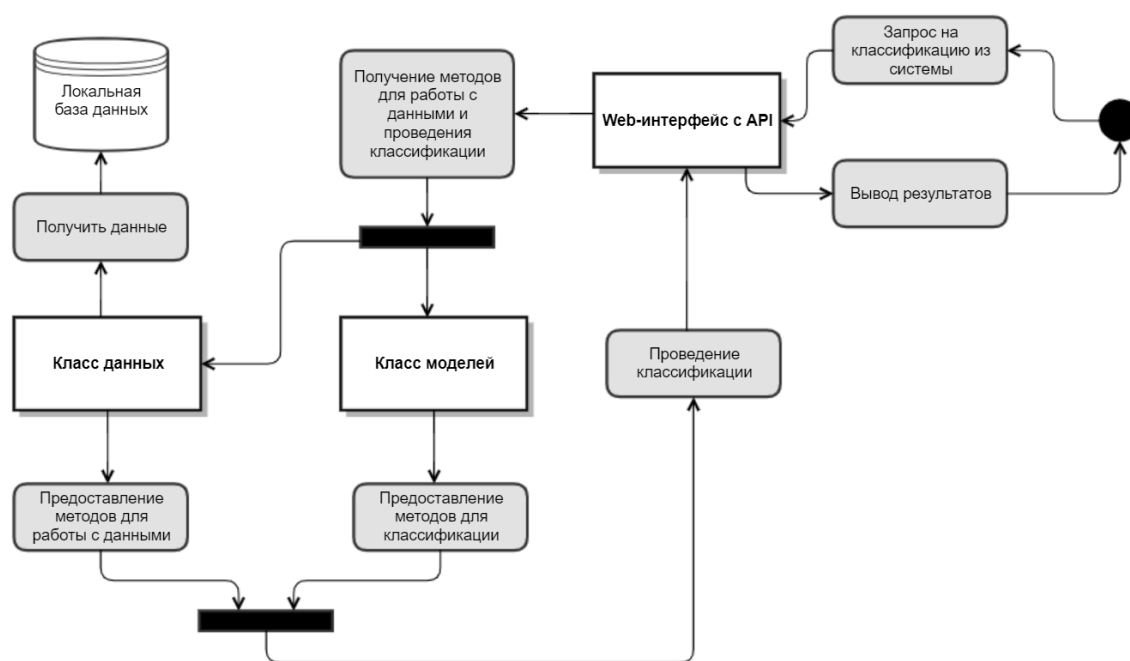


Рис. 2. Логическая схема работы интеллектуального классификатора сообщений

Программная реализация логики и структуры работы интеллектуального классификатора также велась на языке программирования Python 3.6.5, а для разработки и запуска веб-приложения использовался высокоуровневый фреймворк Django.

Обучение алгоритма автоматической классификации производилось на размеченной выборке данных размером 1.5 млн сообщений с количеством категорий проблем равным 262. Для обучения использовались модели на основе нейронных сетей с различным типом архитектур, а также классические методы. По каждому методу была получена метрика F-мера, являющаяся средним гармоническим между точностью (Precision) и полнотой (Recall) и показывающая некоторую степень результативности

(эффективности) определения категории проблемы, выраженную в процентах. Итоговые результаты обучения представлены в таб. 1.

Таблица 1

Результаты обучения с использованием различных методов

Метод машинного обучения	Precision (%)	Recall (%)	F-мера (%)
Сеть прямого распространения (FFNN)	82.0	81.7	81.893
Сверточная нейронная сеть (CNN)	82.3	81.6	82.990
Рекуррентная нейронная сеть (RNN) с LSTM-блоком	80.2	81.4	80.950
Наивный байесовский классификатор	65.0	67.0	65.176
Дерево решений	67.9	65.0	67.643
Метод опорных векторов (SVM)	71.0	70.1	71.008

Наилучшим методом машинного обучения стала сверточная нейронная сеть (CNN), показавшая почти 83% точности определения категории проблемы на основе текста сообщения. В ходе дальнейшей эксплуатации классификатора был использован данный метод.

Для проверки релевантности работы выбранного метода проведено тестирование на основе новых данных, не участвовавших в обучении и валидации алгоритма. Методами интеллектуального классификатора была сформирована выборка из базы данных, содержащая сообщения граждан за двухнедельный период, с 3 по 16 июня 2019 года. Далее был запущен процесс автоматической классификации сообщений с сохранением в сводной таблице определенных категорий проблем. Эти категории были проверены со значениями категорий, представленными службами модерации. На основе анализа результатов проделанной операции выяснилось, что интеллектуальный классификатор показал 95%-ую точность определения категории проблем в сообщениях за указанный период. Таким образом, применение интеллектуального классификатора помогло снизить процент отклонения сообщений по причине неверно выбранной категории с 20–25% до 5%.

Разработка завершена в 2019 г., на 2020 г. запланировано внедрение интеллектуального классификатора в программную инфраструктуру портала «Наш Петербург».

В дальнейших исследованиях планируется измерение и сравнение времени, которое служба модерации тратит на проверку одного сообщения, до и после внедрения разработанного интеллектуального классификатора, а также выявить степень повышения удобства при работе с порталом для граждан и определить уровень влияния решения на продуктивность работы портала.

В перспективе интеллектуальный классификатор, разработанный для портала «Наш Санкт-Петербург», можно использовать в других сферах деятельности, в которых возможно реализовать подходы «обучение с учителем» (Supervised learning) или «полуавтоматическое обучение с учителем» (Semi-supervised learning), например, в СМИ, медицине, градостроительстве, картографии, банковском деле, юриспруденции и т.д. Классификатор настроен на работу преимущественно с текстовой информацией и уже обладает специальными методами по эффективной обработке и преобразованию русскоязычного текста, при этом классификатор способен принимать любые типы данных (изображение, аудиофайл, видео, временной ряд и др.), заранее

преобразованные в числовую последовательность, и выдать необходимый результат классификации.

Важным направлением развития данной темы является возможность применения разработки в проектах развития информационных систем электронного взаимодействия граждан с органами власти, обеспечивающими обратную связь в экосистеме «Умного города». В частности, одной из ключевой задач, обеспечивающей семантическую совместимость различных отраслевых компонентов «Умного города» является необходимость создания и сопровождения классификатора объектов городского хозяйства, описывающего эти объекты во всем многообразии вариантов их использования. Создание инструментария, позволяющего снизить трудозатраты при классификации множества объектов, представляется весьма перспективной задачей.

Литература

1. Чугунов А.В., Рыбальченко П.А. Развитие системы электронного взаимодействия граждан с властями в Санкт-Петербурге: опыт портала «Наш Петербург»: 2014-2018 гг. // Информационные ресурсы России. 2018. № 6. С. 27–34.
2. Барсегян А.А. Анализ данных и процессов: учеб. пособие / Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И.. 3-е изд., перераб. и доп. СПб: БХВ-Петербург. 2009. 512 с.
3. Aggarwal C.C. Data Classification: Algorithms and Applications. Text Classification. Chapman & Hall/CRC, 2014. 705 p.
4. Colas F., Brazdil P. Comparison of SVM and Some Older Classification Algorithms in Text Classification Tasks // IFIP AI 2006: Artificial Intelligence in Theory and Practice / M. Bramer (eds.). 2006. Vol. 217. P. 169–178. DOI: 10.1007/978-0-387-34747-9_18.
5. Prasanna P.L., Rao D.R. Text classification using artificial neural networks // International Journal of Engineering & Technology. 2018. Vol. 7. No. 1.1. P. 603–606. DOI: 10.14419/ijet.v7i1.1.10785.



Богорадникова Дарья Андреевна

Год рождения: 1996

Университет ИТМО,

факультет информационных технологий

и программирования,

студент группы № М41212,

направление подготовки: 09.04.02. – Речевые информационные системы,

e-mail: fanpire0409@yandex.ru



Махныткина Олеся Владимировна

Год рождения: 1982

Университет ИТМО,

факультет информационных технологий

и программирования,

к.т.н., доцент,

e-mail: makhnytchina@ifmo.ru

УДК 004.9

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ИЗВЛЕЧЕНИЯ
ПРИЗНАКОВ ИЗ ТЕКСТА**

Д.А. Богорадникова

Научный руководитель – к.т.н., доцент О.В. Махныткина

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР №619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Аннотация

В работе был проведен сравнительный анализ методов извлечения признаков из текста. В частности, были рассмотрены различные вариации метода Bag of Words, такие как частотный, бинарный мешки слов и мера TF-IDF, и Word2Vec. Для каждого набора признаков были построены модели классификации токсичности текстовых сообщений, что позволило провести анализ влияния методов извлечения признаков на качество классификации.

Ключевые слова.

Классификация текста, методы извлечения признаков, bag of words, word2vec, TF-IDF, машинное обучение.

Извлечение признаков из текста является одним из важнейших этапов при машинной классификации текстов. В связи со стремительным увеличением объемов текстовой информации, которую в настоящее время обрабатывает человек посредством сети Интернет, классификация текстовых сообщений может помочь пользователям избежать так называемых токсичных сообщений, которые несут в себе агрессию, угрозу, оскорбления и т.п. В частности, Google и Jigsaw второй год подряд проводят

конкурс, в рамках которого участникам необходимо создать алгоритм классификации токсичных сообщений методами машинного обучения [1].

В работе осуществлено сравнение таких методов извлечения признаков, как частотный и бинарный мешки слов, мешок слов с мерой TF-IDF и Word2Vec. В качестве набора данных был выбран датасет, предоставленный на платформе Kaggle в рамках соревнования «Jigsaw Unintended Bias in Toxicity Classification». Данный датасет включает 1804874 аннотированных комментария пользователей, собранных со страниц обсуждения Википедии на английском языке. Из них 144334 комментария относятся к классу «токсичные», а оставшиеся 1660540 – к классу «нетоксичные». Для вычислительных экспериментов была осуществлена выборка 10% сообщений из датасета, соотношение классов в полученной выборке составило 1:1.

На первом этапе была проведена предварительная обработка текста, которая заключалась в следующем [2]:

- 1) приведение всех символов к нижнему регистру;
- 2) удаление знаков препинания и стоп-слов, к которым относят служебные слова (союзы, местоимения, частицы и т. д.), слова общей и оценочной лексики, заведомо не являющиеся терминами в рассматриваемой области;
- 3) токенизация – разбиение длинных строк на отдельные слова-токены;
- 4) лемматизация – приведение к смысловой канонической форме – лемме.

Извлечение признаков из предобработанного текста осуществлялось с использованием Bag of Words (мешка слов). Для всех документов, входящих в корпус, строится так называемая матрица термин-документ. Каждая ее строка определяет отдельный термин (слово), а каждый столбец соответствует некоторому документу. Таким образом, на пересечении строк и столбцов будут стоять различные меры, определяющие, входит ли слово в документ, и если входит, то как часто. В данной работе будут рассмотрены три меры: частотная, бинарная и TF-IDF.

Бинарная мера игнорирует, сколько раз слово встречается в документе. В данном случае, на пересечении в матрице будут стоять 0 и 1, где 0 – слово в предложении не встречается, 1 – встречается. Частотная мера включает в себя информацию о количестве вхождений слова в документ и работает по принципу счетчика. Несмотря на свою простоту, данные меры имеют такие недостатки, как игнорирование частотности слова в первом случае и получение чрезмерного влияния слов, не несущих большой информативности, во втором. Для устранения этих недостатков используется мера TF-IDF.

TF-IDF – статистическая мера, используемая для оценки важности слова в контексте документа, являющегося частью коллекции документов или корпуса. Из каждого документа выбирается по одному терму, для которого вычисляется вес на основании частоты его встречаемости как в данном документе, так и во всей анализируемой выборке (т.е. его распространенность). Данная мера позволяет понизить вес слишком часто встречающихся терминов и повысить вес тех терминов, которые встречаются реже [3].

Последним был рассмотрен Word2vec, который ориентирован на статистическую обработку больших массивов текстовой информации. Он собирает статистику о совместном появлении слов в высказываниях, после чего, с использованием нейронных сетей, решает задачу снижения размерности и выдает на выходе компактные векторные представления слов, в максимальной степени отражающие отношения этих слов в обрабатываемых текстах [4]. Существует два алгоритма реализации метода word2vec: CBOW (Continuous Bag of Words) и Skip-gram. CBOW, на основании существующего контекста, определяет слово, которое с наибольшей вероятностью должно встретиться в данном контексте. Skip-gram, в свою

очередь, определяет контекст по исходному слову. Наглядная демонстрация данных алгоритмов представлена на рисунке.

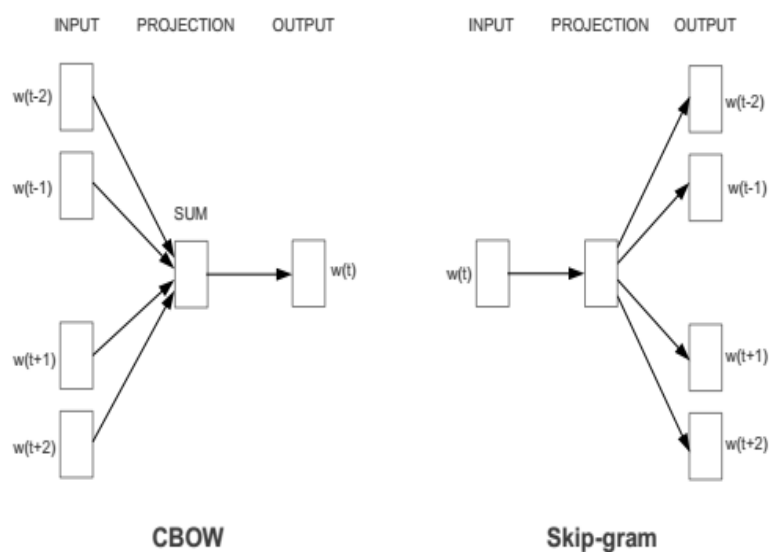


Рисунок. Алгоритмы обучения word2vec

В настоящее время насчитывается большое количество словарей, которые можно использовать при работе с word2vec. В частности, можно использовать предобученный словарь, и словарь, обученный на требуемом наборе данных. В данной работе рассматриваются признаки, полученные на основе обучения word2vec на собственном датасете и полученные из предобученной модели Google news vectors negative300, который был предварительно обучен на статьях из Google.

В качестве классификаторов были выбраны следующие методы:

1) линейный метод опорных векторов (SVM – Support Vector Machine), основная идея которого заключается в построении гиперплоскости, разделяющей объекты выборки наиболее оптимальным способом;

2) дерево решений (Decision Tree Classifier), которое представляет собой иерархическую структуру из узлов и листьев. Узлы содержат решающие правила и проводят проверку соответствия данных этому правилу, после чего данные делятся на подмножества. После процедура повторяется, пока не будет достигнуто условие остановки алгоритма. Последний узел, в котором произошла остановка, становится листом и определяет класс данных [5];

3) случайный лес (Random Forest Classifier) – модель состоит из множества деревьев решений и характеризуется случайной выборкой образцов из набора данных при построении деревьев и наборов параметров при разделении узлов;

4) логистическая регрессия (Logistic regression) – метод линейной классификации, который определяет вероятность принадлежности исходного значения к определенному классу.

5) наивный байесовский классификатор (Naive Bayes Classifier) – простой вероятностный классификатор, основанный на теореме Байеса со строгими (наивными) предположениями о независимости.

Для оценки качества построенных моделей были использованы такие метрики как precision (точность), recall (полнота) и f-мера. Для оценки использованы следующие параметры: TP – истинно-положительное решение, TN – истинно-отрицательное решение, FP – ложноположительное решение, FN – ложноотрицательное решение.

Precision можно интерпретировать как долю объектов, названных классификатором положительными и при этом действительно являющимися положительными:

$$Precision = \frac{TP}{TP + FP}$$

Recall показывает, какую долю объектов положительного класса из всех объектов положительного класса нашел алгоритм:

$$Recall = \frac{TP}{TP + FN}$$

F-мера представляет собой гармоническое среднее между точностью и полнотой. Она стремится к нулю, если точность или полнота стремится к нулю:

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Результаты, полученные в ходе построения моделей классификации текстовых сообщений, обработанных с помощью изученных методов извлечения признаков, представлены в таблице.

Таблица

Оценки методов извлечения признаков

Метод классификации	Метод извлечение признаков	Precision	Recall	F1-score
Linear SVM	W2V(Google)	0.83	0.83	0.83
	W2V	0.53	0.52	0.52
	BoW(частотный)	0.88	0.88	0.88
	BoW(бинарный)	0.88	0.88	0.88
	BoW(TF-IDF)	0.9	0.9	0.9
DTC	W2V(Google)	0.68	0.68	0.68
	W2V	0.53	0.52	0.52
	BoW(частотный)	0.83	0.83	0.83
	BoW(бинарный)	0.83	0.83	0.83
	BoW(TF-IDF)	0.83	0.83	0.83
RFC	W2V(Google)	0.75	0.75	0.75
	W2V	0.53	0.52	0.52
	BoW(частотный)	0.81	0.81	0.81
	BoW(бинарный)	0.86	0.86	0.86
	BoW(TF-IDF)	0.87	0.87	0.87

продолжение таблицы

Метод классификации	Метод извлечение признаков	Precision	Recall	F1-score
LR	W2V(Google)	0.83	0.83	0.83
	W2V	0.53	0.52	0.52
	BoW(частотный)	0.89	0.89	0.89
	BoW(бинарный)	0.9	0.9	0.9
	BoW(TF-IDF)	0.89	0.89	0.89
Bernoulli NB	W2V(Google)	0.74	0.74	0.74
	W2V	0.53	0.52	0.52
	BoW(частотный)	0.76	0.76	0.76
	BoW(бинарный)	0.78	0.78	0.78
	BoW(TF-IDF)	0.78	0.77	0.77
Gaussian NB	W2V(Google)	0.75	0.75	0.75
	W2V	0.56	0.51	0.52
Multinomial NB	BoW(частотный)	0.84	0.83	0.83
	BoW(бинарный)	0.84	0.84	0.84
	BoW(TF-IDF)	0.85	0.85	0.85

Для меры TF-IDF дополнительно было рассмотрено влияние переоценки IDF и ее сглаживания, применение сублинейного TF-масштабирования. Так, при изменении хотя бы одного из этих параметров произошло улучшение всех оценок алгоритма Random Forest Classifier и улучшение recall и f1-меры для Bernoulli Naive Bayes. Данные оценки также улучшились при отсутствии переоценки IDF для Multinomial Naive Bayes.

Таким образом, в ходе всех испытаний наилучший результат был получен при использовании двух алгоритмов. Первый сочетает мешок слов с мерой TF-IDF и SVM-классификатор, второй – бинарный мешок слов с линейным SVM. Точность классификации в данных случаях составила 90%. В перспективе планируется рассмотреть другие методы извлечения признаков из текста, такие как FastText, Glove, Bert, GPT, GPT-2, ELECTRA и ELMO.

Литература

1. Jigsaw Unintended Bias in Toxicity Classification. Detect toxicity across a diverse range of conversations [Электронный ресурс]. Режим доступа: <https://www.kaggle.com/c/jigsaw-unintended-bias-in-toxicity-classification> (дата обращения: 05.02.2020).
2. Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. Автоматическая обработка текстов на естественном языке и анализ данных. – М.: НИУ ВШЭ, 2017. 269 с.

3. Оськина К.А. Оптимизация метода классификации текстов, основанного на TF-IDF, за счет введения дополнительных коэффициентов [Электронный ресурс]. Режим доступа: <https://cyberleninka.ru/article/n/optimizatsiya-metoda-klassifikatsii-tekstov-osnovannogo-na-tf-idf-za-schet-vvedeniya-dopolnitelnyh-koeffitsientov> (Дата обращения: 20.02.2020).
4. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space [Электронный ресурс]. Режим доступа: <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf> (Дата обращения: 21.02.2020).
5. Чудова О.В. Возможность использования алгоритма деревьев решений c4.5 для оценки качества образования // Социологические методы в современной исследовательской практике: Сборник статей, посвященный памяти первого декана факультета социологии НИУ ВШЭ А.О. Крыштановского [Электронный ресурс] / отв. ред. и вступит. ст. О.А. Оберемко; НИУ ВШЭ; РОС; ИС РАН. – М.: НИУ ВШЭ. 2011. С. 60–63.



Васильева Виктория Олеговна

Год рождения: 1996

Университет ИТМО,

факультет прикладной оптики,

студент группы № В42801,

направление подготовки: 16.04.01 – Техническая физика,

e-mail: r-kassi@mail.ru



Быстрянцева Наталья Владимировна

Год рождения: 1981

Университет ИТМО,

факультет прикладной оптики,

канд. архитектуры, доцент,

e-mail: sv.s.lighting@gmail.com

УДК 628.931

**ИССЛЕДОВАНИЕ ПОВЕДЕНЧЕСКОГО ОЖИДАНИЯ
ПОЛЬЗОВАТЕЛЕЙ ОТ СРЕДЫ НА ОСНОВЕ АНАЛИЗА
ПЕШЕХОДНЫХ УЛИЦ В КРОНШТАДТЕ**

В.О. Васильева

Научный руководитель – канд. арх., доцент Н.В. Быстрянцева

Работа выполнена в рамках дисциплины «Разработка и прототипирование световых решений в архитектуре и городской среде».

Аннотация

В работе рассмотрены существующие российские и зарубежные регламенты наружного искусственного освещения, приведено их сравнение. Для комплексного изучения территорий Кронштадта проведено функциональное зонирование города, на базе которого сделан градостроительный анализ территории, исследование пространственной и функциональной организации, пользовательского опыта и коммуникации жителей. На основе анализа регламентов, научных статей и исследования территорий города сформулированы и апробированы гипотезы о пользовательском опыте для разных типов пешеходных улиц в Кронштадте. Исследование для проверки поведенческого ожидания пользователей от улиц проводилось при помощи социологического интернет-опроса. После обработки полученных данных были представлены результаты социологического опроса и даны рекомендации для дальнейшего применения полученных результатов.

Ключевые слова

Пешеходная улица, качество световой среды, освещенность, пользовательский опыт, регламент освещения.

В настоящее время предусмотренные стандартами нормы наружного искусственного освещения являются сугубо утилитарными, и как показывает практика, не отражают специфику климата городов, их пространственную специфику и культуру

восприятия световой среды. В результате такого стандартного подхода возникает низкое качество световой среды, что в свою очередь может привести к ряду факторов негативного влияния на человека, например, к зрительному дискомфорту, к сложностям ориентации в пространстве, к нежеланию человека выходить на улицу в темное время суток и т.д. [1].

Многосторонний учет факторов при проектировании освещения и улучшения качества световой среды пешеходных улиц на сегодняшний день является одной из основных задач для многих городов, и Кронштадт не является исключением. Поиск и формулирование рекомендаций для комплексного подхода к формированию качественной и функциональной световой среды пешеходных улиц является не только оправданным, но и необходимым для раскрытия потенциала Кронштадта и уменьшения ряда существующих проблем как для местных жителей, так и для туристов. Переосмысленный подход способствует повышению интереса к городу, привлекательности его вечерней городской среды и общественных пространств, повышению функции рекреации и общения, особенно, в вечернее время после работы и многому другому.

Основным регламентом, устанавливающим правила на наружное освещение городов, поселков и сельских населенных пунктов, автотранспортных тоннелей и пр. на территории России, является СП 52.13330.2016 (Актуализированная редакция СНиП 23-05-95) [2]. В разделе 7.5.4 «Освещение пешеходных пространств» представлена классификация и нормируемые показатели для пешеходных пространств. Класс объекта по освещению определяется по общему описанию, приведенному в графе Таблицы 26 «Наименование объекта». Такой подход не всегда позволяет точно определить класс объекта, потому что не учитываются параметры окружающей среды. Также стоит отметить, что в регламенте указаны пределы минимальных допустимых значений параметров освещения, а максимальные допустимые не упоминаются.

Подход к наружному искусственному освещению за рубежом сильно отличается от подхода к освещению в России. Например, в своде правил CIE Recommendation для пешеходных зон основу рекомендации составляют шесть классов освещения: от P1 до P6 («P» обозначает пешехода). Для определения классов объекта по освещению вводятся весовые коэффициенты, позволяющие определять класс с учетом окружающих факторов среды. В этом регламенте нормируются такие параметры как:

- 1) средняя горизонтальная освещенность;
- 2) минимальная горизонтальная освещенность;
- 3) горизонтальная средняя максимальная освещенность;
- 4) минимальная вертикальная освещенность;
- 5) полуцилиндрическая освещенность [3].

Рекомендации в European Standard для пешеходных зон схожи со сводом правил CIE Recommendation. Основу рекомендации составляют семь классов освещения: от P1 до P7. Как и в предыдущем стандарте для определения этих классов используются весовые коэффициенты. Рекомендации North American Standard для пешеходных зон гласят, что как горизонтальная, так и вертикальная освещенность должны соответствовать минимальным требованиям. Во всех этих рекомендациях также упоминается адаптивное освещение, которое может изменяться в течение вечера и ночи. Это может позволить сделать периоды с уменьшенным уровнем освещения.

В книге «Road lighting fundamentals technology and application» [3] приведен ряд исследований, доказывающих что каждый светотехнический параметр отвечает за определенные функции. Например, горизонтальная освещенность $E_{ср}$ влияет на *обнаружение пешеходом препятствий*. В лабораторном исследовании (Fotios and Cheal 2013) было выяснено, что для обнаружения препятствия высотой 25 мм, расположенного на расстоянии 6 м впереди от наблюдателя, требуется минимальная

горизонтальная освещенность 1,8 люкс (вероятность обнаружения в этом случае 95%, при условии, что пешеход приспособлен к этому низкому уровню освещения). Ниже этого уровня резко снижается вероятность обнаружения.

Что касается вопроса безопасности, то освещение в населенных пунктах может иметь двойную функцию: *обеспечение ощущения безопасности* (будет стимулировать людей выходить на улицу) и *обеспечение реальной безопасности* (поможет обеспечить безопасность). Яркость улицы и ее окрестностей (пространственная яркость), если она достаточно высока, позволяет сразу увидеть всю территорию, что способствует *ощущению безопасности*. В различных исследованиях (Banister and Bowling 2004; Blöbaum and Henecke 2005; Johansson et al. 2011) собрана информация о том, как пространственная яркость влияет на *чувство безопасности*. В них проводилось анкетирование или интервью с пожилыми людьми, студентами или молодыми женщинами. Полученные данные показали, что многие пешеходы испытывают страх после наступления темноты и более уязвимы для угроз или даже нападений, чем в дневное время, особенно в небезопасных районах. Для некоторых людей это является причиной не выходить на улицу после наступления темноты.

За обеспечение *реальной безопасности* отвечает такой параметр, как полуцилиндрическая освещенность $E_{\text{пл}}$. Она способствует распознаванию лиц; важно обеспечивать не только E_v на уровне лица в качестве основного параметра, но и $E_{\text{пл}}$, ведь человеческое лицо – это не просто плоская вертикальная плоскость: свет, падающий по сторонам лица, способствует его видимости.

Для *визуальной ориентации человека в пространстве*, необходимо освещение фасадов ($E_{\text{фасада}}$) и окружающей среды. Освещение должно освещать не только горизонтальные поверхности, но и вертикальные поверхности фасадов (уличные знаки и номера домов должны быть разборчивыми). Особое внимание здесь следует уделять тому, чтобы избегать «переосвещения» фасадов, чтобы визуально не мешать жильцам в их домах. Исследование (Fotios and Cheal 2013) показало, что идущие по экспериментальному маршруту люди, смотрели на дома, дороги и деревья 50% – 60% времени. Другое исследование (Van Bommel and Van Dijk 1985) проводилось в большом зале, где фасад одного дома визуально оценивался испытуемыми с обочины дороги. Исследование показало, что освещенность фасадов, составляющая всего 1,5 лк, может служить для визуальной ориентации в условиях уличного освещения со средним уровнем освещения 5 лк по горизонтали.

Рассмотренные исследования подтверждают, что перечисленные выше и регламентируемые в зарубежных стандартах светотехнические параметры являются важными для пешеходов, и их стоит учитывать при разработке регламентов, проектировании и улучшении существующей световой среды пешеходных улиц.

В ходе анализа территорий Кронштадта всего было выделено 64 улицы. Более детально была изучена старая (центральная) часть города (рис. 1). Ее выбор для исследования обусловлен тем, что именно там расположена основная часть жилой, производственной, муниципальной и торговой застройки. На этой территории расположено 52 доступные для пешеходов улицы (выделены на рис. 1 (б) синим цветом).

Согласно мнению Н.И. Щепеткова, большинство пешеходных улиц имеют свою преобладающую функцию. Всего можно выделить шесть основных типов пешеходных улиц:

- 1) преимущественно транзитные;
- 2) преимущественно рекреационные;
- 3) транзитно-деятельностные (торговые улицы);
- 4) транзит-общение/рекреация;
- 5) транзит/общение/рекреация/деятельность;

б) улица со смешанным движением) [1].



Рис. 1. Старая (центральная) часть города (а); исследуемая граница старой (центральной) части города (б)

На основе анализа и наблюдений, в Кронштадте можно выделить следующие типы и подтипы пешеходных улиц в зависимости от доминирующих видов деятельности:

1) преимущественно транзитные – служат для перемещения большого потока людей за короткий промежуток времени. С точки зрения световой среды для этих улиц при проектировании важным является создание условий зрительного комфорта и способности ориентации в пространстве;

2) преимущественно рекреационные – служат для реализации функции отдыха и комфортного пребывания человека в среде. Здесь при проектировании стоит создавать минимальное влияние искусственного освещения, тем самым делая комфортную и естественную среду, способствующую человеку создавать «диалог» с природой;

3) транзит/общение/рекреация/деятельность – этот тип пешеходных улиц характерен для знаковых и привлекательных мест в городе. Такие улицы могут нести большую значимость или историю для города, поэтому подход к проектированию световой среды должен быть индивидуален, учитывать специфику среды.

Для комплексного изучения территорий было проведено функциональное зонирование города, на базе которого был сделан градостроительный анализ территории, исследование пространственной и функциональной организации, пользовательского опыта и коммуникации жителей. Все исследования проводились в осенне-зимний период (2019 г.). В результате такого подхода было выделено три конкретные (типичные) улицы для каждого из приведенного выше типа. Дальнейшее исследование строилось на их детальном анализе:

1) улица Всеволода Вишневого (преимущественно транзитные). Выбор обусловлен тем, что она является связующим звеном между двумя улицами городского значения, расположена среди жилой застройки и имеет небольшую протяженность;

2) центральная аллея Петровского парка (преимущественно рекреационные). Выбор обусловлен наличием зеленой зоны, возможностью выхода на пирс, а также тем, что большинство аллей в Кронштадте являются таковыми за счет близлежащих скверов и парков;

3) советская улица (транзит/общение/рекреация/деятельность). Улица имеет большую протяженность, проходит через центр города и несет на себе большую нагрузку.

С целью проверки поведенческого ожидания, для выбранных улиц были сформулированы гипотезы о пользовательском опыте пешеходов (по 3 штуки для

каждой улицы). Для их апробации был составлен социологический интернет-опрос, который проводился в социальной сети Вконтакте. Цель опроса – понять, действительно ли важны рассмотренные светотехнические параметры для пешехода, идущего по улице в условиях исследуемой среды. На рис. 2 представлены сформулированные в исследовании гипотезы и их упрощенные версии для интернет-опроса.

Постановка гипотезы в исследовании	Постановка утверждения в исследовании
Тип улиц: Преимущественно транзитные (Ул. Всеволода Вишневского)	
1. Человеку не достаточно только тех минимальных значений параметров освещения, которые предусмотрены в СП.52.13330.2016, при прохождении пространства для видимости объектов	1. Мне важно просто пройти, поэтому нужно видеть тротуар, по которому я иду и проезжую часть
2. Человеку важна видимость не только объектов в центральном поле зрения, но и периферийных объектов для обеспечения чувства безопасности и создания реальной безопасности	2. Мне важно видеть все что находится по сторонам от тротуара, так я буду чувствовать себя безопаснее
3. Человеку не нужен уникальный дизайн и разнообразие видовых кадров пространства, если это не обусловлено контекстом города и длительностью движения	3. Мне нравится смотреть на жилые дома, мимо которых я иду
Тип улиц: Преимущественно рекреационные (Центральная аллея Петровского парка)	
1. Человеку важна видимость объектов в центральном поле зрения, периферийных объектов, а также предметно-пространственной среды для общения с природой через считываемые им образы	1. Мне важно хорошо видеть всю окружающую меня среду, чтобы отдыхать и прогуливаться по этой аллее
2. Увеличение яркости не является приоритетом для этого типа пространства	2. Мне важно чтобы освещение не доминировало, а лишь дополняло среду
3. Человеку важно чувствовать себя безопасно для комфортного длительного нахождения в пространстве	3. Мне важно чувствовать себя безопасно, чтобы отдыхать и спокойно прогуливаться по этой аллее
Тип улиц: Транзит/общение/рекреация/деятельность (Советская ул.)	
1. Человеку важна видимость объектов в центральном поле зрения, периферийных объектов, других людей, предметно-пространственной среды для реализации функции общения с другими людьми	1. Мне важно хорошо видеть абсолютно всю окружающую меня среду, чтобы комфортно проводить время в пространстве
2. Человеку нужен уникальный дизайн, разнообразие видовых кадров пространства с учетом контекста города	2. Мне важно чтобы освещение на улице было красивым и необычным
3. Человеку нужен световой баланс между всеми видами освещения в пространстве для комфортного нахождения в нем кратковременно или длительно	3. Мне важно чтобы освещение было умеренным и не вызывало дискомфорта

Рис. 2. Формулировка гипотез в исследовании и интернет-опросе

Респондентам было предложено выбрать одно, наиболее близкое им утверждение из каждого блока, после чего им открывался дополнительный блок вопросов. В нем нужно было оценить степень своего согласия или несогласия по шкале от 1 (полностью согласен) до 5 (полностью не согласен) с приведенными утверждениями. Предполагалось, что за каждое утверждение отвечают определенные светотехнические параметры. Если гипотеза подтверждалась, то это означало, что эти параметры важны для пешеходов и на них рекомендуется обращать внимание при проектировании или разработке стандартов. В настоящей работе представлены только данные по гипотезам, которые набрали наибольшее количество ответов по результату опроса. На рис. 3 представлены общие сведения о респондентах. Всего в опросе приняло участие 46 человек.

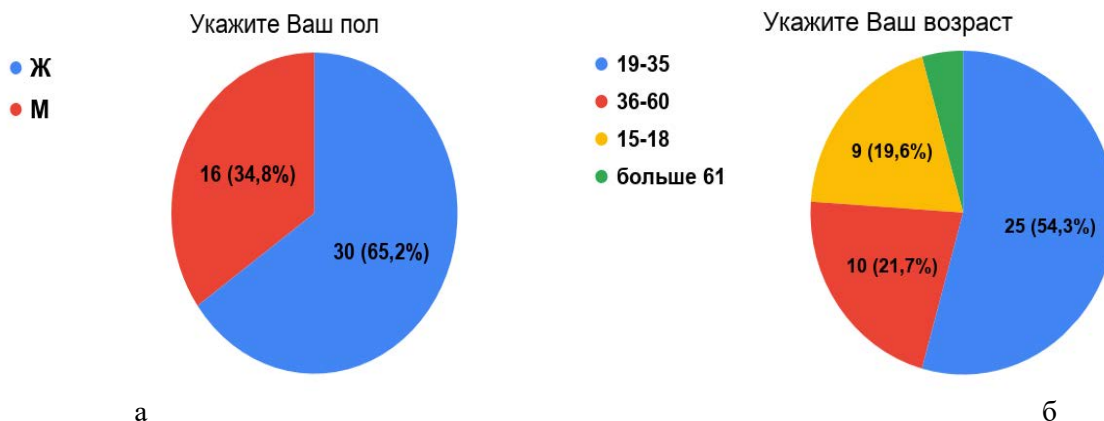


Рис. 3. Общие сведения о респондентах: пол (а); возраст (б)

Среди гипотез при исследовании улицы преимущественно транзитного типа (улицы Вишневого), наибольшее число ответов набрала гипотеза под номером 1 (рис. 2). На рис. 4 приведены фотографии и утверждения, представленные в опросе, а также полученные в результате исследования гистограммы. Исследование показало, что пешеходу для прохождения данного типа пространства важны параметры, выделенные на рис. 4 синей рамкой.



Рис. 4. Результат исследования улицы Всеволода Вишневого для гипотезы №1

Аналогичное исследование было проведено для двух других типов улиц. Результаты представлены на рис. 5 и 6. Для улиц преимущественно рекреационного типа и транзит/общение/рекреация/деятельность, наибольшее число ответов набрали также гипотезы под номером 1 (рис. 2). Исследование показало, что все утверждения подтвердились, и на приведенные в нижней строчке параметры следует обращать внимание при разработке или проектировании среды т.к. они являются важными для пешехода.

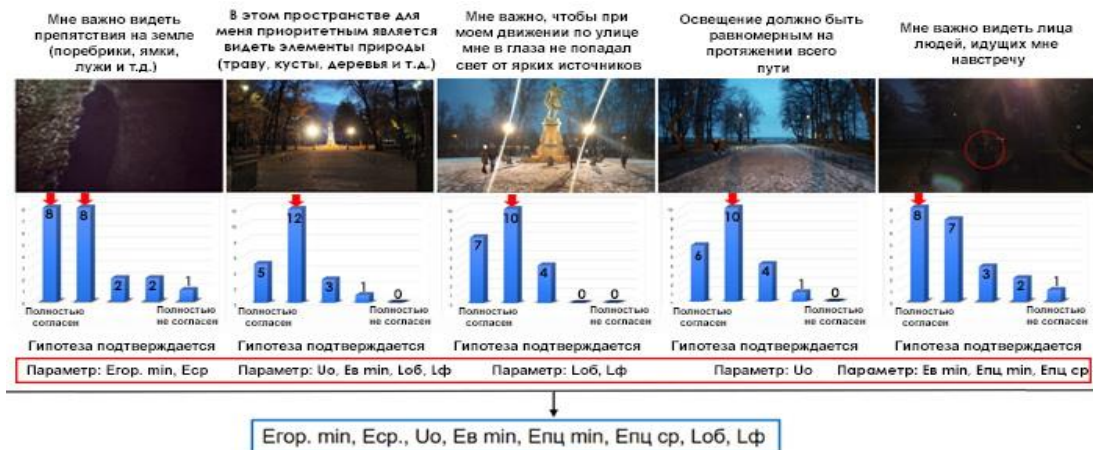


Рис. 5. Результат исследования центральной аллеи Петровского парка для гипотезы №1



Рис. 6. Результат исследования Советской улицы для гипотезы №1

Проведенный социологический опрос показал, что человеку для прохождения каждого рассмотренного типа пешеходных улиц недостаточно одних только норм, регламентируемых в СП 52.1330.2016. Пользователи ожидают от любого типа улиц более ориентированной на человека, комфортной и безопасной световой среды. При улучшении качества световой среды и разработке регламентов стоит учитывать на зарубежный опыт, который ориентирован на создание именно такой среды для пешехода. Полученные в результате социологического опроса данные можно использовать в дальнейшем для предложения новых рекомендаций, направленных на создание более «дружелюбной» и привлекательной для пользователя среды, а также повышения пользовательской активности в темное время суток в городе.

Литература

1. Матовников Г.С. Принципы формирования световой среды пешеходных улиц города (на примере Москвы): дис. на соиск. уч. степени канд. Архитектуры. Москва: Московский Архитектурный Институт. 2017. 226 с.
2. СП 52.13330.2016. Свод правил. Естественное и искусственное освещение. Актуализированная редакция СНиП 23-05-95. Введен 05.08.2017. М.: Стандартинформ, 2012. 135 с.
3. Wout van Bommel Road lighting fundamentals technology and application. Springer, 2015. 334 p.



Новоселов Сергей Александрович

Год рождения: 1985
Университет ИТМО,
кандидат технических наук,
с.н.с. корпоративной лаборатории технологий
человеко-машинного взаимодействия,
e-mail: novoselov@speechpro.com



Газизуллина Алиса Рустамовна

Год рождения: 1996
Университет ИТМО,
факультет информационных технологий,
и программирования,
студент группы № М4121,
направление подготовки: 09.06.01 – Информатика
и вычислительная техника,
e-mail: gazizullina2010@yandex.ru

УДК 004.934

**УЛУЧШЕНИЕ КАЧЕСТВА ТЕКСТОНЕЗАВИСИМОГО
РАСПОЗНАВАНИЯ ДИКТОРА**

А.Р. Газизуллина

Научный руководитель – к.т.н. С.А. Новоселов

Аннотация

Распознавание диктора по индивидуальным акустическим характеристикам его голоса является задачей применимой в сферах информационной безопасности, процессах связанных с высоким уровнем защищенности. Голос является уникальным для каждого человека, что делает его идеальным элементом для идентификации. Однако современные методы верификации по голосу на данный момент требуют улучшений точности. В данной работе мы подходим к решению задачи построения модели распознавания диктора по голосу со стороны перехода от алгоритмов с лидирующей точностью реализованных во фреймворке Kaldi к их реализации в Pytorch. Подавляющее количество научных работ основано на моделях построенных с помощью фреймворка Kaldi, который в свою очередь ограничен в возможностях для индивидуализации. Перенос моделей во фреймворки общего назначения без деградации точности является приоритетной задачей в современном мире индустриальной разработки. Данная работа посвящена описанию экспериментов направленных на перенос X-векторных архитектур оптимизированных под задачу распознавания диктора в микрофонном и телефонном каналах из Kaldi в Pytorch. В качестве тренировочных данных приняты данные из конкурса VoxCeleb (микрофонный канал), телефонной части базы NIST 2019 (телефонный канал), тестовые данные и тестовый протокол взят из валидационного набора базы VOiCES (микрофонный канал), телефонной части NIST 2019 (телефонный канал).

Ключевые слова

X-вектора, распознавание диктора, переход к фреймворку общего назначения, VoxCeleb.

Введение

Распознавание диктора по индивидуальным акустическим характеристикам его голоса является актуальной задачей рассматриваемой в сфере информационной безопасности, а также процессах, связанных с высоким уровнем защищенности. Голос является уникальным для каждого человека, что делает его идеальным элементом для идентификации. Однако качество современных методов верификации по голосу зависит от множества факторов, таких как длительность звукозаписи, особенности акустической обстановки, частоты дискретизации, качества звукозаписи, индивидуальных характеристик микрофона и записывающего устройства, языка, качества обучающей выборки, намерений говорящих выдать себя за других дикторов [1]. Перечисленные факторы формируют пространство гиперпараметров систем распознавания говорящего по голосу большой размерности, что оправдывает постоянную потребность в улучшения качества верификационных систем.

В настоящей работе рассматривается решение задачи построения модели распознавания диктора по голосу на основе X-векторной системы [2] в фреймворке общего назначения. Нами было принято взять за основу исследований архитектуру X-векторов так как она дает наиболее стабильную картину точности распознавания диктора на разных наборах данных. X-вектора появились в процессе эволюции статистических систем верификации, основанных на модели гауссовых смесей таких как i-вектора. X-векторная модель это метод представления голосового сегмента аудиозаписи в сжатой и в то же время богатой индивидуальными для говорящего признаками форме. Простейшая ее конфигурация представляет собой 3 TDNN (нейронная сеть с временной задержкой) слоя с 512 фильтрами и размерами контекста 5, 5, 7, двумя слоями сверток 1x1, статистическим пуллингом, двумя полносвязными слоями и Softmax слой. Изначально X-вектора были реализованы в Kaldi [3], но данный фреймворк неудобен для исследований, ограничен в возможностях кастомизации. Однако, перенос системы в фреймворки общего назначения является нетривиальной задачей. Это связано с тем что методы, с помощью которых Kaldi достигает state-of-the-art результатов, при их воспроизведении для обучения систем в фреймворках общего назначения не дают прироста в качестве. К таким методам можно отнести натуральный градиент [4], метод ансамблирования моделей, метод формирования батчей. В данной работе мы исследуем различные конфигурации X-векторной модели и рассматриваем влияние различных факторов на деградацию качества при переносе системы из Kaldi в Pytorch.

Разработанная нами система позволяет уменьшить значение EER на 1,5% относительно эквивалентной системы в Kaldi по тестам на закрытой корпоративной базе в телефонном канале. Тесты на телефонной части базы NIST 2019 и на базе VOiCES [5] содержащей записи в микрофонном канале демонстрируют, что рассматриваемая в данной работе модель дает результаты сопоставимые с результатами полученными в Kaldi. В то же время мы в отличие от систем в Kaldi используем одну систему для получения результатов, не прибегая к ансамблированию, так же мы оперируем методом оптимизации первого порядка, работая в режиме параллельного вычисления на данных внутри одного батча (Data Parallel).

Подготовка данных

Тренировочные данные: для обучения архитектуры модели нейронных сетей оптимизированной под задачу распознавания диктора в микрофонном канале

применяются речевые сегменты преобразованные в 80-мерные банки фильтров (рассчитанные в полосах мел частотной шкалы на фреймах 20мс со сдвигом 10мс). База обучения состоит из набора склеенных произнесений базы VoxCeleb1,2 [6] с количеством уникальных спикеров 7 140. Произнесения внутри одной сессии сливаются и делятся на неперекрывающиеся отрезки по 4 секунды. Для обучения оптимальной для медиа канала конфигурации X-векторной системы на телефонных данных была использована смешанная база, состоящая из телефонной части датасета NIST SRE 2004-2018 [7], медиа данных баз VoxCeleb 1 и VoxCeleb 2 и части закрытой базы телефонных звонков Telecom. Произнесения были преобразованы в 23-мерные MFCC признаки [8]. При обучении сетей, каждое произнесение разбиваются на сегменты одинаковой длительности 3сек. Количество уникальных дикторов в базе обучения для телефонного канала составило 23844.

Тестовые данные: При выполнении тестовых экспериментов использовался набор данных VOiCES dev 2019 и соответствующий протокол тестирования для оценки стандартных параметров качества верификации. Для оценки качества распознавания в телефонном канале было проведено тестирование на телефонная часть база NIST SRE 2019 eval, тестовой части базы Telecom, тестовой части базы IVR. Произнесения меньше 0,1 сек по длительности были отсечены.

Модель

В данной работе рассматриваются различные конфигурации X-векторной модели, применяемой в гибридных моделях распознавания речи в качестве моделей извлечения признаков из предобработанных и преобразованных в основном в MFCC аудио сегментов. Это глубокая сверточная нейронная сеть, с разными для каждого слоя контекстами по размерности времени входных карт признаков (кадрам). Первые пять слоев нейронной сети обрабатывают кадры речи, с маленьким контекстом по времени центрированным в текущем кадре, последующий слой статистического пулинга [2] агрегирует информацию по всем кадрам, так что следующие два полносвязных слоя оперируют по полным сегментам. Последний слой softmax, присваивает вероятности содержания речи каждого из спикеров в произнесении. В таб. 1, 2 и 3 приведены конфигурации X-векторных систем используемые в данной работе.

Таблица 1

Оригинальная архитектура X-векторной системы (X-vectors)

слои	контекст	кол-во фильтров	слои	контекст	кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-2,t+2}	512	5.кадровый (TDNN-lrelu)	{t}	3*512
2.кадровый (TDNN-lrelu)	{t-2,t+2}	512	статистический пулинг	[0,T)	6*512
3.кадровый (TDNN-lrelu)	{t-3,t+3}	512	6.сегментный 7.сегментный	{0} {0}	512 512
4.кадровый (TDNN-lrelu)	{t}	512	Softmax	{0}	кол-во классов

Таблица 2

Оптимизированная архитектура X-векторной системы (Optimized X-vectors)

слои	контекст	кол-во фильтров	слои	контекст	кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-3,t+3}	512	5.кадровый (TDNN-lrelu)	{t}	1500
2.кадровый (TDNN-lrelu)	{t-2,t+2}	512	статистический пуллинг	[0,T)	3000
3.кадровый (TDNN-lrelu)	{t-1,t+2}	1024	6.сегментный	{0}	512
4.кадровый (TDNN-lrelu)	{t-1, t+1}	1024	Softmax	{0}	кол-во классов

Таблица 3

Расширенная архитектура X-векторной системы (Extended X-vectors)

слои	контекст	кол-во фильтров	слои	контекст	кол-во фильтров
1.кадровый (TDNN-lrelu)	{t-3,t+3}	512	7.кадровый (TDNN-lrelu)	{t-1, t+1}	512
2.кадровый (1x1 conv)	{t}	512	8.кадровый (1x1 conv)	{t}	512
3.кадровый (TDNN-lrelu)	{t-2,t+2}	512	9.кадровый (TDNN-lrelu)	{t}	1500
4.кадровый (1x1 conv)	{t}	512	статистический пуллинг	[0,T)	3000
5.кадровый (TDNN-lrelu)	{t-1,t+2}	512	6.сегментный	{0}	512
6.кадровый (1x1 conv)	{t}	512	Softmax	{0}	кол-во классов

Скоринг

После тренировки сети на задачу классификации спикеров выход последнего полносвязного слоя перед слоем классификации подается в классификатор разделяющий подаваемые латентные представления в метрическом пространстве X-векторных эмбедингов. В качестве такого классификатора нами были выбраны метод основанный на Вероятностном Линейном дискриминантном анализе (PLDA) [9] и модель метрического обучения CSML [10]. Данная архитектура не является end-to-end обучаемой, так как латентные представления выучиваемые X-векторной сетью-экстрактором неразделимы в метрическом пространстве и измерять расстояния по косинусной метрике между представлениями двух дикторов напрямую без дополнительного применения PLDA.

Параметры обучения

Для обучения модели X-векторов представленной в данной работе в качестве функции ошибки была использована кросс энтропия, формула расчета ошибки представлена ниже:

$$\text{loss}(x, \text{class}) = - \log\left(\frac{e^{x[\text{class}]}}{\sum_j e^{x[j]}}\right) = -x[\text{class}] + \log\left(\sum_j e^{x[j]}\right) \quad (1)$$

где x – выход нейронной сети, представляющий собой матрицу предсказаний для каждого примера (под предсказанием имеется ввиду вектор размером $1 \times N$ - общее количество спикеров).

Модели обучаются итеративным методом в несколько эпох. Размер батча фиксирован для всех экспериментов на 128, данный параметр был выявлен оптимальным экспериментально, при увеличении размера батча появляется переобучение, меньший же размер батча вызывает шумы, что замедляет сходимость [11]. Из разделенных на сегменты по 4 с. аудио записи отбирается отрезок длины 3 с. в момент времени выбранный случайно.

Эксперименты с различными методами оптимизации показали что использование AdamW [12] ускоряет сходимость, однако итоговое качество распознавания немного хуже, в отличие от итогов обучения с SGD [13] который сходится медленнее. Для ускорения обучения нами было принято решение последовательно увеличивать размер батча каждую эпоху в два раза, с 128 до 1024 [14]. Первые эпохи подвержены переобучению, так как веса еще не успели стабилизироваться и подача большого количества примеров может привести к тому что модель запомнит все тренировочные примеры, использование относительно маленьких размеров батча на начальных этапах обучения действует по принципу регуляризации и предотвращает переобучение. Последовательное увеличение размера батча позволяет обучать нейронную сеть быстрее без переобучения. Также увеличение размера батча эквивалентно уменьшению шага обучения, следовательно применение агрессивных планировщиков итеративно понижающих шаг обучения каждую последующую эпоху чревато недетерминированным фактическим изменением шага обучения а также падением скорости обучения, поэтому нами было принято использовать планировщик Multistep Learning Rate понижающий шаг обучения в 0.6 раз каждую 5 эпоху.

Эксперименты

В таб. 4 и 5 приведены результаты экспериментов реализации систем на основе X-векторов. В таблице 4 приведено сравнение реализованных нами систем X-векторов, конфигурации которых приведены в таблицах 1, 2, 3, и системы распознавания речи в условиях дальнего микрофона разработанной Университета Технологий BRNO для конкурса VOiCES 2019 [15]. Наша расширенная X-векторная (Extended X-vectors) система распознавания в медиа канале по метрике EER превосходит приведенную систему исследователей из BRNO. В таблице 5 приведены результаты реализованной нами модели X-векторов в Pytorch оптимизированной под задачу распознавания диктора в телефонном канале. EER и minimum DCF нашей системы для телефонного канала не превосходят результаты baseline системы NIST SRE 2019 по данным метрикам, однако являются сопоставимыми и стоит заметить что наша система не достигла насыщения так как училась в 2 эпохи и является сингл системой.

Результаты тестирования модели для записей в микрофонном канале на данных из набора Voices dev set 2019

	cos EER/ minDCF	PLDA EER/ minDCF
X-vectors (weighted)	8.77/0.89	6.44/0.77
Optimized X- vectors	7.2/0.76	4.9/0.67
Extended X-vectors	6.4/0.72	3.98/0.62
BUT Kaldi X- vectors [15]	-	4.68/0.494

Результаты тестирования модели для записей в телефонном канале на данных из тестового набора телефонной части базы конкурса NIST SRE 2019 (NIST SRE 19 eval). NIST tele - телефонная часть базы данных NIST SRE 2019, MIXEDTELE - база содержащая данные из закрытой корпоративной базы данных телефонных разговоров и NIST tele

база для тренировки PLDA	cos EER/ minDCF	PLDA EER/ minDCF	CSML
NIST tele	16.78/0.89	12.82/0.92	7.62/0.80
MIXEDTELE		13.15/0.99	9.02/0.82
NIST baseline [17]	-	6.16/0.440	-

Заключение

Наша работа является второй после [16, 17] описывающей перенос систем из Kaldi в фреймворк общего назначения. Нам удалось построить реализацию алгоритма X-векторов в Pytorch воспроизводящую результаты работы алгоритма на Kaldi. Результаты замеров точности рассмотренных алгоритмов приведены в таблицах 4 и 5. Разработан алгоритм тренировки модели в условиях неопределенности меток и уровня извлекаемых признаков в каждом из слоев. Достигнутая точность позволяет использовать разработанный алгоритм в промышленных системах интеллектуального поиска аудио информации.

Литература

1. Vestman Ville & Kinnunen, Tomi & Hautamäki, Rosa & Sahidullah Md. (2019). Voice Mimicry Attacks Assisted by Automatic Speaker Verification.

2. Snyder David & Garcia-Romero, Daniel & Sell, Gregory & Povey, Daniel & Khudanpur, Sanjeev. (2018). X-Vectors: Robust DNN Embeddings for Speaker Recognition. 5329-5333. 10.1109/ICASSP.2018.8461375.
3. Povey, Daniel & Ghoshal, Arnab & Boulianne, Gilles & Burget, Lukáš & Glembek, Ondrej & Goel, Nagendra & Hannemann, Mirko & Motlíček, Petr & Qian, Yanmin & Schwarz, Petr & Silovský, Jan & Stemmer, Georg & Vesel, Karel. (2011). The Kaldi speech recognition toolkit. IEEE 2011 Workshop on Automatic Speech Recognition and Understanding.
4. Povey Daniel & Zhang, Xiaohui & Khudanpur, Sanjeev. (2014). Parallel training of DNNs with Natural Gradient and Parameter Averaging.
5. Richey Colleen & Barrios, Maria & Armstrong, Zeb & Bartels, Chris & Franco, Horacio & Graciarena, Martin & Lawson, Aaron & Nandwana, Mahesh & Stauffer, Allen & Hout, Julien & Gamble, Paul & Hetherly, Jeffrey & Stephenson, Cory & Ni, Karl. (2018). Voices Obscured in Complex Environmental Settings (VOICES) Corpus. 1566-1570. 10.21437/Interspeech.2018-1454.
6. Nagrani. Arsha & Chung, Joon Son & Zisserman, Andrew. (2017). VoxCeleb: A Large-Scale Speaker Identification Dataset. 10.21437/Interspeech. 2017-950.
7. Greenberg Craig et al. 2010 NIST Speaker Recognition Evaluation Test Set LDC2017S06. Hard Drive. Philadelphia: Linguistic Data Consortium. 2017.
8. Sahidullah Md & Saha, Goutam. (2012). Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. Speech Communication. 54. 543-565. 10.1016/j.specom.2011.11.004.
9. Prince Simon & Elder, James. (2007). Probabilistic Linear Discriminant Analysis for Inferences About Identity. IEEE 11th International Conference on Computer Vision. 1-8. 10.1109/ICCV.2007.4409052.
10. Nguyen Hieu & Bai, Li. (2010). Cosine Similarity Metric Learning for Face Verification. ACCV. 6493. 709-720. 10.1007/978-3-642-19309-5_55.
11. Smith Samuel & Kindermans, Pieter-Jan & Le, Quoc. (2017). Don't Decay the Learning Rate, Increase the Batch Size.
12. Kingma D.P., Adam Ba J. A method for stochastic optimization //arXiv preprint arXiv:1412.6980. 2014.
13. Gower R.M. et al. SGD: General analysis and improved rates //arXiv preprint arXiv:1901.09401. 2019.
14. Goyal Priya & Dollár, Piotr & Girshick, Ross & Noordhuis, Pieter & Wesolowski, Lukasz & Kyrola, Aapo & Tulloch, Andrew & Jia, Yangqing & He, Kaiming. 2017. Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour.
15. Matejka Pavel & Plchot, Oldřich & Zeinali, Hossein & Mošner, Ladislav & Silnova, Anna & Burget, Lukas & Novotny, Ondrej & Glembek, Ondrej. (2019). Analysis of BUT Submission in Far-Field Scenarios of VOICES 2019 Challenge. 2448-2452. 10.21437/Interspeech. 2019-2471.
16. Zeinali H. et al. How to improve your speaker embeddings extractor in generic toolkits //ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2019. C. 6141-6145.
17. Syed Omid Sadjadi, "NIST baseline systems for the 2019 speaker recognition evaluation," NIST. Tech. Rep, 2019.



Галактионова Анастасия Алексеевна

Год рождения: 1985
Университет ИТМО,
Институт дизайна и урбанистики,
аспирант,
направление подготовки: 05.13.10 – Управление
в социальных и экономических системах,
e-mail: aagalaktionova@itmo.ru



Ненько Александра Евгеньевна

Год рождения: 1985
Университет ИТМО,
Институт дизайна и урбанистики,
к.с.н., доцент,
e-mail: al.nenko@itmo.ru

УДК 303.832

**ОБЗОР СУЩЕСТВУЮЩЕЙ ПРАКТИКИ СОВМЕСТНОГО
КАРТИРОВАНИЯ ПРИ ПРОЕКТИРОВАНИИ ОБЩЕСТВЕННЫХ
ПРОСТРАНСТВ**

А.А. Галактионова

Научный руководитель – к.с.н., доцент А.Е. Ненько

Аннотация

В работе рассмотрена практика совместного картирования в развитых зарубежных странах. Акцент сделан на факторах, влияющих на применение практики, количество и качество собранных данных. Несмотря на отсутствие теоретической базы и многообразие факторов, которые влияют на результаты применения метода, специалистам удалось перевести на коммерческую основу созданные онлайн-приложения.

Ключевые слова

Совместное картирование, соучаствующие ГИС, геоинформационные системы общественного участия.

Важнейшей частью городской среды являются общественные пространства. Общественные пространства играют существенную многостороннюю роль в формировании социально устойчивых городов: поощряют людей к пешеходным и велосипедным прогулкам, совместному отдыху и игре. В рамках данной статьи в качестве “общественного пространства” понимаются территории, которыми беспрепятственно пользуется неограниченный круг лиц.

Внимание к общественным пространствам среди российских урбанистов и градостроителей, а также среди городских управленцев в последнее время повысилось, о чем, в частности, свидетельствует реализация федерального проекта “Формирование комфортной городской среды” (2019-2024 гг.), фокусом которого является создание

новых и преобразование существующих общественных пространств, а также мониторинг на федеральном уровне качества городской среды в 1115 городах. Внимание к общественным пространствам иллюстрирует и рост количества новостей в поисковой системе «Яндекс» (рис. 1).

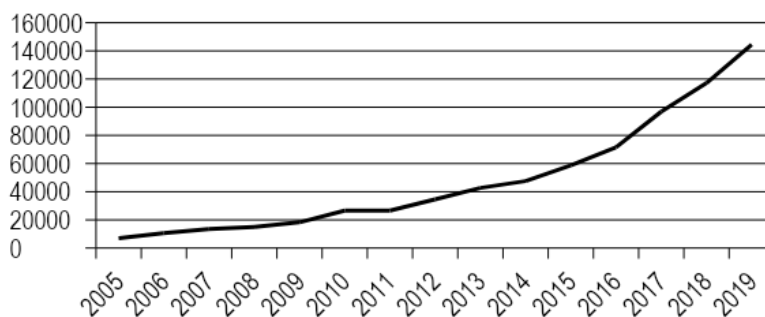


Рис. 1. Количество сообщений по теме поиска: "общественное пространство" в Яндекс.Новостях

В рамках указаний Европейского союза развиваются новые методы совместного проектирования в части развития общественных пространств. Одним из таких методов является метод совместного картирования (*participatory mapping*). Термин включает в себя понятия геоинформационных систем общественного участия (*PPGIS, participatory planning geoinformational systems*), соучаствующих ГИС (*PGIS, participatory geoinformational systems*), геоинформационных систем на основе добровольного участия (*VGI, volunteered geographic information systems*), совместное 3D-моделирование (*participatory three-dimensional modeling, P3DM*) [1]. Данный метод позволяет получить данные о восприятии жителями городской среды с геопривязкой. Вопрос, на который метод может дать ответ, какой вес придают этим ценности жители или пользователи применительно к конкретной территории. Совместное картирование позволяет людям осознать ценность места, где они живут, и осознать свою связь с местом.

Метод активно применяется учеными для сбора информации о качестве среды с точки зрения различных социальных групп: безопасность передвижений из дома в школу для школьников, комфорт городской среды для пожилого населения, оценки среды неформальных поселений (трущоб) в городах, эмоциональную оценку общественных пространств [2]. Метод также применяется в градостроительной практике европейских и российских городов, например, Хельсинки, Тампере, Таллинна, Омска.

Благодаря многолетнему опыту работы с проектами по совместному картированию М. Китты (Университет Аалто, Хельсинки) и Г. Браун (Калифорнийский государственный политехнический университет, США) выделили факторы, влияющие на количество и качество собранных данных, количество участников, а также на применение метода совместного картирования. В таб. 1 эти факторы суммированы на основании опубликованных исследований по четырём типам:

- факторы, влияющие на количество и качество данных;
- факторы, влияющие на количество участников;
- факторы, влияющие на результат применения метода;
- факторы контекста.

Влияние части факторов подтверждено исследованиями, но основная часть выявлена на основе экспертного мнения М. Китты и Г. Брауна.

Факторы, влияющие на применение метода совместного картирования

№ п/п	Факторы	Примечание	Источник
1	Факторы, влияющие на количество и качество данных		
1.1	Тип выборки	У участников случайной выборки ошибки локации встречаются чаще всего (14,5%), затем у тех, кого встретили на территории исследования (7,4%) и потом у тех, кто добровольно вызвался участвовать в опросе (5,9%). Добровольные участники наносят больше данных, чем участники случайной выборки.	[3]
1.2	Социо-демографические характеристики участников	Респондентами в основном являются люди с образованием. Качественные данные в онлайн исследованиях обеспечивают люди, свободно пользующиеся интернетом, а также люди, которые уже получили навык совместного картирования	Мнение эксперта [3]
1.3	Технология картирования (онлайн-опрос или на бумаге)	На бумаге респонденты ставят в несколько раз больше маркеров, чем на онлайн-карте	[4; 5; 6]
1.4	Интерфейс онлайн-карты	Пенсионеры по ошибке нажимали на кнопку «отмена» чаще, когда она была розовой, чем когда кнопку покрасили в серый цвет. Функцию колесика на мышке лучше отключить, чтобы случайное движение не удаляло карту в космос. Перемещение по карте было удобнее со списком районов/улиц, чем скроллингом	[1]
1.5	Использование разных обозначений (точки, линии, полигоны)	Точки наносить проще и эффективнее, чем полигоны, но при этом их нужно примерно в 10 раз больше чем полигонов	Мнение эксперта [3]
1.6	Географическая неграмотность	Трудности ориентирования по карте влияют на желание пройти опрос	[7]
1.7	Знакомство с местом	Люди знакомые с местом, отвечают точнее и подробнее	[3]
1.8	Масштаб карт	На карте мелкого масштаба посетители делают меньше отметок, чем на карте крупного масштаба. При этом респонденты предпочитают делать отметки на карте мелкого масштаба общей территории, чем листать несколько карт частей территории в крупном масштабе	Мнение эксперта [3]

продолжение таблицы

№ п/п	Факторы	Примечание	Источник
2	Факторы, влияющие на количество участников		
2.1	Тип выборки	Респонденты случайной выборки чаще бросают опрос, чем респонденты добровольной выборки	[8]
2.2	Социо-демографические характеристики участников	70% из 100% школьников-участников опроса дают ответы, в отличие от других возрастов. Пенсионерам необходимо человеческое сопровождение онлайн-опроса. Сотрудники (парка) отвечали менее охотно, чем посетители (парка)	[9]
2.3	Технология картирования (онлайн-опрос или на бумаге)	Онлайн-опросы дают 13% ответивших на опрос в противовес бумажным с 30% от количества людей, согласившихся участвовать в опросе	[4; 5; 6]
2.4	Наличие времени	День недели влиял на количество бросивших опрос (на выходных их число было меньше, чем в рабочие дни)	[8]
2.5	Знакомство с местом	Люди знакомые с местом охотнее отвечают, чем незнакомые с местом	
2.6	Необходимость авторизации для онлайн-опроса или оплаты	Необходимость вводить данные для оплаты увеличивает количество тех, кто опрос бросает. Необходимость вводить почтовый адрес побуждала маломотивированных людей сразу бросить опрос, а не в середине процесса, как они бы, вероятно, сделали в случае отсутствия авторизации.	[8]
2.7	Оплата	Оплата (по конкурсу) не увеличивает количество респондентов	[8]
2.8	Доверие участников к организаторам и спонсорам исследования	Конфликт между местными жителями африканской республики и географами Американского географического общества после того, как выяснилось, что исследование и сбор данных оплачивались военным министерством США	
2.9	Наличие заинтересованности	Заинтересованные в опросе лица чаще завершают опрос, чем незаинтересованные	[4; 8]
2.10	Тип участия: сбор данных или участие в принятии решения	Люди не хотят быть просто источниками информации. Им хочется, чтобы предоставленная ими информация влияла на принятие решений.	Мнение эксперта [4]
2.11	Неясность связи между процессом и результатами	Отсутствие видимой связи между участием и принятыми решениями	Мнение эксперта [4]
3	Факторы, влияющие на результат применения метода		
3.1	Тип выборки	Мнение респондентов случайной выборки отличается от мнения респондентов добровольной выборки. Например, на вопрос как следует развивать лесопарки, первые предпочли их сохранение и содержание, а вторые были заинтересованы в развитии утилитарных функций (зон отдыха, туалета, кафе)	

продолжение таблицы

3	Факторы, влияющие на результат применения метода		
3.2	Социо-демографические характеристики участников	Пол, возраст и образование влияют на результаты анализа	
3.3	Вопросы количественные или качественные	Количественные вопросы проще обрабатывать	
3.4	Знакомство с местом	Люди знакомые с местом по-другому его описывают, например, сотрудники парка и посетители пишут о разных потребностях места	
4	Факторы контекста		
4.1	Нерепрезентативность	Совместное картирование не может представить репрезентативную выборку (в Хельсинки на опрос по генеральному плану ответили менее 1% жителей, хотя это значительно больше, чем на публичных слушаниях и воркшопах). В каждом проекте возникает вопрос, какая выборка для решения конкретной проблемы репрезентативна, кто должен представлять заинтересованные стороны, кто защищает общественное благо, в чем заключается мнение «молчаливого большинства» и как его мнение представляет активное меньшинство»	[4]
4.2	Законодательные барьеры	В США запрещено проводить социологические опросы на федеральном уровне без одобрения Комиссии по бюджету	[4]
4.3	Предубежденность специалистов к общественному мнению	Неготовность специалистов перекладывать ответственность за принятие решений на общественное мнение либо недоверие специалистов к общественному мнению	Мнение эксперта [4]
4.4	Ограниченность метода местными интересами в ущерб национальным	Мнение местных жителей не соответствует региональным или национальным интересам (например, по прокладке транспортной инфраструктуры)	Мнение эксперта [4]
4.5	Непредсказуемость результата	Государственные органы не заинтересованы в результатах опроса, которые не соответствуют государственной политике. Некоммерческие общественные организации не заинтересованы в результатах опроса, которые противоречат политике организации (например, экологической направленности деятельности). Производство и бизнес заинтересованы как можно в более узком кругу участников согласования их деятельности	Мнение эксперта [4]
4.6	Неясность цели сбора информации, ее дальнейшее использование	Собранная информация может быть положена на полку или даже утеряна, как в случае с парками Австралии.	[1]

продолжение таблицы

4.7	Уязвимость участников	Участники могут предоставить информацию, которая может быть предъявлена им потом как свидетельство нарушений закона (аборигены рисовали места обитания, а потом выяснилось, что места находятся в границах национальных парков, где эта деятельность запрещена)	[1; 10]
4.8	Отсутствие финансирования, времени и мотивации и сложность применения одного метода в разных городских условиях	Метод картирования необходимо адаптировать к разным условиям, что требует знаний и умений от организаторов	Мнение эксперта [1]
4.9	Отсутствие регламента процедуры	Использование совместного картирования рискует еще более запутать процесс принятия решений, который и так не отличается ясностью и прозрачностью для граждан	Мнение эксперта [4]
4.10	Карьерные риски для ответственных за принятие решений	Поскольку регламенты отсутствуют, нужно искать или обучать чиновников, готовых к апробированию новых инструментов. «Чиновников не награждают за новые методы работы с гражданами. Как раз наоборот» «Не стоит находить злой умысел там, где причиной может являться неопытность или некомпетентность» Неопытность в применении новых методов работы с вовлечением граждан может стоить карьеры и падения доверия граждан к органам власти. При этом медиаторы в виде сторонних организаций это доверие не повышают	Мнение экспертов [1; 4]
4.11	Технические возможности	Онлайн-ресурсы необходимо оптимизировать и обеспечивать респондентов обратной связью	[1]
4.12	Политические риски для организаторов	Политическая ангажированность государственных учреждений, которые оплачивают из бюджета сбор данных. Боязнь политиков революционных настроений в массах	Мнение эксперта [4]

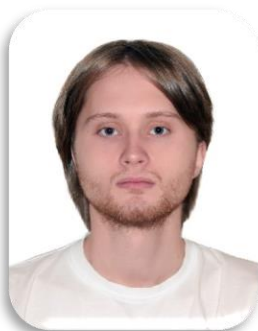
Повторяющимися факторами во всех группах являются тип выборки, социо-демографические характеристики участников, технология картирования. Они влияют на качество и количество данных, количество участников и ответивших на вопросы, результат применения метода. В результате влияния данных факторов метод совместного картирования за десятилетие использования, хотя и проявил себя как удобный инструмент при сборе данных с помощью жителей, не смог явным образом изменить уровень вовлеченности граждан в процесс планирования. Критически настроенные исследования показывают, что метод не может полностью исправить недостатки формального вовлечения граждан.

Тем не менее, метод совместного картирования развивается, поскольку средний онлайн-опрос дает около 200 ответов, что значительно больше стандартных процедур участия граждан в планировании. Также совместное картирование включает больше социальных групп, чем иные виды опросов граждан, хотя пропорционально эти группы не репрезентативны.

Основными задачами для более широкого использования совместного картирования в будущем являются: разработка теоретической базы метода; увеличение количества участников опросов; управление рисками плохого качества данных (предвзятость, репрезентативность выборки и маленький охват участников, географическая неграмотность участников); оценка эффективности использования метода.

Литература

1. Brown G., Kytta M. Key issues and priorities in participatory mapping: Toward integration or increased specialization? // *Applied Geography*. 2018. №95. P. 1–8.
2. Nenko A., Petrova M. Comparing PPGIS and LBSN Data to Measure Emotional Perception of the City. In: Alexandrov D., Boukhanovsky A., Chugunov A., Kabanov Y., Koltsova O., Musabirov I. (eds) *Digital Transformation and Global Society. DTGS 2019 // Communications in Computer and Information Science*. 2019. V. 1038. Springer, Cham https://doi.org/10.1007/978-3-030-37858-5_18. Дата обр. 10.12.2019.
3. Brown G. An empirical evaluation of the spatial accuracy of public participation GIS (PPGIS) data // *Applied Geography*. 2012. №34. P. 289–294.
4. Brown G. Public Participation GIS (PPGIS) for Regional and Environmental Planning: Reflections on a Decade of Empirical Research // *URISA Journal*. 2012. №25(2). P. 5-16.
5. Pocewicz A. An Evaluation of Internet Versus Paper-based Methods for Public Participation Geographic Information Systems (PPGIS) // *Transactions in GIS*. 2012. №16 (1). P. 39–53.
6. Kahila M. Let the Citizens Map—Public Participation GIS as a Planning Support System in the Helsinki Master Plan Process // *Planning Practice and Research*. 2015. №31 (2). P. 1–20. DOI: 10.1080/02697459.2015.1104203.
7. Kahila M. SoftGIS as a Bridge-Builder in Collaborative Urban Planning // *researchgate.net* [Электронный ресурс]. Режим доступа: https://www.researchgate.net/publication/226423814_SoftGIS_as_a_Bridge-BUILDER_in_Collaborative_Urban_Planning. регистр. DOI:10.1007/978-1-4020-8952-7_19 (дата обращения: 10.02.2020).
8. O’Neil K. M. and al. Methodological variables in Web-based research that may affect results: Sample type, monetary incentives, and personal information // *Behavior Research Methods, Instruments, & Computers*. 2001. №33 (2). P. 226–233.
9. Brown G., Kytta M. Key issues and research priorities for public participation GIS (PPGIS): A synthesis based on empirical research // *Applied Geography*. 2014. №46. P. 122–136.
10. Dunn C.E. (2007) Participatory GIS – a people’s GIS? // *Progress in Human Geography*. 2007. №31(5). P.616–637. DOI:10.1177/0309132507081493 /.



Гараев Никита Рустамович

Год рождения: 1997

Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы №М41211,

направление подготовки: 09.04.02 – Информационные
системы и технологии,

e-mail: garaevnick@gmail.com



Мельников Александр Алексеевич

Год рождения: 1986

ООО “ЦРТ”,
ведущий научный сотрудник,
к.ф.-м.н.,

e-mail: melnikov@speechpro.com

УДК 004.622

ОБЗОР МЕТОДОВ РАСШИРЕНИЯ НАБОРОВ ДАННЫХ ДЛЯ ЗАДАЧ КЛАССИФИКАЦИИ

Н.Р. Гараев

Научный руководитель – к.ф.-м.н., в.н.с. А.А. Мельников

Работа выполнена в рамках темы НИР №643202 «Применение генеративно-состязательных сетей к задаче расширения наборов данных для задач классификации».

Аннотация

В настоящее время в задаче многоклассовой классификации широкой популярностью пользуются глубокие нейронные сети. Данные сети сильно зависят от количества и качества данных, подаваемых на вход, в связи с чем приобретают востребованность методы расширения количества данных для обучения. В данной работе проводится обзор методов расширения наборов данных для классификации, обзор архитектур нейронных сетей для расширения, а так же инструментов их имплементации.

Ключевые слова

Биометрия, нейронные сети, датасет, аугментация, генеративные модели, аффинное преобразование.

1. Введение

Задача классификации встречается в практически всех сферах человеческой деятельности. Самым частым примером классификации является биометрия, то есть измерение анатомических, физиологических и поведенческих характеристик, которые используются для идентификации личности. Одна из самых распространённых задач в биометрии – автоматическое распознавание лиц.

В начале XXI века в сфере распознавание лиц широкую популярность получили нейронные сети. В 2011-ом нейронная сеть AlexNet впервые победила другие

алгоритмы в конкурсе ImageNet и тем инициировала бурный рост развития нейронных сетей, что несомненно коснулось и сферы распознавание лиц. В настоящий момент сети продолжают активно развиваться и достигли такого уровня, что на стадии обучения извлекают почти всю информацию из набора данных для обучения. В связи с этим для последующего улучшения качества распознавания нейронными сетями выгоднее расширять наборы данных, нежели улучшать алгоритмы распознавания.

2. Обзор задачи и методов расширения наборов данных

Алгоритмы машинного обучения (в том числе и нейронные сети) как правило имеют несколько стадий работы — обучение и инференс. На этапе обучения как правило выбирается одна из нескольких стратегий [1]: обучение с учителем, обучение без учителя или обучение с подкреплением. Если рассматривать стратегию обучения с учителем, то для реализации такой стратегии требуется подготовленный заранее набор данных из пар объект и относящаяся к нему метка класса этого объекта. Очевидным является тот факт, что качество предсказания алгоритма машинного обучения на стадии инференса зависит от качества обучения этого алгоритма, в параметры, влияющие на качество обучения можно отнести: стратегия обучения, длительность обучения, метрика качества обучения, функция потерь, если она уместна в данной стратегии обучения, размер и качество обучающей выборки.

2.1 Задача расширения обучающего множества

Еще в самом начале появления такой стратегии как обучение с учителем люди заметили, что качество и объем обучающей выборки влияет на конечный результат больше всего [2], поэтому в начале 10-ых годов огромную популярность приобрело такое направление, как DataMining, специализирующееся на сборе больших объемов данных, пригодных для обучения алгоритмов. На рис. 1 приведено сравнение точности предсказания разных алгоритмов машинного обучения в зависимости от размера выборки для обучения [3] (ось абсцисс – размер выборки в логарифмической шкале, ось ординат — точность предсказания).

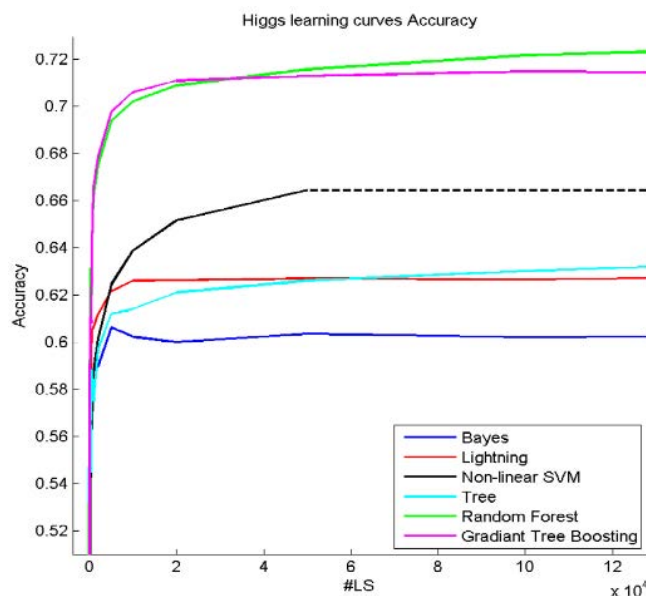


Рис. 1. Зависимость алгоритмов от размера обучающей выборки

На нем видна четкая зависимость, которая имеет место быть для большинства самых популярных алгоритмов машинного обучения: чем больше данных, тем лучше

обучение, так же стоит обратить внимание, на то, что, начиная с некоторого объема, объем выборки перестает влиять на качество обучения большинства алгоритмов. Но стоит отметить, что в отличие от других методов машинного обучения, нейронные сети зависят от качества и объема набора данных сильнее всех и практически не подвержены эффекту того, что начиная с некоторого объема, он перестает влиять на качество, данный эффект начинает проявляться только при огромных объемах выборки.

Исходя из вышеприведенных фактов, можно сделать вывод о том, что для качества работы нейронных сетей в общем, и в задачи распознавание лиц в частности, решающую роль играют объем и качество обучающей выборки. Если говорить о распознавании лиц, то первый датасет в этой области появился в 1965 году и содержал порядка 2000 изображений, после чего до конца XX века порядок объема датасетов оставался прежним.

Например, датасет FERET [4], вышедший в 1993 году содержит в себе 14000 изображений, датасет PIE [5], вышедший в 2003 году содержит в себе 40000 изображений, а датасет Multi-Pie, вышедший в 2008 году содержит в себе 360000 изображений. Но в этот момент встала проблема дальнейшего расширения датасетов, т.к. все датасеты, собранные до этого момента, собирались вручную в контролируемых условиях, что приводило к высоким затратам по времени, деньгам и человеческому труду по сбору и обработке изображений. Но в 2010-ых годах уже начали автоматизировать процесс сборки изображений для датасетов, используя краулеры, а также данные с камер наблюдения и базы данных сотрудников разных учреждений, это позволило увеличить размеры датасетов в разы, например датасет VGGFace, вышедший в 2015 году содержал в себе 2.6 миллиона изображений, а датасет MS-Celeb-1M, вышедший в 2017 – более 10 миллионов изображений. Но такие объемы данных дались человечеству не бесплатно, плата состоит в том, из-за того, что датасеты собраны автоматически, они, во-первых, имеют большое количество шума в метках, то есть часто фотографии одного и того же человека помечены как фотографии разных людей, и наоборот изображения разных людей помечены как изображения одного и того же человека.

В недавнем исследовании учёные выяснили, что «чистота» датасета в плане разметки является критической для качества работы сети, выученной на данном датасете. Они смогли получить более лучший результат работы сети, используя датасет, меньший в 5 раз, чем MS-Celeb-1M, но значительно более чистый относительно шума меток. Второй проблемой таких датасетов является то, что они часто не репрезентативны в плане распределения атрибутов, например, наборы данных лиц, собранные автоматически, имеют плохое разнообразие в эмоциях, позах, освещении, и других атрибутах лица (борода, усы, очки), то есть если база собрана с помощью камер видеонаблюдения, то шанс того что в базу попадет человек с фронтальным лицом мал, так как никто обычно не смотрит в камеру видеонаблюдения или если база собрана с помощью баз данных компаний, то в ней вряд ли будут лица с не фронтальной позой, т.к. фотография на бейдж или на пропуск делается фронтальная. Или, например, если база собиралась с помощью краулеров, то в ней будут в основном фотографии знаменитостей в фотогеничных позах, с накрашенными лицами и побритые, т.к. в основном такие изображения и попадают в интернет. Из-за этих особенностей собирания датасета сеть лучше изучит атрибуты лица, которые лучше представлены в датасете.

Несбалансированность распределений характеристик лица была широко изучена в последнее время, и на основе исследовательских работ по этой тематике можно сделать вывод, что основными проблемами являются несбалансированные

распределения: углов поворота лица (в частности горизонтальных), возрастов, освещений, гендера, расы, наличие очков, бороды, усов.

Третьей проблемой автоматически собранных датасетов являются «длинные хвосты» распределений изображений по людям, то есть такое распределение, при котором только у малого количества людей в датасете есть большое количество изображений, а у большого количества людей имеется малое количество изображений лиц. Подтверждает наличие данной проблемы распределение по количеству изображений для человека для датасета CASIA-WebFace.

В данной работе я хотел бы сделать акцент на последних двух проблемах и методах их решения, обзор которых будет приведен ниже.

2.2 Методы расширения датасета

Для решения проблемы несбалансированности распределений атрибутов лиц (то есть «длинные хвосты» распределений атрибутов лиц) датасетов и проблемы «длинные хвосты» распределений изображений используются различные методы. Некоторые из методов заключаются в том, что функцию ошибки нейронной сети создают таким образом, что она учитывает несбалансированность распределений, то есть уделяет больше внимание (а именно, давая большие градиенты при обратном распространении ошибки) тем лицам, которые в некоем распределении находятся дальше от центра. Другими словами, нейронная сеть при таком подходе будет обращать внимание при обучении, например, на женщину негра в очках и с бородой, голова которого находится в профиль, чем на белого мужчину без атрибутов на лице, голова которого повернута фронтально относительно камеры. Другие же методы решают эту проблему, тем, что выравнивают распределение, просто дублируя изображения из «хвоста» распределения столько раз, чтобы их было столько же сколько и изображений из центра распределения. Третий вид методов, речь о которых и пойдет в данной главе, занимаются тем, что генерируют изображения лиц из «хвоста» распределения, а не просто копируют существующие как в предыдущем методе. Под генерацией из «хвоста» имеется в виду, что, если азиатов в датасете представлено меньше, чем другим рас, значит они находятся в «хвосте» распределения рас для данного датасета, и тогда данный метод не просто скопирует уже имеющиеся изображения азиатов столько раз, сколько нужно для балансировки, а вместо этого сгенерирует новые изображения азиатов, которых еще нет в датасете, столько раз сколько нужно для балансировки. Этот вид методов в свою очередь может быть реализован при помощи разных методов генерации изображений, обзор о которых пойдет дальше.

2.2.1 Аугментация данных

Аугментация данных является простейшим семейством методов генерации данных на основе существующих данных и заключается в применении к изображению аффинной трансформации или отображения относительно какой-либо оси, а также добавление разного рода артефактов к изображению, будь то размытие, окклюзии, блики, смена цифровой палитры, применение цветового фильтра.

Достоинства данного метода заключаются в простоте генерации новых изображений, метод не является затратным ни с точки зрения разработки ни с точки зрения нагрузки на вычислительные ресурсы, но так же он имеет большой недостаток в виду того, что по сути сгенерированное изображение является исходным изображением с незначительными изменениями, и если переборщить с аугментацией данных, то сеть будет сильно переучиваться, так как она просто заучит те части изображения, которые не меняются аугментацией.

2.2.2 Генеративно-состязательные сети

Еще одним методом расширения датасетов являются генеративно-состязательные сети, которые способны генерировать новые изображения из исходного распределения объектов (в моем случае лиц). Смысл данного метода в том, что генеративно-состязательные сети после стадии обучения могут генерировать не изображения из исходного датасета, как предыдущие методы, а изображения из распределения исходного датасета, то есть генерировать новые изображения, по характеристикам аналогичные исходным изображениям.

Важным является тот факт, что генеративно-состязательные сети могут работать с разными видами распределений. Например, с распределением возраста, в таком случае она может генерировать изображения людей из исходного датасета, но с другим возрастом. Такой подход можно распространить на все атрибуты изображения (в моем случае лица). Созданы и широко используются сети для генерации лиц людей с разными возрастными, позами головы (поворотами вокруг оси), освещением, атрибутами лица (усы, борода, очки, шарф), эмоциями, гендером, полом, расой и так далее..

Выбирая из двух представленных выше наборов методов, в своей работе я хотел бы подробнее рассмотреть подход с генеративно-состязательными сетями, в силу их большой изученности мировым сообществом исследователей и наличия у данного метода большого потенциала для применения к задаче расширения датасета.

Так же, стоит отметить, что из множества GAN, решающих разные задачи, можно выделить как менее полезные, так и более полезные относительно моей задачи. Изучив современные работы по данной теме, я пришел к выводу, что наиболее остро для современных датасетов стоит проблема неравномерного распределения горизонтальных поворотов лиц (то есть датасеты не разнообразны в плане разных горизонтальных поворотов лиц, содержа в себе в основном лица во фронтальной позе). Это послужило мотивом для меня углубиться именно в задачу расширения датасета методом генерации лиц под разными углами поворота с помощью GAN.

3. Заключение

Исследование генеративно-состязательных сетей как метода расширения датасета изображений лиц показало широкие возможности данных сетей, что связано с тем, что они работают не с исходными изображениями, а с их распределением, следствием чего является высокая эффективность сетей применительно к задаче расширения датасетов, а также большой спектр возможных сценариев их применения.

Развитием данной работы может послужить исследование и разработка программы для решения смежных задач, таких как генерация лиц под различными углами поворота во всех направлениях, с разным освещением или с разными атрибутами (усы, борода, очки). Так же возможно данную сеть можно использовать для улучшения качества обучения нейронной сети для классификации изображений лиц.

Литература

1. Alan Lapedes, Robert Farber. How Neural Nets Work. <https://papers.nips.cc/paper/59-how-neural-nets-work.pdf>. Дата обр. 10.12 2019.
2. Jason Brownlee. Impact Of Dataset Size On Deep Learning Model Skill And Performance Estimates. <https://machinelearningmastery.com/impact-of-dataset-size-on-deep-learning-model-skill-and-performance-estimates/>. Дата обр. 10.12 2019.
3. Fozoro. Does The Dataset Size Influence A Machine Learning Algorithm? – stackoverflow.com/questions/25665017/does-the-dataset-size-influence-a-machine-learning-algorithm. Дата обр. 10.12 2019.

4. Flanagan Patricia (January 25, 2011). "Face Recognition Technology (FERET)". NIST. Retrieved 2018-07-11.
5. Gross R., Matthews I., Cohn J.F., Kanade T., & Baker S. (2008). Multi-PIE. Proceedings of the Eighth IEEE International Conference on Automatic Face and Gesture Recognition.



Двойникова Анастасия Александровна

Год рождения: 1996

Санкт-Петербургский институт информатики
и автоматизации Российской академии наук, программист,
Университет ИТМО,

факультет информационных технологий
и программирования, студент группы № M41211

направление подготовки: 09.04.02 – Информационные
системы и технологии,

e-mail: anastasia.dvoynikova@gmail.com



Верхоляк Оксана Владимировна

Год рождения: 1991

Санкт-Петербургский институт информатики
и автоматизации Российской академии наук,
младший научный сотрудник,

e-mail: overkholyak@gmail.com



Карпов Алексей Анатольевич

Год рождения: 1978

Санкт-Петербургский институт информатики
и автоматизации Российской академии наук,
д.т.н., доц., г.н.с.,

e-mail: karpov@iias.spb.su

УДК 004.934.2

**СЕНТИМЕНТ-АНАЛИЗ РАЗГОВОРНОЙ РЕЧИ ПРИ ПОМОЩИ
МЕТОДА, ОСНОВАННОГО НА ТОНАЛЬНЫХ СЛОВАРЯХ**

А.А. Двойникова, О.В. Верхоляк, А.А. Карпов

Научный руководитель – д.т.н., доцент, г.н.с. А.А. Карпов

Работа выполнена при поддержке фонда РФФИ (проект №18-07-01407).

Аннотация

В работе предлагается метод, основанный на тональных словарях, для анализа тональности разговорной речи. Для проведения экспериментов используются транскрипции аудиозаписей базы данных RAMAS, полученные с помощью систем автоматического распознавания речи. Машинная классификация данных производится на 3 класса: позитивный, нейтральный и негативный с использованием русскоязычных тональных словарей RuSentiLex, LinisCrowd, WordNetAffect и словаря Белякова.

Ключевые слова

Сентимент-анализ, тональные словари, анализ тональности текста, транскрипция текста, распознавание речи.

Введение

Способами коммуникации людей друг с другом является диалогическое и полилогическое общение, которое происходит с использованием нескольких модальностей: аудио, текстовая, зрительная и тактильная информация. Понимание содержания сообщения и его эмоциональной тональности (сентимента) требует анализа всех модальностей. Как правило, текстовая модальность отображает полярность эмоции, а акустическая модальность - ее интенсивность, поэтому при анализе разговорной речи дикторов, в первую очередь, необходимо учитывать текстовую составляющую высказывания. С помощью диалогического общения люди выражают свое мнение и отношение к различным предметам, людям и ситуациям. Машинный анализ вербальной информации является актуальным как в научной области, так и в коммерческой. За последние годы появилось множество алгоритмов машинного обучения для задачи сентимент-анализа текстов. Многие из них, в особенности нейронные сети, требуют больших объемов данных для обучения, что является проблемой для анализа текстовых данных разговорной речи. Так как диалогическое общение подразумевает под собой свободную речь дикторов на различные тематики, то сбор данных для таких алгоритмов машинного обучения является достаточно трудоемким процессом. Поэтому в статье для анализа тональности текста представлен метод, основанный на тональных словарях. Данный метод не требует больших данных для обучения, что является его главным преимуществом, также к его достоинствам можно отнести высокую скорость работы и простоту реализации алгоритма. Различные исследователи и разработчики используют данный метод для сентимент-анализа текстов. Например, в работе [1] использовался метод, основанный на тональных словарях для классификации отзывов об автомобилях и ресторанах, результат работы алгоритма 40% и 41% F-меры соответственно. Авторы работы [2], анализируя тексты СМИ при помощи словарей, достигли точности (precision) классификации около 90%. В работе [3] решается задача тернарной классификации текстов социально-политической тематики, получена точность (precision) 44%. Обзор литературы показывает, что метод, основанный на тональных словарях, показывает неплохой результат для анализа тональности текстов. В данной статье используется база данных RAMAS для построения тернарного классификатора при помощи русскоязычных тональных словарей.

База данных

В работе использовалась база данных RAMAS [4]. RAMAS – мультимодальный корпус, включающий в себя около 600 аудио- и видеозаписей интерактивных диадических сценариев. В качестве дикторов выступали 10 актеров, которые в парах разыгрывали прописанные по сценарию диалоги. Каждый сценарий включал в себя одну из 6 основных эмоций для каждого диктора: гнев, печаль, отвращение, счастье, страх или удивление, а также некоторые характеристики социального взаимодействия, такие как доминирование и подчинение. Каждое аудиовизуальное высказывание было размечено 5 аннотаторами. Задача аннотаторов состояла в том, чтобы отметить начало и конец каждой эмоции, которая казалась естественной. Если на участке аудио отсутствовала какая-либо эмоция по мнению аннотатора, то такому участку присваивался класс нейтральной эмоции.

Тональные словари

Для анализа тональности текстовой информации в работе использовались тональные словари, такие как RuSentiLex [5], LinisCrowd [3], а также переведенный на русский язык англоязычный словарь WordNetAffect [6] и словарь Белякова [7].

RuSentiLex – упорядоченный по алфавиту тональный словарь русского языка. Он содержит следующие типы сентиментальных слов: слова русского языка на различные тематики, имеющие положительную или отрицательную полярность хотя бы у одного из значений слова, сленг и ругательства из коротких сообщений социальной сети Twitter. Объем RuSentiLex составляет примерно 16 тыс. слов. Каждое слово размечено одной из 4 меток, обозначающей полярность эмоции: позитивная, негативная, нейтральная или смешанная. В словарь включены также слова, имеющие различные значения, при этом каждому значению присваивается своя метка тональности.

LinisCrowd – тональный словарь социально-политической тематики. Авторы словаря создали краудсорсинговый веб-ресурс для разметки слов добровольцами, который позволяет постоянно пополнять словарь. На данный момент LinisCrowd содержит в себе около 10 тыс. слов. Каждое слово имеет метку из множества $\{-2, -1, 0, 1, 2\}$, которые обозначают следующие 5 классов: «сильно отрицательный», «умеренно отрицательный», «нейтральный», «умеренно положительный» и «сильно положительный».

WordNetAffect – русскоязычный лексический ресурс, построенный на основе англоязычного WordNet. WordNet содержит в себе английские слова и словосочетания, размеченные на 6 классов эмоций: радость, страх, гнев, печаль, отвращение и удивление. Для создания WordNetAffect авторы работы [6] перевели каждое слово из WordNet на русский язык вручную, при этом удаляя те переводы слов, которые по значению не были близки к эмоции исходных слов. Общий объем WordNetAffect составил 2,4 тыс. слов.

В работе [7] Беляков М.В. для создания своего словаря использовал новостные ленты сайта Министерства иностранных дел Российской Федерации. Словарь содержит в себе около 700 основ слов, размеченных на 2 класса тональности: положительный и отрицательный.

Предлагаемый подход

Для анализа тональности текстовой орфографической транскрипции разговорной речи в работе использовался метод, основанный на русскоязычных тональных словарях. Классификация текстовых высказываний базы данных RAMAS производилась на 3 класса: позитивный, нейтральный и негативный. Некоторые тональные словари, описанные выше, содержат слова, помеченные более чем на 3 класса. При их использовании слова объединялись для тернарной классификации. Так, например, слова из LinisCrowd, имеющие метку -2 или -1, а также 1 или 2 объединялись в негативный и позитивный классы, соответственно. В словаре WordNetAffect слова, относившиеся к эмоциям типа гнев, страх, печаль и отвращения объединялись в негативный класс, а все остальные - в позитивный.

Для определения тональности текста при помощи словарей необходимо каждое слово из текста найти в тональном словаре и записать его значение. Если слово не присутствует в словаре, то ему присваивается нулевой вес. Далее вычисляется среднее арифметическое весов всех слов и на основе этого принимается решение о принадлежности текста к определенному классу. Если суммарный вес текста больше 0, то текст относится к позитивному классу, если меньше 0, то к негативному, и если вес строго равен 0, то текст является нейтральным.

В словаре RuSentiLex присутствуют слова, которые имеют смешанную оценку, либо слова с несколькими значениями и, соответственно, с различными весами для этих значений. В статье предложен следующий подход к работе с такими словами. Если у слова есть несколько значений с различными весами, то удалялись нейтральные значения в пользу эмоционально окрашенных. Если в списке больше слов с негативными значениями, нежели позитивными, то итоговый вес слова определялся как негативный и наоборот. Если же слово имеет смешанную оценку, то анализируются слова в тексте, стоящие до него. Если перед ним больше позитивных слов, то данное слово относится к позитивному классу, если же больше негативных или совсем нет эмоциональных слов, то данное слово относится к негативному классу.

Предобработка базы данных

Так как RAMAS содержит в себе аудиозаписи диалогов между парой актеров, то первым этапом предобработки базы данных была произведена диаризация дикторов, а также временная сегментация аудиоданных по эмоциям. Каждый сегмент должен содержать какую-либо эмоцию по мнению хотя бы одного аннотатора. Объем всех таких сегментов составил 3,2 тыс. Следующим этапом требовалось получить транскрипцию этих сегментов (перевести речь из аудиосигнала в текст). Транскрибирование выполнено при помощи библиотек автоматического распознавания речи: `speech_recognition` для языка Python от компании Google (<https://pypi.org/project/SpeechRecognition/>) и сервиса `SpeechKit` от компании Яндекс (<https://cloud.yandex.ru/services/speechkit>). Многие аудиосегменты не удалось качественно транскрибировать. Такие сегменты могли содержать в себе аудишумы, не позволяющие выделить речь дикторов, иметь недостаточную длину речевого высказывания, также речь диктора могла быть неразборчивой. При совместном использовании готовых методов распознавания речи от Google и Яндекс суммарно получилось 882 текстовых высказываний.

После получения транскрипции речи производилась предобработка текста (орфографической транскрипции речи), в том числе, токенизация (разделение текста на самостоятельные единицы - слова), приведение всех слов к единому регистру, удаление стоп-слов (слов, не содержащих в себе смысла). Список стоп-слов выгружался из библиотеки `nlk` для Python, предназначенной для работы с естественным языком. Также при помощи этой библиотеки, производился стемминг (выделение основы слова) всех слов. После описанных выше действий удалялись фразы, содержащие в себе менее 3 слов, потому что они не отображали эмоциональную коннотацию. Например «меньше больше», «билет», «забываешь» и т. п.

Так как в работе строится тернарный классификатор, то необходимо было также объединить некоторые эмоциональные состояния в один класс. В негативный класс вошли следующие эмоции: гнев, печаль, отвращение, страх и удивление, в позитивный класс - только эмоция счастья. Эмоция удивления может быть как положительной, так и отрицательной, например, фразы «Ого! Ты купил мне новую машину!» и «Это что получается? Я его больше никогда не увижу?» выражают удивление, но в первом случае фраза носит позитивную тональность, а во втором случае – негативную. Исходя из сценариев, прописанных для диалогов в базе данных RAMAS, было принято решение отнести эмоцию удивления к негативному классу.

Экспериментальные исследования

Для полученной орфографической транскрипции речи из аудиозаписей базы данных RAMAS строился тернарный классификатор при помощи русскоязычных тональных словарей. Эксперименты проводились на каждом словаре по отдельности. Далее все 4 словаря (RuSentiLex, LinisCrowd, WordNetAffect и словарь Белякова) объединялись в один словарь. При этом удалялись повторяющиеся слова с одинаковой

эмоциональной меткой. Объем итогового тонального словаря составил примерно 18 тыс. слов. Если у одного слова встречалось несколько значений с различными метками, то обработка происходила по аналогии со словарем RuSentiLex. Результаты экспериментов с различными тональными словарями с использованием количественных показателей, взвешенных для всех 3 классов точности, полноты и F-меры представлены в таб. 1.

Таблица 1

Результаты экспериментов по sentiment-анализу текстов с различными тональными словарями

Словарь \ Показатель	Точность, %	Полнота, %	F-мера, %
RuSentiLex	60,86	32,77	34,28
LinisCrowd	65,51	26,98	25,99
WordNetAffect	53,89	19,07	12,14
Словарь Белякова	02,68	16,38	04,61
Объединенный словарь	61,57	36,38	39,23

Из таб. 1 видно, что увеличение объема словаря улучшает точность распознавания. Для экспериментов из таб. 1 использовалась система распознавания речи от Google, с помощью которой было получено 712 текстовых высказываний. Также были проведены эксперименты с использованием сервиса SpeechKit от компании Яндекс для транскрибирования аудиозаписей. С помощью него было получено 672 текстовых высказываний. В ходе экспериментов было обнаружено, что некоторые аудиофайлы, которые не смог распознать SpeechKit, смогла транскрибировать библиотека speech_recognition от Google, и наоборот. Поэтому было принято решение использовать оба метода одновременно. Если оба метода от компании Google и Яндекс получали транскрипцию для одного аудиофайла, то вручную экспертом отбирался наилучший вариант. Таким образом, объем используемой базы данных составил 882 текстовых высказываний. Результаты экспериментов с различными системами распознавания русской речи представлены в таб. 2. Классификация производилась при помощи объединенного тонального словаря.

Таблица 2

Результаты экспериментов по sentiment-анализу текстов с использованием различных систем распознавания речи

Система распознавания речи	Точность, %	Полнота, %	F-мера, %
Speech_recognition от Google	61,57	36,38	39,23
SpeechKit от Яндекс	56,21	37,22	39,88
Speech_recognition + SpeechKit	61,51	40,14	43,27

Из таб. 2 можно сделать вывод, что использование одновременно двух систем распознавания русской речи улучшает результат sentiment-анализа текстов согласно взвешенной F-меры на 3,39%.

Выводы

Анализ текстовой модальности разговорной речи дикторов показал результаты для тернарной классификации взвешенной точности (Precision) = 61,51%, взвешенной полноты (recall) = 40, 14% и взвешенной F-меры – 43,27%. По сравнению с другими работами в области анализа текста, анализ разговорной речи показал довольно низкий результат. Это проблема возникает из-за низкого качества некоторых аудиофайлов в базе данных, а также из-за низких результатов работы методов распознавания речи. Приблизительно из 3 тыс. аудиосегментов удалось транскрибировать всего лишь около 800. Также на невысокий результат sentiment-анализа повлиял тот факт, что разметка базы данных происходила только по аудиозаписям, аннотаторы не учитывали текстовую составляющую диалогов между дикторами. Однако анализ тональности текстовой информации разговорной речи, может являться хорошим дополнением к анализу тональности акустических параметров дикторов.

Литература

1. Тутубалина Е.В. и др. Тестирование методов анализа тональности текста, основанных на словарях // Электронные библиотеки. 2015. Т. 18. №. 3-4. С. 138-162.
2. Пазельская А.Г., Соловьев А.Н. Метод определения эмоций в текстах на русском языке //Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 25–29 мая 2011 г.). М.: Изд-во РГГУ. 2011. №. 10. С. 17.
3. Koltsova O.Y., Alexeeva S., Kolcov S. An opinion word lexicon and a training dataset for Russian sentiment analysis of social media // Computational Linguistics and Intellectual Technologies: Materials of DIALOGUE 2016 (Moscow). 2016. С. 277-287.
4. Perepelkina O., Kazimirova E., Konstantinova M. RAMAS: Russian Multimodal Corpus of Dyadic Interaction for Affective Computing //International Conference on Speech and Computer. Springer, Cham. 2018. С. 501-510.
5. Loukachevitch N. and Levchik A. (2016), Creating a general Russian sentiment lexicon, In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), pp. 1171-1176.
6. Sokolova M., Bobicev V. Classification of emotion words in Russian and Romanian languages //Proceedings of the International Conference RANLP-2009. 2009. С. 416-420.
7. Беляков М.В. Анализ новостных сообщений сайта МИД РФ методом sentiment-анализа (статья 2) // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2016. № 4.



Дьякова Валерия Александровна

Год рождения: 1996

Университет ИТМО,

Институт дизайна и урбанистики,
студент группы №С41042

направление подготовки: 09.04.03 – Прикладная
информатика, e-mail: dyakova.valery@yandex.ru



Матросова Евгения Викторовна

Год рождения:

СПб ГБУЗ МИАЦ,

инженер отдела развития информатизации,

e-mail: MatrosovaE@spbmiac.ru



Кононова Ольга Витальевна

Год рождения: 1964

Университет ИТМО,

Институт дизайна и урбанистики,

к.э.н., доцент,

e-mail: kononolg@yandex.ru

УДК 004.9

**АНАЛИЗ ПРАКТИК И СЕРВИСОВ ДИСТАНЦИОННОГО
МОНИТОРИНГА ПОКАЗАТЕЛЕЙ ЗДОРОВЬЯ**

В.А. Дьякова, Е.В. Матросова

Научный руководитель – к.э.н., доцент О.В. Кононова

Работа выполнена совместно с СПб ГБУЗ МИАЦ рамках магистерского исследования «Проектирование сервиса «Самоконтроль за показателями здоровья» в рамках приоритетного проекта Санкт-Петербурга «Электронное здравоохранение» программы обучения «Цифровые технологии умного города».

Аннотация

Работа посвящена исследованию одного из направлений телемедицины – дистанционному мониторингу показателей здоровья пациентов. Проведен анализ практик и сервисов дистанционного мониторинга показателей здоровья в России. Изучены основные точки применения, субъекты и объекты взаимодействия, выявлены преимущества и недостатки данного электронного сервиса, а также изучены нормативные документы, определяющие основные требования и возможности его реализации в Российской Федерации. Подготовлено описание и предложена модель информационного взаимодействия в рамках сервиса дистанционного мониторинга показателей здоровья для дальнейшего проектирования.

Ключевые слова

Дистанционный мониторинг здоровья, телемедицина, цифровая медицина, электронное здравоохранение, цифровые медицинские сервисы, телемедицинское наблюдение, телемедицинский мониторинг.

Сегодня дистанционный мониторинг показателей здоровья является одним из приоритетных направлений развития телемедицины. Цифровые технологии активно внедряются в сферу здравоохранения. В рамках общего государственного вектора в области здравоохранения, направленного на повышение качества оказания медицинских услуг и доступности медицинской помощи, были запущены различные проекты, в том числе по внедрению телемедицины и дистанционного мониторинга показателей здоровья [1]. Первые проекты по осуществлению дистанционного мониторинга показателей здоровья в России начали разрабатываться и внедряться федеральными медицинскими центрами и компаниями-разработчиками с 2015 года. В 2017 году, в результате реализации данного проекта, дистанционным наблюдением было охвачено 1510 пациентов в 11 субъектах РФ. В настоящее время внедрение проекта осуществляется более чем в 25 регионах (в частности, Ленинградская область, Кемеровская область, Сахалинская область, Белгородская область, Рязанская область, Калужская область и др.).

Дистанционным мониторингом здоровья граждан занимаются не только государственные, но и частные организации. Так, например, федеральная сеть медицинских центров Ситилаб предлагает с помощью системы MeDiCase контролировать главные показатели здоровья – давление, пульс, уровень сахара и кислорода в крови. Динамику показателей пациент может отследить в личном кабинете и, при необходимости, получить онлайн-консультацию врача. Компания ООО «АЙ-ФОРС» совместно с профессором В.М. Тавровским и международными экспертами в медицине (Швейцария, Германия, Бельгия, Китай) разработали систему Remsmed, которая предназначена для удаленного мониторинга здоровья и взаимодействия пациента и врача клиники. Встроенные возможности системы позволяют медицинскому специалисту оперативно получать необходимую информацию о состоянии здоровья пациента и своевременно реагировать на изменения в самочувствии, корректируя назначенную программу лечения. Система «Монитор здоровья», разработанная при поддержке ГК «АйТи», Министерства образования и науки РФ, НИУ «Высшая школа экономики», обеспечивает дистанционное взаимодействие врачей, пациентов и медицинской организации, а также автоматизацию процессов сбора и передачи данных с персональных медицинских приборов дистанционного мониторинга [2]. Решения, которые предлагают частные компании, более функциональны, так как, в основном, это полнофункциональные системы, включающие в себя возможность проведения не только дистанционного мониторинга показателей здоровья с помощью диагностических устройств, но и проведение телемедицинской консультации, записи к врачу, ведения онлайн дневника самоконтроля и т.п.

В Санкт-Петербурге был принят к исполнению проект «Электронное здравоохранение», который входит в перечень приоритетных проектов города [3]. Целью проекта является повышение организации оказания медицинской помощи гражданам Санкт-Петербурга за счет создания интегрированной электронной медицинской карты (ИЭМК) петербуржца, которая будет объединять всю информацию о пациенте, полученную от медицинских организаций, а также реализации сервисов для пациентов, медицинских работников и руководителей сферы здравоохранения. Одним из таких сервисов является сервис для дистанционного мониторинга (ДМ) показателей здоровья, подразумевающий дистанционное взаимодействие между медицинским работником и пациентом, которое осуществляется посредством специальных сервисов, как государственных, так и частных, а также диагностических устройств.

Необходимость контроля показателей возможна в различных жизненных ситуациях. Одной из основных точек применения ДМ является контроль

за показателями здоровья при различных хронических неинфекционных заболеваниях, также дистанционный мониторинг показателей здоровья необходим при контроле показателей здоровья работников, контроле протекания беременности и состояния здоровья новорожденных.

Субъектами взаимодействия в рамках дистанционного мониторинга являются врач, пациент, организаторы здравоохранения. Пациент получает возможность дистанционного мониторинга показателей здоровья, врач – возможность дистанционного контроля показателей здоровья пациентов, организаторы здравоохранения – возможность контролировать данный процесс и, впоследствии, повысить качество оказания услуг по дистанционному мониторингу. Объектом взаимодействия в рамках дистанционного мониторинга является информация о состоянии здоровья пациента, контролируемая в рамках взаимодействия субъектов.

Применение телемедицинских технологий при оказании медицинской помощи осуществляется с соблюдением требований, установленных законодательством Российской Федерации в области персональных данных, и соблюдением врачебной тайны. Принятие законов легализовало возможность использования телемедицинских сервисов на государственном уровне, внесло ряд ограничений и определило общий порядок применения данных сервисов при оказании медицинской помощи [4].

Анализируя опыт по внедрению и применению дистанционного мониторинга, необходимо отметить преимущества его использования и возможные недостатки. Среди основных позитивных эффектов удаленного наблюдения стоит отметить повышение доступности и качества оказания медицинской помощи больным. Телемедицинский мониторинг характеризуется высокой простотой и удобством для пациента и увеличивает приверженность к лечению. Эффект дистанционного вмешательства устойчив и, в долгосрочном применении, экономически более эффективен, чем традиционные способы ведения больных. Среди предполагаемых недостатков телемедицинского наблюдения отмечен риск пассивного участия пациента в данном процессе, что, особенно при несвоевременной и неадекватной реакции лечащего врача, способствует неудовлетворительному уровню контроля показателей.

Рассмотрим процесс дистанционного мониторинга показателей здоровья более подробно на примере модели информационного взаимодействия, которая представлена на рисунке. В информационном взаимодействии участвуют пациент, врач и электронные сервисы. На очном приеме врач принимает решение о необходимости ведения дистанционного мониторинга показателей здоровья пациента. Далее, через Портал врача или медицинскую информационную систему (МИС), при условии, что данная информационная система интегрирована с сервисом «Дистанционный мониторинг показателей здоровья», врач назначает пациенту дистанционный мониторинг. Медицинский работник формирует план дистанционного мониторинга пациента, указывает мониторируемые показатели, задает их критичные значения, выбирает необходимые к заполнению пациентом опросники, устанавливает срок мониторинга. Подготовленное назначение обрабатывается сервисом, после чего пациенту в личном кабинете открывается возможность начать дистанционный мониторинг с помощью ведения дневника самоконтроля и/или с помощью автоматической передачи данных с диагностических устройств.

Пациент получает уведомление о том, что ему доступен дистанционный мониторинг. В личном кабинете на портале «Здоровье петербуржца» пациент должен ознакомиться с согласием на обработку персональных данных и с «Памяткой пациента», в которой даны рекомендации по правильному заполнению показателей здоровья в дневнике самоконтроля и использованию диагностических устройств. Полученные в ходе мониторинга данные должны быть предварительно преобразованы в удобный для анализа вид. После автоматической обработки данных врач и пациент,

через свои личные кабинеты, могут просматривать результаты мониторинга. В ходе мониторинга медицинский работник может принять решение о необходимости корректирования лечения, в таком случае он направляет пациенту сведения о рекомендациях, которые отобразятся у пациента в личном кабинете. При необходимости врач может принять решение об изменении параметров дистанционного мониторинга, он отправляет пациенту информацию об этом, после чего пациент продолжает вести дневник самоконтроля показателей здоровья с учетом изменений. Также не исключается вероятность того, что при ведении мониторинга, пациенту может потребоваться внеплановая консультация, для этого он в своем личном кабинете портала создает запрос на проведение телемедицинской консультации, в котором описывает причину и цель внеплановой консультации, после чего врач подтверждает запрос или, при необходимости, назначает пациенту очный прием. Пациенту приходит уведомление о назначении телемедицинской консультации или приглашении на очный прием. Если пришло подтверждение о проведении ТМК, то в назначенное время пациент и медицинский работник подключаются к ТМК в соответствующих разделах своих личных кабинетов. Конечным этапом является завершение дистанционного мониторинга показателей здоровья пациента, в соответствии со сроком установленным медицинским работником.

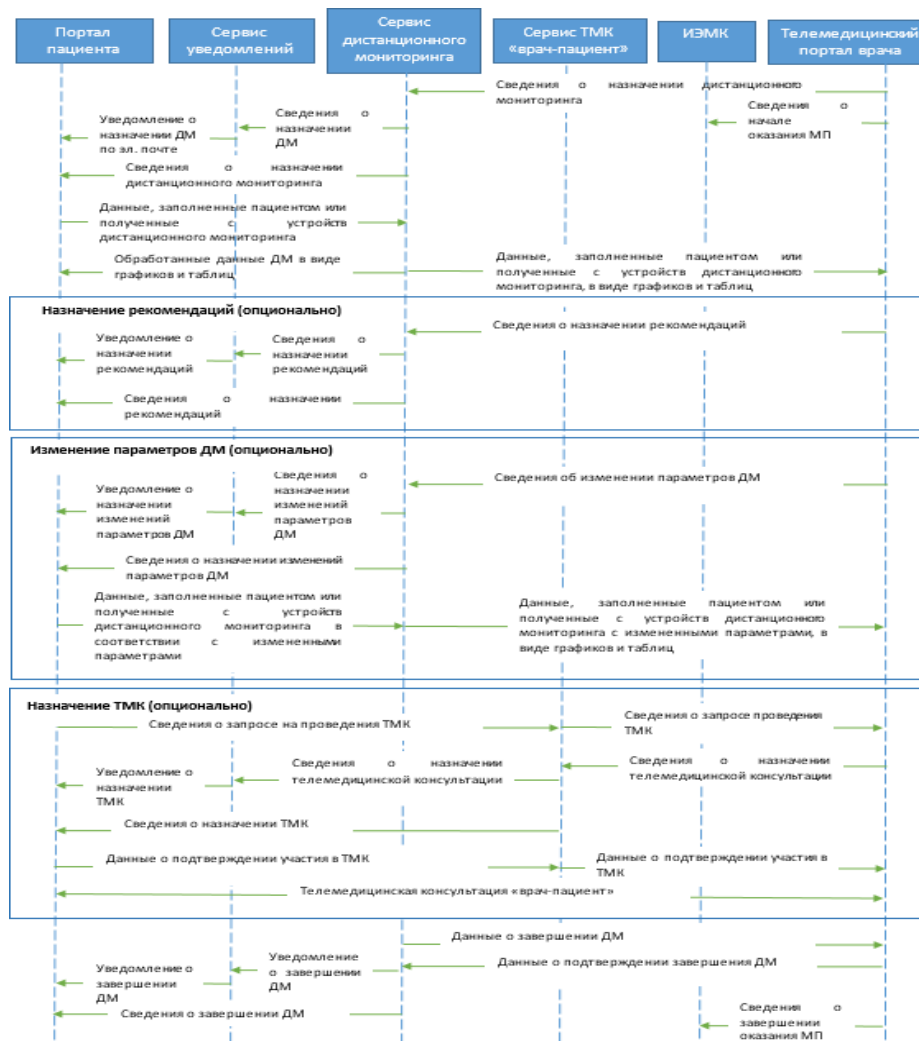


Рисунок. Модель информационного взаимодействия врача и пациента в рамках дистанционного мониторинга показателей здоровья пациентов. Данная модель будет использована при дальнейшем проектировании сервиса дистанционного мониторинга показателей здоровья в рамках приоритетного проекта Санкт-Петербурга «Электронное здравоохранение»

Телемедицинские технологии существенно повышают качество диагностики и мониторинга здоровья граждан, могут применяться для предупреждения заболеваний и повышения качества жизни россиян. Кроме того, их использование повышает доступность получения качественной медицинской помощи: граждане, проживающие в труднодоступных районах, при использовании телемедицинских технологий, имеют возможность получать качественную медицинскую помощь. Это особенно актуально для России с ее огромной территорией, неравномерным распределением населения и концентрацией ведущих специалистов-медиков в крупных городах [5].

На следующем этапе работы планируется, на основе полученных результатов исследования, составить техническое задание для проектирования сервиса для дистанционного мониторинга показателей здоровья и разработать прототипы интерфейсов будущего сервиса.

Литература

1. Григорьева Н.С., Демкина А.Е. Ограничения и возможности для достижения целей Национального проекта «Здравоохранение» в борьбе с сердечно-сосудистыми заболеваниями в условиях современной модели государственного устройства// Государственное управление. Электронный вестник. 2019 г. №76. С. 258.
2. Шалковский А.Г., Купцов С.М., Берсенева Е.А. Актуальные вопросы создания автоматизированной системы дистанционного мониторинга здоровья человека// M-health. 2016 г. С. 67-79.
3. Паспорт приоритетного проекта «Электронное здравоохранение» в редакции протокола от 25 октября 2016 года №9 [Электронный ресурс]. Режим доступа: <http://static.government.ru/media/files/9ES7jBWMiMRqONdJYVLPTyoVKYwgr4Fk.pdf> (дата обращения: 15.12.2019).
4. Прилуков М.Д. Проблемы правового регулирования телемедицины. российский и международный опыт. //Вестник Нижегородского университета им. Н.И. Лобачевского. 2018 (6). С. 136–141.
5. Карпов О.Э., Субботин С.А., Шишканов Д.В., Замятин М.Н. Цифровое здравоохранение. Необходимость и предпосылки (Федеральное государственное бюджетное учреждение «Национальный медико-хирургический Центр имени Н.И. Пирогова» Министерства здравоохранения Российской Федерации, Москва, Россия)// Врач и информационные технологии. 2017 г. №3. С. 7-9.



Евсеева Елена Сергеевна

Год рождения: 1997
Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы №М41211,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: evseevalenoshka@yandex.ru



Шуранов Евгений Витальевич

Год рождения: 1980
Университет ИТМО,
факультет информационных технологий
и программирования,
к.т.н., доцент,
e-mail: shuranov@speechpro.com

УДК 004.93

ДЕТЕКТИРОВАНИЕ АКУСТИЧЕСКИХ СОБЫТИЙ В РАЗЛИЧНЫХ УСЛОВИЯХ ОКРУЖАЮЩЕЙ СРЕДЫ

Е.С. Евсеева

Научный руководитель – к.т.н., доцент Е.В. Шуранов

Аннотация

В данной статье рассматривается детектирование акустических событий на основе задачи детектирования пения птиц, в которой требуется определить наличие или отсутствие птичьего пения в аудиозаписи. Проводится анализ решений данной задачи и приводится решение, основывающееся на использовании метода опорных векторов.

Ключевые слова

Акустические события, детектирование, метод опорных векторов.

Звуки несут в себе большое количество информации о нашей повседневной среде обитания и физических событиях, происходящих в ней, поэтому разработка методов обработки сигналов для автоматического извлечения полезной информации, такой как речь, проезжающий мимо автомобиль, шаги и т.д., имеет огромный потенциал.

Детектирование акустического события – это задача определения временных рамок, когда имеет место акустическое событие в аудиофайле. Различные условия окружающей среды подразумевает под собой то, что данные собирались в различных местах (парки, леса, город и т.п.), записывались на различные устройства в разные погодные условия. Исходя из этого данные будут иметь шумы.

В данной работе будет разобрано детектирование акустических событий на основе одного из заданий конкурса DCASE 2018 года – обнаружение звуков пения птиц.

Обнаружение звуков пения птиц в аудио является важной задачей для автоматического мониторинга дикой природы, а также в управлении гражданской

наукой и аудиотекой. Обнаружение птичьего звука является очень распространенным необходимым первым шагом перед дальнейшим анализом (например, классификация, подсчет) и позволяет проводить работу с большими наборами данных (например, непрерывный 24-часовой мониторинг) путем фильтрации данных до областей интереса.

Задача обнаружения звука пения птиц состоит в том, чтобы спроектировать систему, которая, учитывая короткую аудиозапись, будет возвращать двоичное решение о наличии или отсутствии птичьего звука любого вида [1]. Выходное решение может записываться как просто "0" или "1", или как взвешенные/вероятностные выходы в непрерывном диапазоне $[0, 1]$ для оценки алгоритмов. Для основной оценки используется известный показатель эффективности классификации – площадь под кривой ROC (AUC).

Для решения задачи предоставляется 3 отдельных набора данных для разработки и 3 набора данных для оценки, каждый из которых записан в различных условиях. Наборы имеют различные балансы положительных и отрицательных случаев наличия звуков пения птиц в аудиозаписях, различные виды птиц, различные фоновые звуки, различное записывающее оборудование. Каждый набор данных относится к отдельному проекту мониторинга звука птиц. Все наборы данных содержат 10-секундные WAV-файлы (44,1 кГц моно PCM) и вручную помечены 0 или 1, чтобы указать на отсутствие или присутствие каких-либо птиц в этом 10-секундном аудиофайле.

В существующих решениях данной задачи используются различные модификации и архитектуры нейронных сетей (DCNN, CNN, 3D-CNN+RNN). Для настройки нейронных сетей использовались изображения мел-спектрограмм, получаемые после обработки аудиофайлов. Данные, в частности спектрограммы аудиофайлов, обрабатывались следующими способами [2]:

- получением спектрограмм из фрагментов файла различной длительности;
- преобразованием спектрограммы в мел-спектрограмму;
- удалением низких и высоких частот;
- выполнением нормализации и преобразования спектрограммы мощности в децибелы;
- изменением размера спектрограммы на соответствующий входному размеру нейронной сети.

Нейронные сети хорошо проявляют себя в подобных задачах, результаты имеют высокую точность. Однако их настройка требует большого количества обучающих данных и вычислительных затрат.

Для оптимизации и ускорения настройки реализуем систему детектирования пения птиц на основе метода опорных векторов (SVM). Так как по сути предстоит решать задачу классификации, то для этого подойдет более простой в реализации метод.

Метод Опорных Векторов или SVM (от англ. Support Vector Machines) – это линейный алгоритм используемый в задачах классификации и регрессии [3]. Данный алгоритм имеет широкое применение на практике и может решать линейные и нелинейные задачи. SVM имеет некоторые преимущества перед нейронными сетями:

метод опорных векторов эквивалентен двухслойной нейронной сети, где число нейронов на скрытом слое определяется автоматически как число опорных векторов;

принцип оптимальной разделяющей гиперплоскости приводит к максимизации ширины разделяющей полосы, следовательно, к более уверенной классификации.

Так как детектор пения птиц будет реализовываться с помощью метода опорных векторов, то большое количество данных, которое содержится в картинках мел-спектрограмм, будет сильно влиять на производительность. В качестве данных, подаваемых на вход, будут использоваться мел-кепстральные коэффициенты (MFCC).

Для оценки качества системы для данной задачи будет использоваться оценка AUC-ROC – площадь (Area Under Curve) под кривой ошибок (Receiver Operating Characteristic curve). Также составим матрицу ошибок (confusion matrix) [4].

Для разработки был выбран язык программирования Python, который содержит большое количество библиотек для различных задач, в том числе для обработки и анализа данных различного вида.

Метод опорных векторов был взят из библиотеки `sklearn.svm`. Для обучения модели, основанном на этом методе, требуется подать на вход двумерный массив признаков данных и одномерный массив принадлежности к классам. В качестве входных данных для обучения было использовано 700 аудиофайлов. Набор данных для тестирования состоит из 300 других аудиофайлов.

Входные данные были преобразованы следующим образом. Каждая аудиозапись для более быстрой работы алгоритма была считана с частотой дискретизации 11025 Гц и нарезалась на фрагменты по 1 секунде. Для каждого фрагмента было вычислено 30 мел-кепстральных коэффициентов, и выполнена стандартизация коэффициентов. Таким образом, были получены массивы признаков и значений для обучения метода размеров 700x300 и 700 соответственно.

В процессе тестирования обученной модели для каждой тестовой аудиозаписи были нарезаны на фрагменты по 1 секунде и вычислено 30 мел-кепстральных коэффициентов. Коэффициенты фрагментов объединялись и классифицировались.

После проведения экспериментов, модель оценивалась по AUC-ROC и матрице ошибок. Результат оценки по матрице ошибок приведен в таблице. Оценка AUC-ROC показала результат в 63%, график приведен на рисунке.

Таблица

Матрица ошибок тестирования модели

Матрица ошибок	$y = 1$	$y = 0$
$\hat{y} = 1$	120	83
$\hat{y} = 0$	30	67

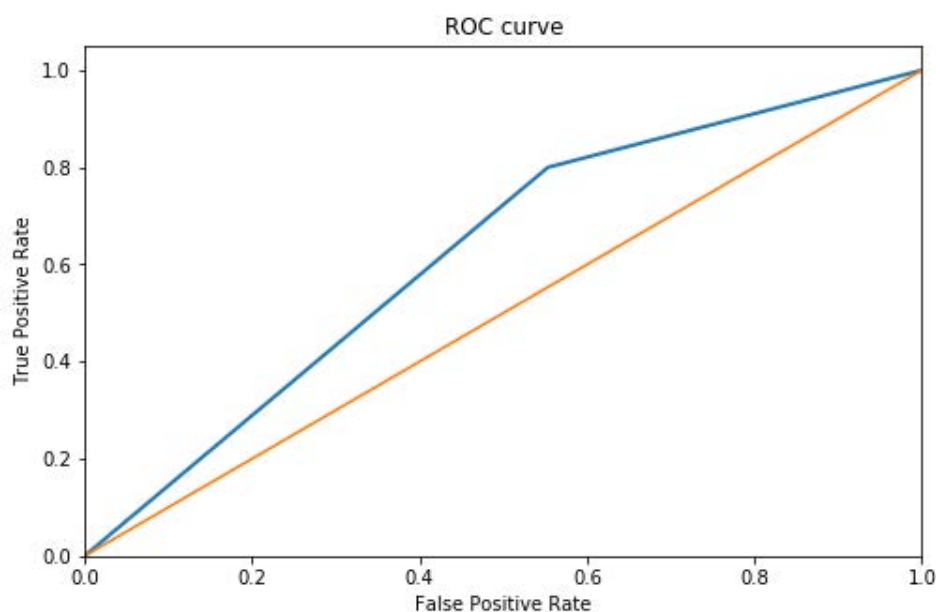


Рисунок. Оценка модели с помощью AUC-ROC

В данной работе было рассмотрено детектирование акустических событий на примере задания детектирования птичьего пения конкурса DCASE 2018. Была предложена реализация детектора пения птиц на основе метода опорных векторов и приведены результаты тестирования и оценок данной модели.

Литература

1. Bird audio detection – DCASE [Электронный ресурс]. Режим доступа: <http://dcase.community/challenge2018/task-bird-audio-detection> (дата обращения: 10.03.2020).
2. Lasseck, Mario. Acoustic bird detection with Deep Convolutional Neural Networks [Электронный ресурс]. Режим доступа: http://dcase.community/documents/challenge2018/technical_reports/DCASE2018_Lasseck_76.pdf (дата обращения: 03.03.2020).
3. Воронцов К.В., Лекции по линейным алгоритмам классификации [Электронный ресурс]. Режим доступа: <http://www.machinelearning.ru/wiki/images/6/68/Voron-ML-Lin.pdf> (дата обращения 12.03.2020).
4. Метрики в задачах машинного обучения [Электронный ресурс]. Режим доступа: <https://habr.com/ru/company/ods/blog/328372/> (дата обращения 10.03.2020).



Захарова Анастасия Алексеевна

Год рождения: 1997

Университет ИТМО,

факультет информационных

технологий и программирования,

студент группы № М41212,

направление подготовки: 09.04.02 – Информационные системы и технологии,

e-mail: nasty.zakharova2016@gmail.com



Махныткина Олеся Владимировна

Год рождения: 1982

Университет ИТМО,

факультет информационных

технологий и программирования,

к.т.н., доцент,

e-mail: makhnytchina@itmo.ru

УДК 004.9

КЛАСТЕРИЗАЦИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ С УЧЕТОМ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ

А.А. Захарова

Научный руководитель – к.т.н., доцент О.В. Махныткина

Работа выполнена в рамках темы НИР №619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Аннотация

Исследование посвящено методам тематического моделирования, направленных на выявление скрытых взаимосвязей между словами и темами в текстовых коллекциях. Цель исследования заключается в том, чтобы провести сравнительный анализ методов тематического моделирования при различных комбинациях предварительной обработки текстовых сообщений.

Ключевые слова

Извлечение информации, обработка естественного языка, тематическое моделирование, латентное размещение Дирихле (LDA), LDA Mallet.

Кластеризация текстовых документов, то есть разбиение документов на определенное, заранее не известное количество подмножеств, помеченных какими – то семантическими описателями одна из распространенных задач, решаемая различными информационными системами.

Актуальность данной темы обусловлена развитием и ростом популярности онлайн – курсов, новостях порталов, социальных сетей, которые содержат различные отзывы людей по различным вопросам. Нарастающим направлением в этой сфере является тематическое моделирование.

Тематическое моделирование – это способ построения модели коллекции текстовых документов, которая определяет, к каким темам относится каждый из документов. Построение тематической модели может рассматриваться, как задача одновременной кластеризации и слов, и документов по их семантической близости. Обычно выполняется нечеткая кластеризация, то есть документ может принадлежать нескольким темам.

В практических разработках широко используется ряд методов тематического моделирования. Среди таких моделей текста, наиболее распространенным является латентное размещение Дирихле. Латентное размещение Дирихле (LDA, Latent Dirichlet Allocation) – применяемая в информационном поиске порождающая модель, позволяющая объяснить результаты наблюдений с помощью неявных групп. Модель принадлежит семейству порождающих вероятностных моделей, в которых темы представлены вероятностями появления каждого слова из заданного набора. Документы в свою очередь могут быть представлены, как сочетания тем. Уникальность модели состоит в том, что темы не обязательно должны быть различными и слова могут встречаться в нескольких темах. Данный метод, дает нам наблюдаемое слово или лексему, по которому можно попытаться определить вероятные темы, распределение слов в каждой теме и сочетание тем в документе. Подход к LDA заключается в том, что каждый документ рассматривается, как набор тем в определенной пропорции [1].

Основными этапы построения тематической модели и оценки ее качества включают следующие этапы:

1. Предобработка текста.

- 1.1 Токенизация – это первый шаг при обработке текста, он заключается в разбиение длинных строк на более мелкие.

- 1.2 Удаление стоп – слов – это слова, которые не несут смысловой нагрузки.

- 1.3 Выявление устойчивых словосочетаний (биграмма) – модель определяет и сохраняет смежные последовательности.

- 1.4 Лемматизация – приводит слово к смысловой канонической форме.

2. Создание слова и корпуса – собрание текстов на данном языке, для того чтобы найти в полном объеме интересующую информацию.

3. Построение модели – к дополнению корпуса и словарю необходимо указать количество тем.

4. Подбор оптимального количества тем – указывается различное количество тем, где каждая тема представляет собой комбинацию ключевых слов, и каждое слово носит определенный вес в тему и на основании значений когерентности и перплексии определяется оптимальное количество тем.

В качестве оценки эффективности работы алгоритмов используются следующие – это когерентность и перплексия. Тема называется когерентной, если термины, наиболее частые в данной теме, неслучайно часто совместно встречаются рядом в документах коллекции. В данной работе использовалась мера когерентности, являющаяся логарифм условной вероятности (log conditional probability, LCP), оценивающая вероятность менее частого слова при условии более частого:

$$LCP(t) = \sum_{i=1}^{k-1} \sum_{j=i}^k \log \frac{N(w_i, w_j)}{N(w_i)},$$

где w_i – i -й термин в порядке убывания, $N(w)$ – число документов, в которых термин w встречается хотя бы один раз, $N(w, w')$ – число документов, в которых термины w, w' встречаются рядом хотя бы один раз [2].

Перплексия является, измерением того, насколько хорошо вероятностная модель предсказывает тему. Так же используется для сравнения вероятностных моделей.

Низкая перплексия указывает на то, что распределение вероятностей хорошо подходит для прогнозирования выборки.

В работе рассматривается датасет русскоязычных отзывов из онлайн – курса по «Машинному обучению», который имеет пятибалльную шкалу оценивания. Данный датасет состоит из 1700 отзывов, в нем 788 отзывов с оценкой «5», 310 с оценкой «4», 262 с оценкой «3», 153 с оценкой «2» и 187 с оценкой «1». Построение тематической модели осуществлялось с использованием реализации метода LDA библиотеке Gensim и библиотеки Mallet. Mallet – это пакет на основе Java для статистической обработки естественного языка, кластеризации, извлечения информации и других приложений машинного обучения для обработки текста. Инструментарий для моделирования тем Mallet содержит эффективные реализации на основе выборок скрытого распределения Дирихле, распределение Патинко и иерархического LDA [3, 4].

Оценки, полученные в результате проведения тематического моделирования с помощью различных методов, представлены в таблице.

Таблица

Оценка работы различных методов

Методы предобработки текста	Модель			
	LDA		LDA Mallet	
	Когерентность	Перплексия	Когерентность	Перплексия
Токенизация	0.36	-7.84	0.32	-6.84
Токенизация + удаление стоп-слов	0.34	-8.76	0.29	-5.86
Токенизация + выявление устойчивых словосочетаний	0.53	-7.83	0.32	-6.83
Токенизация + удаление стоп-слов + устойчивые словосочетания	0.38	-8.77	0.33	-7.15
Лемматизация с токенизацией	0.53	-6.80	0.35	-7.28
Лемматизация с токенизацией + удаление стоп-слов	0.47	-7.36	0.37	-7.64
Лемматизация + токенизация + выявление устойчивых словосочетаний	0.47	-6.87	0.36	-7.49
Лемматизация + токенизация + удаление стоп-слов + выявление устойчивых словосочетаний	0.36	-7.49	0.38	-8.77

Наилучшие результаты показали методы «Токенизация + выявление устойчивых словосочетаний» и «Лемматизация с токенизацией», где оценка когерентности составляет 0.53. Оптимальное количество тем в модели LDA составляет 20 тем, при когерентности 0.39. На рисунке представлена интерактивная диаграмма модели LDA с помощью ruLDAvis.

На левой части графика круги представляют собой тему. Чем больше круг, то тема считается распространенной. На правой части представлены слова, они являются ключевыми словами и они формируют выбранную тему. На примере, во втором кругу ключевыми словами являются «курс», «очень», «хороший». А в пятом кругу будут «вопрос», «плохо», «объяснить».

Таким образом, лучшие результаты получены с использованием модели LDA с использованием токенизации, лемматизации и выявления устойчивых словосочетаний. В дальнейшем планируется провести исследование нейросетевых моделей для кластеризации текстовых документов.

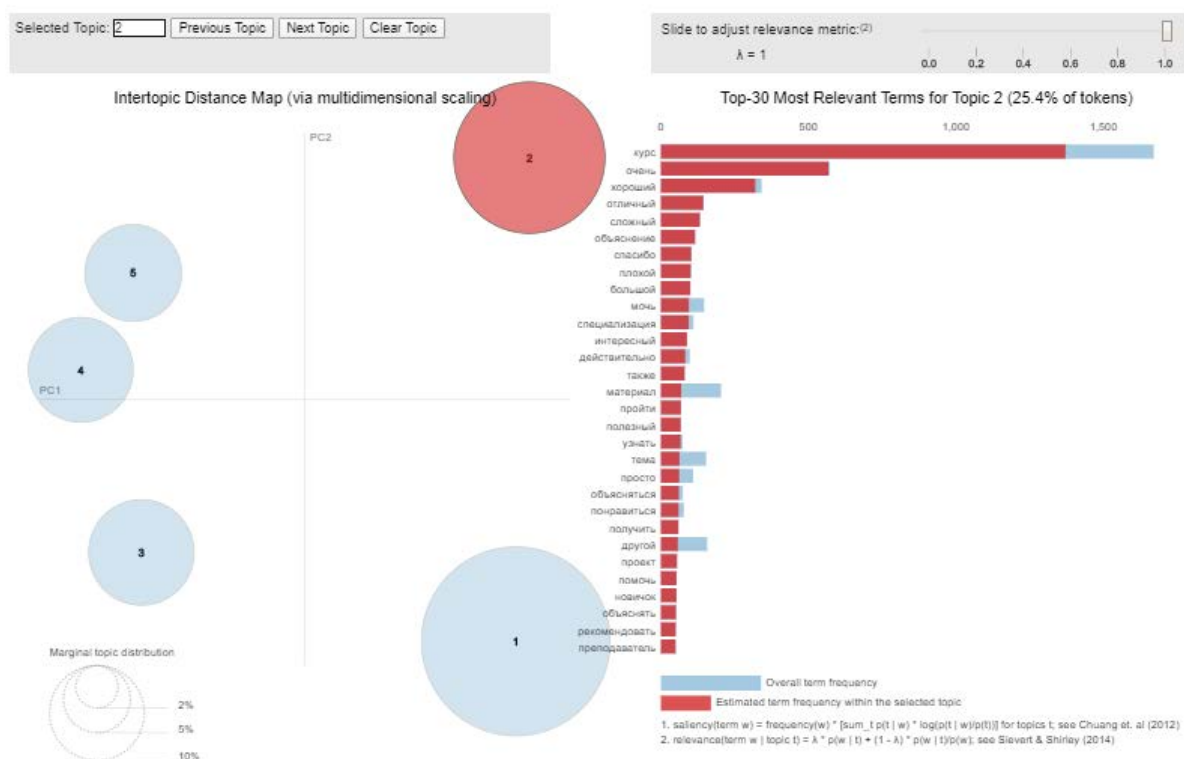


Рисунок. Модель LDA

Литература

1. Коршунов А., Гомзин А. Тематическое моделирование текстов на естественном языке // Труды Института системного программирования РАН. 2012. С. 215-242.
2. Воронцов К. В. Вероятностное тематическое моделирование. 2013 [Электронный ресурс]. Режим доступа: <http://www.machinelearning.ru/wiki/images/2/22/Voron-2013-ptm.pdf> (дата обращения: 10.01.2020).
3. MALLET [Электронный ресурс]. Режим доступа: <http://mallet.cs.umass.edu/topics.php> (дата обращения: 8.01.2020).
4. Егоров А.В., Куприянова Н.И. Особенности методов кластеризации данных // Известия ЮФУ. Технические науки. 2011. С. 174-178.

Ильина Екатерина Ростиславовна

Год рождения: 1983

Университет ИТМО,

Институт дизайна и урбанистики,

аспирант,

направление подготовки: 09.06.01 – Информатика

и вычислительная техника,

e-mail: ilyinaer@gmail.com

УДК 338.49

ОБЗОР МЕТОДИК ОПРЕДЕЛЕНИЯ ДОСТУПНОСТИ СОЦИАЛЬНЫХ ОБЪЕКТОВ И ОБЕСПЕЧЕННОСТИ НАСЕЛЕНИЯ СОЦИАЛЬНЫМИ ОБЪЕКТАМИ

Е.Р. Ильина

Научный руководитель – к.т.н. С.А. Митягин

Аннотация

Работа направлена на решение проблемы управления планированием размещения и использования объектов социальной инфраструктуры в крупных городах и посвящена анализу основных понятий и методик, которые используются для решения данной задачи. В частности, рассматриваются понятия обеспеченности и доступности социальными объектами, проводится анализ нормативных документов и обзор методик определения доступности социальных объектов и обеспеченности населения социальными объектами на примере школьных образовательных учреждений Санкт-Петербурга.

Ключевые слова

Доступность, обеспеченность, объекты социальной инфраструктуры, управление планированием.

В последнее время в России остро проявилась проблема нехватки социальной инфраструктуры, которая наблюдается во многих городах страны. Это связано с рядом причин: увеличением числа населения, износом имеющихся зданий, недостаточными темпами строительства новых объектов инфраструктуры. Для решения данной проблемы необходимо не только грамотно планировать размещение новых учреждений и использование имеющихся, но и точно оценивать ситуацию на данный момент, что невозможно без проведения анализа существующих методик и выявления оптимальной методики.

Целью данной работы является выявление методик определения доступности и обеспеченности, оптимальных для оценки потребности в объектах социальной инфраструктуры. В качестве примера рассматривается определение доступности и обеспеченности объектов школьного образования.

Как показал анализ нормативно-правовых документов и научной литературы, точные определения обеспеченности и доступности отсутствуют. Под обеспеченностью понимается достаточность, применимо к общеобразовательным учреждениям – достаточное количество мест. Под доступностью понимается равенство образовательных возможностей независимо от социально-экономического статуса, места проживания, возможностей здоровья, причем зачастую в исследованиях упор делается на пространственную доступность, в том числе пешеходную.

Изучение публикаций по доступности и обеспеченности социальной инфраструктуры позволило выявить следующие методики определения пешеходной доступности:

– определение доступности по времени, затрачиваемому на путь к объекту, в сравнении с нормативным временем. В данной методике время может быть определено как непосредственно путём измерения, так и путём вычисления, исходя из данных о длине действительно пройденного пути (рис. 1) и средней скорости пешехода;



Рис. 1. Длина пройденного пути от школы до жилого дома

– определение доступности на основе радиусов обслуживания, т.е. путём сравнения расстояния между объектами, измеренного по прямой (рис. 2), с нормативным радиусом обслуживания, который может варьироваться в зависимости от региона.



Рис. 2. Длина прямого расстояния от школы до жилого дома

Анализ публикаций по теме исследования обеспеченности социальной инфраструктурой и, в частности, общеобразовательными учреждениями, выявил несколько различных методик определения обеспеченности:

1) определение обеспеченности с использованием региональных нормативов градостроительного проектирования. В данной методике проводится сравнение соотношения фактического числа мест в общеобразовательных учреждениях к числу жителей, измеренному в тысячах, с нормативным показателем – числом мест на 1 тыс. жителей [1];

2) определение обеспеченности по сменам, то есть путём сравнения с нулевым значением процентного соотношения суммы обучающихся во вторую и третью смены к общей численности обучающихся [2];

3) определение обеспеченности по фактическому числу мест, то есть путём сравнения с единицей соотношения общего числа мест в общеобразовательных учреждениях к общей численности учащихся [3];

4) определение обеспеченности по числу общеобразовательных учреждений. В данной методике соотношение числа общеобразовательных учреждений к числу административных образований сравнивается с средним количеством общеобразовательных учреждений на одно административное образование [4].

На основании проведенного исследования можно сделать вывод, что оптимальными методиками для оценки потребности в объектах социальной инфраструктуры являются методика определения доступности по времени, затрачиваемому на путь к объекту, и методика определения обеспеченности по фактическому числу мест. Достоинством методики определения доступности по времени, в отличие от методики определения доступности на основе радиусов обслуживания, является учёт фактически существующих пешеходных путей, длина которых может значительно отличаться от длины пути, построенного напрямую между объектами. Достоинством методики определения обеспеченности по фактическому числу мест является сравнение фактически существующих данных (число мест и число учащихся), а также возможность применения методики не только для выявления дефицита мест, но также для выявления возможного избытка мест в школах. В то же время, нужно учитывать, что для применения данных методик требуется предварительная работа с исходными данными, формирование наборов которых может потребовать значительного времени.

Литература

1. Перькова М.В., Бутко О.В. Анализ обеспеченности образовательными учреждениями территории Белгородской агломерации // Научные технологии и экономика инноваций (XXII научные чтения): материалы Междунар. научно-практ. конф. (Белгород. 6–7 окт. 2016 г.). 2016. С. 167–173.
2. Хубулова В.В. Типология, особенности и факторы регионального развития социальной сферы // Вестник НГУЭУ. 2013. № 3. С. 118–131.
3. Приказ Минрегиона России от 09.09.2013 №371 «Об утверждении методики оценки качества городской среды проживания» [Электронный ресурс]. Режим доступа: <http://docs.cntd.ru/document/499077345> (дата обращения: 28.02.2020).
4. Илюхин А.А., Илюхина С.В. Обеспеченность дошкольными и общеобразовательными учреждениями городских округов и муниципальных образований Свердловской области // Материалы IV Всероссийского симпозиума по региональной экономике. 2017. С. 20–23.



Карлов Борис Александрович

Год рождения: 1977
Университет ИТМО,
факультет инфокоммуникационных технологий,
студент группы № К42201,
направление подготовки: 11.04.02 – Программное
обеспечение в инфокоммуникациях,
e-mail: bkarlov@gmail.com



Ананченко Игорь Викторович

Год рождения: 1968
Университет ИТМО,
факультет инфокоммуникационных технологий,
к.т.н., доцент,
e-mail: anantchenko@yandex.ru

УДК 004.891.2

**ИСПОЛЬЗОВАНИЕ BLACKHAT-ИНСТРУМЕНТА
ДЛЯ ЗАДАЧ ИНВЕНТАРИЗАЦИИ СЕТЕВОЙ ИНФРАСТРУКТУРЫ**

Б.А. Карлов, И.В. Ананченко

Научный руководитель – к.т.н., доцент И.В. Ананченко

Аннотация

В работе рассмотрено тестовое практическое развёртывание веб-сервиса 3WiFi и возможное его использование в частной сетевой инфраструктуре, например, для задач её инвентаризации или аудита безопасности.

Ключевые слова

Веб-сервис 3WiFi, blackhat-инструмент, Router Scan, безопасность Wi-Fi.

Введение

Технология беспроводной локальной сети Wi-Fi на протяжении вот уже почти четверти века развивается если и не стремительно, то уж точно последовательно и неуклонно. Регулярный выпуск на протяжении этого времени спецификаций стандарта IEEE 802.11 (рис. 1) и планы Wi-Fi Alliance выйти за пределы «домашней» сети (рис. 2), а также востребованность Wi-Fi у пользователя – сейчас такой модуль есть в каждом смартфоне и ноутбуке – позволяют предположить, что у технологии есть стабильное будущее [1].

Вместе с тем, бурный рост приносит и проблемы – технология уже настолько привычна обычному пользователю, не связанному с миром инфокоммуникаций, что он не задумывается о рисках, которые заложены в ней по природе или возникли в процессе её эволюции. Бум Wi-Fi породил целое новое поколение хакеров, специализирующихся на изобретении всё новых и новых способов взлома беспроводки и атаки пользователей и корпоративной инфраструктуры. Ещё с 2004 года Gartner предупреждали, что безопасность WLAN будет одной из основных проблем – и прогноз оправдывается [2].

Особенности Wi-Fi с точки зрения безопасности – это трудно контролируемая среда с общим доступом, лёгкость развёртывания инфраструктуры, мобильность пользователей, удобство и незаметность атак на сеть. При этом для беспроводной сети традиционные средства защиты зачастую уже не могут обеспечить достаточную эффективность (рис. 3) [3].



Рис. 1. Эволюция спецификаций стандарта IEEE 802.11

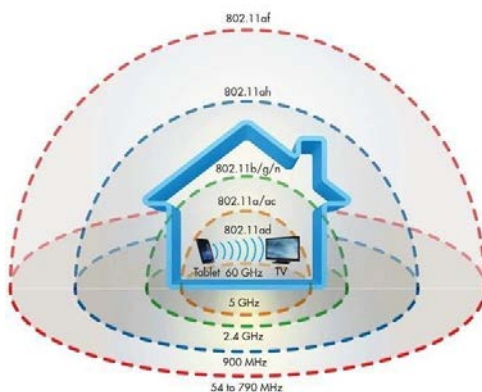


Рис. 2. Стандарты Wi-Fi нынешних и следующих поколений

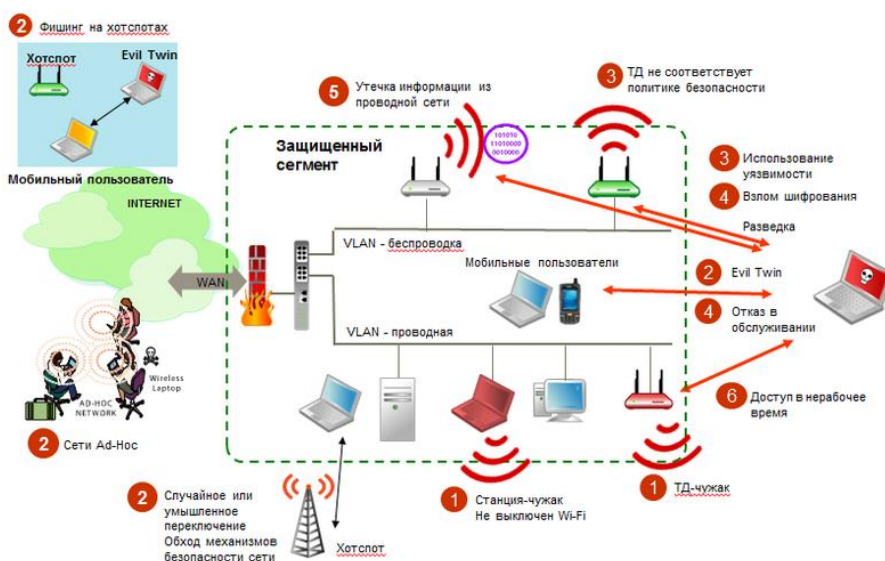


Рис. 3. Основные риски безопасности сетей Wi-Fi

Изучению одного из упомянутых инструментов, эксплуатирующих известные уязвимости и некорректные конфигурации, и – даже в большей степени – связанного с ним веб-сервиса и посвящена данная работа.

Описание и функциональность ПО Router Scan и сервиса 3WiFi

Как пишут сами авторы изучаемого программного обеспечения, «Router Scan умеет находить и определять различные устройства из большого числа известных роутеров/маршрутизаторов и, что самое главное, - вытаскивать из них полезную информацию, в частности характеристики беспроводной сети: способ защиты точки доступа (шифрование), имя точки доступа (SSID) и ключ точки доступа (парольная фраза). Также получает информацию о WAN соединении (удобно при сканировании локальной сети) и выводит марку и модель роутера. Получение информации происходит по двум возможным путям: программа попытается подобрать пару логин/пароль к маршрутизатору из списка стандартных паролей, в результате чего получит доступ. Либо будут использованы неразрушающие уязвимости (или баги) для конкретной модели маршрутизатора, позволяющие получить необходимую информацию и/или обойти процесс авторизации» [4]. Также есть упоминание о возможностях беспроводного аудита сетей, но всё же в большей степени ПО ближе к blackhat-инструментам, к утилитам для взлома устройств и получения несанкционированного доступа (рис. 4).

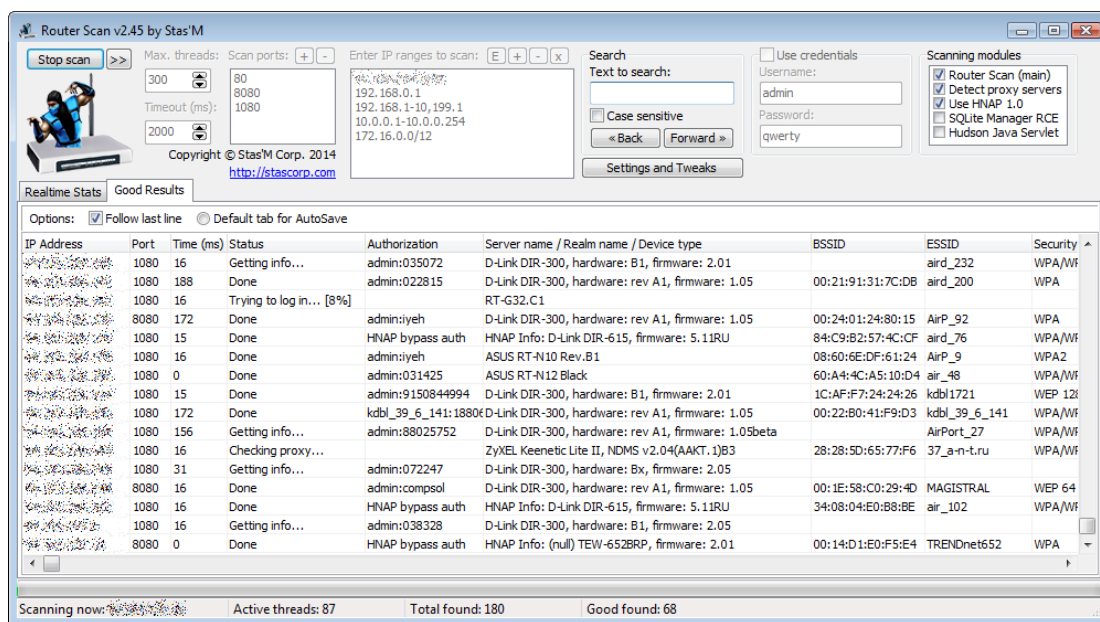


Рис. 4. Демонстрация процесса сканирования

Получить данный вывод заставляет и веб-сервис 3WiFi – свободная база точек доступа, информация для которого собирается пользователями именно с помощью ПО Router Scan, а также его упоминание на ресурсах хакерской направленности [5].

Веб-сервис 3WiFi собирает в единой базе данных информацию об уязвимых роутерах, для удобства использования отображает её на карте, позволяет находить точки доступа по идентификаторам беспроводной сети и MAC-адресам, а также по местоположению (и даже в некотором радиусе вокруг – до 25 км!). Сервис хорошо интегрирован с ПО Router Scan: отчёты, формируемые программой при сканировании сети, полностью готовы для загрузки в базу данных проекта и не требуют дополнительной обработки. Данные о координатах устройств сервис получает от нескольких провайдеров геолокации (Wi-Fi positioning system).

В сентябре 2019 года создатели сообщили, что база данных насчитывает более 11 миллионов записей. Ресурс в основном ориентирован на Россию и соседние государства, поэтому большая часть точек доступа проекта сосредоточена в крупных городах России. И тем внушительнее выглядит база данных 3WiFi – ведь это только уязвимые и скомпрометированные устройства! Для сравнения: у крупных провайдеров геолокации базы данных насчитывают 1-2 миллиарда записей уникальных сетей Wi-Fi (рис. 5).

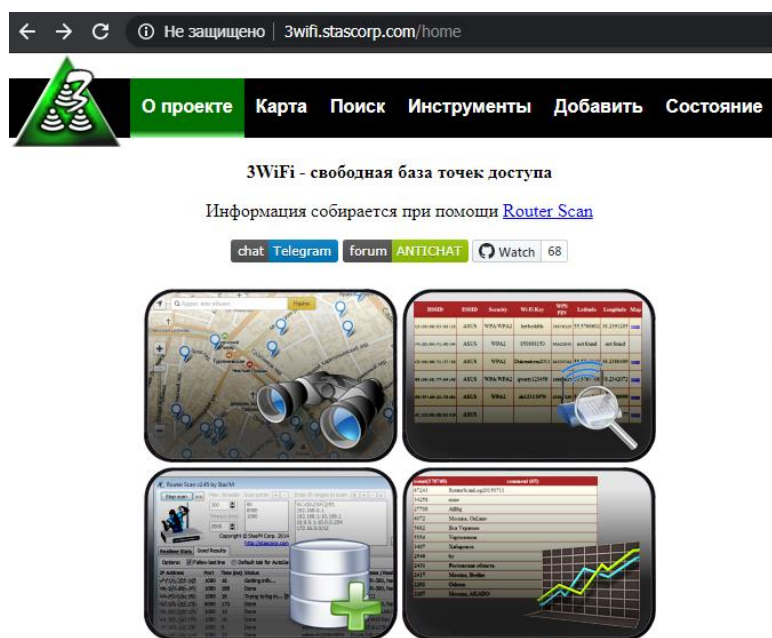


Рис. 5. Веб-сайт сервиса 3WiFi

Примечательно, что проект опубликовал исходный код сервиса на GitHub. Основная разработка шла в 2015-2018 годах силами нескольких программистов, но репозиторий не лишён поддержки – последнее изменение датировано 25.10.2019. Этой возможностью можно воспользоваться для развёртывания собственной копии сервиса, изучения исходного кода и особенностей работы и применения «в мирных целях».

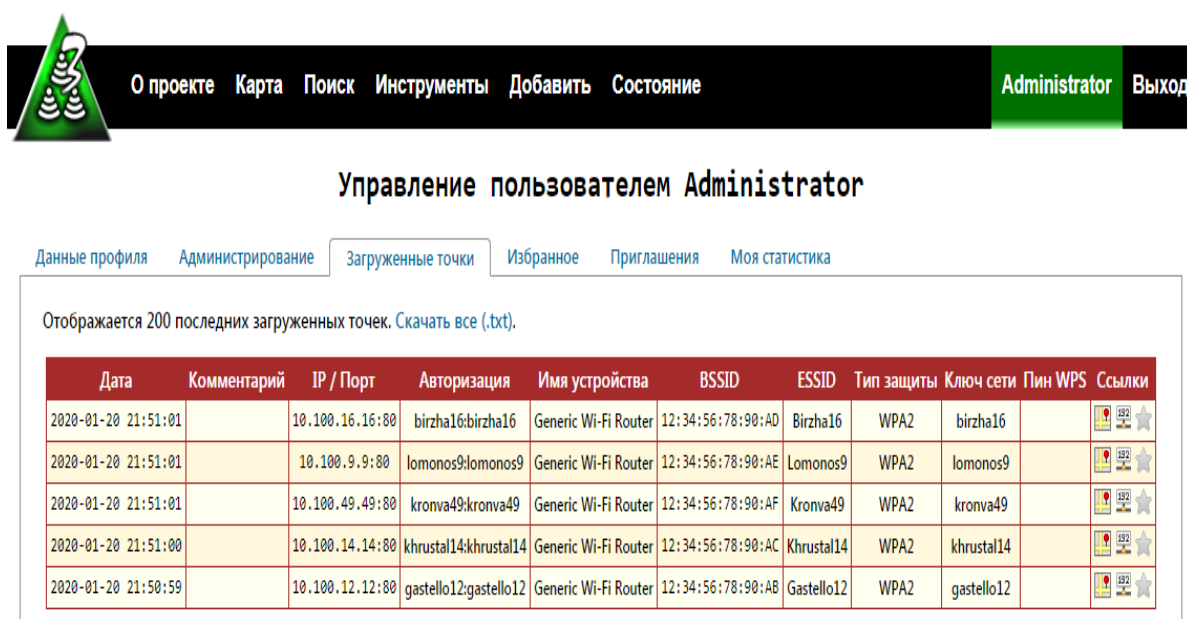
Развёртывание сервиса 3WiFi и использование его для сбора информации о сетях Wi-Fi

Для создания тестового окружения достаточно взять любой набор XAMP (X – любая подходящая операционная система, A – веб-сервер Apache, M – СУБД MariaDB/MySQL, P – язык программирования PHP). В этой работе используется GNU/Linux Debian 9.9 (Linux kernel 4.9) с соответствующими этой версии пакетами (без обновлений): Apache 2.4.25, MariaDB 10.1.38, PHP 7.0.33. Процесс развёртывания сервиса кратко описан на репозитории проекта <https://github.com/binarymaster/3WiFi>: скопировать файлы с исходным кодом PHP в корневой директорий веб-сервера, создать базу данных с помощью файла 3wifi.sql и настроить веб-сервер и конфигурационный файл сервиса для доступа к базе данных и работы с ней.

Сервис создавался для пятой версии PHP, поэтому потребуется некоторая адаптация для PHP7. В первую очередь это касается способа передачи POST-данных: в основном управляющем файле 3wifi.php вместо предопределённой переменной \$HTTP_RAW_POST_DATA следует использовать php://input.

Кроме того, для работы php-функций сервиса требуются следующие расширения/модули PHP: MySQL, Bcmath, CURL, JSON, XML/SimpleXML.

Для демонстрации в базу были загружены данные о нескольких точках доступа – с помощью искусственно созданного файла отчёта Router Scan, содержащего информацию о якобы уязвимых роутерах. И в этот момент здесь возникает принципиальный вопрос о целесообразности использования сервиса в частных целях (при наличии уже довольно большой базы проекта 3WiFi) и критериев его применения. Дело в том, что сервис может быть использован для разных целей – например, для сбора и хранения информации о сетевой инфраструктуре, при этом в качестве полезного дополнения выступает карта с нанесёнными на неё устройствами. Среди функциональности ПО Router Scan есть получение доступа к роутерам с параметрами доступа из predeterminedного словаря, поэтому программа при сканировании составит отчёт, готовый для загрузки, если заранее «обучить» её стандартным парам «логин-пароль», используемым в частной сети. С другой стороны, ещё одним вариантом применения может быть проверка устройств частной сети на наличие известных уязвимостей для своевременного их устранения (рис. 6).



The screenshot shows the web interface of Router Scan. At the top, there is a navigation bar with a logo on the left and menu items: "О проекте", "Карта", "Поиск", "Инструменты", "Добавить", "Состояние", "Administrator", and "Выход". Below the navigation bar is the title "Управление пользователем Administrator". There are several tabs: "Данные профиля", "Администрирование", "Загруженные точки" (selected), "Избранное", "Приглашения", and "Моя статистика". The main content area displays "Отображается 200 последних загруженных точек. Скачать все (.txt)." and a table with the following data:
















Дата	Комментарий	IP / Порт	Авторизация	Имя устройства	BSSID	ESSID	Тип защиты	Ключ сети	Пин WPS	Ссылки
2020-01-20 21:51:01		10.100.16.16:80	birzha16:birzha16	Generic Wi-Fi Router	12:34:56:78:90:AD	Birzha16	WPA2	birzha16		  
2020-01-20 21:51:01		10.100.9.9:80	lomonos9:lomonos9	Generic Wi-Fi Router	12:34:56:78:90:AE	Lomonos9	WPA2	lomonos9		  
2020-01-20 21:51:01		10.100.49.49:80	kronva49:kronva49	Generic Wi-Fi Router	12:34:56:78:90:AF	Kronva49	WPA2	kronva49		  
2020-01-20 21:51:00		10.100.14.14:80	khkrustal14:khkrustal14	Generic Wi-Fi Router	12:34:56:78:90:AC	Khkrustal14	WPA2	khkrustal14		  
2020-01-20 21:50:59		10.100.12.12:80	gastello12:gastello12	Generic Wi-Fi Router	12:34:56:78:90:AB	Gastello12	WPA2	gastello12		  

Рис. 6. Точки доступа в веб-интерфейсе управления

Определение координат производится через провайдеров геолокации и поэтому может не дать результатов для частных сетей (и особенно – для только что введённых в эксплуатацию точек доступа), но в сервисе есть возможность задать координаты вручную. Другой способ – создать небольшой php-скрипт, обновляющий координаты в базе данных сервиса, принимая таблицу с ними в качестве входных параметров (рис. 7).

После загрузки через веб-интерфейс файлы отчётов обрабатываются демонами 3wifid.php, отвечающими за загрузку данных в базу, получение координат и завершение обработки заданий. Итоговый результат – отображение точек доступа на карте и возможность использовать их в поиске (рис. 8).

Заключение

Router Scan – это довольно мощный инструмент сканирования сетей, организации атак, пентестинга, аудита безопасности, он может стать универсальным способом проверки устройств во время их внедрения и эксплуатации, а вместе с сервисом 3WiFi – удобной платформой для инвентаризации сетевой инфраструктуры.

```

MariaDB [3wifi]> select * from base;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| id | time | cmtid | IP | Port | Authorization | name | RadioOff | Hidden |
| NoBSSID | BSSID | ESSID | Security | WiFiKey | WPSPIN | LANIP | LANMask | WANIP | WANMask |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 15 | 2020-01-20 21:50:59 | NULL | 174328844 | 80 | gastello12:gastello12 | Generic Wi-Fi Router | | | |
| 0 | 20015998341291 | Gastello12 | 193 | gastello12 | 1 | -1062728703 | -256 | 174328844 | -1 |
| NULL | 174325813 | 174390837 | NULL |
| 16 | 2020-01-20 21:51:00 | NULL | 174329358 | 80 | khrustal14:khrustal14 | Generic Wi-Fi Router | | |
| 0 | 20015998341292 | Khrustal14 | 193 | khrustal14 | 1 | -1062728191 | -256 | 174329358 | -1 |
| NULL | 174325813 | 174390837 | NULL |
| 17 | 2020-01-20 21:51:01 | NULL | 174329872 | 80 | birzha16:birzha16 | Generic Wi-Fi Router | | |
| 0 | 20015998341293 | Birzha16 | 193 | birzha16 | 1 | -1062727679 | -256 | 174329872 | -1 |
| NULL | 174325813 | 174390837 | NULL |
| 18 | 2020-01-20 21:51:01 | NULL | 174328073 | 80 | lomonos9:lomonos9 | Generic Wi-Fi Router | | |
| 0 | 20015998341294 | Lomonos9 | 193 | lomonos9 | 1 | -1062729471 | -256 | 174328073 | -1 |
| NULL | 174325813 | 174390837 | NULL |
| 19 | 2020-01-20 21:51:01 | NULL | 174338353 | 80 | kronva49:kronva49 | Generic Wi-Fi Router | | |
| 0 | 20015998341295 | Kronva49 | 193 | kronva49 | 1 | -1062719231 | -256 | 174338353 | -1 |
| NULL | 174325813 | 174390837 | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)

MariaDB [3wifi]> select * from geo;
+-----+-----+-----+-----+
| BSSID | latitude | longitude | quadkey |
+-----+-----+-----+-----+
| 20015998341291 | 59.85950089 | 30.32410049 | 26828383880804 |
| 20015998341292 | 59.91076660 | 30.40186501 | 26828292826590 |
| 20015998341293 | 59.94440842 | 30.29515648 | 26828281563720 |
| 20015998341294 | 59.92728424 | 30.33834457 | 26828285359542 |
| 20015998341295 | 59.95634842 | 30.30995560 | 26828280189995 |
+-----+-----+-----+-----+
5 rows in set (0.00 sec)
    
```

Рис. 7. Точки доступа в базе данных

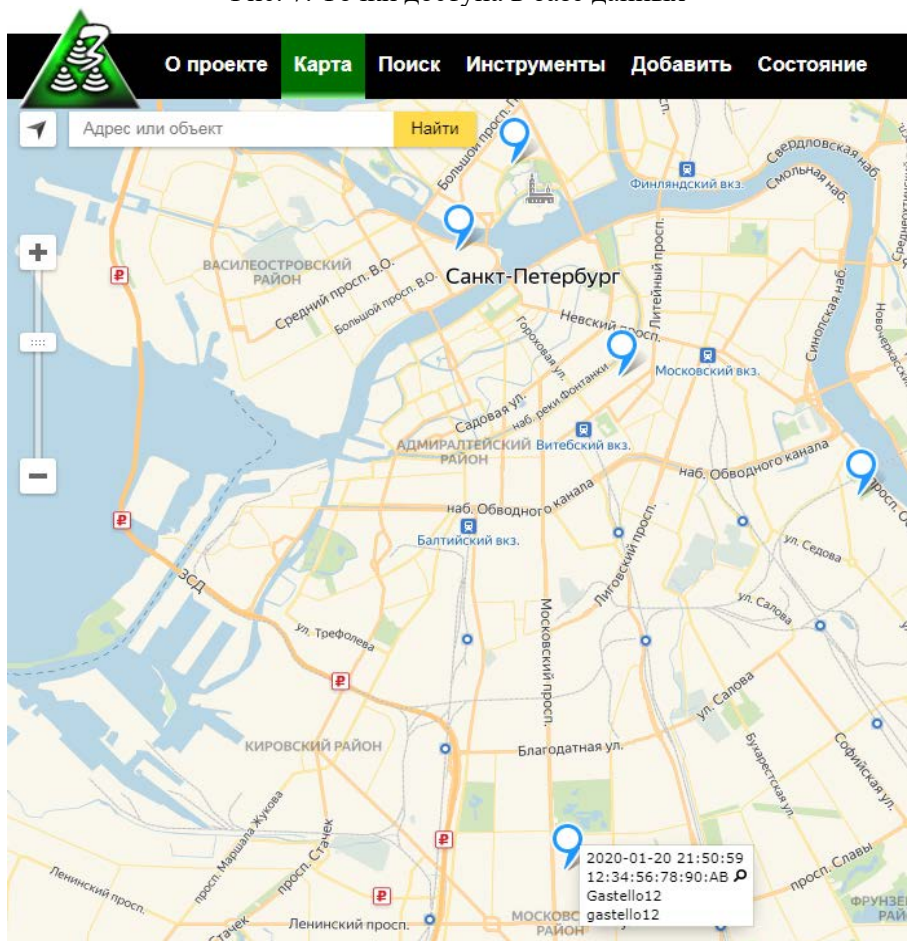


Рис. 8. Точки доступа на карте

Литература

1. Будущее технологии Wi-Fi. Д. Денисов, 25.09.2017, [Электронный ресурс]. Режим доступа: <https://nag.ru/articles/article/32434/buduschee-tehnologii-wi-fi.html> (дата обращения: 03.02.2020).
2. Безопасность сетей 802.11. основные угрозы. Бандурян А., 24.09.2012, [Электронный ресурс]. Режим доступа: <https://habr.com/ru/post/151126/> (дата обращения: 03.02.2020).
3. Веревкин С.А., Ананченко И.В. Уязвимости современных беспроводных Wi-Fi сетей //В сборнике: Инновационные научные исследования: теория, методология, практика. Сборник статей XII Международной научно-практической конференции: в 2 частях. 2018. С. 32-35.
4. Router Scan v2.60 Beta by Stas'M (build 04.02.2019). 04.02.2019, [Электронный ресурс]. Режим доступа: <http://stascorp.com/load/1-1-0-56> (дата обращения: 03.02.2020).
5. Сервис 3WiFi: пароли от точек доступа Wi-Fi. [Электронный ресурс]. Режим доступа: <https://hackware.ru/?p=4474> (дата обращения: 03.02.2020).



Коробова Полина Ивановна

Год рождения: 1997
Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы № m41211,
направление подготовки: 09.04.02 – Речевые
информационные системы,
e-mail: korobovapolina@list.ru



Мамаев Никита Константинович

Год рождения: 1996
Университет ИТМО,
факультет информационных
технологий и программирования,
студент группы № m42211,
направление подготовки: 09.04.02 – Речевые
информационные системы,
E-mail: nikita.mamaev1@gmail.com



Махныткина Олеся Владимировна

Год рождения: 1982
Университет ИТМО,
факультет информационных
технологий и программирования,
к.т.н., доцент,
e-mail: makhnytkina@corp.ifmo.ru

УДК 004

**МЕТОДЫ ИЗВЛЕЧЕНИЯ ТЕРМИНОВ
ИЗ ТЕКСТОВ НА РУССКОМ ЯЗЫКЕ**

П.И. Коробова, Н.К. Мамаев

Научный руководитель – к.т.н. О.В. Махныткина

Работа выполнена в рамках НИР №619423 «Разработка виртуального диалогового помощника для поддержки проведения дистанционного экзамена на основе аргументационного подхода и глубокого машинного обучения».

Аннотация

Данная статья посвящена решению задачи извлечения терминов из текстов на русском языке. Рассмотрены этапы предварительной обработки, описаны и исследованы алгоритмы и методы решения данной задачи, проведен сравнительный анализ библиотек для извлечения терминов. В качестве набора данных в работе используются небольшие лекции, так как именно в учебных материалах содержится большое количество терминов.

Ключевые слова

Извлечение терминов, обработка естественного языка, предварительная обработка текста.

Извлечение терминов является важной задачей, которая используется в областях знаний, таких как интеллектуальный анализ текста, информационный поиск и обработка естественного языка. Под термином в данной работе будет пониматься слово или словосочетание, которое в совокупности с другими терминами будет представлять текст [1, 2]. Выделение терминов также может использоваться для создания и развития терминологических ресурсов, а также для обработки документов [4]. Задача автоматического определения ключевых терминов представляет собой необходимый этап обработки текста для решения различного рода задач, таких как автоматический информационный поиск, классификация и др.

Большой проблемой обработки русского языка является недостаточность необходимых словарей, наборов данных и программного обеспечения, именно поэтому, работа с русскоязычными текстами для извлечения терминов является актуальной.

Целью данной работы является проведение сравнительного анализа алгоритмов и библиотек для извлечения терминов на русском языке. Для достижения цели была проведена предобработка текста, проанализированы методы к извлечению терминов на русском языке и протестированы программные библиотеки на Python для извлечения терминов из конспектов лекций по дисциплине «Машинное обучение» на русском языке.

Базовая предобработка как правило включает выявление и удаление списка стоп слов, токенизацию, лемматизацию, также дополнительно можно использовать исправление орфографических ошибок. Список стоп слов должен быть качественно составлен, чтобы исключить слова, которые не относятся к тематике текста. Токенизация используется для того, чтобы разбить текст на более мелкие части, токены. К токенам относятся как слова, так и знаки препинания. Для токенизации могут быть использованы различные функции и инструменты, такие как регулярные выражения или набор функций токенизации в библиотеках Python. Лемматизация используется для приведения слова к его базовой форме. Для лемматизации используется морфологический анализатор *py morphology2*.

После предварительной обработки применяется алгоритм для извлечения информации. Ниже рассмотрим подходы и библиотеки, которые применяются для извлечения терминов:

1) графоориентированный подход;

Выполняется построение взвешенного неориентированного графа в виде $(V; E)$, где V - множество слов, E - множество связей между ними. В качестве V можно принять множество всех уникальных лемм исходного текста. Множество E строится путём последовательного сканирования текста заданным окном из слов. Строится взвешенный неориентированный граф $(V; E)$ на основе исходного текста. Ребро в графе соответствует наличию семантической информации между двумя словами. Приблизённо вычисляется PageRank для построенного графа. Алгоритм PageRank чаще используется в поисковых системах, но также применяется в информационном поиске для задачи ранжирования. Вершины с наибольшим значением веса TextRank считаются терминами. TextRank является приложением алгоритма PageRank, который применяется к задачам обработки текста, в данном случае для извлечения терминов. Для реализации данного метода была использована упрощённая графовая модель на основе библиотеки *Summa* метода TextRank, позволяющая осуществлять обобщение текста, извлечение ключевых слов и терминов [3];

2) подход на основе правил.

Сначала составляется описание шаблонов, которые потребуются для извлечения объектов из текста. Затем шаблон применяется к документу, после чего выявляются кусочки текста, которые подходят под данное правило. В настоящее время также существуют библиотеки, которые содержат уже готовые правила для определенного рода задач. Для исследования данного подхода была использована библиотека *rutermextract*, позволяющая извлекать ключевые слова на основе заранее заданных правил.[5] На данный момент это единственный возможный вариант для русского языка, так как не существует открытого синтаксического корпуса, который можно использовать для обучения синтаксических моделей. Данная библиотека имеет ряд ограничений. Например, сейчас не извлекаются ключевые слова, содержащие предлоги («обучение без учителя», «обучение с учителем»). Также некоторые извлеченные фразы могут не являться на самом деле ключевыми. Размера текста не всегда бывает достаточно для того, чтобы отличить важные для текста слова от неважных, основываясь только на количестве употреблений. Поэтому необходимо использовать сторонние модели (например, *tf-idf*) для определения важности ключевых слов;

3) подход на основе машинного обучения.

Подходу, который основан на машинном обучении необходим большой объем входных данных. Данные должны быть размечены, где для каждого слова должна быть размечена морфология, синтаксис, и также связи между словами. К достоинствам данного подхода можно отнести минимум ручной работы, кроме создания размеченного корпуса. Так же то, что такую систему легко перенастроить и нет необходимости создания правил и онтологий. Для исследования возможностей данного подхода была использована библиотека *Gensim*. С ее помощью можно обрабатывать тексты, работать с векторными моделями слов (такими как *Word2Vec*, *FastText* и т. д.) и создавать тематические модели текстов.

Тематическое моделирование – это метод извлечения основных тем, которым посвящен обрабатываемый текст. В пакете *Gensim* реализованы основные алгоритмы тематического моделирования: скрытое распределение Дирихле (*LDA*) и скрытое семантическое индексирование (*LSI*). Многие алгоритмы доступны и в других пакетах, таких как *scikit*, *R* и т. д. но скорость обработки, качество результата и оценки тематических моделей не имеют аналогов с *gensim*. Еще одно существенное преимущество *gensim*: он позволяет обрабатывать большие текстовые файлы, не загружая весь файл в память.

В качестве набора данных были использованы лекции по дисциплине «Машинное обучение». Была произведена разметка текста, выделены термины. Рассмотрим пример термина из текста: «Естественный язык (ЕЯ)– язык, используемый для общения людей и не созданный целенаправленно». В представленном предложении словосочетание «естественный язык» является термином.

В качестве оценки эффективности работы используются: полнота(*recall*), точность(*precision*), F-мера(*f1-score*). Оценивание производится на основе следующих показателей: *TN* – истинноотрицательное решение, *TP* – истинно-положительное решение *FP* – ложноположительное решение, *FN* – ложноотрицательное решение.

F-мера основывается на результатах оценки точности и полноты. В ходе тестирования вышеописанных подходов и библиотек были получены результаты, представленные в таблице.

Стоит отметить, что библиотека *Summa* показала хороший результат. С помощью нее было найдено наибольшее количество терминов, но кандидатов в термины было определено ею много, большая часть которых составляют ложные термины.

Результаты тестирования

Библиотека	Precision	Recall	F1-score
Rutermextract	0,134	0,9	0,23
Gensim	0,269	0,566	0,35
Summa	0,306	0,633	0,41

В работе были рассмотрены методы предварительной обработки текстов и проведен сравнительный анализ библиотек, которые используют различные методы для извлечения терминов, такие как: метод на основе правил, метод на основе машинного обучения и метод на основе графовых моделей. Для улучшения качества результатов планируется составление дополнительных правил для извлечения терминов, использование ансамблей моделей и использование более широкого спектра методов предварительной обработки текста.

Литература

1. Красавина В.Д., Мирзагитова А.Р. Оптимизация поиска в системе LeadScanner с помощью автоматического выделения ключевых слов и словосочетаний // Труды международной конференции «Корпусная лингвистика–2015». СПб. 2015. С. 296–306.
2. Москвина А.Д., Митрофанова О.А., Ерофеева А.Р., Харabet Я.К. Автоматическое выделение ключевых слов и словосочетаний из русскоязычных корпусов текстов с помощью алгоритма RAKE // Труды международной конференции «Корпусная лингвистика–2017». СПб. 2017. С. 268–277.
3. Barrios F., López F., Argerich L., Wachenchauzer, R.: "Variations of the Similarity Function of TextRank for Automated Summarization". Anales de las 44JAIIO. Jornadas Argentinas de Informática, Argentine Symposium on Artificial Intelligence, 2015.
4. Roberto Ortiz, David Pinto, Mireya Tovar , Н'ector Jim'enez-Salazar . BUAP: An Unsupervised Approach to Automatic Keyphrase Extraction from Scientific Articles, 2010.
5. Описание библиотеки rutermextract [Электронный ресурс]. Режим доступа: <https://pypi.org/project/rutermextract/>. Дата обр. 10.12 2019.



Котов Денис Олегович

Год рождения: 1987
Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы №М41211,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: preductor@gmail.com



Шуранов Евгений Витальевич

Год рождения: 1980
Университет ИТМО,
факультет информационных технологий
и программирования,
к.т.н., доцент,
e-mail: shuranov@speechpro.com

УДК 004.93

**ПРИМЕНЕНИЕ ТЕХНОЛОГИИ РЕЛЕЙБЛИНГА
К РАЗМЕТКЕ БАЗ ДАННЫХ**

Д.О. Котов

Научный руководитель – к.т.н., доцент Е.В. Шуранов

Работа выполнена в рамках темы НИР «Использование машинного обучения для детектирования эмоций».

Аннотация

В работе рассмотрено применение технологии обучения с частичным привлечением учителя в приложении к разметке баз данных эмоций человека.

Ключевые слова

Машинное обучение (ml), нейронные сети (nn), релейблинг.

Последние несколько лет в разработке систем человеко-машинного взаимодействия стали активно применяться различные технологии машинного обучения и нейросетевые подходы в частности. Это позволило значительно повысить эффективность таких систем и подступиться к таким сложным и доселе невыполнимым задачам, как детектирование объектов, синтез речи и распознавание эмоций человека.

Распознавание эмоций – это очень молодая область исследования. И в силу этого обстоятельства пока не существует единого стандарта сбора данных для дальнейшего обучения моделей и проведения различных экспериментов. Это касается и набора эмоций, содержащихся в базе, и способа разметки оценщиками, и даже способа и качества сбора самих данных. Это приводит к ситуации, когда работать с такими базами довольно сложно - не представляется возможным их объединять или делать оценку модели, обученной на данных одной базы, при помощи инференса на другой, поскольку распределения данных в этих базах могут слишком сильно отличаться [1-4].

Именно поэтому была предпринята попытка применения технологии обучения с частичным привлечением учителя. Обучение с частичным привлечением учителя (semi-supervised learning) - способ машинного обучения, разновидность обучения с учителем, которое также использует неразмеченные данные для тренировки — обычно небольшое количество размеченных данных и большое количество неразмеченных данных.

Суть ее применения в данном контексте заключается в том, что имея некоторую эталонную базу данных, то есть базу, качеству данных которой и качеству разметки мы доверяем, мы можем разметить большое количество данных с некачественной или некорректной разметкой. Процесс обучения представлен на рис. 1.

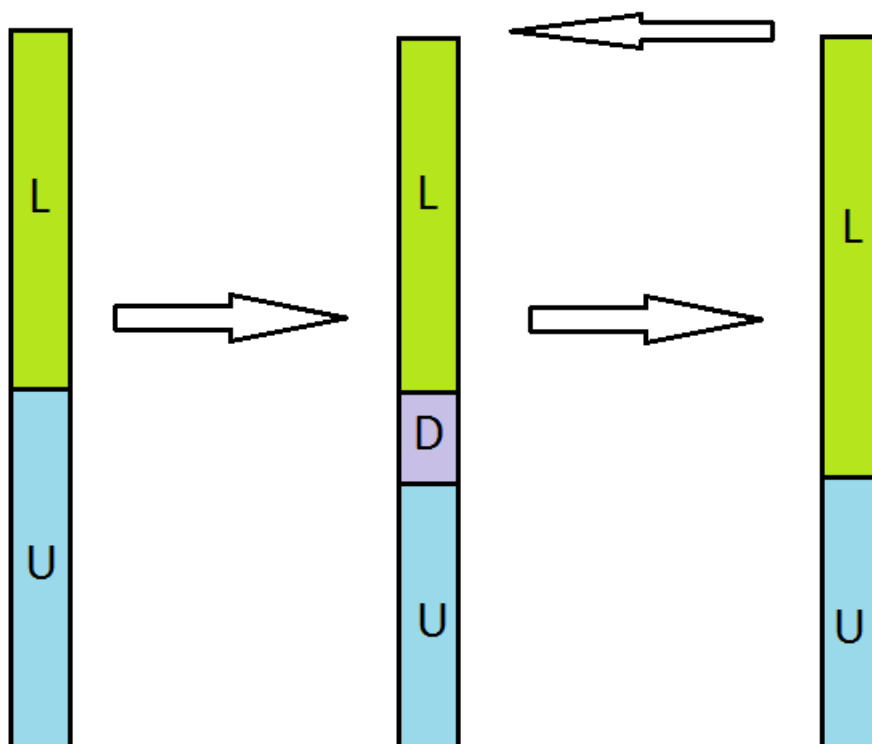


Рис. 1. Процесс обучения с частичным привлечением учителя

Обозначим за L данные проверенной базы, которые мы берем за основу обучения модели, за U данные, которые надо разметить. Тогда обучив модель машинного обучения на данных из базы L на первом шаге мы делаем прогноз на базе U , далее получив некоторое распределение вероятностей по классам, отсекаем те данные, в разметке которых наш алгоритм наиболее уверен, с вероятностью выше некоторого порогового значения, которое мы задаем как гиперпараметр. Эти данные обозначим как D . Далее добавляем их к выборке L и повторяем обучение уже на новых данных. Это повторяется до тех пор пока уверенность прогноза не станет ниже порогового значения.

Описание эксперимента

В качестве эталонной базы была взята база IEMOCAP, составленная Стэнфордским университетом. Она представляет собой отыгрыши эмоциональных сцен актерами. Запись производилась как в аудио, так и видео формате, также прилагается текстовая аннотация. Разметка производилась перекрестным способом, это означает, что каждый файл просматривало не менее трех человек и в финальную версию попадали только те данные, согласованность в оценке которых среди разметчиков была выше порогового значения. На эту базу часто цитируются и в научных кругах, что лишний раз подчеркивает ее высокое качество. В качестве базы, разметку которой

переразмечали была выбрана база OMG, которая плохо себя показала при обучении различных моделей и при ручной проверке.

При обучении использовалась тренировочная выборка базы IEMOCAP, но для чистоты эксперимента инференс проводился не только на тестовой выборке той же базы, но и на тестовой выборке базы MOSEI. Эта база представляет собой набор видео данных с сайта Youtube.com, размеченные при помощи аутсорсинговой платформы Amazon Mechanical Turk. В отличие от базы IEMOCAP, данные базы MOSEI лишены некоторых специфичных особенностей, связанных с тем, что все записи в IEMOCAP сделаны ограниченным количеством людей, то есть предполагается, что и набор голосов со свойственными ему характеристиками будет ограниченным. За ключевую метрику качества была принята макро f-мера. Точность и полнота инференса тоже приведены на рис. 2.

Результаты на IEMOCAP test				Результаты на MOSEI test			
номер итерации	точность	полнота	f-мера	номер итерации	точность	полнота	f-мера
до первой итерации	0.844	0.816	0.823	до первой итерации	0.352	0.403	0.265
1	0.844	0.816	0.823	1	0.352	0.403	0.265
2	0.844	0.816	0.823	2	0.358	0.408	0.266
3	0.844	0.816	0.823	3	0.352	0.402	0.263
4	0.844	0.818	0.825	4	0.347	0.404	0.258
5	0.848	0.826	0.832	5	0.354	0.411	0.259
6	0.834	0.824	0.826	6	0.351	0.405	0.251
7	0.816	0.805	0.807	7	0.357	0.410	0.255
8	0.812	0.797	0.801	8	0.356	0.397	0.253
9	0.812	0.798	0.801	9	0.362	0.414	0.259
10	0.826	0.811	0.816	10	0.363	0.409	0.258

Рис. 2. Результаты эксперимента по релейблингу базы OMG

По полученным результатам можно сказать, что оценка модели, обученной на тренировочной выборке IEMOCAP при помощи тестовой выборки той же базы дает значительно более высокие показатели f-меры, чем модель, обученная на тренировочной выборке базы IEMOCAP с инференсом на базе MOSEI. Что подтверждает предположение о том, что ограниченный набор голосов обучающей и тестовой выборки приводит к переобучению модели.

Для наглядности продемонстрирую те же результаты в виде графика на рис. 3.

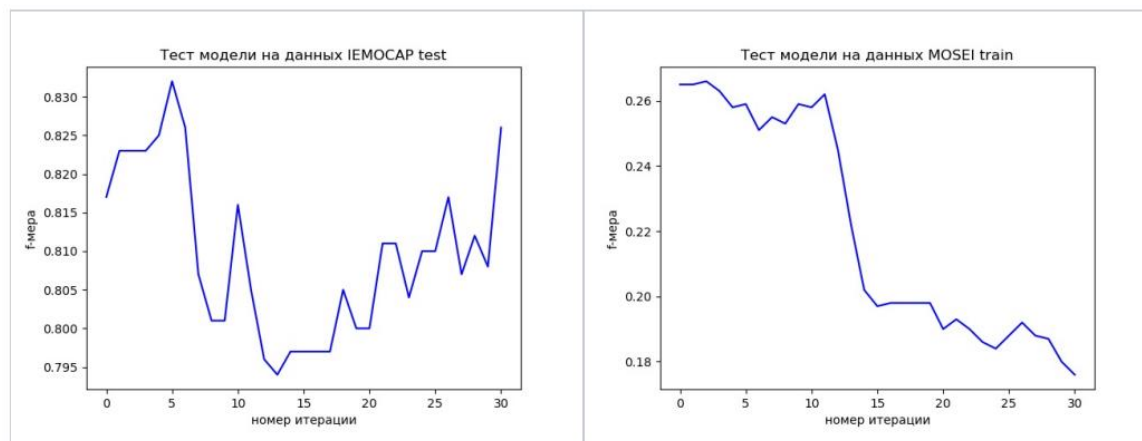


Рис. 3. Графики зависимости f-меры от итерации

Касаемо целей текущего эксперимента наблюдается небольшой прирост показателей метрики на первых пяти итерациях, далее наблюдается ее падение. Это объясняется тем, что проблемы обучающего множества только множатся, статистическая погрешность запоминается моделью. Но главная цель - получение более устойчивой модели, достигнута, так как модель видит больше данных.

Литература

1. AudioSet, URL: <https://research.google.com/audioset/> (дата обращения 27.02.2020);
2. IEMOCAP[2] (The Interactive Emotional Dyadic Motion Capture) URL: <https://sail.usc.edu/iemocap/> (дата обращения 27.02.2020).
3. CMU-MOSEI или MOSEI [3] (CMU Multimodal Opinion Sentiment and Emotion Intensity) URL: <https://github.com/A2Zadeh/CMU-MultimodalSDK> (дата обращения 27.02.2020).
4. OMG [4] (Russian Multimodal Corpus of Dyadic Interaction for Affective Computing) URL: <https://neurodatalab.com/science/publications/ramas-russian-multimodal-corpus-of-dyadic-interaction-for-affective-computing-for-speccom-2018/> / (дата обращения 27.02.2020).



Куан Чонг Тхе

Год рождения: 1984
Университет ИТМО,
факультет информационных технологий
и программирования,
аспирант,
направление подготовки: 09.06.01 – Информатика
и вычислительная техника,
e-mail: quantrongthe@itmo.ru



Столбов Михаил Борисович

Год рождения: 1952
Университет ИТМО,
факультет информационных
технологий и программирования,
к.т.н, доцент,
e-mail: stolbov@speechpro.com

УДК 621.391.8

**ИССЛЕДОВАНИЕ АЛГОРИТМА MVDR ДЛЯ ДВУХЭЛЕМЕНТНЫХ
МИКРОФОННЫХ РЕШЕТОК С ПРОДОЛЬНОЙ И ПОПЕРЕЧНОЙ
АРХИТЕКТУРОЙ В УСЛОВИЯХ ПОМЕЩЕНИЯ**

Куан Чонг Тхе

Научный руководитель – к.т.н., доцент М.Б. Столбов

Аннотация

В статье, рассмотрен эксперимент по подавлению широкополосного шума аудиоколонки в условиях замкнутого помещения при использовании адаптивного алгоритма минимума дисперсии шума (MVDR) для двухэлементных микрофонных решеток (MP2). Сравнялось подавление шума при использовании MP2 с поперечной и продольной архитектурой.

В ходе экспериментов было определено максимально допустимое значение коэффициента сглаживания. Эксперименты показали приблизительно одинаковую для обеих архитектур степень подавления когерентного шума и большую эффективность продольной архитектуры при подавлении диффузного шума.

Ключевые слова

Двухэлементная микрофонная решетка, алгоритм минимума дисперсии шума, подавление шума, поперечная архитектура, продольная архитектура.

Введение

Выделение речи в шумовой акустической обстановке представляется важную задачу для многих приложений. Микрофонные решетки (MP) используют пространственную информацию для выделения полезного сигнала [1]. Применение MP позволяет ослаблять когерентные помехи, диффузный шум и реверберацию. Двухэлементная решетка (MP2) [2-4] является самым простым типом MP. Благодаря простоте и компактности MP2 могут быть использованы во многих приложениях.

Принято различать поперечную и продольную архитектуру МР. Целью эксперимента было сравнение алгоритма MVDR для МР2 с поперечной и продольной архитектурой.

Двухэлементные микрофонные решетки состоят из двух идентичных ненаправленных микрофонов, разнесенных на расстояние d . На рис. 1 представлена схема обработки сигналов МР2 в частотной области.

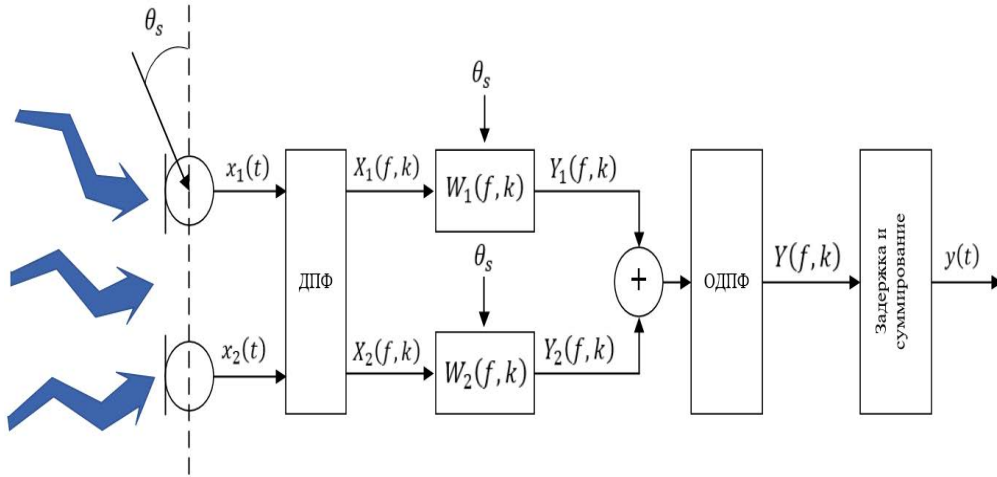


Рис. 1. Схема обработки сигналов МР2 в частотной области

Рассмотрим схему обработки сигналов МР2 в частотной области. Сигналы $x_1(t)$ и $x_2(t)$ с микрофонов 1 и 2, сегментируются на пересекающиеся временные кадры, взвешиваются временным окном и преобразуются с помощью дискретного преобразования Фурье (ДПФ):

$$\begin{aligned} \text{ДПФ} \quad \text{ДПФ} \\ x_1(t) &\rightarrow X_1(f, k), \\ x_2(t) &\rightarrow X_2(f, k), \end{aligned}$$

где f, k – индексы частоты и номера кадра, соответственно.

Сигналы на каждой частоте взвешиваются с использованием комплексных весовых коэффициентов $W_1(f)$, $W_2(f)$, после чего суммируются:

$$Y(f, k) = W_1(f) X_1(f, k) + W_2(f) X_2(f, k) = W^H(f) X(f, k)$$

Полученный сигнал преобразуется во временную область с использованием обратного дискретного преобразования Фурье (ОДПФ) и алгоритма задержки и суммирования.

В зависимости от ориентации МР относительно источника целевого сигнала различают два типа архитектуры: поперечная и продольная МР. В случае поперечной архитектуры (broadside, BS) источник целевого сигнала расположен вблизи направления нормали к оси МР2 ($\theta_s = 90^\circ$). В случае продольной архитектуры (endfire, EF) источник целевого сигнала расположен вблизи направления оси МР2 ($\theta_s = 0^\circ$).

Алгоритм минимум дисперсии шума

Направление прихода целевого сигнала и шума задаются углами θ_s от оси, проходящей через микрофоны. В представлении кратковременного преобразования Фурье сигнал $S(f, k)$ акустического целевого источника с направления θ_s и

когерентный широкополосный шум $V(f, k)$ формирует на микрофонах вектор сигналов:

$$X(f, k) = S(f, k)D_s(f) + V(f, k),$$

где f, k – индекс частоты и номера кадра, $X(f, k) = [X_1(f, k) \ X_2(f, k)]^T$ $D_s(f) = [e^{+j\Phi_s} \ e^{-j\Phi_s}]^T$ – вектор фазовых сдвигов сигналов микрофонов относительно центральной точки между ними, и $()^T$ – символ транспонирования, $\Phi_s(f)$ – фазовые сдвиги:

$$\Phi_s(f) = \pi d \cos(\theta_s) / \lambda = \pi f \tau_0 \cos(\theta_s)$$

где d – расстояние между микрофонами, c – скорость звука в воздухе, $\tau_0 = d/c$ время прохождения звука по оси МР2 между микрофонами. θ_s – направление прихода сигнала.

Алгоритм MVDR основывается на ограниченной критерии неискаженного полезного сигнала с целевого направления θ_s и минимума выходной мощности шумов. Весовые коэффициенты являются решением следующей оптимизационной задачи [1]:

$$\min_W E\{|Y(f, k)|^2\} = \min_W [W^H(f)P_{vv}(f)W(f)]$$

при условии $W^H(f)D_s(f) = 1$

где $E\{.\}$ – символ математического ожидания, $P_{vv}(f)$ – ковариационная матрица спектров шума на микрофонах:

$$P_{vv}(f) = E\{V(f, k)V^H(f, k)\}$$

Решение оптимизационной задачи приводит к следующему соотношению для вектора оптимальных весов [1]:

$$W_0(f) = \frac{P_{vv}^{-1}(f)d_s(f)}{d_s^H(f)P_{vv}^{-1}(f)d_s(f)}$$

При решении практических задач ковариационная матрица спектров шума оценивается по текущим данным с использованием алгоритма экспоненциального сглаживания:

$$P_{X_i X_j}(f, k) = (1 - \beta)\alpha \times P_{X_i X_j}(f, k - 1) + \beta \times X_i^*(f, k)X_j(f, k), \quad i, j = 1, 2$$

где β коэффициент сглаживания.

Веса фильтра MVDR вычисляются на основе текущей оценки матрицы:

$$W(f, k) = \frac{P_{VV}^{-1}(f, k)D_s(f)}{D_s^H(f)P_{VV}^{-1}(f, k)D_s(f)}$$

Эксперименты

Эксперименты проводились в помещении $5,2 \times 8,4 \times 2,8$ м. Расстояние между двумя микрофонами составляло 5 см. Широкополосный шум транслировался через аудиокolonку. Схема эксперимента представлена на рис. 2. Исследовалась степень подавления шума в зависимости от архитектуры МР2 и расстояния между МР2 и аудиокolonкой. Расстояние составляло 0,5 м, 1 м, 2 м, 4 м и 6 м.

При реализации алгоритма MVDR, было определено значение коэффициента сглаживания $\beta \leq 0,1$. Степень подавления шума для двух разных архитектур MP2 представлена в таблице.

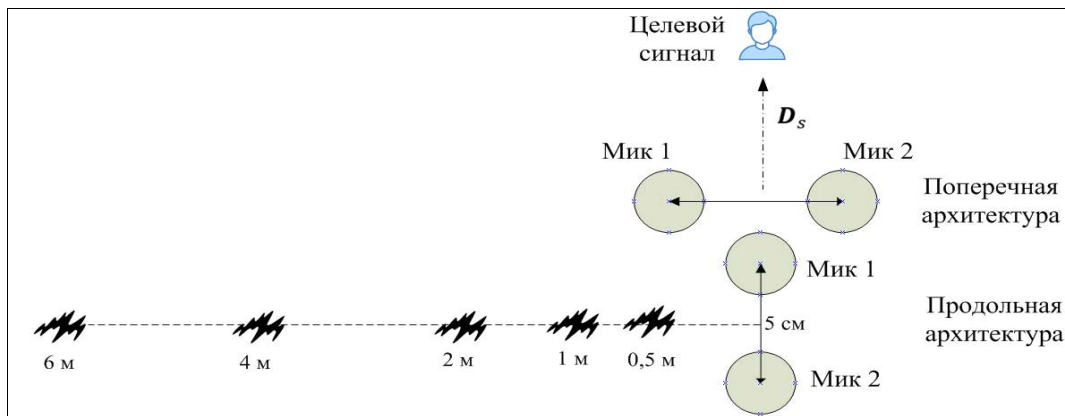


Рис. 2. Схема эксперимента

Таблица

Степень шумоподавления (дБ) алгоритмом MVDR для двух архитектур

Расстояние (м)	0,5	1	2	4	6
Продольная архитектура	11	9,5	9	7	10
Поперечная архитектура	10,3	11,1	8,2	7,4	4,4

Выводы

По мере увеличения расстояния между колонкой и MP2 возрастает доля диффузного шума при этом алгоритм MVDR позволяет подавить как когерентный, так и диффузный шум. Эксперименты показали приблизительно одинаковую для обеих архитектур степень подавления когерентного шума и большую эффективность продольной архитектуры при подавлении диффузного шума. Практическая значимость данной работы заключается в уточнении возможности подавления широкополосных когерентных и диффузных шумов при обработке сигналов MP2 с использованием алгоритма MVDR.

Литература

1. Brandstein M. and Ward D. (Eds.). *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer. 2001.
2. Lockwood M. et al. Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms. // *J. Acoust. Soc. Am.* 115 (1), 2004. P. 379-391.
3. Столбов М.Б., Куан Ч.Т. Исследование двухканального алгоритма MVDR для выделения речи из когерентного шума // *Научно-технический вестник информационных технологий, механики и оптики*. 2019. Т. 19. № 1. С. 180–183. doi: 10.17586/2226-1494-2019-19-1-180-183.
4. Куан Ч.Т., Столбов М.Б. Алгоритм формирования диаграммы направленности микрофонных решеток с продольной архитектурой для выделения речевых сигналов // *XLVII Научной и учебно-методической конференции университета ИТМО, С.-Петербург*. 2017.



Краснова Ольга Александровна

Год рождения: 1996
Университет ИТМО,
Институт Дизайна и Урбанистики,
студент группы С42012,
направление подготовки: 27.04.07 – Наукоемкие
технологии и экономика инноваций,
e-mail: olia.krasnova26.04.1996@gmail.com



Лымарь Варвара Владимировна

Год рождения: 1989
Университет ИТМО,
Институт Дизайна и Урбанистики,
преподаватель практики, аспирант,
направление аспирантуры: 05.13.10 – Управление
в социальных и экономических системах,
e-mail: lymar.varvara@gmail.com

УДК 711.55

**МЕТОДЫ ОБРАБОТКИ РАЗНОРОДНЫХ ГОРОДСКИХ ДАННЫХ,
ПРИМЕНИМЫЕ К ИНДУСТРИАЛЬНЫМ ТЕРРИТОРИЯМ**

В.В. Лымарь, О.А. Краснова

Научный руководитель - преподаватель практики В.В. Лымарь

Аннотация

В работе была поставлена гипотеза о сложности развития территорий «серых поясов» в постсоветских городах. Для ее проверки было проведено исследование «серого пояса» Красноярска, и была выявлена неравномерность расположения организаций, которая описывается моделью А.А. Высоковского. Были обнаружены территории со схожими параметрами в разных частях города, что говорит о возможности выделения морфологических типов в рамках «серого пояса». Эти типы могут быть выявлены с использованием современных средств параметризации.

Ключевые слова

«Серый пояс», морфология городской застройки, Spacematrix, MXI model.

В России существует проблема, связанная с «Серыми поясами», описанная в исследовании Всемирного банка [1]. Из-за отсутствия земельных рынков в СССР возникли крупные промышленные территории в непосредственной близости от центральных городских районов. Эта проблема характерна для городов всех стран, где, перерыв в существовании земельных рынков составил порядка 70 лет.

В научной литературе была выдвинута гипотеза о неравномерности и сложности территорий «Серого пояса», что препятствует их комплексному развитию [2, 3]. Для проверки этой гипотезы было выполнено исследование с использованием данных обо всех организациях, полученных из рекламно-справочной системы 2ГИС. Полученный массив данных содержал машиночитаемую информацию о 42,3 тысячах организаций с координатами и профилем деятельности каждой, представляющей собой категорию.

Эти категории были классифицированы в соответствии с такими параметрами как: среднее количество времени, проводимое посетителем в компании, с привязкой рабочему дню (менее 1 часа, 1-4 часов, 4-8 часов, более 8 часов); количество посетителей на одного сотрудника; а также наличие производства. В результате было получено 13 категорий компаний: кафе, магазины, офисы с посетителями, медицинские учреждения, курсы, досуговые учреждения, офисы, образовательные учреждения, мастерские, склады-заводы, транспортные компании, гостиницы, приюты. В дальнейшем планируется уточнение списка организаций, в том числе на основании общероссийского классификатора экономической деятельности (ОКВЭД).

Далее на основании совместного визуального анализа карт города из открытых источников и спутниковых снимков были определены границы «серого пояса» Красноярска, включающего промышленные предприятия, в том числе преобразованные под иное назначение; объекты инженерной инфраструктуры, зоны железнодорожного транспорта с примыкающей к ним инфраструктурой, и стоянки грузового транспорта.

При совмещении границ территорий и полученных данных было выявлено неравномерное расположение организаций в пределах «Серого пояса» (рисунок). Была обнаружена обратная пропорциональная зависимость плотности организаций от площади земельных участков Росреестра (Федеральная служба государственной регистрации, кадастра и картографии). В пределах наиболее крупных предприятий города плотность организаций была наименьшей. Территории с наибольшей плотностью организаций располагались, как правило, вблизи крупных городских магистралей на границе с жилыми зонами, территории с меньшей плотностью располагались на перекрестках внутри промышленных территорий.

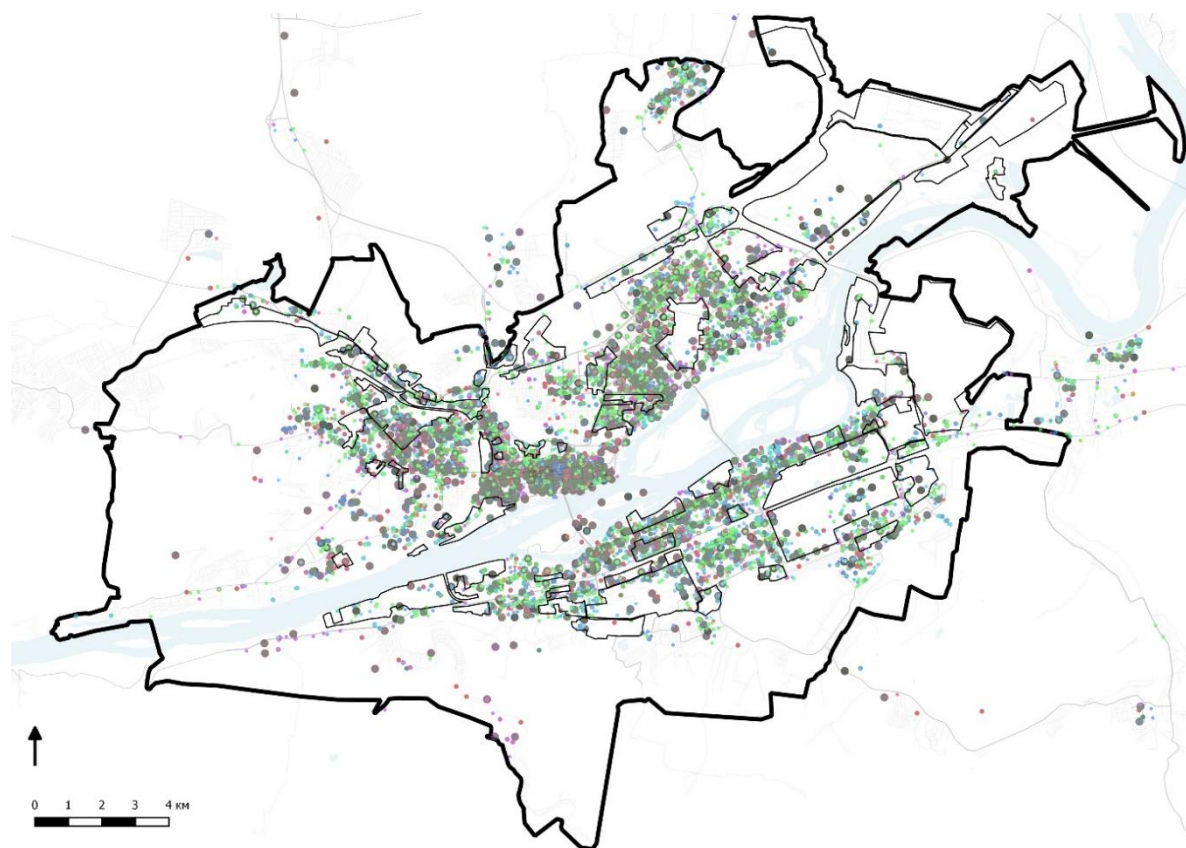


Рисунок. Плотность организаций в пределах "Серого пояса".
Толстыми линиями обозначены границы города, тонкими – границы "серого пояса",
точками – организации.

Полученное распределение организаций может быть описано несколькими пространственными моделями. В частности, были рассмотрены такие модели как модель неравномерного районирования А.А. Высоковского и модель узлового района по Б. Б. Родмана [4]. Модель неравномерного районирования А.А.Высоковского [5] наиболее точно описывает полученные результаты.

Основной задачей модели является определение центров коммерческой активности и тяготеющих к ним монофункциональных районов. Центральные места являются наиболее значимыми и привлекательными для горожан, так как они характеризуются разнообразием видов деятельности, высокой интенсивностью протекающих процессов и большой степенью публичности поведения людей. Монофункциональные районы связаны с индивидуальным пребыванием субъектов в городском пространстве. Ядра активности и монофункциональные районы в совокупности формируют «пространственные единицы».

Для «серого пояса» Красноярска в качестве многофункциональных ядер выступают территории с большой концентрацией организаций, ориентированных на прием посетителей, и мелким размером земельных участков. Примером монофункциональных территорий можно назвать крупные промышленные предприятия с участками значительной площади.

Также было выявлено, что для территорий с одинаковой плотностью организаций характерно разное сочетание их категорий. К примеру, территории с высокой плотностью представляли собой кварталы, занятые преимущественно магазинами (рынки и торговые центры), магазинами в сочетании с мастерскими (строительные базы и автосервисы), а также всеми видами организаций (офисные центры). Схожее сочетание категорий организаций в разных частях города было визуально идентичным, при этом для территорий с разными сочетаниями организаций был характерен разный визуальный ландшафт. Подобные закономерности описываются понятием морфологии городской застройки.

Городская морфология («urban morphology») - изучение структуры города и истории ее изменения. Морфология городов относится к «изучению физической ткани городской структуры, а также людей и процессов, формирующих ее» [6]. Термин также можно понимать, как изучение физических и пространственных характеристик всей городской структуры. К различным морфологическим типам могут быть применены различные стратегии развития. Морфологические типы могут быть выделены с использованием средств параметризации, таких, как методика Spacematrix, модель смешанности функций (MXI) и плотности земельных участков [7]. Далее будет изучена возможность применения этих методик к территориям «Серого пояса».

Методика Spacematrix устанавливает корреляцию между различными показателями застройки, такими, как плотность, компактность, соотношение открытого пространства к застроенному, средняя этажность и плотность улично-дорожной сети. Она разработана для выявления морфологических типов в жилой и многофункциональной застройке, где использование зданий зависит от их физических параметров. При этом для «Серого пояса» характерна ситуация, при которой старые корпуса промышленных предприятий, занимающие значительную площадь, изменяют свое функциональное назначение, либо разделяются между множеством собственников. Это приводит к таким последствиям, как превращение территории крупного промышленного предприятия в локальный подцентр. Поэтому выделение морфологических типов на основании физических параметров зданий является некорректным. Для территорий «Серого пояса» к определяющим параметрам относятся плотность и компактность расположения организаций. Для использования методики Spacematrix в данном исследовании необходимы такие параметры, как количество организаций на гектар и среднее расстояние между ними.

Модель Spacematrix успешно дополняется моделью смешанности функций (MXI) [7]. Она отображает степень смешения основных функций: деловой, жилой и рекреационной, выраженная в процентном соотношении. При 33% каждой функции территория считается абсолютно многофункциональной.

Для территорий «Серого пояса» Красноярска на основании анализа тепловых карт всех категорий организаций было выделено 4 основные комбинации функций: магазины, мастерские, заводы (склады-заводы и транспортные компании), и офисы (кафе, офисы с посетителями, медицинские учреждения, курсы, досуговые учреждения, офисы, образовательные учреждения, гостиницы). Таким образом, для расчета этого индекса следует считать абсолютно многофункциональной территорию с 25% каждой функции.

Средняя плотность земельных участков определяется как количество участков на гектар, и используется для уточнения получившихся морфологических типов и выявления их границ. Дополнительная проверка морфологических типов может быть проведена выполнением натурного обследования визуального ландшафта.

Таким образом, гипотеза о сложности использования территорий «Серого пояса» была подтверждена неравномерностью расположения организаций и земельных участков. Тем не менее, для территорий в разных частях города характерны схожие черты, включающее в себя сочетание организаций, размер земельных участков и визуальный ландшафт. Поэтому в дальнейшем планируется выявление морфологических типов при помощи описанных выше параметрических методик. Это позволит учитывать существующее использование территорий при разработке планов их развития.

Литература

1. Bertaud. "The Spatial Structures of Central and Eastern European cities: more European than Socialist?," М. 1999. С. 127-131.
2. Алексахина В.В., «Градостроительный аспект реорганизации производственных территорий мегаполиса на примере Москвы» Academia. Архитектура и строительство. pp. 54-62, 2010.
3. Энгель Б., Хорн К., Зинглер Ф., Рогге Н., Сендер А., «Трансформация наследия промышленного строительства Санкт-Петербурга. Варианты будущего развития южного промышленного пояса,» Проект Байкал. pp. 134-141, 2016.
4. Потапенко А.А., «Построение неравномерно-районированной модели (на примере г.Владивостока),» Информационные технологии и архитектура. pp. 403-404. 2018.
5. Высоковский А.А. Правила землепользования и застройки: руководство по разработке. Опыт введения правового зонирования в Кыргызстане, Бишкек: «Ега-Басма». 2005. 329 с.
6. [В Интернете]. Available: <http://www.urbanform.org/>.
7. M. Akkelies van Nes, «Combination of Space syntax with Spacematrix and the Mixed use index. The Rotterdam South test case». pp. 3-8, 2012.



Мамаев Никита Константинович

Год рождения: 1996
Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы М42211,
направление подготовки: 04.09.02 Информационные
системы и технологии,
e-mail: nikita.mamaev1@gmail.com



Махныткина Олеся Владимировна

Год рождения: 1982
Университет ИТМО,
факультет информационных
технологий и программирования,
к.т.н., доцент,
e-mail: makhnytkina@itmo.ru

УДК 004.89

АНАЛИЗ НОВЫХ ПОДХОДОВ К РАЗРАБОТКЕ ДИАЛОГОВЫХ СИСТЕМ НА ОСНОВЕ ПОИСКА

Н.К. Мамаев

Научный руководитель – к.т.н., доцент О.В. Махныткина

Работа выполняется при финансовой поддержке Минобрнауки России, Соглашение 14.575.21.0178 (Уникальный идентификатор проекта: RFMEFI57518X0178).

Аннотация

Актуальной проблемой в сфере автоматической обработки языка сегодня является разработка автоматической диалоговой системы, взаимодействующей с человеком посредством естественного языка. Один из подходов к построению такой системы – использование диалоговых моделей на основе глубокого машинного обучения. В этой статье представлен анализ некоторых диалоговых моделей на основе глубокого обучения, разработанных за последние несколько лет.

Ключевые слова

Диалоговая система, чатбот, обработка естественного языка.

За последнее десятилетие многие зарубежные академические работы в сфере обработки естественного языка были направлены на построение диалоговых систем. Это направление является востребованным и в России, поскольку диалоговые системы позволяют упростить процесс взаимодействия человека с базами данных, оптимизировать процесс предоставления клиентской поддержки, ускорить проведение рабочих процессов.

Современные диалоговые модели подразделяются на две категории: генеративные, составляющие ответ “с нуля”, и модели на основе поиска, которые в качестве ответа предоставляют некоторую реплику из базы, опираясь на статистику диалогов из обучающей выборки. Несмотря на существенное ограничение, модели на

основе поиска более надёжны, потому что более предсказуемы. В то же время генеративные диалоговые модели, как и другие генеративные языковые модели, часто составляют чересчур общие, однообразные, а иногда и вовсе бессмысленные ответы; а также требуют гораздо больше обучающих данных. По этой причине мы составили обзор диалоговых систем на основе поиска, а также сравнили качество, достигаемое на базе данных с соревнования Dialogue System Technology Challenge 8.

Диалоговые модели на основе поиска в свою очередь подразделяются на модели без сегментации контекста и модели с сегментацией контекста. Модели без сегментации контекста обрабатывают его как целостную единицу текста, в то время как модели с сегментацией контекста делят его на высказывания.

Одним из более простых примеров модели **без сегментации контекста** является Dual Encoder [1]. Эта модель основана на рекуррентной “сиамской” нейронной сети-кодировщике (vanilla RNN/LSTM), которая формирует векторное представление контекста и ответа, снимаемые с последнего скрытого состояния. Показатель уверенности для ответа вычисляется как матричное умножение векторов контекста и ответа слева и справа соответственно на матрицу, которая является обучаемым параметром модели. Помимо Dual Encoder на основе рекуррентной сети также были исследованы варианты на основе свёрточной сети [2], а также двунаправленной LSTM-сети [3].

Ещё одной диалоговой моделью без сегментации контекста является QA-LSTM [4].

В отличие от предыдущей, эта модель формирует векторные представления контекста и ответа независимо, для этого используются две двунаправленные LSTM-сети. Агрегация представлений также происходит иначе – для этого используется average- или max-pooling векторов с каждого скрытого состояния сети. Последним шагом является вычисление косинусного сходства для агрегированных представлений. Авторы также исследуют дополнения архитектуры: во-первых, слой свёртки над слоем кодирования, чтобы получить более “целостное” представление реплики; а также слой attention, улучшающий способность системы запоминать информацию на долгое время.

В качестве последнего примера модели без сегментации контекста мы приводим модель Enhanced Sequential Inference Model, адаптированную под диалоговую задачу [5] и изначально предназначавшуюся для решения задачи логического вывода на естественном языке (natural language inference) [6].

Эта модель включает “сиамскую” Bi-LSTM нейросеть, которая используется для получения представлений контекста и ответа аналогично тому, как это делается в Dual Encoder. Затем к ним применяется механизм перекрёстного внимания для моделирования семантических отношений между контекстом и ответом. Получившиеся представления агрегируются посредством average- или max-pooling, и для выбора одного из ответов-кандидатов применяется классификатор-перцептрон (MLP).

Одним из примеров модели **с сегментацией контекста** является модель Sequential Matching Network [7].

Каждая реплика из контекста попадает в GRU-нейросеть, с помощью которой формируются словные представления. Затем ответ пословно попарно сопоставляется с каждой репликой контекста, генерируются матрицы сопоставления, которые подаются на слой свёртки, и производится операция max-pooling. Входами к предпоследнему слою сети для каждой пары “реплика-ответ” (каждая реплика из контекста с одинаковым ответом) являются векторы, по которым вычисляется показатель уверенности для каждой пары. Таким образом, извлечение информации из каждой реплики происходит на разных уровнях, и с опорой на истинный ответ, с целью выявить наиболее полезные семантические структуры для выбора ответа.

Ещё одним примером модели с сегментацией контекста является модель Deep Attention Matching Network, вдохновлённая моделью Transformer [8].

Здесь используются кодирующие блоки из Transformer. Словные представления с блоков агрегируются в матрицы сопоставления по принципу self-attention и cross-attention между репликой и ответом. Затем матрицы конкатенируются в трёхмерный тензор, по которому производится трёхмерная свёртка и max-pooling; показатель уверенности для данной пары “контекст-ответ” вычисляется с помощью MLP.

В таб. 1 приведены значения метрики R@1 на корпусе Ubuntu [1], опубликованные авторами.

Таблица 1

Значения R@1 (in 10) на Ubuntu Dialogue Corpus, тестовая выборка

Модель	R@1
Dual Encoder, RNN	40.3%
Dual Encoder, LSTM	60.4%
QA-LSTM	–
ESIM	79.6%
SMN	72.6%
DAM	76.7%

Поскольку ESIM показывает наилучший результат на этом корпусе, в дальнейших экспериментах было решено использовать эту модель. В экспериментах мы также использовали модель Dual Encoder, используя результаты, полученные с её помощью, как бейзлайн.

Корпус, предоставляемый в рамках соревнования DSTC-8 [9], также содержит диалоги из Ubuntu IRC, как и Ubuntu Dialogue Corpus. Отличается он, во-первых, размером: 255 тыс. положительных обучающих примеров против 80 тыс. для UDC. Также для каждого контекста имеется 100 вариантов ответа вместо 10, и для некоторых контекстов верный ответ отсутствует.

Для инициализации векторных представлений слов из словаря модели мы использовали подходы word2vec, GloVe. Мы также использовали модель BERT [10] без дообучения, извлекая с её помощью векторные представления слов. Наконец, мы экспериментировали с label smoothing, а также варьировали количество дистракторов (неправильных вариантов ответов) в обучении, выбирая 5, 50 или 100. Мы обнаружили, что большее число дистракторов приводит к повышению R@1 во всех случаях.

Наилучшие результаты эксперимента можно наблюдать в таб. 2. Примечательно, что применение label smoothing, а также использование сложной модели для формирования векторных представлений слов (BERT), не дали прироста по метрике на нашей задаче.

Таблица 2

Значения R@1 (in 100) на корпусе Ubuntu DSTC-8, валидационная выборка

Модель	R@1
w2v+DE	26%
w2v+ESIM	63.4%
GloVe+ESIM	59.7%
Конкатенация (w2v, GloVe)+ESIM	65%
BERT+ESIM	45%
w2v+ESIM, label smoothing	57.3%

Литература

1. Ryan Lowe, Nissan Pow, Iulian Serban, Joelle Pineau. The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue system. Режим обращения: <https://arxiv.org/abs/1506.08909> (дата обращения: 23.12.2019).
2. Yann LeCun, Leon Bottou, Yoshua Bengio, Patrick Haffner. Gradient-based learning applied to document recognition. Режим обращения: <https://www.bioinf.jku.at/publications/older/2604.pdf> (дата обращения: 23.12.2019).
3. Rudolf Kadlec, Martin Schmid, Jan Kleindienst. Improved deep learning baselines for ubuntu corpus dialogs. Режим обращения: <https://arxiv.org/abs/1510.03753> (дата обращения: 23.12.2019).
4. Ming Tan, Cicero dos Santos, Bing Xiang, Bowen Zhou. LSTM-based deep learning models for non-factoid answer selection. Режим обращения: <https://arxiv.org/abs/1511.04108> (дата обращения: 23.12.2019).
5. Qian Chen, Wen Wang. Sequential attention-based network for noetic end-to-end response selection. Режим обращения: <https://arxiv.org/abs/1901.02609> (дата обращения: 23.12.2019).
6. Qian Chen, Xiaodan Zhu, Zhen-Hua Ling, Diana Inkpen, Si Wei. Neural natural language inference models enhanced with external knowledge. Режим обращения: <https://arxiv.org/abs/1711.04289> (дата обращения: 23.12.2019).
7. Yu Wu, Wei Wu, Chen Xing, Ming Zhou, Zhoujun Li. Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. Режим обращения: <https://arxiv.org/abs/1612.01627> (дата обращения: 23.12.2019).
8. Xiangyang Zhou, Lu Li, Daxiang Dong, Yi Liu, Ying Chen, Wayne Xin Zhao, Dianhai Yu, Hua Wu. Multi-turn response selection for chatbots with deep attention matching network. Режим обращения: <https://www.aclweb.org/anthology/P18-1103/> (дата обращения: 23.12.2019).
9. DSTC-8 homepage [Электронный ресурс]. Режим обращения: <https://sites.google.com/dstc.community/dstc8/home> (дата обращения: 23.12.2019).
10. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Режим обращения: <https://arxiv.org/abs/1810.04805> (дата обращения: 23.12.2019).



Маркитантов Максим Викторович

Год рождения: 1995

Санкт-Петербургский институт информатики и автоматизации,
Российской академии наук,
младший научный сотрудник,
Университет ИТМО,
факультет информационных
технологий и программирования,
аспирант,

направление подготовки: 05.13.17 Теоретические
основы информатики,

e-mail: m.markitantov@yandex.ru



Карпов Алексей Анатольевич

Год рождения: 1978

Санкт-Петербургский институт информатики и автоматизации
Российской академии наук,
главный научный сотрудник,
Университет ИТМО,
факультет информационных технологий
и программирования,
д.т.н., профессор,

e-mail: karpov@iias.spb.su

УДК 004.934.2

**АНАЛИТИЧЕСКИЙ ОБЗОР АУДИОВИЗУАЛЬНЫХ
КОРПУСОВ РЕЧИ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ
ВОЗРАСТА ДИКТОРА**

М.В. Маркитантов

Научный руководитель – д.т.н., профессор А.А. Карпов

Работа выполнена при финансовой поддержке Российского научного фонда (проект №18-11-00145).

Аннотация

Работа содержит аналитический обзор аудиовизуальных корпусов речи AFEW, AFEW-VA, MODALITY, SEWA, VoxCeleb и VoxCeleb 2, которые содержат представительный набор различных дикторов с указанием пола и возраста в метаданных. Многомодальная автоматическая система для распознавания пола и возраста диктора, построенная на основе этих корпусов, позволит значительно увеличить точность распознавания в сравнении с системой, которая основана только на голосе или на изображении лица человека, так как низкая точность алгоритма классификации по одной модальности может компенсироваться высокой точностью по другой.

Ключевые слова

Компьютерная паралингвистика, машинное обучение, многомодальный корпус, распознавание возраста диктора, речевые технологии, компьютерное зрение.

Введение

Современные технологии машинного обучения ориентированы на выявление различных свойств наблюдаемого человека, которые затем могут использоваться в

системах верификации и идентификации. В частности, для улучшения человеко-машинного взаимодействия, а также в работе телефонных контакт-центров, учреждений здравоохранения и для повышения эффективности целевой рекламы. Автоматическая система распознавания пола и возраста может быть полезным инструментом в судебно-медицинских приложениях; она может помочь сузить список подозреваемых [1].

Так как современные алгоритмы компьютерного зрения достигли хорошего результата, при распознавании возраста и пола диктора можно использовать изображения лиц [2]. При отсутствии прямого контакта с клиентом (пользователем) может использоваться речь человека, которая содержит паралингвистическую информацию, включающую в себя различные характеристики диктора (пол, возраст, эмоциональное состояние) [3].

Комбинирование двух модальностей позволит значительно увеличить точность распознавания пола и возраста диктора в сравнении с системой, которая основана только на голосе или на изображении лица человека, так как низкая точность алгоритма классификации по одной модальности может компенсироваться высокой точностью по другой. Следовательно, аудиовизуальный анализ характеристик диктора улучшит эффективность существующих систем верификации и идентификации.

Целью работы является анализ существующих аудиовизуальных корпусов речи, содержащих представительный набор различных дикторов с указанием пола и возраста в метаданных.

Речевые корпуса

Важную роль при реализации алгоритмов машинного обучения играют данные. Существует несколько аудиовизуальных корпусов речи с указанием пола и возраста дикторов в метаданных.

Аудиовизуальный корпус AFEW (Acted Facial Expressions in the Wild) [4] состоит из 1426 коротких видеоклипов, которые содержат эмоциональные высказывания 330 дикторов с возрастом от 1 до 70 лет, извлеченные из различных фильмов. Каждое высказывание сохранено в формате AVI и аннотировано следующими данными: эмоциональным состоянием диктора (гнев, отвращение, страх, счастье, грусть, удивление, нейтральное состояние), длиной высказывания, а также информацией о дикторе (поза, пол, возраст). Данные высказывания содержат выражения переменной длины, продолжительность фраз находится в диапазоне от 300 до 5400 миллисекунд. Пример изображений корпуса AFEW представлен на рис. 1.



Рис. 1. Изображения кадров корпуса AFEW

Корпус AFEW-VA [5] содержит 600 видеороликов, извлеченных из художественных фильмов. Количество видео кадров варьируется от 10 до 120. В общей сложности содержит более 30000 аннотированных кадров валентностью и интенсивностью возбуждения эмоции от -10 до 10. Содержит записи 240 дикторов с возрастом от 8 до 76 лет, 52% из которых женщины. Каждый фрейм аннотирован ориентирами лица, валентностью, интенсивностью высказывания, полом и возрастом диктора.

Многомодальный корпус MODALITY [6] состоит из 30 часов английской речи, которая содержит высказывания 42 дикторов (33 мужчины и 9 женщин) с возрастом от 14 до 60 лет (средний возраст: 34 года). База данных содержит стереоскопические видеопотоки с высоким разрешением и высокой частотой кадров, а также аудио сигналы, полученные с помощью массива микрофонов. Высказывания корпуса содержат выражения переменной длины: командные слова, названия месяцев и дней, записанные за 12 сеансов. Пример изображений корпуса представлен на рис. 2. Каждое высказывание аннотировано в соответствии с полом и возрастом диктора. Аудиофайлы хранятся в формате WAV с частотой дискретизации 16 кГц, видеофайлы – в формате MKV с разрешением 1080p и частотой кадров 100. Размер корпуса составляет 2,5 ТБ. Данный корпус может быть использован для систем AVSR (Audio-video speech recognition), а записи в шумных условиях могут использоваться для проверки надежности систем распознавания речи.



Рис. 2. Примеры кадров видео из корпуса MODALITY

База данных SEWA [7] состоит из 30 часов аудиовизуальной речи, которая содержит высказывания 398 дикторов, 50% из которых составляют женщины в возрасте от 18 до 65 лет. Каждое высказывание аннотировано ориентирами лица, валентностью, интенсивностью высказывания, полом и возрастом диктора. Данный корпус содержит эпизоды взаимодействия дикторов разного возраста, пола и культурного происхождения (британцы, немцы, венгры, греки, сербы и китайцы). База данных состоит из 6 групп добровольцев по 66 дикторов в каждой группе, равномерно распределенных по пяти возрастным группам: 18-29, 30-39, 40-49, 50-59, 60-65. Аудиофайлы хранятся в формате WAV, видеофайлы – в формате AVI с разрешением 480x360 и частотой кадров от 20 до 30. Точно аннотированные данные являются основой разработки аудиовизуальных алгоритмов распознавания, отслеживания и интерпретации поведения человека, которые позволят добиться надежной производительности в реальных условиях.

Корпус VoxCeleb [8, 9] состоит из 352 часов речи, которая содержит 153516 высказываний 1251 диктора. Данные высказывания были получены из видеозаписей

знаменитостей, загруженных на YouTube. Набор данных охватывает широкий спектр разных национальностей, акцентов, профессий, возрастов, сбалансированных по полу, 55% высказываний из которых составляют мужчины. Длина высказываний составляет от 3 до 20 секунд. В отличие от предыдущих корпусов речи, высказывания данного набора данных аннотированы только полом диктора.

Набор данных VoxCeleb 2 [9, 10] является второй версией корпуса VoxCeleb. Он содержит высказывания общей длительностью 2442 часа от 6112 дикторов, 61% из которых – мужчины. Число высказываний в данном корпусе превышает 1000000. Данные VoxCeleb 1 и VoxCeleb 2 не пересекаются. Данный корпус был подготовлен автоматически с помощью алгоритмов компьютерного зрения, без использования ручной разметки. Представленный набор данных аннотирован только полом диктора. Распределение национальной принадлежности дикторов в двух представленных корпусах приведено на рис. 3.

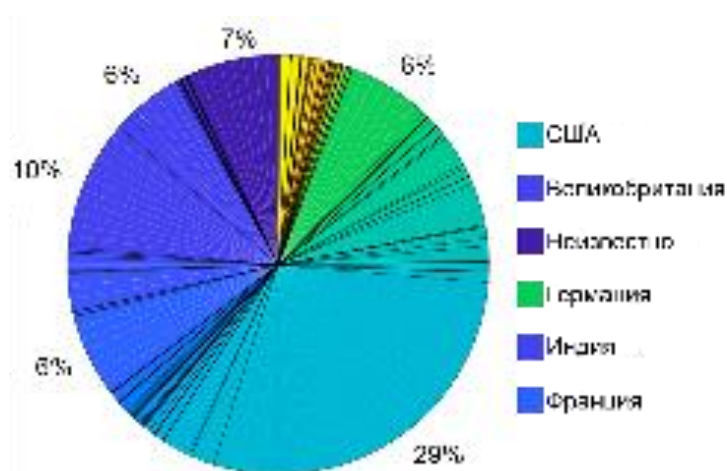


Рис. 3. Распределение национальной принадлежности дикторов

Заключение

В ходе исследования был проведен анализ аудиовизуальных речи, содержащих представительный набор различных дикторов с указанием пола и возраста в метаданных.

Из аналитического обзора существующих корпусов можно сделать вывод, что для разработки многомодальной системы автоматического распознавания пола и возраста диктора можно использовать корпуса AFEW, AFEW-VA, MODALITY, SEWA. Корпуса VoxCeleb и VoxCeleb 2 можно использовать только для систем распознавания пола диктора, так как в метаданных представленных корпусов не содержится возраст дикторов.

Литература

1. Markitantov M., Verkholyak O. Automatic Recognition of Speaker Age and Gender Based on Deep Neural Networks // Lecture Notes in Computer Science, Springer LNAI 11658, SPECOM 2019. 2019. pp. 327-336.
2. Angulu R., Tapamo J. R., Adewumi A. O. Age estimation via face images: a survey // EURASIP Journal on Image and Video Processing. 2018.
3. Маркитантов М.В., Карпов А.А. Автоматическое распознавание возраста и пола по голосу на основе глубоких нейронных сетей // Информационно-измерительные и управляющие системы. 2019. Т. 17. № 5. С. 76-83.
4. Dhall A., Goecke R., Lucey S., Gedeon T. Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark // 2011 IEEE International

- Conference on Computer Vision Workshops (ICCV Workshops), Barcelona. 2011. pp. 2106-2112.
5. Kossaifi J., Tzimiropoulos G., Todorovic S., Pantic M. AFEW-VA database for valence and arousal estimation in-the-wild // *Image and Vision Computing*. 2017. V. 65. pp. 23-36.
 6. Czyzewski A., Kostek B., Bratoszewski P. An audio-visual corpus for multimodal automatic speech recognition // *Journal of Intelligent Information Systems*. 2017. pp. 167–192.
 7. Kossaifi J., Walecki R., Panagakis Y., Shen J., Schmitt M., Ringeval F., Han J., Pandit V., Schuller B., Star K., Hajiyev E., Pantic M. SEWA DB: A Rich Database for Audio-Visual Emotion and Sentiment Research in the Wild // *In IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2019. P. 17.
 8. Nagrani A., Chung J. S., Zisserman A. VoxCeleb: a largescale speaker identification dataset // *In INTERSPEECH 2017*. 2017. pp. 950-954.
 9. Nagrani A., Chung J. S., Xie W., Zisserman A. Voxceleb: Large-scale speaker verification in the wild // *Computer Speech & Language*. 2020. V. 60. P. 15.
 10. Chung J.S., Nagrani A., Zisserman A. VoxCeleb2: Deep Speaker Recognition // *In INTERSPEECH 2018*. 2018. pp. 1086-1090.



Михайлова Ирина Алексеевна

Год рождения: 1995
Университет ИТМО,
Институт дизайна и урбанистики,
группа № С42551,
направление подготовки: 09.04.03 – Прикладная информатика,
e-mail: m.irka@mail.ru



Архипов Алексей Алексеевич

Год рождения: 1988
СПб ГБУЗ МИАЦ,
начальник отдела мониторинга показателей эффективности
электронных сервисов в сфере здравоохранения,
e-mail: Aarhipov@spbmiac.ru

УДК 004.415.2

**ПРОЕКТИРОВАНИЕ СЕРВИСА О СОСТОЯНИИ ЗДОРОВЬЯ
ПАЦИЕНТА ДЛЯ ЕДИНОЙ ГОСУДАРСТВЕННОЙ
ИНФОРМАЦИОННОЙ СИСТЕМЫ В СФЕРЕ ЗДРАВООХРАНЕНИЯ
САНКТ ПЕТЕРБУРГА**

И.А. Михайлова, А.А. Архипов

Научный руководитель – к.экон.наук О.В. Кононова

Работа выполнена в рамках темы ВКР «Проектирование сервиса о состоянии здоровья пациента для единой государственной информационной системы в сфере здравоохранения Санкт-Петербурга».

Аннотация

В статье представлены результаты проектирования сервиса «О состоянии здоровья пациента», с помощью которого пациент сможет самостоятельно вносить свои персональные данные о состоянии здоровья. Врач, с помощью этого сервиса, опираясь на данные внесенные пациентами, сможет более точно ставить диагнозы. Показатели, отклоняющиеся от нормы, помогут врачу своевременно оказывать помощь пациентам, и это может существенно увеличить продолжительность жизни больных и предотвратить развитие многих осложнений.

Ключевые слова

Ключевые показатели здоровья, электронный сервис, пациент, медицинская карта, интегральный анамнез.

Опыт научных наблюдений свидетельствует, что у большинства людей при соблюдении ими гигиенических правил есть возможность жить до 100 лет и более. Многие люди не соблюдают простейших, обоснованных наукой норм здорового образа

жизни [1]. Одни страдают от малоподвижности (гиподинамии), другие излишествуют в еде с почти неизбежным в этих случаях развитием ожирения, склероза сосудов, сахарного диабета, третьи не умеют отдыхать, страдают бессонницей [2].

Во всех этих случаях людям могут помочь ИТ. Опросы и наблюдения показали, что пациенты с хроническими заболеваниями, которым нужно постоянное наблюдение врача, а также пациенты, у которых присутствуют разного рода недомогания, требующие контроля показателей здоровья, заинтересованы в цифровых сервисах, позволяющих иметь доступ к данным об основных показателях здоровья. Поэтому Санкт-Петербургским медицинским информационно-аналитическим центром разработан сервис «Ключевые данные о здоровье пациента», где каждый пациент в случае необходимости может просмотреть свою медицинскую карту [3]. В процессе внедрения сервиса выяснилось, что пациенты готовы осуществлять контроль ряда показателей самостоятельно. В связи с этим было принято решение о проектировании дополнительного сервиса «О состоянии здоровья пациента», предназначенного для организации самоконтроля и удалённого контроля ключевых показателей здоровья человека.

В ходе опроса специалистов профилактической медицины были выявлены основные показатели для наполнения сервиса: артериальное давление, вес, уровень глюкозы в крови, аллергия, сердечный ритм, рост, холестерин, температура. Показатели здоровья, включенные в сервис, заполняются самим пациентом, не являются обязательными для внесения. Сервис создаётся для персонального пользования пациента, и только пациент решает, какие показатели нужно добавить в свою медицинскую карту. Порядок внесения данных в систему представлен на (рис. 1).

Сервис даёт множество преимуществ, первое – врач может обратиться к этим данным, он вправе не доверять им. В то же время показатели могут сыграть ключевую роль в понимании текущего состояния и динамики развития состояния больного. Для самого пациента – это возможность, заполняя показатели здоровья, отслеживать изменения значения показателей и принимать своевременные решения обратиться к врачу. Сервис удобен для людей, страдающих хроническими заболеваниями.

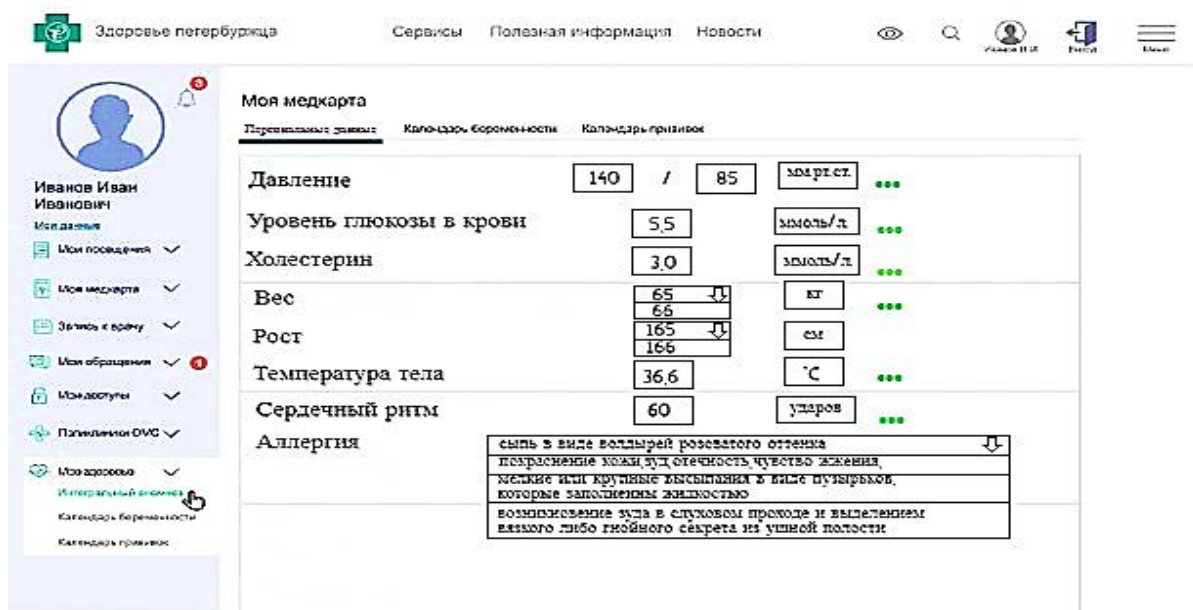


Рис. 1. Сервис «О состоянии здоровья пациента»: введение персональных записей о состоянии здоровья

На основании внесенных данных будет формироваться отчет для врача. Шапка отчета создается автоматически, ФИО и дата рождения будут заполняться из «Личной карты пациента», наименования показателей и дата заполнения будут указываться автоматически из сервиса «Ведение персональных записей о состоянии здоровья пациента».

Наполнением отчета будут служить данные, внесенные пациентами (рис. 2). Эти данные будут отображаться в сервисе «Интегральный анамнез пациента» предназначенном для врачей, во вкладке «Персональные записи пациента».


ФИО	Дата рождения	Дата заполнения
Шипков Павел Юрьевич	12.05.1971	20.12.2019

Давление	140/85 мм.рт.ст.
Уровень глюкозы в крови	5,5 ммоль/л.
Холестерин	3,0 ммоль/л.
Вес	88
Рост	180
Температура тела	36,8
Сердечный ритм	-
Аллергия	Сыпь в виде волдырей розоватого оттенка

Рис. 2. Формирование отчета для врача

В целях контроля за состоянием своего здоровья, пациент может ответить на вопросы анкеты «Самоконтроль здоровья пациента», и если имеются показания обратиться к доктору [3]. Анкета находится в сервисе «Ведение персональных записей о состоянии здоровья пациента», и состоит из 7 вопросов. Пример вопросов анкеты представлен на рис. 3.

Укажите ваше артериальное давление



от 110/70 до 130/85

от 140/90 до 180/110

Рис. 3. Пример представления вопроса и ответов на него в анкете для пациента

В зависимости от выбранного ответа, пациенту будут предложены рекомендации (рис. 4).

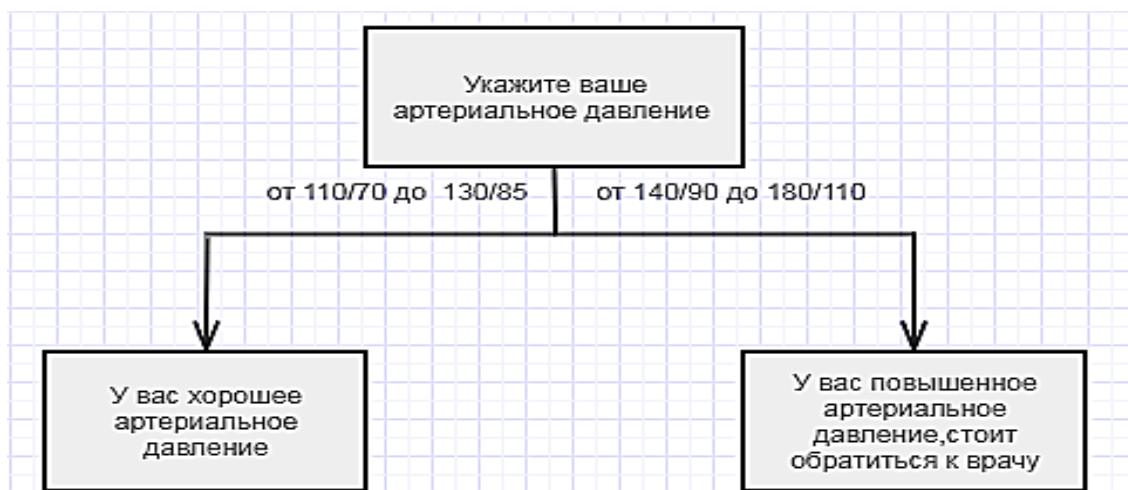


Рис. 4. Пример рекомендаций

После того, как пациент ответит на все вопросы анкеты, на экране появится заключение, и если имеются показания, он может своевременно обратиться к врачу [4]. (рис. 5).

Шипков Павел Юрьевич, рекомендации для вас:

Артериальное давление	Норма
Индекс массы тела	23.23 Норма
Уровень глюкозы в крови	Средняя форма диабета, советуем обратиться к врачу
Уровень общего холестерина	Советуем обратиться к врачу и сдать анализы
Норма сна	В пределах нормы

Мы рекомендуем вам пройти несколько обследований и консультаций, их вы можете получить у вашего участкового врача-терапевта Козловой Т.Ю.

Записаться вы можете перейдя по ссылке: <https://www.gorzdrav.spb.ru/signup/free/?>

Рис. 5. Пример отчета для пациента

В результате исследования были решены следующие задачи:

- выявлены основные данные для наполнения сервиса ведения персональных записей о состоянии здоровья пациента в рамках единой государственной информационной системы в сфере здравоохранения Санкт-Петербурга;
- терминологически описано представление данных, предложен и составлен образец, как будут представлены данные в сервисе;
- составлена и детально описана анкета «Самоконтроль здоровья пациента».
- выявлены функциональные требования к сервису.

Результаты исследования будут использованы при разработке сервиса «О состоянии здоровья пациента» в рамках единой государственной информационной системы в сфере электронного здравоохранения Санкт-Петербурга.

Литература

1. Кочорова Л.В., Пенюгина Е.Н., Клюковкин К.С. Основные направления реформирования системы здравоохранения Санкт-Петербурга // Вестник Санкт-Петербургского университета. Медицина. 2007. № 1. С. 131-135.
2. Леванов В.М. Основные направления развития региональных систем электронного здравоохранения // Социальные аспекты здоровья населения. 2018. № 5 (27). 9 с.
3. МИАЦ. Проект «Электронное здравоохранение». Общее описание [Электронный ресурс]. Режим доступа: <http://spbmiac.ru/ehlektronnoe-zdravookhranenie/proekt-ehlektronnoe-zdravookhranenie/obshhee-opisanie/> (дата обращения 20.02.2020).
4. Шапцев В.А. Теоретические основы создания информационного общества. Тюмень: Издательство Тюменского государственного университета. 2015. 356 с.



Мурадян Арман Хачатурович
Университет ИТМО,
факультет инфокоммуникационных технологий,
студент группы №К42402,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: arm.murr@gmail.com

УДК 004.056

КЛАССИФИКАЦИЯ СПОСОБОВ ДЕТЕКТИРОВАНИЯ И БЛОКИРОВАНИЯ АВТОМАТИЗИРОВАННОГО СБОРА ДАННЫХ С ВЕБ-РЕСУРСОВ

А.Х. Мурадян

Научный руководитель – к.т.н., доцент Н.Н. Горлушкина

Работа выполнена в рамках темы НИР «Способы детектирования и блокирования автоматизированного сбора информации с вебсайтов».

Аннотация

В статье приведены результаты проведенного анализа, способов детектирования и блокирования автоматизированного интернет трафика. Произведена первичная оценка эффективности и надёжности приведенных методов. На основе полученных результатов произведена их классификация и градация. Сделаны выводы о возможности применения этих способов для борьбы против автоматизированного трафика и создания специализированных систем, основанных на анализе множества факторов.

Ключевые слова

Детектирование, блокирование, парсинг, автоматизированный сбор данных, защита сайта.

Автоматизированный сбор информации с веб-ресурсов, иными словами – парсинг, приносит на данный момент много проблем различным компаниям и организациям. Компании хотят защитить свою интеллектуальную собственность, снизить нагрузку на оборудование, повысить качество трафика и просто усложнить жизнь конкурентам. Есть множество средств и способов детектирования и блокировки парсинга, эффективность, быстрдействие, вид парсинга, от которого они защищают, у всех различны. Поэтому появляется актуальная проблема классификация этого разнообразия инструментов и методов. Такая классификация поможет веб мастерам точнее и быстрее подобрать необходимые средства против определенных видов ботов. Также на основе данной классификации, можно разработать модель и сервис для автоматического детектирования и блокирования различных ботов, используя наиболее эффективные средства [1-2].

Рассмотрим и проанализируем основные способы обнаружения и блокирования подозрительного трафика, произведя оценку каждого из них на предмет эффективности и точности. Проведём классификацию рассмотренных методов.

Способы детектирования

1. Не работающий JavaScript. JavaScript обычно используется как встраиваемый язык для программного доступа к объектам приложений. Наиболее широкое применение находит в браузерах как язык сценариев для придания интерактивности веб-страницам.

На данный момент js используется почти на всех интернет ресурсах. Многие из них попросту не работают без него. Количество реальных посетителей не использующих js сводится к минимуму, а вот автоматизированным системам намного проще собирать данные, не интерпретируя js-код. Отсюда можно прийти к выводу что, если js у пользователя отключён, значит вероятность того, что это не человек крайне велика.

2. WebDriver. Selenium WebDriver — это инструмент для автоматизации действий веб-браузера. В большинстве случаев используется для тестирования Web-приложений, но этим не ограничивается. В частности, он может быть использован для решения рутинных задач администрирования сайта или регулярного получения данных из различных источников (сайтов).

Если сайт ограничивает доступ к работе без js, то чтобы обойти эту защиту против сайта может быть использован WebDriver. Чтобы детектировать его использование можно получить данные с navigator.webdriver() и если он будет равен true, это будет означать, что пользователь использует автоматизированное ПО, что также говорит о высокой вероятности того, что это вовсе не человек.

3. Нестандартный User-Agent. User-Agent – это текстовая часть запроса, которую веб-приложения используют для сообщения сайту информации о себе. User-Agent браузера содержит название и версию приложения, а также данные об операционной системе компьютера: версия, разрядность, язык по умолчанию и другие параметры.

Поисковые роботы, такие как Google и Yandex используют специальный UserAgent, нечистые на руку разработчики могут подменить UserAgent своего бота, дабы притвориться ботом популярной сети. Для детектирования таких случаев у Google и Yandex есть способы проверки, принадлежности к их сетям. Использование таких проверок позволит со стопроцентной вероятностью выявлять автоматизированный нежелательный трафик.

Также можно блокировать все нетипичные UserAgent, которые принадлежат другим специальным сервисам (например, Postman UserAgent).

4. Ловушки для ботов. Некоторые парсинг-боты бездумно обходят все ссылки на сайте. Для обнаружения таких ботов можно поставить на сайте ссылки-ловушки – это невидимые для человека, но видимые роботу ссылки. Детектируя несколько подобных переходов одного пользователя, можно будет с высокой вероятностью говорить о том, что это бот.

5. Короткое время между запросами с одного ip. Характерная черта бота — это малое время между запросами. Обычно человек при переходах по страницам проводит на каждой странице хотя бы несколько секунд, периодически останавливаясь на более длительное время. Бот же «атакует» сайт запросами с интервалом менее секунды. Заметив такую активность, можно с некоторой вероятностью говорить о том, что это бот. Если активность и переходы не прекращаются долгое время, то вероятность возрастает.

6. Большое количество запросов с одного ip. Аналогично предыдущему пункту, обычный пользователь вряд ли будет посещать большое количество страниц (от нескольких сотен). Здесь важным моментом будет убедиться в том, что это действительно один и тот же пользователь. Для этого необходимо использовать специальные инструменты аналитики, запись fingerprint браузера, cookie и другие параметры.

7. Детектирование по отпечатку браузера [3-4]. Fingerprint или отпечаток компьютера (браузера) – информация, собранная об удалённом устройстве для дальнейшей идентификации, фингерпринтинг — сбор этой информации. Отпечатки могут быть использованы полностью или частично для идентификации, даже когда cookie выключены. На данный момент появились способы однозначного детектирования не только браузера, но и компьютера пользователя, основанные на особенностях работы графической подсистемы компьютера.

8. Нетипичные запросы, порядок запросов [5]. Анализируя поведение пользователя, выявляя нетипичное поведение, переходы по страницам, можно с помощью средств аналитики построить модель действий добросовестного пользователя и при отклонении от поведения можно будет с определенной долей вероятности сделать вывод о нестандартном поведении. В данном случае будет полезно сформировать модели не только пользователей, но и поисковых роботов, чтобы не мешать их работе.

9. Наличие ip в спам базах, ip принадлежит прокси. Есть множество ресурсов и баз данных, в том числе свободных, которые хранят списки ip адресов, которые зарекомендовали себя не с лучшей стороны. При заходе пользователя на сайт можно проверять входит ли его ip в эти списки, и в какое количество, в зависимости от этого фактора в совокупности с другими делать выводы о автоматизированном трафике.

10. Отсутствие активности мыши. Ещё один из факторов, что был обнаружен во время эксперимента, при автоматизации трафика. Если происходят переходы по страницам, а движения мыши не фиксируются, ли они есть но наводятся не на те элементы, в которых находится следующая ссылка, то это может говорить о том, что используется автоматизированное ПО.

11. Ip адреса рядом идущих запросов принадлежат одной подсети. Проверка принадлежности рядом идущих запросов пользователей к одной подсети может являться одним из фактов парсинга, который нужно учитывать.

12. Прямые заходы. Большое количество прямых заходов на разные страницы сайта одним и тем же пользователем может говорить о высокой вероятности того, что это бот, потому что в таких системах при переходе по страницам обычно не поставляется параметр referer.

Рассмотренные методы можно классифицировать по следующим параметрам:

- по степени однозначности детектирования;
- методы, с высокой вероятностью детектирующие бота;
- не работающий JavaScript;
- Webdriver;
- подмена User-Agent;
- отсутствие активности мыши;
- ловушки для ботов;
- прямые заходы;
- методы детектирующие с некоторой вероятностью;
- нетипичный порядок запросов;
- наличие ip в спам базах, списках прокси;
- принадлежность ip адресов к одной подсети;
- короткое время между запросами с одного ip;
- большое количество запросов с одного ip;
- по времени необходимому на детектирование;
- методы мгновенного детектирования;
- не работающий JavaScript;

- WebDriver;
- подмена User-Agent;
- короткое время между запросами с одного ip;
- большое количество запросов с одного ip;
- наличие ip в спам базах, списках прокси;
- методы отложенного детектирования (те, для которых нужно собрать некоторую статистику);
- принадлежность ip адресов к одной подсети;
- нетипичный порядок запросов;
- ловушки для ботов;
- отсутствие активности мыши;
- прямые заходы.

Способы блокировок

13. Блокировка по ip. Один из самых простых способов ограничения и блокировки – это закрыть доступ к конкретному ip адресу, чей трафик был отмечен системой аналитики как автоматизированный. Данный способ является эффективным лишь против простейших, любительских парсеров, профессиональные системы, скорее всего будут использовать пул прокси и быстро смогут поменять ip адрес и продолжить работу до следующей блокировки.

Также нужно будет учесть то, что под одним ip в сеть может выходить множество пользователей, особенно это актуально для мобильной сети, поэтому есть риск таким образом заблокировать реальных пользователей.

14. Блокировка по отпечатку браузера. Об этой теме говорилось ранее в способах детектирования трафика. Получив однажды отпечаток браузера или компьютера, можно будет блокировать злоумышленника независимо от его ip адреса и других систем обхода блокировок.

15. Ограничивание информации по логину. Может выступать как дополнительная мера ограничения парсинга, увидев подозрительную активность от конкретного пользователя, можно быстро его заблокировать. Также можно привязать к аккаунту номер телефона, тогда злоумышленнику сложно будет создать большое количество новых и новых аккаунтов.

16. Вывод капчи [6, 7]. Один из действенных способов успокоить трафик, или на время его заморозить. Однако сейчас в сети появились сервисы, которые нанимают людей расшифровывать капчи, за маленькие деньги. Поэтому наиболее эффективным будет создать свой собственный алгоритм капчи, использующий нестандартные техники ее решения, в этом случае решать такую капчу будет экономически не выгодно. Вряд ли сам злоумышленник будет это делать, ровно, как и система решения капч не будет ради одного сайта разрабатывать API.

17. Подстановка ошибочных данных. Данный способ подойдет если на сайте хранится большое количество ценной информации, которую конкуренты желают скопировать. Обнаружив парсинг, сайт не будет внешне никак об этом говорить, а будет незаметно подставлять неверные данные, либо говорить, что таких данных нет. Конкурент не скоро заметит такую защиту, а понять какие условия учитываются при детектировании парсинга будет намного сложнее.

18. Отображение важных данных в виде картинок. Может быть применено как дополнительная мера защиты важных данных. Например, номера телефонов пользователей площадки интернет торговли, немногие захотят внедрять в систему парсинга сложные системы по распознаванию изображений.

Рассмотренные методы можно классифицировать по следующим параметрам:

1. По однозначности блокировки бота:
 - однозначное блокирование;
 - блокировка по отпечатку браузера;
 - неоднозначное блокирование (под блокировку могут попасть обычные пользователи);
 - блокировка по ip;
 - отображение важных данных в виде картинок (все пользователи испытывают неудобства);
2. По эффективности действия против ботов:
 - легко обойти защиту;
 - отображение важных данных в виде картинок;
 - блокировка по ip;
 - ограничивание информации по логину;
 - обход защиты требует значительных средств;
 - вывод капчи (при условии использования собственной капчи);
 - блокировка по отпечатку браузера/ компьютера;
 - подстановка ошибочных данных.

Вывод

Проанализировав эти методы детектирования и блокировки автоматизированного трафика, можно сделать вывод, что построив систему глубокого поведенческого и параметрического анализа пользователей веб-ресурса, и используя наиболее эффективные против каждого вида атак схемы блокировки, можно достичь высоких результатов при борьбе с парсингом сайтов, очень сильно затруднив сбор данных с сайта и сделав её (парсинга) стоимость слишком высокой и не выгодной.

Литература

1. Менщиков А.А., Гатчин Ю.А. Методы обнаружения автоматизированного сбора информации с веб-ресурсов // Кибернетика и программирование. 2015. № 5. С. 136–157.
2. Менщиков А.А., Комарова А.В., Гатчин Ю.А. Изучение поведения средств автоматизированного сбора информации с веб-ресурсов // Вопросы кибербезопасности. 2017. №3. С. 49-54.
3. Московченко В.М., Столяров Д.О., Горбунов А.А., Белянин В.И. Анализ технологий защиты от идентификации веб-браузеров // NBI-technologies. 2018. – №1. С. 34-39.
4. (Cross-)Browser Fingerprinting via OS and Hardware Level Features. [Электронный ресурс]. Режим доступа: <https://www.ndss-symposium.org/ndss2017/ndss-2017-programme/cross-browser-fingerprinting-os-and-hardware-level-features/> (дата обращения: 03.02.2020).
5. Менщиков А.А., Гатчин Ю.А. Метод обнаружения веб-роботов на основе анализа графа пользовательского поведения // Программные продукты и системы. 2019. №4. С. 607-612.
6. Угниченко Д. Эффективные способы защиты от парсинга сайта [Электронный ресурс]. Режим доступа: <https://ru.megaindex.com/blog/stop-bots> (Дата обращения: 05.02.2020).
7. Дурнев А. Противодействие парсингу, или, как защитить свои данные от копирования. // [Электронный ресурс]. Режим доступа: <https://www.seonews.ru/analytics/protivodeystvie-parsingu-ili-kak-zashchitit-svoi-dannye-ot-kopirovaniya/> (Дата обращения: 04.02.2020).



Петров Олег Евгеньевич

Год рождения: 1989

Университет ИТМО,
факультет информационных
технологий и программирования,
аспирант,

направление подготовки: 05.13.11– Математическое и программное
обеспечение вычислительных машин, комплексов и сетей,

e-mail: petrov-o@speechpro.com



Кабаров Владимир Иосифович

Год рождения: 1959

Университет ИТМО,
факультет информационных
технологий и программирования,
старший преподаватель,

e-mail: kabarov@speechpro.com



Матвеев Юрий Николаевич

Год Рождения: 1955

Университет ИТМО,
факультет информационных
технологий и программирования,
д.т.н., профессор,

e-mail: matveev@speechpro.com

УДК 004.89

**ИСПОЛЬЗОВАНИЕ ВЕКТОРНЫХ ПРОСТРАНСТВ
ДЛЯ РЕАЛИЗАЦИИ ТЕКСТОВОГО ПОИСКА ПО БОЛЬШИМ
ОБЪЕМАМ МЕДИА ДАННЫХ**

О.Е. Петров, В.И. Кабаров

Научный руководитель – д.т.н. Ю.Н. Матвеев

Аннотация

В статье рассматривается подход с использованием векторного пространства для слов и предложений для организации семантического поиска по речевым данным с использованием систем автоматического распознавания речи. Предложенный в статье подход позволяет использовать существующие решения для хранения поисковых индексов, а также осуществлять по ним запросы, заданные с помощью естественного языка, используя предварительно обученные модели.

Ключевые слова

Распознавание речи, речевые информационные системы, семантический поиск, обработка естественного языка.

Обработка естественного языка применяется в различных задачах [1], в том числе в реализации голосовых помощников, осуществления автоматических переводов с одного языка на другой и классификации текстов. В качестве источника данных для обработки можно также использовать результаты работы системы автоматического распознавания речи. Комбинируя технологии автоматического распознавания речи и обработку естественного языка, можно строить системы индексации и поиска для обработки больших объемов медиа-данных.

Для построения индекса поиска на основе результатов работы системы автоматического распознавания речи можно использовать лучшую гипотезу, полученную из системы распознавания [2], применяя к ней подходы обработки текста [3]. Это возможно, только для систем распознавания, дающих высокий показатель точности. В качестве альтернативы можно использовать более полную информацию, получаемую в процессе распознавания речи:

- N -лучших гипотез распознавания, упорядоченных по мере уверенности результата;
- словные сети распознавания, в которых веса у каждого слова могут представлять пару акустического и языкового веса или их сумму на заданном полукольце;
- словные сети спутывания, полученные с выполнением MBR-декодирования словных сетей распознавания [4].

В настоящей работе рассматривается вариант с использованием лучшей гипотезы распознавания речи на русском языке, так как для этого языка точность распознавания речи достаточно высокая.

Предварительная обработка данных

Перед индексацией лучшая гипотеза, полученная из системы автоматического распознавания речи, отождествляется с текстом. Текст дополнительно разделяется на абзацы и отдельные произнесения:

- по записям, в каждой из которых получен свой набор текстов;
- по каналам в многоканальных записях, где каждый диктор записан в свой канал;
- по метainформации самих записей, в которых разные события могут быть отмечены специальными временными метками.

По просодическим признакам. Например, по паузе, обнаруженной детектором речи или непосредственно системой распознавания речи.

По разметке диаризации, если фразы, распознанные в одном канале, отнесены к различным дикторам.

По лексическим признакам. Например, предложения могут разделяться по знакам препинания, расставленным соответствующим компонентом.

Для каждого выделенного сегмента речи следующим шагом выполняется предобработка, переводящая его в формат удобный для работы. Предобработка может состоять из каскада этапов, которые могут отличаться в зависимости от задачи и реализации. Далее приведен один из возможных набор этапов, который использовался при проведении экспериментов:

- 1) перевод всех символов в тексте к одному регистру;
- 2) нормализация цифр, чисел, дат и времени. Приведение их к одному виду;
- 3) удаление символов пунктуации;
- 4) удаление пробельных символов и токенизация. Для результатов распознавания речи может быть не актуально, так как результат уже разбит на слова.

Векторное представление данных

Модель векторного представления слов позволяет представить слово в виде вектора чисел фиксированной размерности. Свойства и размерность векторов подбираются таким образом, чтобы сохранить семантические свойства слова — слова, векторы которых в пространстве находятся близко друг к другу, должны быть схожи по семантическому значению. Направления в векторном пространстве привязываются к различным аспектам значения слова. Например, вектор для слова "Россия" может быть близок к "СССР" в одном направлении и к слову "Санкт-Петербург" в другом.

Исследователи уже довольно давно проявляют интерес к векторным представлениям слов. За последние несколько лет во множестве задач обработки естественного языка удалось получить хорошие результаты при использовании нейронных сетей для формирования векторных пространств слов. Был разработан целый ряд алгоритмов векторного представления слов, в том числе Word2vec и GloVe. Эти подходы используют большие корпуса текстов и исследуют контекст, в котором появляется каждое слово, чтобы определить его векторное представление.

1. В Word2vec специальный алгоритм (skip-gram) позволяет обучать нейронную сеть предсказывать контекстные слова вокруг слова в предложении. После обучения такая сеть позволяет получать векторное представление каждого слова.

2. В GloVe близость слов зависит от того, как часто они появляются с другими контекстными словами. Алгоритм обучает простую линейную модель по подсчетам совпадений слов.

Многие исследовательские группы распространяют модели, которые были предварительно обучены на больших текстовых корпусах, таких как Wikipedia или Common Crawl. Предварительно обученные модели можно использовать для решения прикладных задач, но часто используется дополнительное обучение и настройка модели для целевого набора данных и задач.

Некоторое время назад стали появляться подходы, которые позволяют строить векторные пространства не только для слов, но и более длинных частей текста: предложений или целых абзацев. Большинство современных подходов основаны на сложных архитектурах нейронных сетей и часто включают в себя специально размеченные данные для обучения для более качественного выделения семантической информации.

Обученная модель позволяет для выделенного сегмента текста получить как векторное представление для каждого слова в контексте, так и вектор для всего сегмента. Как и в случае со словными моделями, доступны предварительно обученные модели, что позволяет пропустить дорогостоящий процесс обучения. Несмотря на то, что процесс обучения может быть очень ресурсоемким, получение векторов по обученной модели гораздо менее требователен к ресурсам – векторное представление для текстов и предложений можно получать в приложениях реального времени.

Среди распространенных методов реализации векторных пространств для текстов можно выделить InferSent, ELMo и BERT [5]. На сегодняшний день активно ведутся исследования в этой области, поэтому можно ожидать появления новых подходов и более качественных моделей.

Семантический поиск и индексация

Для реализации поиска по медиа-данным необходимо обеспечить два процесса:

- 1) индексацию и хранение данных;
- 2) Формирование запроса и обход индекса поиска.

В качестве инструмента хранения индекса можно использовать готовое решение, поддерживающее хранение векторов чисел. В решениях на базе поискового

движка Lucene, таких как Elasticsearch или Solr, такая возможность есть. Тогда поисковой индекс может быть построен следующим образом:

- 1) из медиа-данных выделяются звуковые каналы, которые обрабатываются с помощью системы автоматического распознавания речи;
- 2) полученные данные предварительно обрабатываются и нарезаются на сегменты текста с сохранением информации о том, к какому медиа-файлу относится сегмент;
- 3) с помощью выбранной модели для каждого сегмента текста строится его векторное представление;
- 4) полученные вектора вместе с текстом и другой метаинформацией добавляются в индекс поиска.

Процесс поиска выполняется следующим образом:

1. пользователь формирует произвольный текстовый запрос.
2. с помощью выбранной модели для текста запроса строится его векторное представление.
3. для поиска и ранжирования результата, попарно вычисляется мера сходства между каждым индексированным вектором и вектором запроса.

В качестве меры сходства векторов обычно используется косинусное сходство. Для двух векторов признаков, A и B , косинусное сходство, $\cos(\theta)$, может быть вычислено через скалярное произведение и норму следующим образом:

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Вычисления большого количества косинусных сходств для пар векторов может быть оптимизировано предварительным расчетом нормы векторного представления текста запроса. Нужно отметить, что ускорение вычисления косинусного сходства и поиск конечного количества самых близких векторов само по себе является вопросом для исследования и в рамках данной статьи не рассматривается.

Выводы

Методы построения векторных пространств для текстов дают большие возможности для извлечения лингвистической и семантической информации из текстов. Индексируя векторные представления и оценивая расстояния между векторами, мы можем сравнивать предложения и абзацы, используя понятие сходства, которое выходит за рамки их пересечения на уровне слов. Опираясь в качестве входных данных результатами распознавания, можно вводить еще и понятие акустической близости слов, которая помогла бы нивелировать возможные ошибки системы автоматического распознавания речи.

Большое количество готовых инструментов и предварительно обученных моделей на каждом этапе вычислений позволяют быстро проводить эксперименты и оперативно внедрять решения, основанные на предлагаемом подходе, в реальные продукты.

Литература

1. Nugmanova Aigul, Bulusheva Anna, Chernykh Irina, Matveev Yuriy. Unsupervised Training of Automatic Dialog Systems for the Customer Support Service. Proc. 2019

- International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS). Sochi, Russia, Russia.
2. George Saon, Daniel Povey, Geoffrey Zweig. Anatomy of an extremely fast LVCSR decoder. Eurospeech, 9th European Conference on Speech Communication and Technology, Lisbon, Portugal, September 4-8, 2005.
 3. Justin Zobel, Alistair Moffat. Inverted files for text search engines. ACM Computing Surveys (CSUR), Volume 38 Issue 2, 2006.
 4. Lidia Mangu, Eric Brill, Andreas Stolcke. Finding consensus in speech recognition: word error minimization and other applications of confusion networks. *Computer Speech & Language* 14(4): 373-400 (2000).
 5. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proc. of NAACL. 2018.



Пилясова Дарья Дмитриевна

Год рождения: 1996

Университет ИТМО,

Институт Дизайна и Урбанистики,

магистрант группы № С42551,

направление подготовки: 09.04.03 – Прикладная информатика.

Управление государственными информационными системами,

e-mail: pilyasova.d@gmail.com

УДК 004.415.2

АНАЛИЗ ДАННЫХ И МОДЕЛИРОВАНИЕ ЦИФРОВЫХ РЕШЕНИЙ ГОРОДА В СФЕРЕ ОБРАЗОВАНИЯ И КУЛЬТУРЫ

Д.Д. Пилясова

Научный руководитель – к.э.н., доцент О.В. Кононова

Работа выполнена в рамках темы НИР 3 «Создание банка цифровых решений города в сферах образование и культура».

Аннотация

В работе представлены основные этапы, методы и результаты анализа данных и проектирования банка цифровых решений города в сферах образования и культуры. Данный сервис обеспечивает возможность поиска решений в данных сферах при реализации проектов «Умный город», а также цифровизации регионов в целом.

Ключевые слова

Умный город, банк решений, образование в умном городе, культура в умном городе, цифровизация, архитектурный подход, типологизация городов, базы данных.

В настоящее время в России реализуется проект построения «Умных городов». Она направлена на повышение конкурентоспособности российских городов, формирование эффективной системы управления городским хозяйством, создание безопасных и комфортных условий для жизни горожан и базируется на 5 ключевых принципах:

- ориентация на человека;
- технологичность городской инфраструктуры;
- повышение качества управления городскими ресурсами;
- комфортная и безопасная среда;
- экономическая эффективность, в том числе, сервисная составляющая городской среды.

Проект цифровизации городского хозяйства «Умный город» реализуется Министерством строительства и жилищно-коммунального хозяйства Российской Федерации в рамках национальных проектов «Жилье и городская среда» и «Цифровая экономика».

Сферы образование и культура, хоть и не являются обязательными для участвующих в проекте городах, но вызывают высокий общественный интерес. Проведенное проектным офисом «Умный Санкт-Петербург» социологическое исследование показало, что жители Санкт-Петербурга высказали свою

заинтересованность в развитии человеческого капитала при реализации проекта «Умный город Санкт-Петербург» [1].

Для понимания вектора развития сфер образование и культура города необходимо определить его характеристики, особенности и потребности, а также рассмотреть уже имеющийся опыт решений, реализуемых в рамках проектов «Умный город». Определить особенности города, влияющие на выбор пути развития, можно воспользовавшись классификацией городов [2].

Классификация позволяет подбирать лишь те решения, которые будут соответствовать стратегии города, и будут подходить возможностям городской среды. Также для определения особенностей города можно воспользоваться архитектурным подходом. Исследователи работают над определением общей архитектуры умного города, чтобы облегчить разработку и внедрение информационных и коммуникационных технологий умных городов в будущем. Проведя анализ различных методик к рассмотрению города с использованием архитектурного подхода, и принимая во внимание цели умных городов и показатели, обозначенные в стандартах, предложенных министерством строительства [3], автором были выделены следующие необходимые элементы архитектуры для последующего рассмотрения сфер образования и культуры:

- слой технических элементов;
- слой передачи данных;
- слой работы с данными;
- слой интерфейсов и технологий;
- слой потребителей.

При этом важным является защита информации на каждом из приведенных слоев. Модель предложенной автором архитектуры представлена на рис. 1.



Рис. 1. Архитектура Умного города

Отличительной особенностью предложенной архитектуры является методология описания Умного города: заполнение таблиц производится «сверху вниз с 1 по 5 слой», т.е. в первую очередь выявляются заинтересованные стороны (слой потребителей), далее выявляются предлагаемые для каждого из них существующие или внедряемые на рынке интерфейсы и подходы, при этом интерфейсы включают в себя один или

несколько подходов, затем описывается работа с данными в каждом из них, и в последнюю очередь выявляются применяемые технологии передачи данных и технические элементы. Дублирующийся элемент не записывается повторно.

Термин «слой» в данном контексте и далее, используется в значении слова домен. Слой технических элементов включает в себя всевозможные физические устройства, способные генерировать данные, такие как датчики, пользовательские устройства, объекты интернета вещей, и т.д. Слой передачи данных выступает в качестве основы любой архитектуры умного города, связывая источники данных со станциями управления. Слой передачи данных — это конвергенция различных сетей связи. Слой передачи данных состоит из устройств связи и передачи данных, в том числе, использующих проводные, беспроводные и спутниковые технологии.

Слой работы с данными – это важный элемент любого умного города, который находится между слоем технических элементов и слоем интерфейсов. Этот слой выполняет различные задачи по обработке, организации, анализу, хранению данных и принятию решений. Основной задачей слоя работы с данными является поддержание жизнеспособности и консистентности данных, которые сосредоточены на очистке данных, их эволюции, ассоциации и обслуживании. Уровень управления данными может быть далее разделен на категории объединения, анализа, обработки и хранения данных, а также управления событиями и решениями.

Слой интерфейсов и технологий отображает требования ко всей экосистеме и интерфейсам цифровой экономики города со стороны жителей, бизнеса, научного сообщества и органов муниципалитета. Слой потребителей содержит потребности жителей города, организаций, сообществ. Также этот слой отвечает за контроль гражданами качества предоставляемых цифровых услуг и их участие в управлении городом.

Сравнив архитектуру умных городов в контексте образования в России и мире, можно сделать выводы о том, что:

– в России технологии VR находятся на стадии изучения возможностей использования в образовании. Поэтому нет ни одного реализованного или находящегося в завершающей стадии реализации. По этой же причине в России не используются устройства захвата движения [4];

– на сегодняшний день в России только изучается, но не применяется технология Big Data применительно к образованию, поэтому отсутствует возможность управления событиями и решениями на основе анализа данных.

Стоит отметить, что все подходы в той или иной мере используются в отечественных проектах, однако большинство проектов являются коммерческими. Это говорит о том, что частные организации охотнее применяют инновационные подходы, чем государственные учреждения.

Проведя анализ состояния сферы культуры в России и мире, можно сделать выводы о том, что:

В России технологии VR используются менее активно, чем в мире. На данном этапе посетителю не представляется возможности погрузиться полностью в виртуальную реальность, используя устройства захвата движения, которые позволяют взаимодействовать с объектами в среде виртуальной реальности.

В России не используются технологии Big Data, поскольку не существует тенденции сбора больших объёмов неструктурированных данных в сфере культуры.

Технология геймификации используется только для оффлайн-экскурсий, а также, как и в сфере образования, большинство проектов в этой области являются коммерческими.

Выявленные различия могут быть связаны с тем, что сферы образования и культуры не являются приоритетными и обязательными при реализации проектов Умных городов в России.

Изучив потребности и особенности необходимо сформировать рекомендации для реализации проектов «Умный город». Для того, чтобы понять какие решения были правильно внедрены при реализации проектов «Умный город» можно воспользоваться «Банком решений», предназначенным для хранения и систематизации успешных практик, которые были применены при внедрении и реализации проектов.

На этапе проектирования «Банка цифровых решений умного города в сферах образования и культуры» был рассмотрен опыт реализации банков решений в России. Проведя их анализ, было выявлено, что ни один из них не удовлетворяет потребностям, возникающим при реализации проектов «Умный город» в сферах образования и культуры [5].

В таблице представлен сравнительный анализ банков решений.

Таблица

Сравнительный анализ банков решений

Название	Назначение	Фильтры		Тип базы данных	Владелец
		Сферы деятельности	Типы городов		
Электронная база данных кейсов политической геймификации	Хранение и анализ информации о кейсах использования геймификации в политической деятельности	Нет	Нет	Реляционная, Microsoft Access	Санкт-Петербургский государственный университет
«Банк решений Умного города» [6]	Сбор данных о решениях «Умного города», изучение международных и российских практик, а также поиск новых идей для развития существующих проектов	Нет	Нет	Нет информации	Министерство строительства и жилищно-коммунального хозяйства Российской Федерации

продолжение таблицы

Название	Назначение	Фильтры		Тип базы данных	Владелец
		Сферы деятельности	Типы городов		
База практик умного города [7]	Сбор данных о Российских реализованных решениях «Умного города»	Нет	Нет	Нет информации	ICT.MOSCOW и ДИТ Москвы
«Банк цифровых решений в сферах образование и культура»	Сбор и анализ данных о цифровых решениях в сфере образование и культура, поиск решений по определенным характеристикам города	Только сферы образование и культура	Да	Реляционная, PostgreSQL	Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики.

Проведя анализ существующих в России и мире решений, в сферах образования и культуры, и принимая во внимание опыт реализации существующих банков цифровых решений, были выделены следующие необходимые элементы для описания цифровых решений в сферах образования и культуры:

- название цифрового решения;
- описание цифрового решения;
- информация о реализации цифрового решения;
- затраты на реализацию решения;
- эффекты, достигнутые при реализации решения;
- направление, к которому относится решение.

Предлагаемая структура, сущности и связи в архитектуре базы данных позволяют осуществлять поиск решений по характеристикам города, а также производить выборку решений по сферам. Стоит отметить что технология поиска может быть применена к любым сферам, не только к образованию и культуре. Предложенная структура позволит хранить необходимую и полную информацию о решении [8].

Для построения информационно-логической модели данных было использовано CASE-средство – ComputerAssociates ErWin7.3. Информационно-логическая модель разработана в нотации IDEF1X. Спроектированная модель отображает сущности базы данных, состоящие из атрибутов, и приведена на рис. 2.

Проекты, составляющие собой банк решений, содержат большое количество несистематизированной информации. Такой как текстовое описание, информация о реализации, затраты на реализацию, эффекты, достигнутые после внедрения решения и т.д. Хранить ее в полях реляционной базы данных нецелесообразно, поэтому обычно для таких целей используют не реляционные базы данных, например, MongoDB [9], и хранят информацию как объект или документ. Однако, принимая во внимание

необходимость работы со списками, поиском и тегами, лучше использовать реляционную базу данных, поскольку не реляционные базы делают это медленнее. Поэтому планируется использовать PostgreSQL. Это решение позволит использовать тип данных json и jsonb. Тип данных json и jsonb - стандарт javascript object notation позволяет хранить произвольные данные в формате ключ-значение. «Произвольные», в данном контексте, означает, что мы можем сохранять неструктурированные в едином формате данные, а затем, с помощью встроенного функционала PostgreSQL производить по ним следующие операции: полнотекстовый поиск, выстраивание связи many-to-many, one-to-many между различными колонками формата json и jsonb, в том числе и между различными таблицами базы данных, а также осуществлять лёгкие и производительные операции, объединения, разъединения и создания таких типов данных, как массивы, объекты. Наличие указанных типов данных колоночных значений PostgreSQL фактически объединяет в себе преимущества реляционных и нереляционных (NoSQL) баз данных. Например, таких как - MySQL и MongoDB.

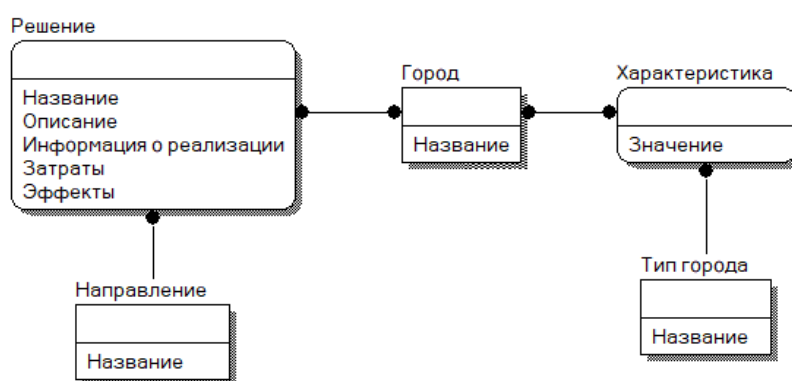


Рис. 2. Полная атрибутивная модель

В дальнейшем необходимо сформировать требования к разрабатываемой системе и передать их в разработку. Для формирования наиболее точных требований необходимо охватить и проанализировать больше примеров различных «Банков решений». Во время разработки системы необходимо собрать данные об удачных решениях в сферах образования и культуры, чтобы после публикации банка решений заполнить его готовыми данными. Последним этапом планируется провести тестирование системы и её внедрение.

Литература

1. Приложение №2 к концепции внедрения технологий умного города в Санкт-Петербурге «Умный Санкт-Петербург» Проблемный анализ ситуации в Санкт-Петербурге // Санкт-Петербург. 2018.
2. Бабюх В.А., Кайсарова Ж.Е. Понятие «Город» и типологизация городов: проблемы соотношения при определении даты возникновения городских поселений // Вестник Казанского технологического университета. 2014. №20.
3. Базовые и дополнительные требования к умным городам (стандарт «Умный город») [Электронный ресурс] // Режим доступа: [https:// russiasmartcity.ru /documents/18](https://russiasmartcity.ru/documents/18) (дата обращения: 13.03.2020).
4. Smart learning environments of the future [Электронный ресурс]. Режим доступа: [https:// forumvirium.fi/en/smart-learning-environments-of-the-future](https://forumvirium.fi/en/smart-learning-environments-of-the-future) (дата обращения: 28.01.2020).

5. Pro IoT: ICT.MOSCOW и ДИТ Москвы создали базу практик умного города [Электронный ресурс]. Режим доступа: <http://pro-iot.pro/materials/iot-rynok/news/ict-moscow-i-dit-moskvy-sozdali-bazu-praktik-umnogo-goroda/> (дата обращения: 28.01.2020).
6. «Банк решений Умного города» [Электронный ресурс]. Режим доступа: <https://russiasmartcity.ru/> (дата обращения: 13.03.2020).
7. База практик умного города [Электронный ресурс]. Режим доступа: <https://ict.moscow/projects/smart-cities/> (дата обращения: 13.03.2020).
8. Базы данных. Практическое применение СУБД SQL и NoSQL-типа для проектирования информационных систем: учеб. пособие / Мартишин С.А., Симонов В.Л., Храпченко М.В. М.: ИД «ФОРУМ»: ИНФРА-М. 2018. 368 с.
9. The database for modern applications – официальный сайт компании [Электронный ресурс]. Режим доступа: <https://www.mongodb.com/> (дата обращения 28.01.2020).



Рюмина Елена Витальевна

Год рождения: 1991

Санкт-Петербургский институт информатики
и автоматизации Российской академии наук,
программист,

Университет ИТМО,

факультет информационных технологий

и программирования,

студент группы №М41211,

направление подготовки: 09.04.02 – Информационные
системы и технологии,

e-mail: riumina_ev@mail.ru



Карпов Алексей Анатольевич

Год рождения: 1978

Санкт-Петербургский институт информатики
и автоматизации Российской академии наук,

д.т.н., доцент,

e-mail: karпов@iias.spb.su

УДК 004.932.2

МЕТОД ИЗВЛЕЧЕНИЯ ИНФОРМАТИВНЫХ ВИДЕОПРИЗНАКОВ ДЛЯ РАСПОЗНАВАНИЯ ЭМОЦИЙ

Е.В. Рюмина

Научный руководитель – д.т.н., доцент А.А. Карпов

Работа выполнена при финансовой поддержке Российского научного фонда РФФ №18-11-00145 «Разработка и исследование интеллектуальной системы для комплексного паралингвистического анализа речи».

Аннотация

В работе рассмотрено применение метода Active shape model для извлечения информативных видеопризнаков выражений лиц. Использование Active shape model предложено также в качестве детектирования областей лиц. Применение Active shape model для извлечения признаков позволяет рассчитывать расстояние между ключевыми точками и оценивать их зависимость в динамике. Использование метода позволяет адаптировать классификатор к разного рода окклюзиям. Представлено сравнение точности классификации эмоций человека при использовании различных методов детектирования лица, различных извлеченных признаках и современных результатов, достигнутых другими авторами.

Ключевые слова

Распознавание выражений лица, классификация, извлечение признаков, active shape model, лицевые ориентиры.

Распознавание эмоций по выражениям лица человека является важной научно-исследовательской проблемой, которая находит свое применение в таких областях как компьютерное зрение, искусственный интеллект и когнитивные науки. На сегодняшний день многие ученые и исследователи уделяют большое внимание решению задач автоматического распознавания выражений лица, однако сталкиваются с рядом проблем, таких как угол поворота головы, освещение, окклюзии, внутриклассовые различия и межклассовые сходства и др.

Рассматриваемый в данной статье метод *Active shape model* (далее *ASM*) [1] обходит такие проблемы, как освещение и окклюзии, что делает метод *ASM* широко используемым для решения задач распознавания выражений лица.

ASM обнаруживает края объектов при помощи лицевых ориентиров, которые представляют собой цепочку последовательности 68-ми ключевых точек, очерчивающих форму и все подвижные части лица, необходимые для распознавания выражений эмоций.

Преимуществами метода *ASM* являются: простота реализации, скорость и точность детектирования ключевых точек.

К недостаткам метода *ASM* можно отнести необходимость предварительного детектирования области лица.

В данном исследовании в качестве метода предварительного детектирования области лица выбран метод *Histogram of oriented gradients* (далее, *HOG*) [2], так как реализация данного метода проста, осуществляется через библиотеку *Dlib* и имеет хорошую скорость детектирования лица.

Для реализации метода *ASM* использовалась также библиотека *Dlib*. По координатам ключевых точек можно вычислить положение точек относительно друг друга [3]. Для этого необходимо вычислить среднее значение по обеим осям, результатом будет нахождение центра масс всех ключевых точек. Затем измеряется расстояние от центра масс до каждой точки.

Формула для нахождения координат центров масс (1):

$$x_{mean} = \frac{\sum_{i=1}^n x_i}{n}; y_{mean} = \frac{\sum_{i=1}^n y_i}{n}, \quad (1)$$

где, n – 68-мь ключевых точек;

x_i и y_i – координаты 68-ми ключевых точек.

Формула для нахождения расстояния между центром масс и ключевыми точками (2):

$$x'_i = x_i - x_{mean}; y'_i = y_i - y_{mean}, \quad (2)$$

Формула для нахождения евклидова расстояния имеет вид (3):

$$d_i = \sqrt{(x_i^2 - x_{mean}^2) + (y_i^2 - y_{mean}^2)}, \quad (3)$$

Так как лица могут быть наклонены, следует учитывать угловое смещение β ключевых точек по формуле (4):

$$\beta = \frac{atan2(x'_i, y'_i) * 180}{\pi}, \quad (4)$$

где, *atan2* – возвращает угол θ между лучом в точке (x, y) и положительной осью x , ограниченной $(-\pi, \pi]$.

Так как пропорции лица у разных людей различны, то изображения локализованных областей лиц будут иметь разное расширение, а расстояниями между координатами ключевых точек будут смещены в зависимости от пропорций лица, поэтому необходимо осуществлять масштабирование изображений лиц и найденных координат. Масштабирование изображений осуществляется по формуле (5) к разрешению 48×48 пикселей.

$$x_{norm} = \frac{x * size_{norm}}{width_{image}}; y_{norm} = \frac{y * size_{norm}}{height_{image}}, \quad (5)$$

где, x, y – координаты ключевых точек или лица;

$size_{norm}$ – размер, к которому необходимо осуществить масштабирование;

$width_{image}$ и $height_{image}$ – ширина и высота исходного изображения.

Метод ASM сравнивается с методами *Single shot multibox detector (SSD)* [4] и HOG по точности и скорости детектирования областей лиц, так как на данный момент эти методы показывают высокую точность и скорость обнаружения областей лиц.

Применяя для вычисления формулы (1-5), были получены:

1) векторы признаков, содержащие информацию о координатах точек, евклидовым расстоянием между ними и центром масс, а также угловое смещение ключевых точек;

2) изображения в оттенках серого, где на черном фоне расположены координаты лица, центр масс и расстояния между координатами и центром масс (далее – маски);

3) локализованные области лиц в оттенках серого, извлеченных методами HOG, HOG+ASM и SSD. Локализация областей методами HOG и SSD осуществлялась по найденным координатам, тогда как использование метода ASM позволило учитывать область лица по максимальным и минимальным значениям координат найденных ключевых точек.

Исследования проводились на базе данных CREMA-D [5], которая содержит 7442 аудио-видео записей с разрешением 480×360 пикселей 91-го актера (48 мужчин и 43 женщины). База данных содержит шесть эмоций: гнев, отвращение, страх, радость, грусть, нейтральность.

Для дальнейшего анализа из видео извлекался каждый пятый кадр, всего было получено 111 184 выражений лиц 91-го актера. Так как в базе данных CREMA-D не предусмотрено деление данных на обучающий и тестируемый наборы, то извлеченные изображения выражений лиц были поделены на обучающий (80%) и тестируемый (20%) наборы.

В таблице представлены результаты применения предложенного метода детектирования области лица по точности и скорости, а также результаты, полученные с помощью методов SSD и HOG. Точность находится по формуле (6):

$$\text{Точность} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (6)$$

где, TP – истинно-положительные;

FP – ложно-положительные;

TN – истинно-отрицательные;

FN – ложно-отрицательные.

Как видно из таблицы, скорость детектирования лица значительно выше у метода HOG. Комбинация методов HOG и ASM, уступает по скорости методу HOG, так как необходимо дополнительное время для обнаружения ключевых точек. Метод SSD

показывает точность детектирования лица ненамного выше методов HOG и HOG+ASM.

Таблица

Сравнение методов детектирования области лица

Метод	Скорость, FPS	Количество найденных лиц	Точность, обнаружения области лица, %
SSD	7.7	111 161 (23 FN)	99.98
HOG	14.2	111 161 (23 FN, 26 FP)	99.96
HOG + ASM	13.8	111 161 (23 FN, 31 FP)	99.96

Для оценки точности классификации эмоций по выражениям лиц масштабированные изображения областей лиц, полученные методами SSD, HOG и HOG+ASM, а также векторы признаков и маски признаков, полученные методами HOG+ASM, подавались на сверточную нейронную сеть (*Convolutional neural network, CNN*), предложенную авторами статьи [6]. CNN обучалась на 40 эпохах, размер батча 64, оптимизатор AdamW со скоростью обучения 0.001 и сокращением веса 0.00005.

На рисунке представлены результаты экспериментов, полученные за счет комбинирования извлеченных признаков из локализованных областей лиц, векторов и масок признаков и CNN, а также результаты, полученные авторами статей [5, 7].

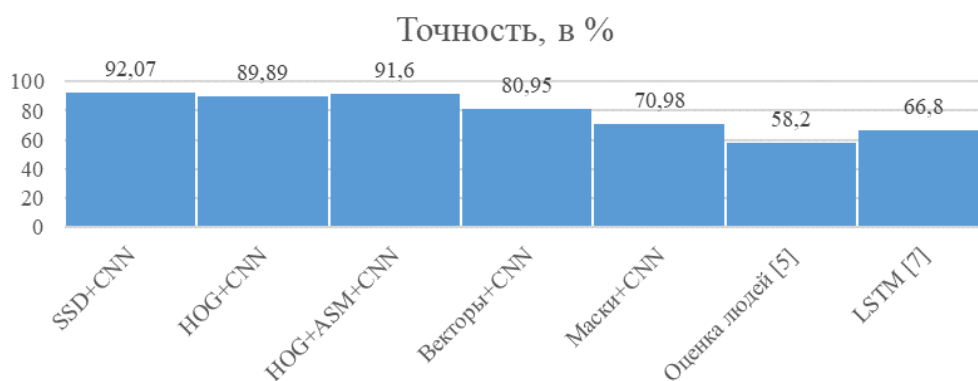


Рисунок. Результаты экспериментов

По рисунку можно заметить, что точность классификации эмоций по извлеченным признакам из локализованных областей лиц, детектируемых методами HOG, HOG+ASM и SSD, выше, чем извлеченные признаки методом ASM, сохраненные в векторах и масках признаков. Несмотря на это, точность классификации по векторам и маскам признаков превышает точность, достигнутую авторами в статьях [5, 7]. По полученным результатам можно заметить, что извлечение признаков из локализованных областей лиц методом SSD, показывает лучшую точность классификации эмоций, чем метод HOG и ASM на 0.47%. А также комбинация методов HOG+ASM позволяет увеличить точность классификации эмоций по сравнению с методом HOG на 1.71%. Можно сделать вывод: несмотря на то, что точность классификации эмоций по извлеченным признакам из локализованных областей лиц с использованием метода SSD выше точности, полученной при использовании комбинации методов HOG и ASM, применение метода ASM в дополнении к методу HOG, за счет своей простоты реализации и скорости детектирования ключевых точек, является более эффективным методом для локализации областей лиц.

Извлечение таких признаков, как расстояние между ключевыми точками и центром масс, а также угловое смещение, в сравнении с извлеченными признаками из локализованных областей лиц, являются недостаточными признаками для классификации эмоций, так же, как и маски признаков. В связи с этим необходимо проводить дальнейшие исследования в определении зависимостей между ключевыми точками лица для последующей классификации эмоций.

Литература

1. Cootes T.F., Taylor C.J., Cooper D.H., Graham J. Active shape models-their training and application // *Computer vision and image understanding*. 1995. V. 61. №1. P. 38–59.
2. Déniz O., Bueno G., Salido J., De la Torre F. Face recognition using histograms of oriented gradients // *Pattern Recognition Letters*. 2011. V. 32. № 12. P. 1598–1603.
3. Van Gent P. Emotion recognition using facial landmarks python DLib and OpenCV // *A tech blog about fun things with Python and embedded electronics*. 2016.
4. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.Y., Berg A.C. SSD: single shot multibox detector // *14th European conference on computer vision*. Amsterdam, The Netherlands. 2016. V. 9905. P. 21–37.
5. Cao H., Cooper D.G., Keutmann M.K., Gur R.C., Nenkova A., Verma R. CREMA-D: Crowd-sourced emotional multimodal actors dataset // *IEEE transactions on affective computing*. 2014. V. 5 № 4. P. 377–390.
6. Talegaonkar I., Joshi K., Valunj S., Kohok R., Kulkarni A. Real time facial expression recognition using deep learning // *Proc. of International Conference on Communication and Information Processing (ICCIP)*. 2019. [Электронный ресурс]. Режим доступа: <https://ssrn.com/abstract=3421486> (дата обращения 05.02.2019)
7. Ghaleb E., Popa M., Asteriadis S. Multimodal and Temporal Perception of Audio-visual Cues for Emotion Recognition // *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 2019. P. 552–558.



Тимофеева Ангелина Олеговна

Год рождения: 1997

Университет ИТМО,

Институт дизайна и урбанистики,

магистрант группы № С41042,

направление подготовки: 09.04.03 – Прикладная информатика,

e-mail: linatim@yandex.ru



Кононова Ольга Витальевна

Год рождения: 1964

Университет ИТМО,

Институт дизайна и урбанистики,

к.э.н., доцент,

e-mail: kononolg@yandex.ru



Матросова Евгения Викторовна

Год рождения: 1986

СПб ГБУЗ МИАЦ,

ведущий инженер отдела

развития информатизации,

e-mail: MatrosovaE@spbmiac.ru

УДК 004.9

**АНАЛИЗ БИЗНЕС И ИНФОРМАЦИОННЫХ
ПРОЦЕССОВ В РАМКАХ ТЕЛЕМЕДИЦИНСКИХ
КОНСУЛЬТАЦИЙ «ВРАЧ-ПАЦИЕНТ»**

А.О. Тимофеева, Е.В. Матросова, О.В. Кононова

Научный руководитель – к.э.н., доцент О.В. Кононова

Работа выполнена в рамках темы НИР «Анализ бизнес и информационных процессов в рамках телемедицинских консультаций «Врач-пациент» магистерской программы «Цифровые технологии умного города».

Аннотация

В статье представлены результаты оценки степени вхождения телемедицины в концепцию электронного здравоохранения. Выделены и описаны формы информационного взаимодействия, а также, описаны бизнес-процессы в телемедицинских консультациях «Врач-пациент».

Ключевые слова

Телемедицина, телемедицинские технологии, телемедицинская консультация, электронное здравоохранение, цифровое здравоохранение.

Сегодня дистанционные консультации с врачами воспринимаются как неотъемлемая часть оказания медицинской помощи. Для успешного функционирования телемедицинских систем важно, чтобы врачи научились удаленно работать с пациентами. Также, пациенты должны представлять, что они получают от телемедицинской консультации, какие у них есть возможности и ограничения при получении такой медицинской помощи [1].

Существует ряд требований, ограничивающих проведение телемедицинских консультаций на территории Российской Федерации нормативно-правовыми актами [2]. Указанные требования в обязательном порядке должны быть использованы при описании и проектировании сервиса телемедицинских консультаций «Врач-пациент»:

1. Первичный осмотр пациента в удаленном формате, а также, удаленная постановка диагноза запрещены на уровне законодательства.

2. Необходимо наличие регистрации медицинской организации в Федеральном реестре медицинских организаций Единой государственной информационной системы в сфере здравоохранения.

3. К осуществлению консультирования могут привлекаться только действующие сотрудники – медицинские работники консультирующей медицинской организации, сведения о которых внесены в Федеральный регистр медицинских работников.

4. Врач должен обладать правом усиленной квалифицированной электронной подписи, а пациент и/или его законный представитель правом простой электронной подписи, посредством применения Единой системы идентификации и аутентификации.

30 января 2018 года на заседании Комиссии по обеспечению устойчивого развития экономики и социальной стабильности в Санкт-Петербурге был рассмотрен и утвержден приоритетный городской проект – «Электронное здравоохранение». Одной из задач проекта является реализация 45 электронных сервисов для врачей, пациентов и руководителей здравоохранения, в том числе – сервис телемедицинских консультаций «Врач-пациент» [3].

Для оценки степени вхождения телемедицины в концепцию электронного здравоохранения был проанализирован поток публикаций из отечественных СМИ за период с 2008 по 2018 год включительно, которые содержатся в цифровом информационном ресурсе Интегрум, а также из научных публикаций за этот же период (научная электронная библиотека eLibrary). Было выполнено два расширенных запроса: «Цифровое здравоохранение» OR «Электронное здравоохранение»; «Телемедицина». Результаты запросов за 2008-2018 представлены на рис. 1.

Полученные данные позволяют сделать следующие выводы:

– резкий рост как общественного, так и научного интереса к тематике электронного здравоохранения, цифровой медицины наблюдается с 2015 года, что объясняется сначала широким обсуждением, а затем принятием нормативно-правовых актов, регулирующих оказание медицинской помощи с применением телемедицинских технологий;

– рост научного интереса к проблемам телемедицины начался с 2013 года в отличие от общественного интереса (2015), сохраняет устойчивую положительную динамику, в абсолютных значениях превышает показатели электронного и цифрового здравоохранения;

– вопросы электронного здравоохранения в большей степени обсуждаются научным сообществом, а телемедицина находит больший отклик в СМИ;

– телемедицина является ключевым направлением в концепции электронного здравоохранения.

Экспертный анализ показал, что наибольший процент работ относится к предметным областям: Законы и юриспруденция (31%), Медицина и здравоохранение (30%). Анализ результатов запросов также выявил, что в обсуждении путей развития

электронного здравоохранения принимают наиболее активное участие представители высших учебных заведений, академического сообщества. В меньшей степени авторами научных публикаций являются представители медицинских организаций, бизнеса и государственных структур.



Рис. 1. Поток публикаций из отечественных СМИ и поток научных публикаций

Запрос «Телемедицинская консультация» OR «Telemedicine consultation» в сделанных ранее подборках материалов (запросы «Электронное здравоохранение» OR «E-health», «Цифровое здравоохранение» OR «Электронное здравоохранение») показал незначительный интерес к тематике электронных медицинских консультаций за весь период наблюдений: в 2008 году – 4 научных публикаций, в 2019 году число публикаций возросло до 29, что говорит о том, что данное научное направление находится на раннем этапе развития, с тенденцией повышения значимости тематики. Таким образом, задача проектирования и реализации электронных сервисов информационного взаимодействия в рамках телемедицинской консультации является не только актуальными, но и малоизученным направлением Телемедицины.

В рамках телемедицинской консультации «Врач-пациент» можно выделить следующие формы взаимодействия:

1. Врач – пациент (и/или его законный представитель);
2. Пациент – администратор;
3. Врач – администратор;
4. Пациент (и/или его законный представитель) – программный агент;
5. Программный агент – программный агент.

Чтобы обеспечить информационное взаимодействие врача и администратора системы (который также является медицинским работником) друг с другом, или с пациентом, им необходимо наличие компьютера, подключенного к защищенной сети. Также, необходим доступ к телемедицинскому portalу через личный кабинет, или доступ к Медицинской информационной системе, эксплуатируемой в медицинской организации, предоставляющей пациентам услуги телемедицинских консультаций (при условии, что Медицинская информационная система интегрирована с сервисом телемедицинских консультаций).

В ходе очной консультации с пациентом врач принимает решение, о необходимости проведения в дальнейшем телемедицинской консультации. Врач назначает дистанционную консультацию на определённое время в электронном виде, а также оставляет комментарий с «Памяткой пациенту» (при необходимости). В памятке содержится информация о том, какие документы необходимо приложить пациенту на телемедицинскую консультацию, а также, информация о том, как нужно подготовиться к консультации.

Пациент из личного кабинета получает уведомление о назначенной консультации, в соответствии с 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации», знакомится с добровольным информированным согласием на получение консультативной медицинской помощи в формате телемедицинской консультации и принимает его условия, подтверждает готовность к получению назначенной врачом телемедицинской консультации (или может изменить время консультации до определённого момента).

Врачу и пациенту перед началом консультации приходит уведомление (по электронной почте и в личный кабинет) о предстоящей консультации. Пациент и врач в своём личном кабинете подключаются к виде-конференц-связи в назначенное время. Во время консультации пациент и врач могут воспользоваться чатом для передачи текстовой информации и обмена документами. По результату проведённой консультации, врач формирует консультативное заключение, которое пациент может просмотреть и скачать в соответствующем разделе личного кабинета. Пациент может запросить (оставить заявку) дополнительную консультацию, указав желаемую дату и время.

Пациент может получить консультацию не только в режиме реального времени, но и в режиме отложенных консультаций (при условии, если медицинская организация поддерживает данный режим), т. е. пациент при формировании заявки, описывает необходимые сведения и прикрепляет необходимые документы, результаты медицинских обследований и анализов или снимки (прикрепление осуществляется, как в виде ссылок на документы из электронной медицинской карты пациента, так и в виде отдельных файлов). Врач-консультант получает заявку, знакомится с предоставленными сведениями и, как и в случае с онлайн консультацией, формирует консультативное заключение.

По результатам консультации врач может сформировать для пациента электронное направление на оказание специализированной медицинской помощи (как на определенное время, так и без указания времени), которое будет доступно пациенту в личном кабинете. Также, по результатам консультации, пациент может получить в личном кабинете, сформированный врачом, электронный рецепт или медицинскую справку, которые можно сохранить и распечатать.

Во взаимодействии между врачом и пациентом отдельную роль может играть администратор. Он может взять на себя часть обязанностей, таких как согласование с пациентом даты и времени телемедицинской консультации, распределение полученных от пациентов дополнительных заявок на телемедицинскую консультацию[4].

На рис. 2 представлена схема информационного взаимодействия в телемедицинской консультации «Врач-пациент» в режиме реального времени.

Для назначения дистанционной консультации врач обращается к сервису «Телемедицинская консультация» (ТМК) из портала врача, а «ТМК» передает информацию в подсистему уведомлений. Пациент получает уведомление и через личный кабинет портала пациента обращается к «ТМК», где принимает условия проведения консультации и подтверждает её (а также, может сделать запрос на дополнительную консультацию). Далее, информация отправляется в подсистему

уведомлений, после чего, врачу приходит уведомление о подтверждённой консультации.

В сервисе «ТМК» врач и пациент могут подключиться к видео-конференц-связи и/или связаться в чате. Там же, врач формирует и отправляет пациенту заключение, которое отправляется пациенту и в Электронную медицинскую карту Петербуржца.

Исходя из проведенного исследования, можно сделать вывод, о том, что, на сегодняшний день, развитие направления телемедицины – телемедицинские консультации, актуально и востребовано. Нормативно-правовая база Российской Федерации позволяет проводить дистанционные консультации «Врач-пациент», при этом, накладывает ряд требований и ограничений, которые необходимо учитывать при описании и проектировании сервиса.

Результаты исследования бизнес и информационных процессов в телемедицинских консультациях «Врач-пациент» будут использованы в СПб ГБУЗ «Медицинский информационно-аналитический центр» для создания нового электронного сервиса, в рамках приоритетного проекта Санкт-Петербурга «Электронное здравоохранение».

На следующем этапе работы планируется сформулировать функциональные требования для сервиса телемедицинских консультаций «Врач-пациент».

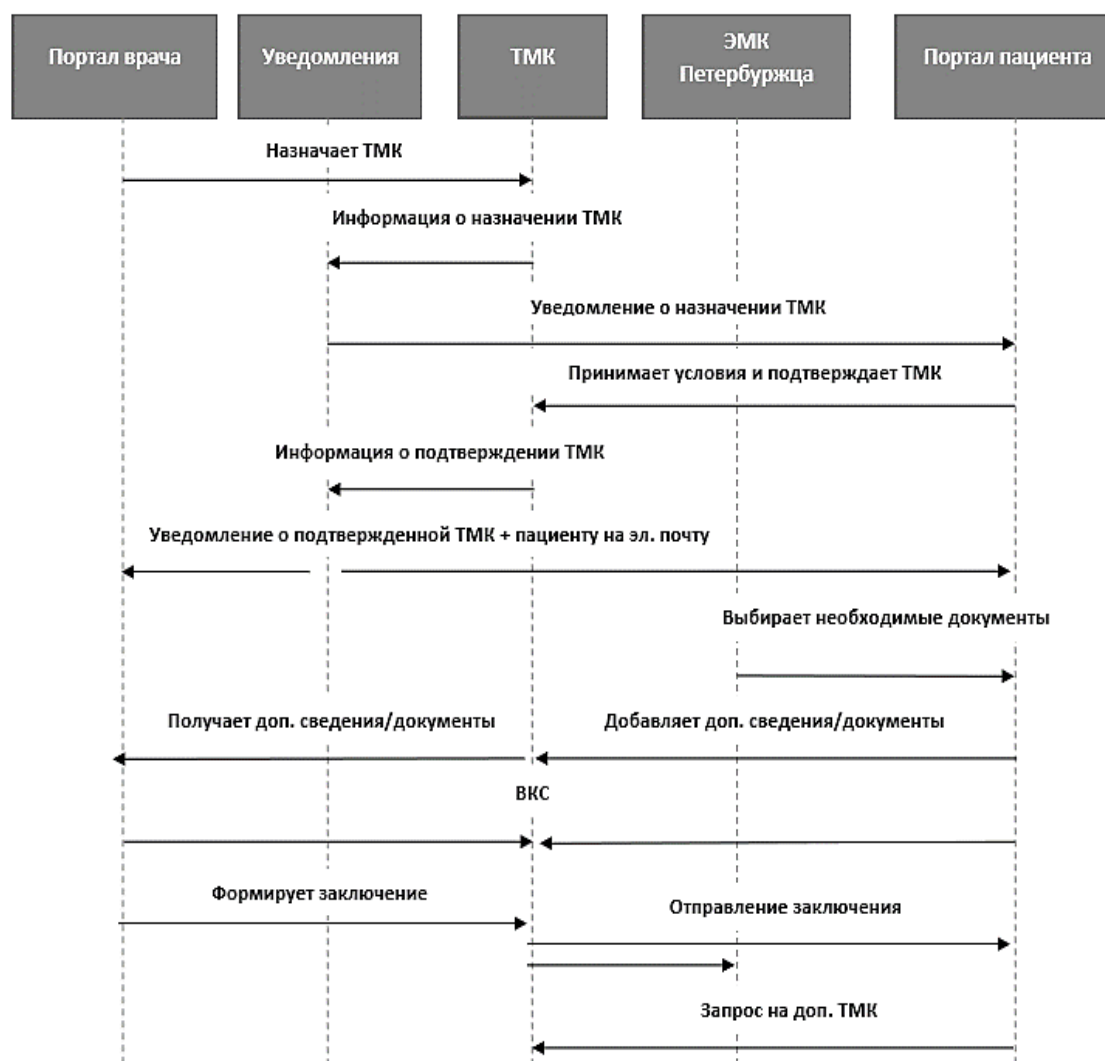


Рис. 2. Схема информационного взаимодействия в телемедицинской консультации «Врач-пациент» в режиме реального времени

Литература

1. Пузин С.Н., Сертакова О.В., Решетов Д.Н. Телемедицина как вектор инновационного развития системы оказания услуг в сфере здравоохранения // Вестник всероссийского общества специалистов по медико-социальной экспертизе, реабилитации и реабилитационной индустрии. Москва: Общероссийская общественная организация “Всероссийское общество специалистов по медико-социальной экспертизе, реабилитации и реабилитационной индустрии ВРОСЭРРИ” 2018. С. 65–73.
2. Прилуков М.Д. Проблемы правового регулирования телемедицины. Российский и международный опыт // Вестник Нижегородского университета им. Н. И. Лобачевского. Нижний Новгород: Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, 2018. С. 136–141.
3. Борисов Д.Н., Иванов В.В. Организационная телемедицина // Врач и информационные технологии. Москва: Издательский дом “Менеджер здравоохранения” 2017. С. 112–120.
4. Владзимирский А.В. Первичная телемедицинская консультация «пациент-врач»: первая систематизация методологи // Журнал телемедицины и электронного здравоохранения. Москва: ГБУЗ “Научно-практический центр медицинской радиологии ДЗМ” 2017. С. 109–120.



Тимофеева Елена Павловна

Год рождения: 1997
Университет ИТМО,
факультет информационных технологий
и программирования,
студент группы №М41212,
направление подготовки: 09.04.02 – Информационные
системы и технологии,
e-mail: etimofeeva277@gmail.com



Шуранов Евгений Витальевич

Год рождения: 1980
Университет ИТМО,
факультет информационных технологий
и программирования,
к.т.н.,
e-mail: evshuranov@itmo.ru

УДК 004.93

ДЕТЕКТИРОВАНИЕ АКУСТИЧЕСКИХ СОБЫТИЙ В БЫТОВЫХ УСЛОВИЯХ

Е.П. Тимофеева

Научный руководитель – к.т.н. Е.В. Шуранов

Аннотация

В работе выполнен обзор метода детектирования акустических событий в бытовых условиях, в основе которого лежит применение рекуррентных сверточных нейронных сетей, а также были изучены способы постобработки данных для улучшения детектирования.

Ключевые слова

Детектирование, акустические события, сверточные нейронные сети, рекуррентные нейронные сети.

В настоящее время продолжают активно развиваться технологии анализа речевых сигналов, классификации акустических событий, фильтрации, а также голосовые интерфейсы человеко-машинного взаимодействия, решающие задачи распознавания речи и голосовой биометрии. Для соответствующих систем важную роль играют детекторы акустических событий. На сегодняшний день известно множество методов детектирования речи, однако, разработка новых подходов по-прежнему сохраняет свою актуальность.

Чаще всего для решения задач детектирования применяют сверточные нейронные сети. Так как они используются для распознавания образов, однако в этом случае невозможно учитывать влияние паттернов друг на друга. Для этого придётся анализировать полученные паттерны другой сетью и с другой архитектурой – например, рекуррентной нейронной сетью. Рекуррентные нейронные сети чаще используются для анализа таких данных как текст или аудиофайлы, где структура данных предполагает зависимости данных от значений друг друга. То есть нейронная

сеть анализирует и запоминает такие зависимости, обучаясь находить закономерности в них. Объединение этих нейронных сетей даст устойчивость модели и преимущество в поиске закономерностей при обучении. Такая модель эффективна при классификации и при расставлении меток.

Эффективным методом для получения признаков, является метод *log-mel energies*. Так как пики в спектре *log Mel-filter bank* являются важными признаками в характеристике звуков речи. Однако низкоэнергетические возмущения в спектре мощности могут стать значимыми после *log* сжатия. Даже если пики в спектре остаются константными, возмущениями с малой энергией в спектре мощности могут создавать огромные вариации *cepstral* коэффициентов. Возведение в степень спектра *log Mel-filter bank* перед вычислением *cepstral* коэффициентов может значительно снизить чувствительность этих коэффициентов к ненужным возмущениям низкой энергии. Спектр модуляции *Mel-cepstrum* рассчитывается на основе обработанного *cepstral* коэффициента, что приводит к устойчивости к шуму. В экспериментах с речевыми сигналами было показано, что предлагаемые методы, основанные на функциях, дают значительное увеличение эффективности распознавания речи в условиях нестационарного шума [1].

В данной работе рассматривается один из методов детектирования, предложенный участниками конкурса *DCASE 2019*, в основе которого лежит применение рекуррентных сверточных нейронных сетей. Задача конкурса заключается в детектировании акустических событий в бытовых условиях с использованием слабо маркированных данных (без отметок времени) (рис. 1). Целью систем является предоставление не только класса события, но и границ времени события, учитывая, что в аудиозаписи могут присутствовать несколько событий, а также, что эти события могут пересекаться. Так как есть пересекающиеся события, участниками были предложены некоторые улучшения. В основном это оптимизация постобработки данных [2].

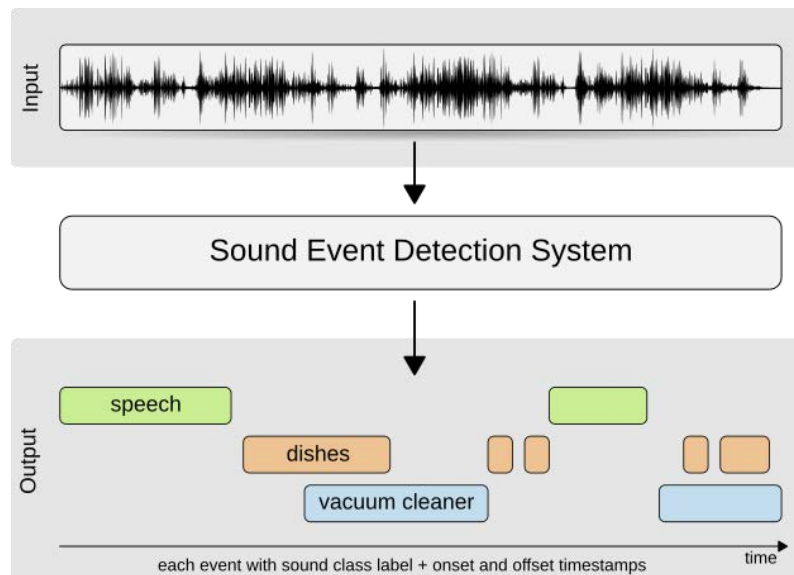


Рис. 1. Система детектирования акустических событий

В этой работе сделан обзор системы *PELLEGRINI_IRIT_task4_1* по отчету участника «*multi-task learning and post processing optimization for sound event detection*» [3, 4].

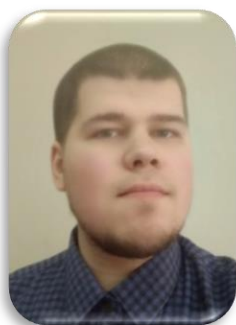
В отчете участниками были предложены три модели для обнаружения акустических событий. Эти методы были адаптированы для использования различных видов подмножеств, в том числе сильно и слабо аннотированных. Был использован многозадачный подход к обучению, чтобы использовать эти смешанные аннотированные данные. Кроме того, в основном усилия были сосредоточены на изучении методов постобработки для оценки наиболее эффективных пороговых значений для принятия решений об обнаружении и локализации акустических событий. Были представлены различные алгоритмы оптимизации, которые значительно повысили производительность модели. Лучшая модель выполняет локализацию и аудиометки в одной ветке. В наборе данных для проверки правильности выполнения задачи 4 DCASE 2019 было получено 39,9% баллов, по которым базовая система дала 23,7%. Без пороговой оптимизации для мечения или локализации та же модель достигла F-показателя 15,9%. Результаты представлены на рис. 2.

Class	Baseline		Model 1		Model 2		Model 3	
	Eb	Sb	Eb	Sb	Eb	Sb	Eb	Sb
Alarm bell ringing			42.3%	76.5%	31.8%	71.3%	25.5%	74.2%
Blender	-	-	40.4%	57.9%	27.5%	45.3%	30.1%	52.9%
Cat	-	-	45.9%	57.8%	28.2%	52.9%	19.4%	54.4%
Dishes	-	-	26.6%	47.9%	21.6%	49.0%	24.7%	46.7%
Dog	-	-	17.8%	31.8%	19.3%	49.5%	13.4%	56.6%
Electric shaver toothbrush	-	-	48.5%	60.6%	52.3%	64.5%	49.1%	65.4%
Frying	-	-	35.5%	61.4%	47.9%	63.1%	44.3%	60.4%
Running water	-	-	30.7%	61.6%	17.3%	54.1%	18.5%	59.3%
Speech	-	-	39.9%	79.9%	40.0%	78.2%	44.9%	78.8%
Vacuum cleaner	-	-	68.7%	69.7%	51.8%	66.9%	52.1%	66.2%
Macro-F1	23.7%	55.2%	39.9%	60.5%	33.8%	59.5%	32.2%	61.5%

Рис. 2. Результаты для трех моделей. Eb означает «на основе событий», а Sb - на основе сегментов макро-F1-оценки

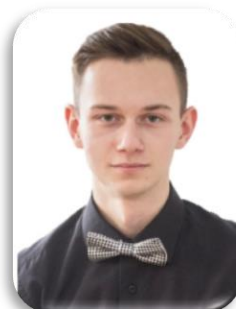
Литература

1. Tyagi V., Wellekens C. On desensitizing the Mel-Cepstrum to spurious spectral components for Robust Speech Recognition, in Acoustics, Speech, and Signal Processing. // IEEE International Conference. 2005. Vol. 1. P. 529-532.
2. DCASE2019 challenge, Sound event detection in domestic environments, Results [Электронный ресурс]. Режим доступа: <http://dcase.community/challenge2019/task-sound-event-detection-in-domestic-environments-results> (дата обращения: 10.12.2019).
3. Leo Cances, Patrice Guyot, Thomas Pellegrini. Multi-task learning and post processing optimization for sound event detection // Technical Report, Sound event detection in domestic environments, DCASE2019. 2019.
4. Leo Cances, Patrice Guyot, Thomas Pellegrini. Evaluation of post-processing algorithms for polyphonic sound event detection. arXiv:1906.06909.



Тропников Александр Сергеевич

Год рождения: 1996
Университет ИТМО,
Институт дизайна и урбанистики,
студент группы № С42551,
направление подготовки: 09.04.03 – Прикладная информатика,
e-mail: petyabegen@mail.ru



Беген Петр Николаевич

Год рождения: 1996
Университет ИТМО,
Институт дизайна и урбанистики,
студент группы № С42551
направление подготовки: 09.04.03 – Прикладная информатика,
e-mail: astropnikov@corp.ifmo.ru



Низомутдинов Борис Абдуллохонович

Год рождения: 1990
Университет ИТМО,
Институт дизайна и урбанистики,
e-mail: boris-wels@yandex.ru



Углова Анна Борисовна

Год рождения: 1989
РГПУ им. А.И. Герцена,
кафедра психологии профессиональной деятельности,
к.психол.н.,
e-mail: anna.uglova@list.ru

УДК 004.023

**ПРИМЕНЕНИЕ ЦИФРОВЫХ ОТПЕЧАТКОВ
ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ ДЛЯ
ПРОГНОЗИРОВАНИЯ ИХ ИНФОРМАЦИОННОГО ОБРАЗА
А.С. Тропников, П.Н. Беген, Б.А. Низомутдинов, А.Б. Углова
Научный руководитель – к.полит.н. А.В. Чугунов¹
1 – Университет ИТМО**

Аннотация

Работа выполнена с использованием промежуточных результатов третьего этапа магистерского исследования «Проектирование автоматизированной системы выявления психоэмоциональных особенностей информационного образа пользователя социальных сетей».

В статье представлен краткий обзор возможностей использования автоматизированных средств анализа цифровых отпечатков пользователей социальных сетей с целью прогнозирования их личностных черт.

Ключевые слова

Социальные сети, психоэмоциональные особенности, социально-психологические характеристики, цифровые отпечатки, информационный образ.

Все большее количество социальных интеракций и общественных действий фиксируется в сети Интернет самими пользователями. Люди фотографируют любимые места во время отпуска, комментируют записи в блогах, оставляют контактные данные в социальных сетях. Совершая подобные действия, пользователи оставляют «цифровые отпечатки» – набор разнообразных данных, начиная от «лайков» и фотографий, заканчивая данными геолокации и MAC-адресов устройств, с которых осуществлялся доступ в Интернет.

«Цифровые отпечатки» используются в: таргетинговой рекламе и маркетинге [1], рекомендательных системах, персонализированных поисковых запросах, скорринговых банковских системах, системах рекрутинга персонала.

В последние годы проведены несколько крупных исследований, посвященных определению различных "атрибутов" пользователей социальных сетей (пол, возраст, семейный статус, образование и т.п.), где в качестве исходных данных использованы только оставленные пользователями «цифровые отпечатки» [2, 3]. Благодаря этому, мы имеем возможность с высокой точностью предположить пол или семейный статус человека на основе его «лайков», даже в том случае, если он сам не вносил информацию о себе [4, 5]. Стало возможным на основе пользовательских музыкальных предпочтений или семантики комментариев установить отдельные психологические черты пользователя.

Наше исследование посвящено расширению списка «цифровых отпечатков» путем включения в собираемый массив данных новых данных, полученных машинным анализом изображений. А также поиску новых корреляционных связей между психологическими особенностями людей и их данных в профиле социальных сетей

При помощи психодиагностического инструментария был разработан авторский тест, состоявший из более чем 200 вопросов и нацеленный на определение уровня зависимости пользователя от социальных сетей, его интересов, желаний, социального круга, психологических особенностей, самооценки (ценностный опросник С. Шварца; методику исследования самоотношения С.Р. Пантелеева; тест диагностики степени удовлетворенности потребностей Маслоу).

В тестировании приняли участие 253 человека, большинство из которых являлись студентами университета. В ходе обработки результатов тестирования, были отобраны 180 человек, предоставивших доступ к своим профилям в социальной сети ВКонтакте. Используя средства парсинга, был собран массив данных, содержащих всю общедоступную информацию, размещенную на странице профиля: имя, фамилия, количество подписок на сообщества, количество аудиозаписей и видеозаписей, аватары, статус о работе и семейном положении, возраст, количество друзей и т.д.

Применяя облачный сервис машинного анализа изображений Azure, собранные аватары также были проанализированы с целью получения новых данных для добавления их в массив. Таким образом, была получена информация об

отображаемых эмоциях на аватаре, определенных предметах, общей композиции, интенсивности цвета, количестве людей и их возрасте. Эта информация была добавлена в общий массив данных, публикуемый пользователями в своих профилях.

На финальном этапе исследования, данные, полученные в ходе автоматической выгрузки, были объединены в более крупные категории, описывающие базовые компоненты информационного образа пользователя: информация об имени пользователя (самоименование), информация о возрасте, пространственная локализация, информация о близких отношениях, о профессиональном статусе, ассоциированность в сети, информация об образовании, конфиденциальности, информация о количестве друзей, подписчиков, фотографий, видеозаписей, аудиозаписей, групп, подписок, а также эмоции и объекты, выявленные на фотографиях пользователей (рис. 1).

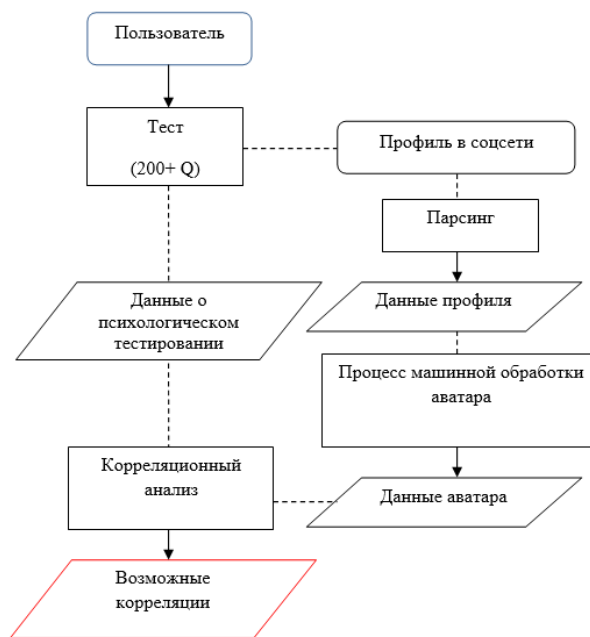


Рис. 1. Методология исследования

В процессе обработки данных профилей с помощью непараметрического однофакторного дисперсионного анализа (критерий Краскела-Уолисса) нами был проверен ряд гипотез о влиянии независимых переменных (пол, возраст, уровень образования) на зависимые переменные, в качестве которых выступали данные, публикуемые пользователями в своих профилях, составляющие информационный образ пользователя. Были выявлены достоверно значимые различия в дисперсиях по следующим параметрам (таблица 1, 2, 3).

В результате сравнительного анализа информационных образов мужчин и женщин, можно говорить о том, что информационный образ женщин более многогранный из-за достоверно большего количества, используемых социальных сетей, в то время как информационный образ мужчин обладает большим семантическим разнообразием, так как они выкладывают больше визуального контента и личных сообщений. В тоже время обнаружили достоверно значимые различия в информационном содержании фотографий пользователей разного пола, что может быть использовано в процессе создания автоматизированного анализа информационного образа пользователей.

Анализ информационного образа пользователей разных возрастных групп показал, что более молодые пользователи достоверно больше времени проводят в

социальных сетях, что соотносится с основной задачей юношеского возраста развитием навыков организации своей социальной жизни и приобретением новых социальных контактов. Вместе с тем в информационном образе пользователей раннего зрелого возраста достоверно в большем количестве представлена фотографическая информация, что, по-видимому, связано с возрастной задачей обретения собственной идентичности, утверждением Я-концепции и является определенным хранилищем опыта. У испытуемых раннего зрелого возраста также достоверно больше друзей в социальной сети, что с одной стороны указывает на более широкий круг социальных связей, а с другой может косвенно указывать на тот факт, что изучаемая нами социальная сеть теряет свою популярность у молодежи.

Сравнение групп испытуемых с разным уровнем образования показало, что в информационном образе людей с высшим образованием достоверно больше визуальной информации и информации о социальных контактах, также они реже пользуются социальными сетями, что с одной стороны связано с возрастными особенностями, а с другой может объясняться более высокой социальной ответственностью, которая не дает возможности тратить время на общение в социальных сетях. При анализе фотографий пользователей было выявлено, что у пользователей с более высоким уровнем образования на аватарах достоверно чаще встречаются предметные теги, что, по-видимому, является отражением уже сложившейся социальной идентичности пользователей, которую они презентуют через релевантные для них объекты социокультурного опыта.

В результате проведенного исследования были найдены различного рода корреляции, демонстрирующие возможность определения отдельных психологических черт пользователя на основе его цифровых отпечатков. Можно сказать, что изучение взаимосвязей компонентов информационного образа и личностных особенностей показало довольно высокие прогностические возможности анализа данных социального профиля.

Дальнейшее направление исследований будет связано с расширением выборки и анализом данных с учетом социально-демографических характеристик пользователей, а также особенностями их самопрезентации в сети.

Таблица 1

Достоверно значимые различия в информационном образе мужчин и женщин

Компоненты информационного образа пользователя	Средние показатели в группе мужчин	Средние показатели в группе женщин	H (Критерий Краскела-Уолиса)	p (Уровень доверительной вероятности)
Количество используемых социальных сетей	2,17	2,54	5,06	0,026
Количество видеозаписей	320,25	140,76	4,43	0,037
Количество заметок	2,98	0,57	4,16	0,043
1 подкластер, объединяющий теги на фотографиях (black, black and white, black hair, white, lip, lipstick, eyes, long hair, portrait, hair)	0,38	1,23	7,57	0,007
4 подкластер тегов (design, logo, minimalist, text)	0,52	0,21	4,02	0,047
5 подкластер тегов (clothing, face, fashion, fashion accessory, girl, human face, outdoor, person, smile, woman, female)	2,65	5,46	53,87	0,001

Таблица 2

Достоверно значимые различия в информационном образе пользователей периода юности (18-23 года) и первого периода зрелости (24-35)

Компоненты информационного образа пользователя	Средние показатели в группе 18-23 года	Средние показатели в группе 24-35 лет	Н (Критерий Краскела-Уолиса)	р (Уровень доверительной вероятности)
Количество времени, проводимое в социальных сетях	5,14	3,81	5,45	0,021
Количество друзей	131,80	342,33	4,49	0,036
Количество фотографий	62,25	194,84	6,73	0,010
Количество фотоальбомов	0,47	1,59	9,51	0,002

Таблица 3

Достоверно значимые различия в информационном образе пользователей с различным уровнем образования

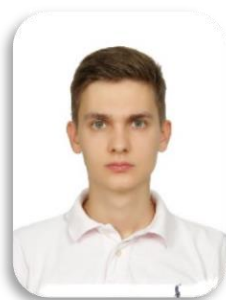
Компоненты информационного образа пользователя	Средние показатели в группе с высшим образованием	Средние показатели в группе без высшего образования	Н (Критерий Краскела-Уолиса)	(Уровень доверительной вероятности)
Количество времени, проводимое в социальных сетях	3,81	4,96	3,98	0,0495
Количество друзей	379,29	140,16	5,38	0,0216
Количество фотографий	203,84	75,13	5,82	0,0169
Количество фотоальбомов	1,64	0,59	7,76	0,0060
1 подкластер тегов на фотографиях (black, black and white, black hair, white, lip, lipstick, eyes, long hair, portrait, hair)	0,48	1,10	3,88	0,0495
2 подкластер тегов (car, drawing, dress, flower, ground, plant, tree, footwear, standing, hiking, sky, jacket, wall)	1,14	0,84	3,93	0,0497

Инструментарий, использованный в данной работе, был отобран на основании сравнительного анализа, основанного на ряде критериев, включая: максимальный объем обрабатываемых данных, скорость обработки данных и формат выгрузки данных.

Наш дальнейший интерес на последующих этапах будет сконцентрирован на проведении качественных исследований с целью выявления паттернов использования групп, а также публикуемого контента.

Литература

1. Chen Y, Pavlov D., Canny JF. 2009 Large-scale behavioral targeting. International Conference on Knowledge Discovery and Data Mining. pp 209–218.
2. Kosinski M., Matz S., Gosling S. et al. Facebook as a social science research tool: Opportunities, challenges, ethical considerations and practical guidelines // American Psychologist. 2015. Vol. 70. N 6. P. 543–556.
3. Xenos S., Ryan T. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage // Computers in Human Behavior. V. 27. N 5. 2011. - P. 1658-1664.
4. Kosinski M., Stilwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior // Proc. the National Academy of Science of the United State of America. 2013. Vol. 110. P. 5802-5805. DOI: 10.1073/pnas.1218772110.
5. Ross C., Orr E.S., Sisic M., Arseneault J.M., Simmering M.G., Orr R.R.: Personality and motivations associated with facebook use // Computers in Human Behavior. 2009. Vol. 25. P. 578-586. DOI: 10.1016/j.chb.2008.12.024.



Фоменко Богдан Михайлович

Год рождения: 1997
Университет ИТМО,
факультет инфокоммуникационных технологий,
студент группы № К42211,
направление подготовки: 11.04.02 – Программное
обеспечение в инфокоммуникациях,
e-mail: fomenko_bogdan@mail.ru



Осипов Никита Алексеевич

Год рождения: 1972
Университет ИТМО,
факультет инфокоммуникационных технологий,
к.т.н., доцент,
e-mail: nikita@ifmo.spb.ru

УДК 004.6

ВАЖНОСТЬ АНАЛИЗА МЕТОДОВ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В УПРАВЛЯЮЩИХ СИСТЕМАХ

Б.М. Фоменко

Научный руководитель – к.т.н, доцент Н.А. Осипов

Аннотация

В данной работе проведен анализ существующих методов обработки больших данных в управляющих системах, показана важность выбора нужного метода анализа больших данных на конкретном примере.

Ключевые слова

Большие данные, методы обработки больших данных, искусственный интеллект.

В наше время существует множество разнообразных методов анализа массивов данных, основу их составляет инструментарий, заимствованный из статистики и информатики (например, машинное обучение). Необходимо так же понимать, что исследователи продолжают создавать новые методики и усовершенствовать существующие. Применение таких методик не ограничивается только большими данными, их так же успешно используют с меньшими по объему массивами. Очевиден тот факт, чем более объемный и диверсифицируемый массив мы анализируем, тем более точные и релевантные данные мы сможем получить на выходе [1-4].

Какие же методы обработки больших данных существуют:

1. Методы класса или глубинный анализ. Анализ структурированных данных с помощью математических моделей, основанных на статистических, вероятностных и оптимизационных методах, является основным предназначением технологии Data Mining. Такую технологию следует использовать с целью выявления в них ранее неизвестных закономерностей, зависимостей и извлечения неожиданной информации.

2. Краудсорсинг. Данная методика позволяет получить данные одновременно из большого, практически неограниченного количества источников. Краудсорсинг

наиболее применим, если необходимо систематизировать значительный объем необработанных данных, которые весьма сложно автоматически отсортировать.

3. А/В-тестирование. Суть данного метода в том, чтобы из всего объема данных выбрать контрольный набор элементов и поочередно сравнить с другими похожими наборами, где было произведено изменение одного из элементов.

4. Прогнозная (предиктивная) аналитика. Данный метод анализа содержит в себе большое количество различных методов статистики, анализа данных и теории игр. Весь этот набор можно использовать для анализа текущих и исторических данных или событий, а также для прогноза таковых в будущем.

5. Машинное обучение (искусственный интеллект). Такая методика анализа и обработки больших данных основывается на эмпирическом исследовании информации и дальнейшем построении алгоритмов самообучения систем. Главная идея машинного обучения в том, чтобы минимизировать участие человека в выявлении закономерностей и принятия решений аналитическими системами.

6. Сетевой анализ. Наиболее популярный метод для исследования социальных сетей – после получения статистических данных необходимо анализировать созданные в сети узлы, то есть взаимодействия между отдельными пользователями и их сообществами.

В общем и целом, все вышеперечисленные методы обработки данных подходят к управляющим системам. Так как разновидностей управляющих систем в мире огромное количество, логично будет предположить, что не все из методов одинаково хорошо будут работать с различными задачами, проще говоря, не всегда метод, который прекрасно показывает себя в работе с одной задачей, будет так же хорошо показывать себя и с другой. Важность анализа методов обработки заключается в выборе наиболее подходящего. Но как же выбрать тот самый метод? На самом деле вначале для этого следует определиться, чем же именно будет заниматься наша управляющая система. Постараемся определить наиболее удачный метод обработки больших данных на примере системы управления потоками движения при помощи светофора.

Ежедневно сотни и тысячи городских камер анализируют трафик пешеходов и автомобилей. Делается это для того, чтобы понять, как наиболее эффективно использовать регулировку движения. Никого не удивит тот факт, что утром, когда наибольшее число жителей городов идут на работу или учебу движение на перекрестках весьма оживленно. Скорее всего, каждый из нас замечал, что утром светофоры любимого перекрестка имеют более длительный временной интервал, чем в обед. Делается это для того, чтобы максимально оптимизировать поток движения, как автомобилей, так и пешеходов. Если просто сделать одинаковое время переключения светофора, скорее всего это приведет к тому, что автомобили утром будут стоять в пробке, а в обед пешеходы будут подолгу ожидать нужного сигнала. Для всего этого и используется анализ данных с городских камер. Объем таких данных колоссальный, а ведь его еще нужно проанализировать и выбрать суть.

Для того, чтобы сделать это наиболее рационально нужно определить более подходящий метод обработки данных. По моему мнению наиболее подходящим вариантом для данной управляющей системы является метод машинного обучения. На вход в нашу систему поступает большое количество различных параметров: поток машин, людей, день недели, погодные условия, время дня и прочие. Далее с этими данными можно поступить разными способами, например, использовать глубокое обучение, чтобы дать системе все имеющиеся параметры и рычаги влияния, а в дальнейшем стимулировать правильные решения наградой и отвергать не правильные «наказывая». Такой метод обучения так же называется «Обучение с подкреплением».

На основе таких ответов нейросеть будет обучаться, что в конечном итоге позволит максимально оптимизировать поток транспорта и пешеходов на заданном участке.

Литература

1. Технологии обработки Больших Данных [Электронный ресурс]. Режим доступа: <https://moluch.ru/archive/183/46957> (дата обращения 2.02.2020). своб.
2. Что такое Big Data: характеристики, классификация, примеры [Электронный ресурс]. Режим доступа: <https://neurohive.io/ru/osnovy-data-science/big-data/> (дата обращения 03.02.2020). своб.
3. Самойлова И.А. Технологии обработки больших данных // Молодой ученый. 2017. №49. С. 26-28. (дата обращения: 02.02.2020). своб.
4. Big Data и актуальность подхода сегодня [Электронный ресурс]. Режим доступа: http://www.aviant.ru/about/articles/big_data_and_relevance/ (дата обращения 03.02.2020). своб.



Хватова Алена Валерьевна
Университет ИТМО,
факультет прикладной оптики,
магистрант группы №_В42801,
направление подготовки: 16.04.01 – Техническая физика,
e-mail: alenachvatova@gmail.com



Лекус Елена Юрьевна
Университет ИТМО,
факультет прикладной оптики
к.культурологии, доцент,
e-mail: lekus_elen@mail.ru



Быстрянцева Наталья Владимировна
Университет ИТМО,
факультет прикладной оптики,
к.архитектуры, доцент,
e-mail: sv.s.lighting@gmail.com

УДК 535.242.3

ОРГАНИЗАЦИЯ СВЕТОВОЙ СРЕДЫ В ЗАВИСИМОСТИ ОТ ФУНКЦИОНАЛЬНОГО НАЗНАЧЕНИЯ ПРОСТРАНСТВА

А.В. Хватова

**Научные руководители – к. культурологии, доцент Е.Ю. Лекус;
к. архитектуры, доцент Н.В. Быстрянцева**

Работа выполнена в рамках дисциплины «Разработка и прототипирование световых решений в архитектуре и городской среде» и разработки «Дизайн-кода освещения зеленых зон Санкт-Петербурга».

Аннотация

Объектом исследования являются зеленые зоны в структуре архитектурного ансамбля исторической части Санкт-Петербурга. В данной работе проанализирована существующая ситуация городской среды в вечернее время суток; разработаны сценарии освещения для трех зеленых зон: Исаакиевский сквер, Екатерининский сад, Летний сад; с помощью социологического опроса выявлены приемы освещения, которые являются предпочтительными для пользователей в структуре ансамбля исторического центра.

Ключевые слова

Зеленые зоны, озелененные территории, архитектурный ансамбль, приемы освещения, пространственная организация, задачи пользователей, контрастность освещения, Исаакиевский сквер, Екатерининский сад, Летний сад.

Сохранение исторически сформированного ансамбля является ключевым элементом стратегии городского обновления Санкт-Петербурга. Поскольку это большой город с круглосуточным функционированием, дневной образ города и вечерний являются важными компонентами городской идентичности. Искусственный свет является композиционным и художественным средством формирования образа города в вечернее время суток.

Световую композицию исторических ансамблей предпочтительно выстраивать за счет иерархии световых доминант и ориентиров, соотношения цветоцветовых контрастов крупных элементов, с учетом сохранения характера застройки. Освещение зеленых зон в структуре исторического ансамбля подчиняется светокомпозиционной иерархии [1, 3].

При «конструировании» светового образа объекта возможны два способа: ассоциативное подобие дневному образу (дневной образ выступает как прототип) и создание нового «альтернативного» образа, лишённого прототипа в городской среде [5]. Первый способ более приемлем для использования в исторической среде, так как привычный (дневной) образ уже сложился в общественном сознании и, его нецелесообразно радикально изменять. Второй способ применяется для среды, образ которой требует трансформации и искажений признаков дневного образа для повышения образно-художественной выразительности в вечернее время суток.

В конструировании светового образа наибольшим светомодулирующим эффектом обладает контрастность освещения. Эта характеристика влияет на восприятие объемно-пространственной композиции, формирует иерархию всех элементов в среде. Контрастность освещения – это контраст между затемненными и освещенными элементами, необходимый для восприятия особенностей пространственной организации и художественных особенностей среды [3]. Можно выделить 4 степени контрастности:

1. Высокая степень контрастности (яркая архитектура, темные деревья).
2. Высокая степень контрастности (яркие деревья, темная архитектура).
3. Средняя степень контрастности.
4. Низкая степень контрастности (деревья той же яркости, что и архитектура).

Создание освещения возможно по средствам применения световых приемов, выбор которых определяют архитектурно-пластической системы архитектурной среды: классическая, модернистская и постмодернистская. Можно выделить следующие световые приемы при работе с элементами озеленения [2]:

1. Заливающее освещение, которое может работать на выявление объемов, периметров и границ, отдельных элементов, контрформы (когда сам объект остается затемненным на высвеченном фоне (рис. 1, а).

2. Локальное освещение, при котором светом выявляются отдельные элементы озеленения: кроны, стволы (рис. 1, б).

3. Световая графика, выстраивающая объем с помощью отдельных точек, линий и пятен, уделяя главное внимание четкости, модульности, структурности (рис. 1, в).

4. Световая живопись представляет собой крупные освещенные участками с активным использованием цветного света или проекционные многоцветные изображения статического или динамического характера (рис. 1, г).

Световая графика и живопись могут работать на сохранение основной формы (тектоничность) или на ее разрушение (атектоничность).

В исследовании рассматривается исторический центра Санкт-Петербурга. За границы рассмотрения выбрана территория объединенной зоны охраны объектов культурного наследия на территории Санкт-Петербурга (Закон Санкт-Петербурга от 19.01.2009 №820-7). В охранной зоне располагается 33 озелененные территории, которые можно разделить на три типа по признакам пространственной организации в зависимости от пропорционального соотношения массы озеленения и архитектурных объектов [4, 6].



Рис. 1. Приемы освещения: заливающее освещение (а); локальное освещение (б); световая графика (в); световая живопись (г)

1. Тип 1. Доминанта архитектурных объектов.
2. Тип 2. Равнозначность архитектурных объектов и озеленения.
3. Тип 3. Доминанта озеленения над массой архитектурных объектов (или полное их отсутствие последних) (рис. 2).

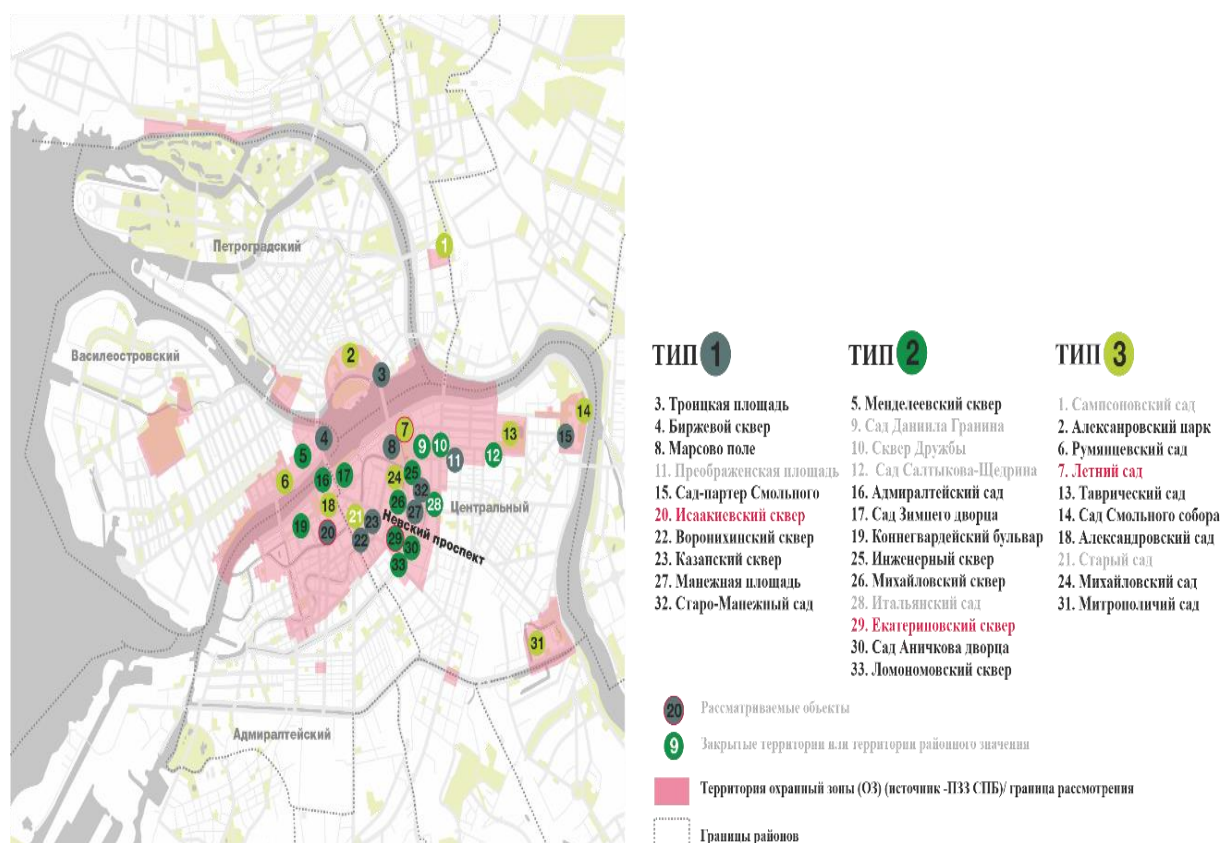


Рис. 2. Ситуационный план

Для каждого типа пространственной организации была определена наиболее популярная территория и точка визуального раскрытия для анализа на основе тепловой карты популярных мест для фото (рис. 3, а), тепловой карты пешеходной активности (рис. 3, б), ментальной карты (рис. 3, в) и социологического опроса (рис. 4).

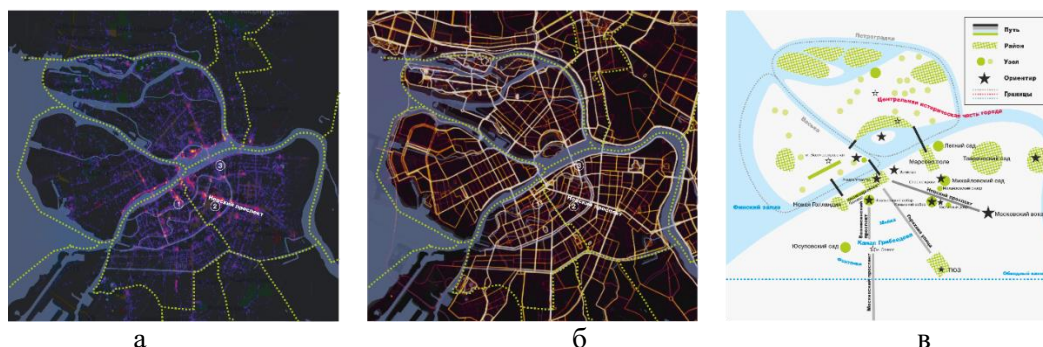


Рис. 3. тепловая карта популярных мест для фото (а); тепловая карта пешеходной активности (б); ментальная карта(в)

На основании анализа можно выделить 3 объекта из каждой категории, которые наиболее интересны пользователям. Типу 1 соответствует Исаакиевский сквер, типу 2 - Екатерининский сад, типу 3 - Летний сад.

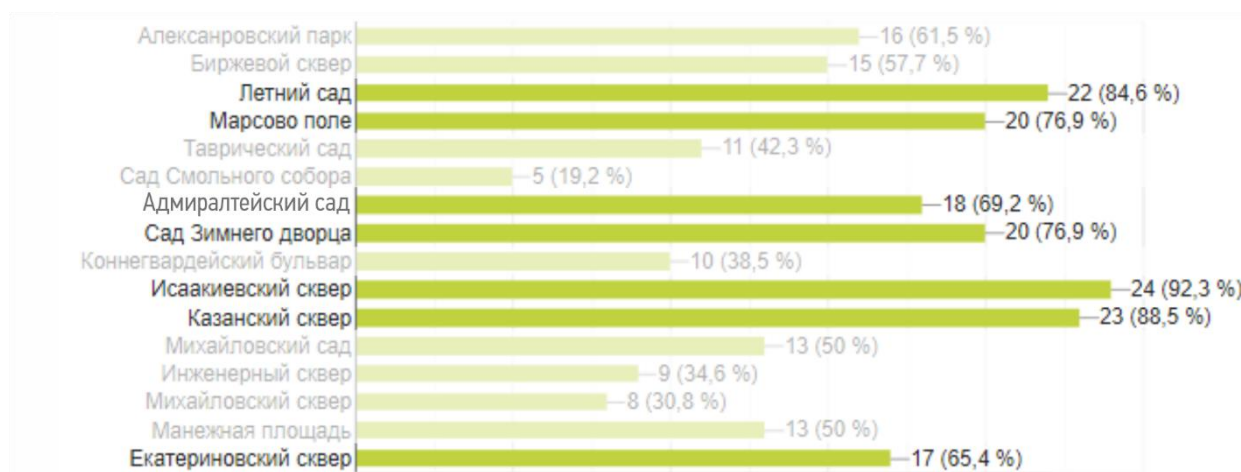


Рис. 4. Результаты опроса. Зеленые зоны исторического центра, наиболее интересные пользователям

Поведенческие задачи пользователей в городских пространствах можно разделить на три типа [5]:

движение:

сквозное движение через или вдоль территории;

сквозное или циклическое движение на велосипедах (или другом транспорте);

отдых:

активный отдых;

спокойный отдых;

общение:

общение с человеком или формирование городских сообществ.

Основная задача пользователей вблизи трех рассматриваемых территорий это сквозное движение вдоль территории. С помощью социологического опроса выявлены приемы освещения, которые являются предпочтительными для пользователей в

структуре ансамбля исторического центра, и которые простимулируют смену транзитного движения на циклическое замедленное движение с целью отдыха, изучения и погружения в среду. Задачами социологического опроса является:

1. Определение пользовательского интереса к существующим зеленым зонам исторического центра Санкт-Петербурга.

2. Определение световых приёмов, использование которых, по мнению пользователей, будет уместным в историческом центре Санкт-Петербурга (на примере трех территорий: Исаакиевского сквера, Екатерининского сквера, Летнего сада) и повлияет на формирование новых поведенческих задач.

По результатам социологического опроса можно сделать следующие выводы:

1. Приоритетной степенью контрастности для Исаакиевского собора является высокая степень контрастности (яркая архитектура, темные деревья); для Екатерининского сквера - низкая степень контрастности (деревья и архитектура равнозначны по яркости); для Летнего сада – средняя степень контрастности (рис. 5).

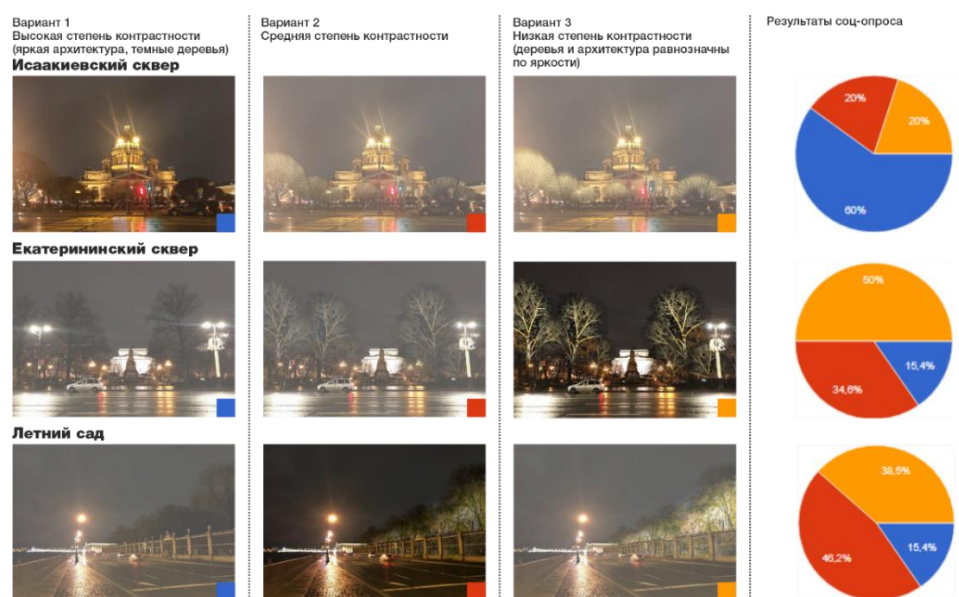


Рис. 5. Результаты опроса. Предпочтительная степень контрастности

2. Из предложенных вариантов цветового оформления респонденты выбрали решение с высоким индексом цветопередачи для всех трех территорий: Исаакиевский сквер, Екатерининский сквер, Летний сад (рис. 6).



Рис. 6. Результаты опроса. Предпочтительные приемы освещения

3. Из предложенных художественных приемов освещения, основанных на световой графике и живописи (тектоничные и атектоничные решения), респонденты выбрали (рис. 7):

- для Исаакиевского сквера: приемы графика/тектоника и живопись/атектоника (с предметными элементами изображения) как возможные решения для особых городских событий, таких как праздники или фестивали;
- для Екатерининского сквера: приемы графика/тектоника, графика/атектоника и живопись/атектоника (с предметными элементами изображения) как возможные решения для особых городских событий, таких как праздники или фестивали;
- для Летнего сада: приемы графика/тектоника как возможные решения для особых городских событий, таких как праздники или фестивали.

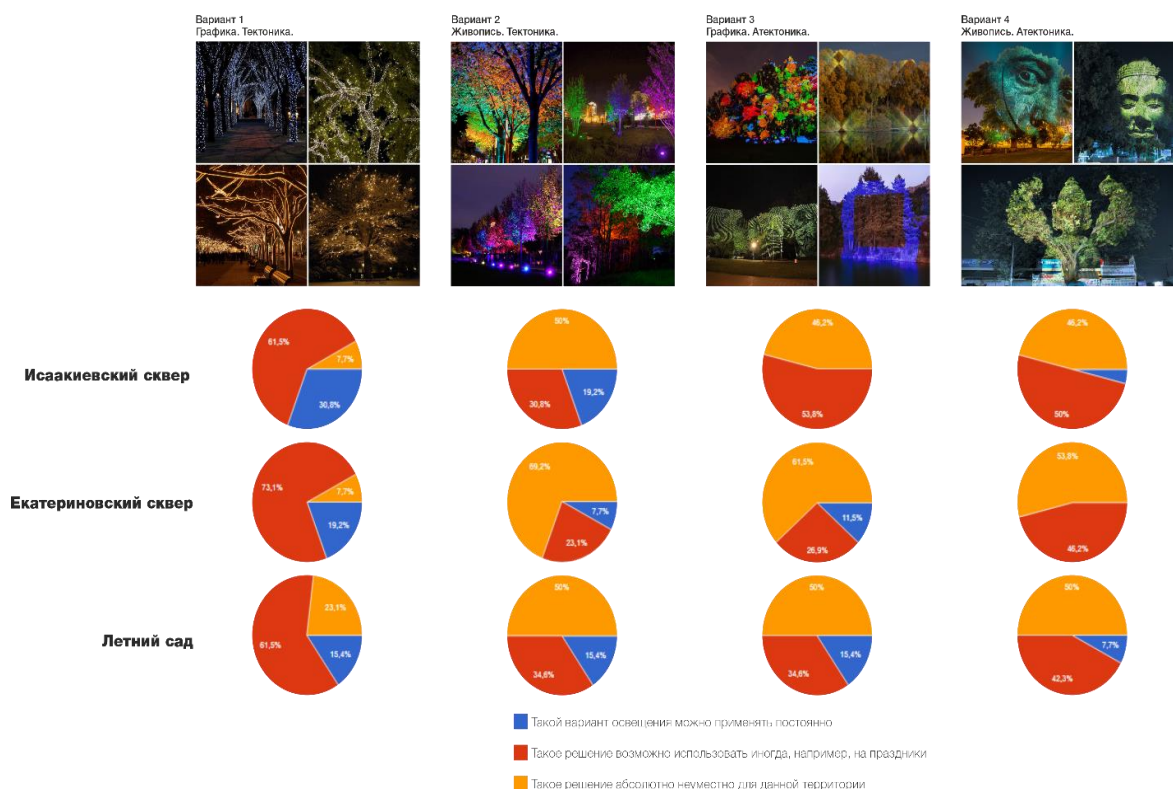


Рис. 7. Результаты опроса. Способ использования приемов освещения

Проведенный социологический опрос показал, что пользователи предпочитают для каждой из трех рассмотренной озелененной территории, соответствующей определенному типу пространственной организации, в темное время суток воссоздание за счет освещения дневного образа с помощью световых приемов, которые сохраняют тектоничность фронтов зелёных насаждений. Это предпочтительный вариант освещения в историческом контексте. Респонденты сошлись во мнении, что использование тех решений, которые формируют новый «альтернативный» образ, возможно в качестве инструмента, создающего кратковременную событийность (в рамках фестивалей и городских праздников). Атектоничные решения не допустимо использовать в историческом контексте, (как показывает социологический опрос, респонденты даже в периферийных районах не готовы видеть атектоничные решения).

Вывод: применение художественных приемов освещения возможно только при условии разработки и соблюдения системы ограничений нормативной базы и

предпочтений пользователей, обеспечивающих сохранение ключевых элементов исторической среды и иерархию всех объектов.

Литература

1. Бутыревская И.Н. Формирование основных типов светопространств как объектов ствetoурбанистического проектирования / Бутыревская И.Н., Орлова Л.Н. // Приволжский научный журнал. 2012. № 1. С. 127-132.
2. Быстрянцева Н.В. Комплексный подход в создании световой среды вечернего города: диссертация. Москва. 2015. 126 с.
3. Карпенко В.Е. Световое проектирование городской среды // Вестник Инженерной школы Дальневост. федеральн. ун-та. 2016. № 1(26). 121 с.
4. Линч К. Образ города / Линч К. М.: Стройиздат. 1982. 328 с.
5. Щепетков Н.И. Световой дизайн города: учебное пособие / Щепетков Н.И. М.: Архитектура-С. 2006. 320 с., ил.
6. Хасиева С.А. Архитектура городской среды / Хасиева С.А. М.:Стройиздат. 2001. 200 с.



Чунькова Екатерина Владимировна

Год рождения: 1982
Университет ИТМО,
факультет прикладной оптики,
магистрант группы №_B42801,
направление подготовки: 16.04.01 – Техническая физика,
e-mail: 4unkova@mail.ru



Лекус Елена Юрьевна

Год рождения: 1976
Университет ИТМО,
факультет прикладной оптики,
к.культурологии, доцент,
e-mail: lekus_elena@mail.ru

УДК 535.242.3

**ВЛИЯНИЕ ДАЛЬНОСТИ ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ
НА ВОСПРИЯТИЕ СВЕТОВОЙ СРЕДЫ**

Е.В. Чунькова

Научный руководитель – к.культурологии, доцент Е.Ю. Лекус

Работа выполнена в рамках дисциплины «Разработка и прототипирование световых решений в архитектуре и городской среде» и разработки «Дизайн-кода набережных Санкт-Петербурга».

Аннотация

Данная работа посвящена проблеме привлечения пешеходов на набережные в вечернее время в осенне-зимний период. Для решения проблемы были рассмотрены исследования, посвященные влиянию световой среды на пользователя (формирование чувства безопасности, реальная безопасность, комфорт, дальность зрительного восприятия) в городской среде. В работе приводятся результаты изучения трех видовых точек на набережных Санкт-Петербурга в границах исторического центра, выбранных по принципу дальности восприятия (камерный, ансамблевый, ландшафтный масштабы). В ходе проведенного исследования были выделены качественные и количественные критерии оценки световой среды (безопасность, комфорт, дальность восприятия), а также осуществлены измерения горизонтальной и полуцилиндрической освещенностей на выбранных участках и проведён социологический опрос с целью выявления закономерностей во влиянии дальности зрительного восприятия на восприятие световой среды.

Ключевые слова

Масштаб восприятия, дальность восприятия, световая среда, городское освещение, панорама города, безопасность, комфорт, пользователь пешеход, набережные.

В данной работе объектом исследования являются набережные и их использование в вечернее время пользователем-пешеходом.

В Санкт-Петербурге в осенне-зимний период, помимо изменений климатических условий, значительно укорачивается световой день, в связи с этим, уменьшается количество людей, совершающих пешие прогулки вдоль набережных. В данной работе предпринята попытка исследования влияния световой среды на восприятие архитектурного пространства и природного ландшафта пользователем-пешеходом, а также на его поведение в темное время суток с целью улучшения качества световой среды и повышения ее привлекательности в вечерни часы в осенне-зимний период.

В мировой практике вопросам регенерации прибрежных территорий и «возврата» реки жителям уделяется большое внимание. При формировании общественного пространства набережных учитываются такие факторы как: 1) положение в структуре города, 2) интенсивность движения, 3) обеспечивает ли данное пространство потребности всех категорий пользователей, 4) удобно ли осуществлять его эксплуатацию [1].

Проводя сравнительный анализ исследований городской световой среды, а также российских [2] и зарубежных [3] норм освещения, можно отметить тот факт, что зарубежные нормы базируются на указанных выше критериях оценки и на количественных показателях, то есть прежде, чем обращаться к количественным показателям за рубежом детально анализируют и оценивают условия и качество среды. В российских нормах учитывается гораздо меньше критериев, основным критерием является класс объекта по освещению, который определяется по общему описанию, параметры окружающей среды никак не учитываются.

Для определения критериев оценки световой среды набережных автором учитывался тип набережных, тип пользователя, выполняемые пользователем функции в пространстве (общение, движение, отдых) [4], особенности световой среды города (дальность восприятия, действующие нормы освещения).

Опираясь на зарубежный опыт и на научные исследования (изучение влияния света на такие факторы как реальная безопасность, чувство безопасности (Бойс, Экланд, Гамильтон и Бруно), визуальный комфорт [3], ориентация в пространстве [3], видимость пути (Рэйнхам, Давудян)) и выделяя определяющие факторы для пешехода (функция [4] и дальность восприятия [5]) были определены следующие критерии оценки, которые опираются на качественные и количественные показатели рис. 1.

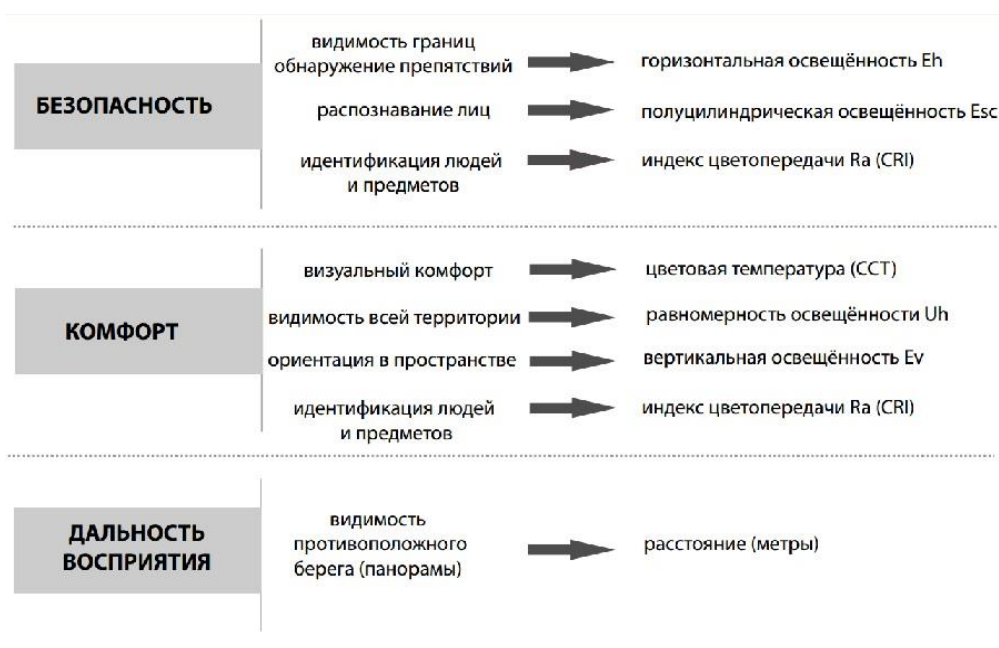


Рис. 1. Качественные и количественные критерии оценки

В ходе исследования была выдвинута гипотеза: для решения проблемы привлечения пешеходов на набережные города в вечернее время необходимо расширение рекреационных возможностей набережных (функции общения, движения, отдыха) за счет создания световой среды, обеспечивающей чувство безопасности и комфорта у пешехода с учетом дальности восприятия.

Для выявления закономерностей между влиянием дальности зрительного восприятия и восприятием световой среды было решено выбрать три участка набережных, обладающих точками обзора с разной дальностью восприятия:

– ландшафтный масштаб (характерен при восприятии крупных градостроительных образований извне со значительных расстояний, когда отсутствуют непосредственные контакты человека с объектом наблюдения, а основное значение в световом решении имеют крупномасштабные панорамы с общими силуэтными очертаниями [4]. Дальность восприятия более 500 метров (объект воспринимается, как часть общего пейзажа; при восприятии ведущую роль играют не детали, а игра архитектурных масс, цвета начинают сливаться; отсутствуют четкие переходы от света к тени) [5];

– ансамблевый масштаб (свойственен восприятию архитектурных комплексов со средних дистанций, когда контакты с архитектурой ещё разорваны, опосредованы пространством, а её оценка связана с предварительным зрительным прочтением композиционных особенностей объёмно-пространственного построения и функциональной специфики ансамбля, с распределением масс в архитектурном пространстве [4]. Дальность восприятия 100-500 метров. На этом расстоянии чётко воспринимаются детали, размеры, насыщенность) [5];

– камерный масштаб (непосредственный контакт (направленно или произвольно) с архитектурной средой и людьми. Внимание человека обращено в поле центрального зрения, заполненное в основном, фасадными поверхностями и воспринимаемое фрагментарно в деталях и различных ракурсах [4]. Дальность восприятия менее 100 метров).

Исходя из приведенной классификации для анализа были выбраны фрагменты Дворцовой и Адмиралтейской набережных и набережной канала Грибоедова. Эти три участка являются точками притяжения для туристов и жителей города, как часть туристического маршрута с основными историческими достопримечательностями и панорамой города.

1. Фрагмент Дворцовой набережной у Дворцового моста. Общие характеристики: пользователь – пешеход; дальность восприятия - от 400 до 4100 м. (ландшафтный масштаб восприятия); считываются только силуэтные очертания города; на данном участке преобладают функции движения, общения.

На выбранном участке Дворцовой набережной были проведены измерения горизонтальной и полуцилиндрической освещённости в тёмное время суток. Для проведения измерений освещённости использовался люксметр, измерения проводились по методам измерений нормируемых параметров Гост Р 5508-2013 и по методам расчёта нормируемых параметров Гост Р 5507-2013. Анализ существующей световой среды показал, что нормам соответствуют количественные показатели полуцилиндрической освещённости, оказывающие влияние на «реальную безопасность» (распознавание лиц) и соблюдены (даже чуть превышены) количественные показатели горизонтальной освещённости, влияющие на безопасность среды (видимость пути, распознавание препятствий на дороге) рис. 2

2. Фрагмент Адмиралтейской набережной у здания Адмиралтейства в зоне пешеходного пространства вдоль границы с водой. Общие характеристики: пользователь – пешеход; дальность восприятия - от 300 до 1500 м. (ансамблевый масштаб восприятия); на данном участке открывается панорама города с видом на

городские доминанты и архитектурную застройку на противоположном берегу; четко просматриваются детали, размеры и цветовая насыщенность архитектурной среды; преобладают функции общения, движения, отдыха.

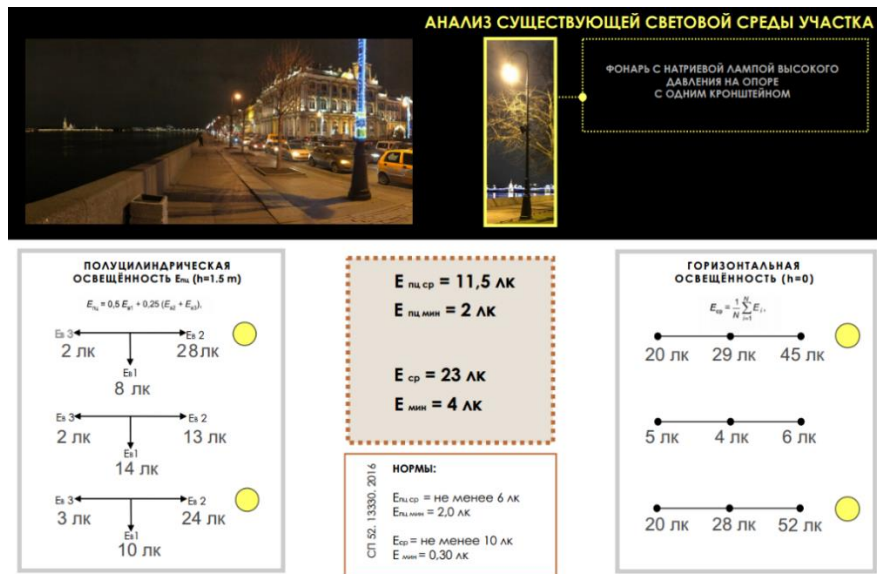


Рис. 2. Анализ существующей световой среды участка на Дворцовой набережной

На указанном участке Адмиралтейской набережной были проведены измерения горизонтальной и полуцилиндрической освещённостей в тёмное время суток. Для проведения измерений освещённостей использовался люксметр, измерения проводились по методам измерений нормируемых параметров ГОСТ Р 5508-2013 и по методам расчёта нормируемых параметров Гост Р 5507-2013. Анализ существующей световой среды показал, что нормам соответствуют количественные показатели полуцилиндрической освещённости, оказывающие влияние на «реальную безопасность» (распознавание лиц) и соблюдены (даже чуть завышены) количественные показатели горизонтальной освещённости, оказывающие влияние на безопасность среды (видимость пути, распознавание препятствий на дороге) рис. 3.

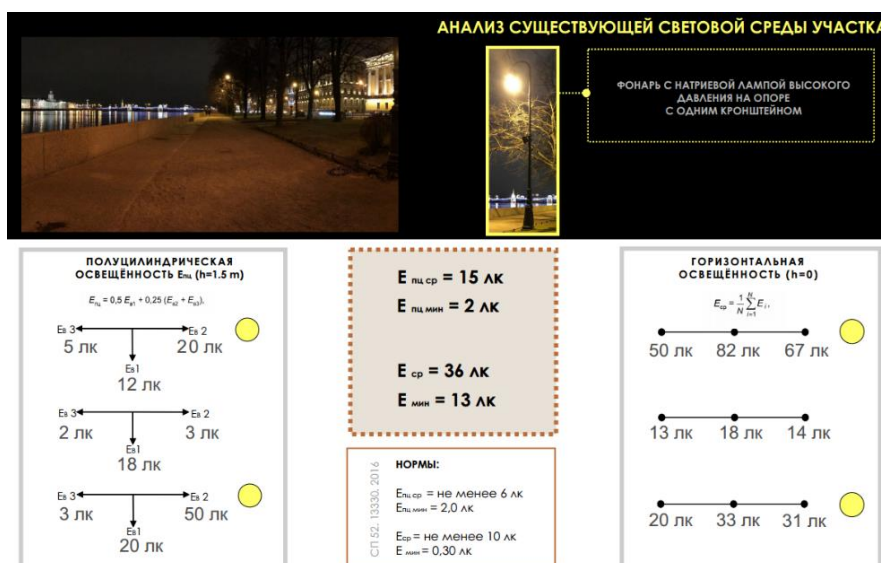


Рис. 3. Анализ существующей световой среды участка на Адмиралтейской набережной

3. Фрагмент набережной канала Грибоедова у МИСП (Музей искусства Санкт-Петербурга XX-XXI вв). Общие характеристики: пользователь – пешеход; дальность объектов составляет 27 м. (камерный масштаб восприятия); с набережной вдоль границы с водой открывается вид на архитектурную застройку на противоположном берегу; четко просматриваются детали; преобладают функции движения, общения.

На выбранном участке набережной канала Грибоедова были проведены измерения горизонтальной и полуцилиндрической освещённости в тёмное время суток. Для проведения измерений освещённости использовался люксметр, измерения проводились по методам измерений нормируемых параметров Гост Р 5508-2013 и по методам расчёта нормируемых параметров Гост Р 5507-2013. Анализ существующей световой среды показал, что нормам не соответствуют количественные показатели полуцилиндрической освещённости, оказывающие влияние на «реальную безопасность» (распознавание лиц) и соблюдены минимумы по количественным показателям горизонтальной освещённости, оказывающих влияние на безопасность среды (видимость пути, распознавание препятствий на дороге) рис. 4

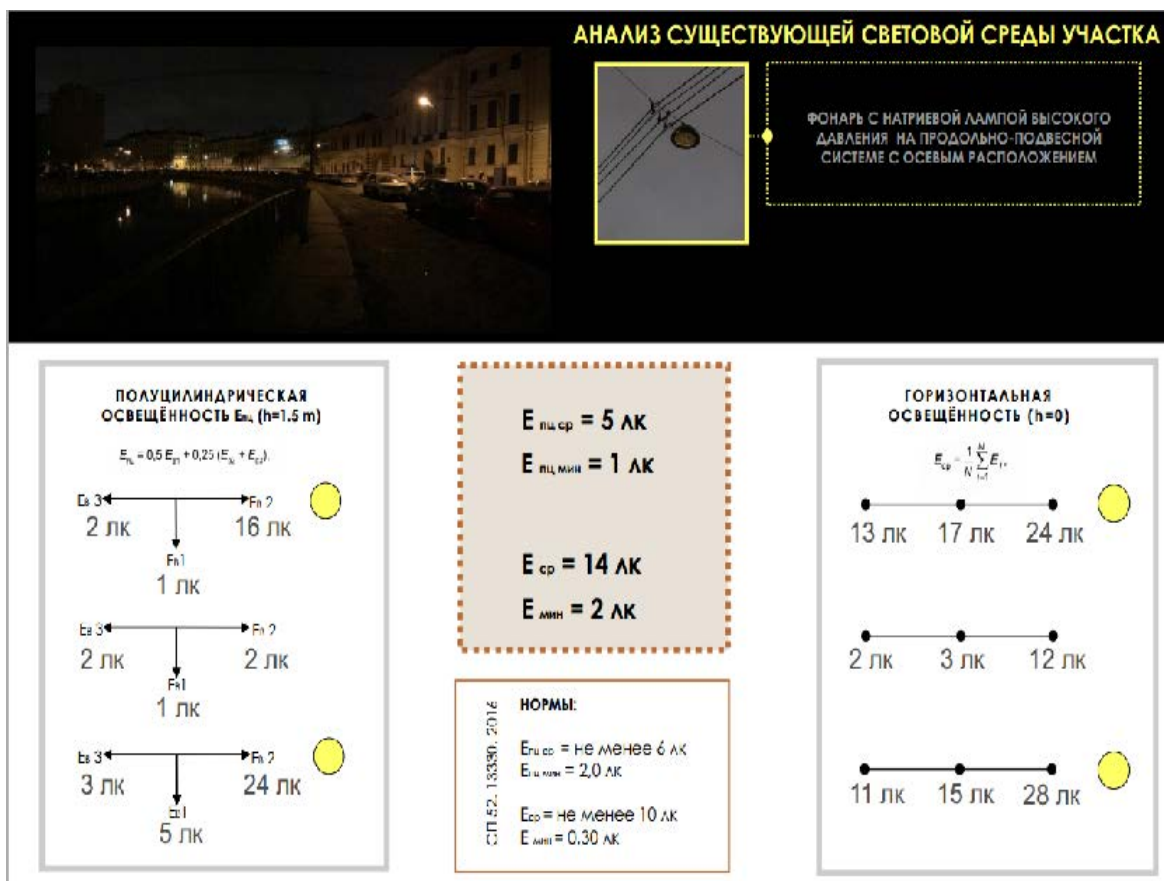


Рис. 4. Анализ существующей световой среды участка на набережной канала Грибоедова

Для выявления закономерностей между влиянием дальности восприятия и восприятием световой среды набережных был проведён соцопрос среди жителей Санкт-Петербурга по методу онлайн-анкетирования. Опрос проводился в социальных сетях в группах, имеющих непосредственное отношение к событиям Адмиралтейского района. Респондентам было предложено представить, что они идут по местам, представленным на фото и ответить на вопросы (рис. 5. и рис. 6).

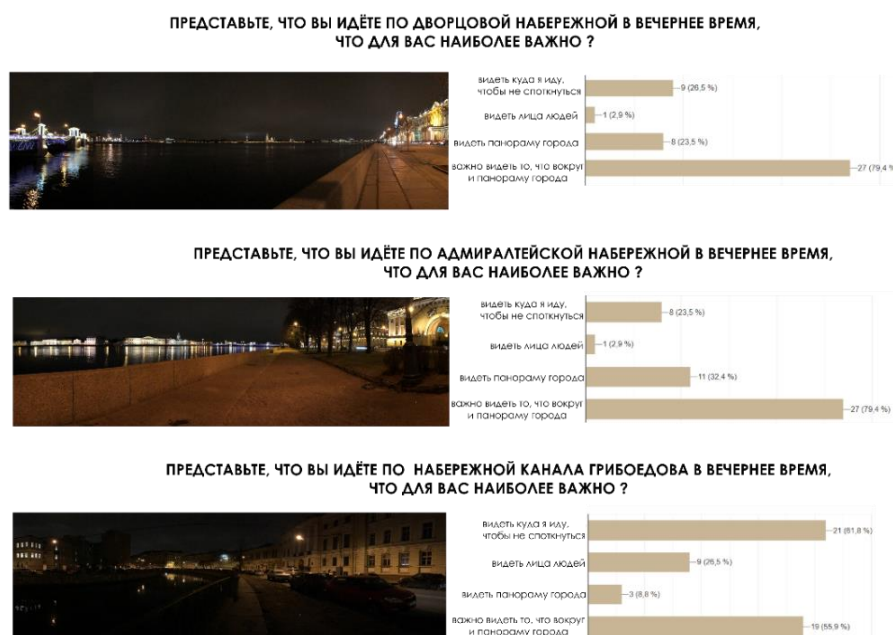


Рис. 5. Социальный опрос, 1 часть

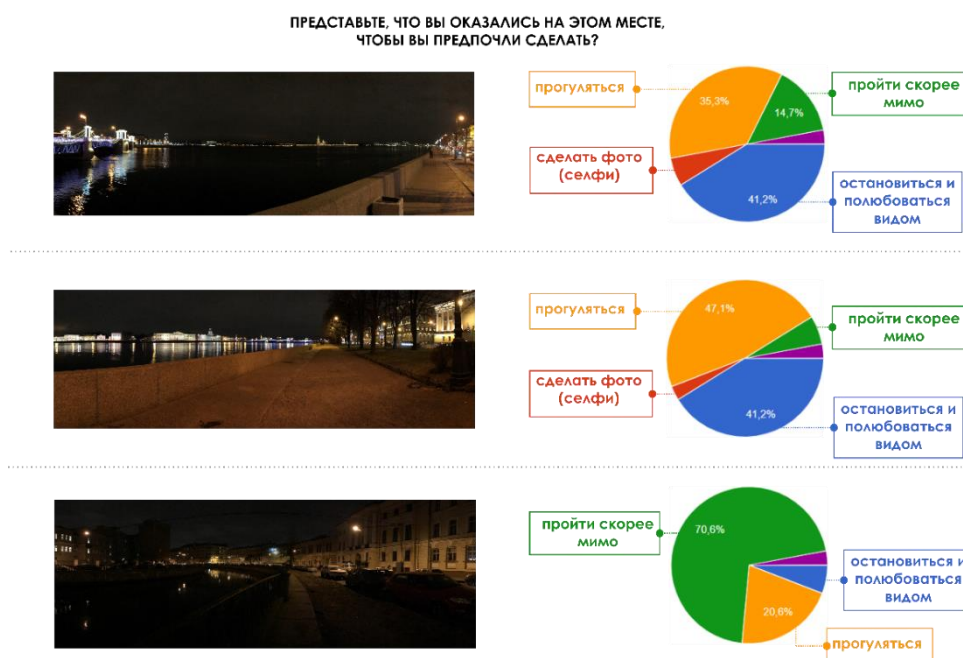


Рис. 6. Социальный опрос, 2 часть

Результаты опроса показали, что большинство респондентов обращали внимание на панораму города там, где соблюдены и чуть завышены нормы по количественным показателям горизонтальной и полуцилиндрической освещённости, влияющие на безопасность. На основании соцопроса был сделан вывод, что без удовлетворения потребности в безопасности пешеход не может полноценно воспринимать силуэт города: пространственное очертание, воспринимаемое контурами или массой (доминанты – характерные ориентиры, нюансные силуэтные акценты, архитектурно-художественная композиция города). На основании проведенных исследований световой среды с учетом дальности восприятия были выявлены закономерности во влиянии дальности зрительного восприятия на восприятие световой среды набережных:

1. При ландшафтном масштабе восприятия пешеходу важно видеть препятствия на дороге, ориентироваться в пространстве (среда вокруг), а затем панораму города (силуэт города).

2. При ансамблевом масштабе восприятия пешеходу важно ориентироваться в пространстве (среда вокруг), видеть панораму города (силуэт города), а затем видеть препятствия на дороге.

3. При камерном масштабе восприятия пешеходу важно видеть препятствия на дороге и распознавать лица идущих навстречу людей, затем ориентироваться в пространстве (видимость среды вокруг) и только потом силуэт города.

Заключение: данное исследование является первым шагом к исследованию проблемы привлечения пешеходов на набережные в вечернее время в осенне-зимний период с помощью изменения световой среды. В ходе исследовательской работы автор предпринял попытку выявления закономерностей во влиянии дальности восприятия на восприятие световой среды.

В данном исследовании не затрагивался социокультурный контекст, который планируется учитывать в дальнейших исследованиях.

Литература

1. Альбом типовых решений по комплексному благоустройству набережных Москвы-реки/ по заказу Комитета по архитектуре и градостроительству города Москвы. М. 2016.564 с.
2. Справочная книга по светотехнике/ Под. Ред. Айзенберга Ю.Б., Бооса Г.В. С74 4-е изд. перераб. И доп. М. 1892
3. Wout van Bommel // Road lighting fundamentals technology and application, 2015.
4. Щепетков Н.И. Световой дизайн города: учебное пособие / Н.И. Щепетков. М. Архитектура-С. 2006. 320 с.
5. Хасиева С.А. Архитектура городской среды / Хасиева С.А. М.: Стройиздат. 2001. 200 с.



Белый Владислав Александрович
Университет ИТМО,
Институт дизайна и урбанистики, аспирант,
направление подготовки: 23.00.02 – Политические институты,
процессы и технологии,
e-mail: vladislav@itmo.ru



Чугунов Андрей Владимирович
Год рождения: 1956
Университет ИТМО,
Институт дизайна и урбанистики, к.полит.наук,
e-mail: chugunov@itmo.ru

УДК 328:316.77

**К ВОПРОСУ О РАЗВИТИИ ЭЛЕКТРОННОГО УЧАСТИЯ
ГРАЖДАН В КОНТЕКСТЕ УПРАВЛЕНИЯ ПОЛИТИЧЕСКИМИ
ПРОЦЕССАМИ В СОВРЕМЕННОЙ РОССИИ**

В.А. Белый

Научный руководитель – к.полит.н. А.В. Чугунов

Работа выполнена при поддержке РФФ, проект №18-18-00360 «Электронное участие как фактор динамики политического процесса и процесса принятия государственных решений».

Аннотация

В работе рассмотрены тенденции гражданского участия в современной России. Автором анализируются ключевые аспекты социально-политического развития общества, предложена возможная концепция развития электронного участия.

Ключевые слова

Электронное участие, политика, гражданское общество, государство, партиципаторная демократия.

Управление политическими процессами в современной России зависит от особенностей сложившейся политической системы. Так, политическая среда в стране находится в состоянии непрерывной флуктуации, сильно зависящей от конъюнктуры решений политических элит. В свою очередь, политическая элита срощена с экономической, что позволяет экспертам говорить о формировании в стране олигархической системы, в которой решения принимаются самодостаточными политико-финансовыми кланами [1].

При этом специалисты отмечают, что современная экономическая, а значит и политическая модель России не может работать в рамках долгосрочных интересов, и ориентирована лишь на внешние рынки [2]. Об этом свидетельствуют частота политических реформ и активная внешняя политика. Однако такая ситуация лишь подчеркивает необходимость стабилизирующего центра. Сама политическая система регулярно пытается декларировать, что управление политическими процессами зависит исключительно от роли одного человека – президента. Тем не менее, в современном обществе необходима более гибкая политическая система, способная оперативно реагировать на изменения, учитывающая мнения сторон, запросы населения.

Отметим, что современные технологии позволяют использовать множество механизмов гражданского участия и управления политическими процессами. Внедрение электронных технологий в политическую сферу возможно как в контексте реального развития гражданского участия в рамках электронной демократии, так и в контексте создания фиктивных механизмов для контроля и снижения социальной напряженности в авторитарном государстве.

В данной работе гражданское участие рассматривается исходя из трактовки партиципаторной демократии. Гражданское участие – активная деятельность граждан в решении вопросов общественной жизни путем наделения населения полномочиями принятия решений и контроля наряду с сохранением принципов элитарной концепции демократии, характеризующейся политическим участием населения через компетентных представителей, избираемых путем выборов представителей власти. Однако существует ли спрос на развитие гражданского, в частности, электронного участия в современном российском обществе?

Анализируя потенциал гражданского участия, эксперты отмечают слабость институтов общественного саморегулирования, политических механизмов и неуверенность населения в завтрашнем дне на фоне сокращения реальных доходов и ухудшения государственного регулирования общественной сферы. По мнению экспертов, гражданам некогда заниматься общественным благом и 70 % населения не готово активно участвовать в политике [3]. Однако в пользу развития гражданского участия на локальном уровне говорит уверенность населения в то, что оно может повлиять на происходящие процессы. Так, доля тех, кто считает, что ни на что не может повлиять в масштабах страны – более 50%, доля же тех, кто уверен, что не может повлиять на происходящее в своем районе, городе составляет около 40%. Процент тех, кто считает, что не может повлиять на то, что происходит в их доме, дворе, где они живут – менее 20% [4]. При этом с 2017 по 2019 год наблюдается значительный рост спроса населения на гражданские права: право на свободу слова, социальную защиту, справедливый суд, свободу собраний [5]. Исходя из анализа примеров гражданского участия, можно отметить, что незначительно за последний год снизилась доля населения, готовая участвовать в выборах, а вот доля тех, кто согласен работать волонтером в общественных и политических организациях несколько выросла. Тем не менее, лидирующей остается готовность населения поддерживать институты элитарной демократии, а механизмы партиципаторной демократии для населения все еще отходят на второй план.

Власть должна быть заинтересована в развитии каналов гражданского участия, поскольку они становятся контролируемым клапаном для регуляции социального недовольства [6]. Гражданское участие может рассматриваться, как способ упорядочить противоречия и обезопасить общество от деструктивных конфликтов. Необходимо понимать, что происходит рост информированности населения, что позволяет направлять ряд общественно-политических вопросов на местный, муниципальный уровень. Это способствует более эффективному решению локальных

вопросов, повышает контроль над распределением бюджетных средств, противодействует коррупции.

Электронная форма гражданского участия способна избавить от лишней бюрократии, сделать легитимной и прозрачной саму процедуру принятия решений, позволит повысить уровень доверия к власти, усилить заинтересованность в гражданском участии молодежи. Традиционно гражданское участие сталкивается с проблемами бюрократизма согласований, ограниченности ресурсов и экономических возможностей [7]. Привлечение широкой общественности в электронном формате способно разрешить большинство из этих проблем. Также необходимо учитывать и смену поколений, российская молодежь будет воспринимать технологичные формы участия намного позитивнее, чем старшие поколения. В целом, исторический контекст российского общества одновременно позволяет говорить о возможностях кооперации населения, а реципрокность общества в условиях кризиса и вовсе может рассматриваться как одно из решений проблем экономики и капиталистического отчуждения человека от человека.

В целом, сложившаяся ситуация позволяет говорить одновременно о перспективах развития делиберативной демократии, основная идея которой состоит в том, что политические решения, влияющие на жизнь граждан, должны приниматься самими гражданами и электронного участия – как наиболее отвечающей требованиям современного общества форме. Тем не менее, ключевую роль в развитии электронного участия будет играть государство.

Предложенная концепция делиберативной демократии может быть использована при реализации проекта «Умный город» – проекта интеграции инфокоммуникационных систем с использованием передовых технологий в масштабах города с целью высокотехнологичного управления происходящими в нем процессами. Управление политическими процессами в «Умном городе» возможно с помощью механизмов электронного участия граждан: развития проектов инициативного бюджетирования и открытых бюджетов, создания порталов обсуждения и решения городских проблем, порталов для голосований по политическим решениям и инициативам населения. При этом данная концепция не повлияет на возможность власти контролировать политические процессы и не должна вызывать опасений у политических лидеров, стремящихся к контролю политического пространства. Сегодняшняя Россия нуждается в инновационных изменениях и реформировании политической системы для обеспечения стабильного и устойчивого развития в условиях конкуренции на мировой арене.

Таким образом, в России имеются предпосылки для развития системы электронного участия граждан и уже реализуются проекты, на основе которых электронное участие может развиваться. Существование и развитие федеральных программ цифровизации должно обеспечивать успешность внедрения подобных систем, что позволит проработать нормативно-правовую базу развития электронного участия на федеральном уровне. При этом данная концепция управления политическими процессами вписывается в развитие как демократического, так и авторитарного общества.

Литература

1. Гаман-Голутвина О.В. Бюрократия или олигархия? [Электронный ресурс]. Режим доступа: <http://ecsocman.hse.ru/data/936/680/1219/021Gaman-Golutvina.pdf> (дата обращения: 12.12.2019)
2. Абгаджав Д.А. Факторы политической конфликтности современной России // Вестник СПбГУ. 2014. Вып. 4. С. 103–108.

3. Политическое сознание [Электронный ресурс]. Режим доступа: <https://www.levada.ru/2019/07/15/politicheskoe-soznanie> (дата обращения: 09.01.2020).
4. Ответственность и влияние [Электронный ресурс]. Режим доступа: <https://www.levada.ru/2019/10/22/otvetstvennost-i-vliyanie-3/> (дата обращения: 09.01.2020).
5. В России растёт спрос на гражданские права [Электронный ресурс]. Режим доступа: <https://www.levada.ru/2019/11/20/v-rossii-rastet-spros-na-grazhdanskie-prava/> (дата обращения: 09.01.2020).
6. Coser L. Social Conflict and the Theory of Social Change // *The British Journal of Sociology*. 1957. V. 8. № 3. P. 197–207.
7. Ключникова Т.Н. Социальный потенциал гражданского участия в местном самоуправлении как метод развития муниципального управления // *Вопросы управления*. 2014. Вып. 4. С. 162–168.



Каштанов Кирилл Михайлович
Университет ИТМО,
факультет инфокоммуникационных технологий,
студент группы №К3421,
направление подготовки: 11.03.02 – Интеллектуальные
инфокоммуникационные системы,
e-mail: k.kashtanov98@gmail.com



Иванов Сергей Евгеньевич
Университет ИТМО,
факультет инфокоммуникационных технологий,
к.ф.-м.н., доцент,
e-mail: 45is@mail.ru

УДК 004.01

ТЕХНОЛОГИИ РАЗРАБОТКИ НАВИГАЦИОННОГО ПРИЛОЖЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ МАРШРУТА ТРАНСПОРТА С УЧЁТОМ АВТОЗАПРАВОК

К.М. Каштанов, С.Е. Иванов

Научный руководитель – к.ф.-м.н., доцент С.Е. Иванов

Аннотация

В статье описывается ход разработки навигационного приложения, производящего расчет автомобильных маршрутов с учетом расхода топлива и расположения автозаправок, определенных по заданным пользователем критериям. Рассматриваются необходимые методы и подходы выявления и документирования бизнес- и функциональных требований, сценариев использования и состояний системы, описание которых должно обеспечить наиболее эффективный ход разработки и полное представление проектной команды о разрабатываемой системе. Дается описание алгоритму расчета маршрутов.

Ключевые слова

Навигация, навигационные приложения, RUP, Rational Unified Process, бизнес-анализ, требования, сценарий использования, диаграмма активности, алгоритм Дейкстры, расчет маршрутов.

На сегодняшний день в России на тысячу человек приходится более 300 автомобилей [1], что свидетельствует о повышенной нагрузке российских дорог. Не удивительно, что в связи с этим и в век информационных технологий разрабатывается множество решений для автомобильной навигации. Если раньше для этих целей использовались специализированные устройства, то сейчас большую часть этого рынка занимают смартфоны и прочие многофункциональные мобильные устройства [2].

На сегодняшний день существует множество навигационных приложений, предоставляющих тот или иной функционал. В таб. 1 приведено сравнение функциональных возможностей наиболее популярных навигационных приложений.

Функциональное сравнение навигационных приложений

Функциональные возможности	Навигационные приложения		
	Google	Яндекс*	2Гис
Поиск адреса	+	+	+
Расчет маршрута от точки до точки	+	+	+
Добавление промежуточных остановок на маршруте	+	+	+
Информация о заправках (имеющееся топливо, услуги, цены и пр.)	-	+	+
Поиск автомобильных заправок по критериям (тип топлива, услуги, цены и пр.)	-	-	+/- (поиск только по типу топлива)
Расчет расхода автомобильного топлива	-	-	-
Оплата заправки из приложения	-	+	-
Система отзывов и оценок	+	+	+

Как можно видеть, имеющиеся приложения хоть и удовлетворяют потребностям навигации, но при этом не реализуют значительную часть возможного функционала, касающегося автомобилей и автомобильных заправок. **Целью работы** являлось создание новой системы, которая бы учитывала отмеченные недостатки существующих решений.

Для достижения поставленной цели требуется определиться с методологией будущей разработки. Планирование хода работ позволит оптимизировать труды и повысить их эффективность. Так, было принято решение следовать принципам гибкой методологии Rational Unified Process, которая предполагает итерационный подход. Следование такому подходу позволяет принимать во внимание изменяющиеся требования, вероятность которых крайне высока. Более того, унифицированный процесс предполагает постепенное наращивание продукта путем повторения пройденных стадий разработки [3].

Так, разработка упомянутого навигационного приложения была разделена на следующие стадии:

- 1) планирование – включает анализ текущего рынка, проведение опросов потенциальных пользователей, составление плана работ;
- 2) анализ – выявление всех требований программного обеспечения;

3) дизайн – детальное описание процессов и взаимодействий системы и пользователей, создание макета системы.

На стадии инициации главной целью являлось проведение опроса потенциальных пользователей приложения для определения необходимости и актуальности создания нового решения [4]. Так, из 20 опрошенных автомобилистов лишь 25% респондентов либо удовлетворены функционалом описанных выше приложений, либо вовсе не нуждаются в описываемом расширении функциональных возможностей навигационных приложений. 20% опрошенных считают новые функции полезными, но не необходимыми, и оставшиеся 55% нуждаются в функциональных возможностях по поиску заправок в приложении и расчету топлива. Результаты наглядно приведены на рис. 1 ниже.

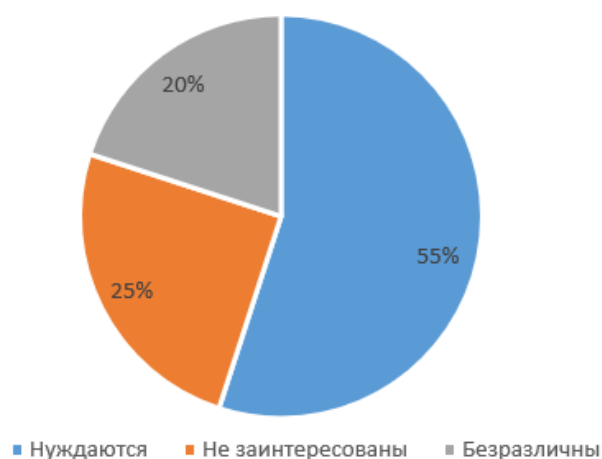


Рис. 1. Результаты опроса

Результаты опроса подтвердили целесообразность разработки нового решения, которое бы учитывало потребности автомобилистов.

Далее требовалось приступить к планированию будущих работ. Было принято решение ограничить итерацию одной-двумя неделями с проработкой стадий анализа и дизайна. Такой подход позволил проводить максимальный объем работ за малое количество времени.

Важнейшей частью разработки любого программного обеспечения является выявление и документирование требований будущего решения, чему и посвящена была вся стадия анализа [5]. Выявление требований позволяет:

- определить потребности пользователей;
- функциональные возможности программного обеспечения;
- границы программного обеспечения.

Хорошо описав требования на этой стадии, можно получить максимально качественный и полный продукт. Для этого все требования были разбиты на:

- бизнес-требования;
- функциональные требования.

Бизнес-требования описывают высокоуровневые цели, которые должны быть удовлетворены разрабатываемым программным обеспечением. Для формализации и последующего удобства следует каждому требованию выдать уникальный идентификатор и присвоить приоритет:

- О – обязательный, система не будет выдана без реализации этого требования;
- В – высокий, отсутствие реализации этих требований будет иметь существенное влияние на бизнес;

– Ж – желаемый, отсутствие реализации этих требований будет иметь умеренное влияние на бизнес;

– Н – необязательный, отсутствие реализации этих требований будет иметь минимальное влияние на бизнес;

Стоит упомянуть, что было принято решение выявлять и документировать только необходимый функционал приложения. В таб. 2 представлены все выявленные и задокументированные бизнес-требования (все представленные требования имеют обязательный приоритет).

Таблица 2

Бизнес-требования

№	Описание
БТ-1	Пользователи должны иметь возможность входить в систему
БТ-2	Система должна предоставлять пользователям актуальную информацию о заправках и ценах
БТ-3	Система должна обеспечивать поиск заправочных станций на основе следующих параметров: <ul style="list-style-type: none">– тип заправочной станции;– тип топлива;– цена топлива;– зона поиска;– дополнительные услуги на станции
БТ-4	Пользователи должны иметь возможность использовать навигационные карты
БТ-5	Зарегистрированные пользователи должны иметь возможность оставлять отзывы и оценки заправок
БТ-6	Зарегистрированные пользователи должны иметь возможность редактировать свои отзывы и оценки
БТ-7	Система должна позволять зарегистрированным пользователям хранить карты лояльности заправочных станций
БТ-8	Система должна предоставлять пользователям настраивать приложение (валюта, язык)
БТ-9	Система должна предоставлять результаты поиска заправок
БТ-10	Система должна производить расчет расхода топлива

На основе составленных высокоуровневых бизнес-требований происходит выявление и документация уже функциональных требований, каждое из которых должно покрывать хотя бы одно описанное бизнес-требование. Функциональные требования служат для описания непосредственно функциональности программного обеспечения, которая должна быть реализована разработчиками, чтобы пользователи смогли выполнять свои задачи в рамках бизнес-требований.

Для повышения качества разрабатываемой системы все документируемые требования должны обладать следующими характеристиками:

- полнота;
- корректность;
- осуществимость;
- необходимость;
- недвусмысленность [5].

Описанные характеристики позволят создать список максимально исчерпывающих функционал требований.

Также хорошей практикой является создание так называемой Traceability Matrix – таблицы отслеживания [6]. Она наглядно устанавливает соответствие между описываемыми функциональными требованиями и ими реализуемыми бизнес-требованиями. Это позволит отследить, что все бизнес-требования покрываются функциональными, а все функциональные требования нужны системе (таб. 3).

Таблица 3

Traceability Matrix

№	БТ-1	БТ-2	БТ-3	БТ-4	БТ-5	БТ-6	БТ-7	БТ-8	БТ-9	БТ-10
ФТ – <БТ-1> – 01	+									
ФТ – <БТ-1> – 02	+									
ФТ – <БТ-3> – 05		+	+							
ФТ – <БТ-3> – 07		+	+							
ФТ – <БТ-4> – 11				+						
ФТ – <БТ-4> – 12				+						
ФТ – <БТ-5> – 25					+					
ФТ – <БТ-6> – 29						+				
ФТ – <БТ-6> – 30						+				
ФТ – <БТ-2> – 34		+								
ФТ – <БТ-2> – 35		+								
ФТ – <БТ-7> – 39							+			
ФТ – <БТ-7> – 40							+			
ФТ – <БТ-8> – 43								+		
ФТ – <БТ-9> – 50									+	
ФТ – <БТ-10> – 51										+

Примечание. В таблице отслеживания приведена лишь часть функциональных требований для наглядности.

По завершении трех итераций были выявлены и задокументированы более 50 функциональных требований, которые наиболее полно описывают разрабатываемую систему. В таб. 4 представлены некоторые из них (все описанные требования имеют обязательный приоритет).

После стадии анализа в каждой итерации наступала стадия дизайна, которая отводилась для документирования сценариев, процессов и правил.

На этой стадии происходило моделирование посредством унифицированного языка UML, который позволяет представить систему со всех точек зрения, относящихся к ее разработке и внедрению [7].

Таблица 4

Функциональные требования

№	Описание
ФТ – <БТ-1> – 01	система должна позволять пользователям регистрироваться в системе по адресу электронной почты
ФТ – <БТ-3> – 06	система должна обеспечивать поиск заправочных станций по цене бензина
ФТ – <БТ-3> – 07	система должна обеспечивать поиск заправочных станций по списку дополнительных услуг: – туалет; – магазин; – кафе; – мойка
ФТ – <БТ-4> – 11	система должна производить расчет маршрута от текущего местоположения пользователя до местоположения заправки
ФТ – <БТ-5> – 25	система должна позволять зарегистрированным пользователям оставлять текстовые отзывы о заправках
ФТ – <БТ-5> – 26	система должна запрещать незарегистрированным пользователям оставлять текстовые отзывы о заправках
ФТ – <БТ-2> – 37	система должна выводить актуальную информацию о ценах на топливо по всем заправкам, имеющимся в системе
ФТ – <БТ-8> – 44	система должна позволять пользователю настраивать валюту отображения всех цен
ФТ – <БТ-10> – 51	система должна рассчитывать расход топлива пользователя исходя из преодоленной дистанции

Для моделирования динамических аспектов разрабатываемой системы следовало построить диаграмму вариантов использования. Использование этой диаграммы позволяет:

- визуализировать возможности системы;
- визуализировать поведение системы;
- визуализировать внешнее поведение субъектов системы [8].

Диаграмму вариантов использования применяют в двух случаях: для моделирования контекста субъекта и для моделирования требований к субъекту. Так как требования к разрабатываемой системе уже были задокументированы в формате естественного языка, то было принято решение использовать диаграмму вариантов использования для моделирования контекста субъектов.

Для такой диаграммы требуется:

- идентифицировать действующие лица – в разрабатываемой системе это «неавторизованный пользователь» и «авторизованный пользователь», который полностью наследует сценарии первого;
- определить сами сценарии и их отношения между собой.

Плохой практикой является отображение всех сценариев обособлено друг от друга – многие сценарии могут находиться в отношениях наследования, расширения или

дополнения. Отображение этих отношений может наглядно показывать границы и возможности разрабатываемой системы.

На рис. 2 можно видеть составленную диаграмму вариантов использования.

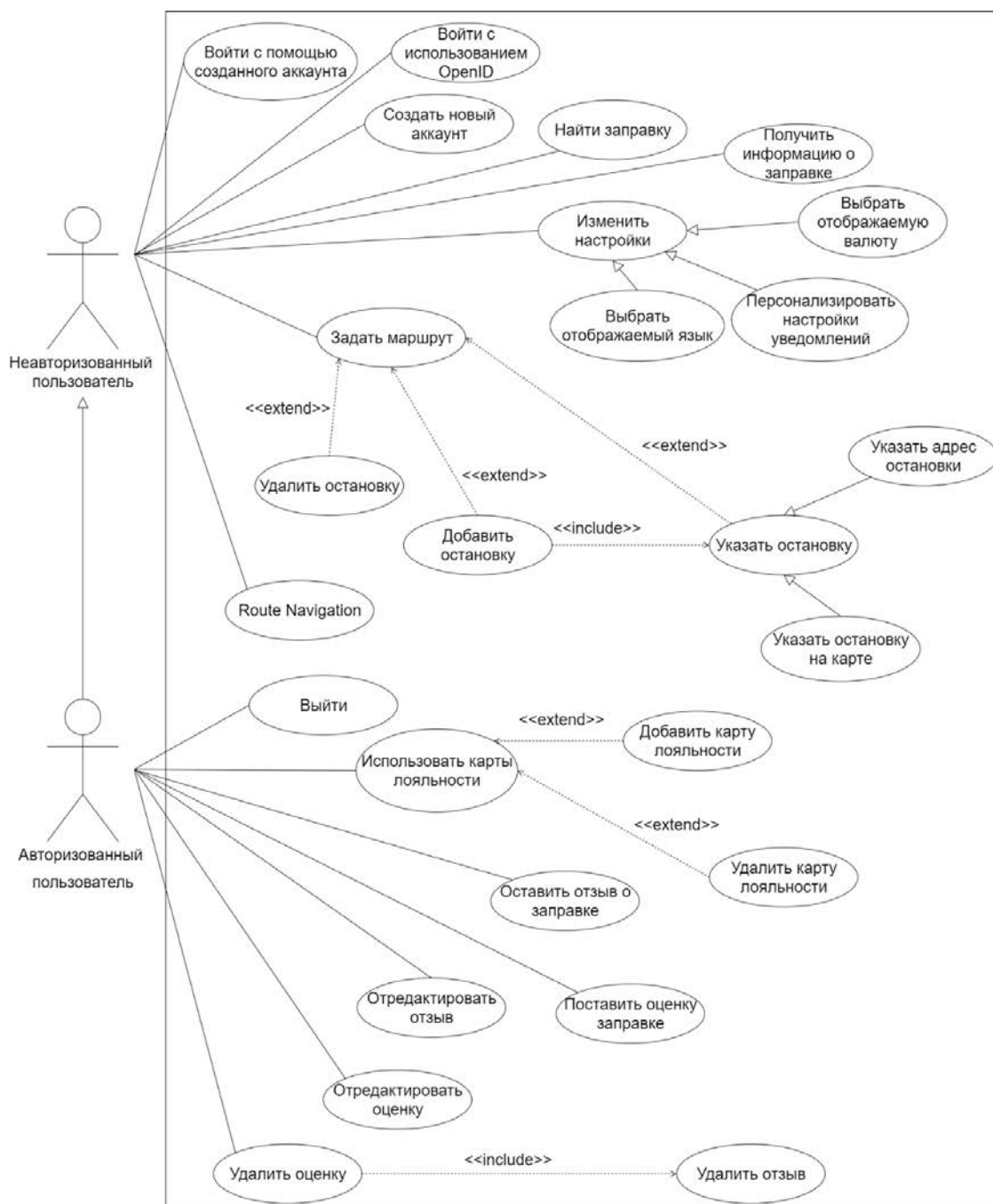


Рис. 2. Диаграмма вариантов использования

Составив диаграмму вариантов использования, можно приступить к документированию сценариев использования (use case). Сценарии должны содержать E2E-описание (конечное) каждого варианта использования. Для документирования сценариев требуется использовать таблицу, в которой указываются все актеры, участвующие в сценарии, стандартный ход сценария и альтернативный ему (таб. 5).

Для полной детализации сценария также следует документировать предусловия, триггеры и постусловия, что позволяет более полно описать состояния системы в ходе ее использования и подготовиться к моделированию диаграмм активностей и

пользовательского интерфейса. На них также следует указывать ссылки при составлении описательной таблицы сценария, чтобы упростить навигацию по объемной документации. Таким образом было детализировано 26 вариантов использования.

Следующим этапом дизайна являлось моделирование уже упомянутых диаграмм активностей. Такой тип диаграммы позволяет моделировать динамические аспекты поведения разрабатываемой системы. На ней показаны состояния системы в разный момент времени и триггеры, которые приводят к смене текущего состояния системы. Диаграммы активностей являются необходимым дополнением к сценариям использования, так как они вместе детализируют один процесс, рассматриваемый с разных сторон (первые – с точки зрения системы, вторые – с точки зрения пользователя) [9].

Таблица 5

Пример сценария использования. Добавить остановку на маршруте

Use Case ID и название	UC – ADSTP – 001. Добавить остановку на маршруте
Описание	Описывает процесс добавления остановки на маршруте
Актеры	– Авторизованный пользователь – Неавторизованный пользователь
Предусловия	3. Пользователь находится на экране «Маршрут»
Результат/постусловия	1. Форма «Добавить остановку» отображается 2. Кнопка «Удалить остановку» отображается
Триггеры	Пользователь нажимает кнопку «Добавить остановку»
Стандартный сценарий	1. Пользователю отображается экран «Маршрут» со следующими формами и кнопками: – “От” форма – “До” форма – “Добавить остановку” кнопка 2. Пользователь нажимает кнопку “Добавить остановку” 3. Система отображает пользователю форму “Добавить остановку” и кнопку “Удалить остановку”
Альтернативный сценарий	Альтернативный сценарий 1: 1a Количество добавленных остановок равно 98 2a Пользователь не может добавить больше остановок. Система не отображает кнопку «Добавить остановку»
Ссылки на UI	– Экран «Маршрут»
Ссылки на диаграммы активностей	– Диаграмма активности «Добавление остановки» – Диаграмма активности «Получить маршрут»

Составление диаграмм активностей посредством языка UML позволяет более точно описать разные типы процессов: ветвления, параллельные процессы, циклы и прочие. Это особенно хорошо для систем, в которой предполагаются сложные и многоуровневые переходы между состояниями системы. Примером этого является процесс «Поиск заправки», диаграмма активности которого представлена на рис. 3. У

этого процесса возможны множественные выходы (выходы с текущего экрана и переходы на другие) и циклы внутри процесса, что всё вместе крайне трудно описать посредством естественного языка или простейших алгоритмов.

Стоит отметить, что необязательно на каждый описанный сценарий использования должна быть составлена диаграмма активности. Это обусловлено тем, что некоторые процессы могут быть просты и элементарны и не нуждаться в дополнительном описании. Также это объясняется тем, что одна диаграмма активности может соответствовать более одному сценарию, как, например, сделано со сценариями «Найти заправку» и «Получить информацию о заправке», которым соответствует одна диаграмма активности «Поиска заправки», так как она завершается демонстрацией экрана «Информации о заправке». В итоге, было составлено 12 диаграмм активностей.

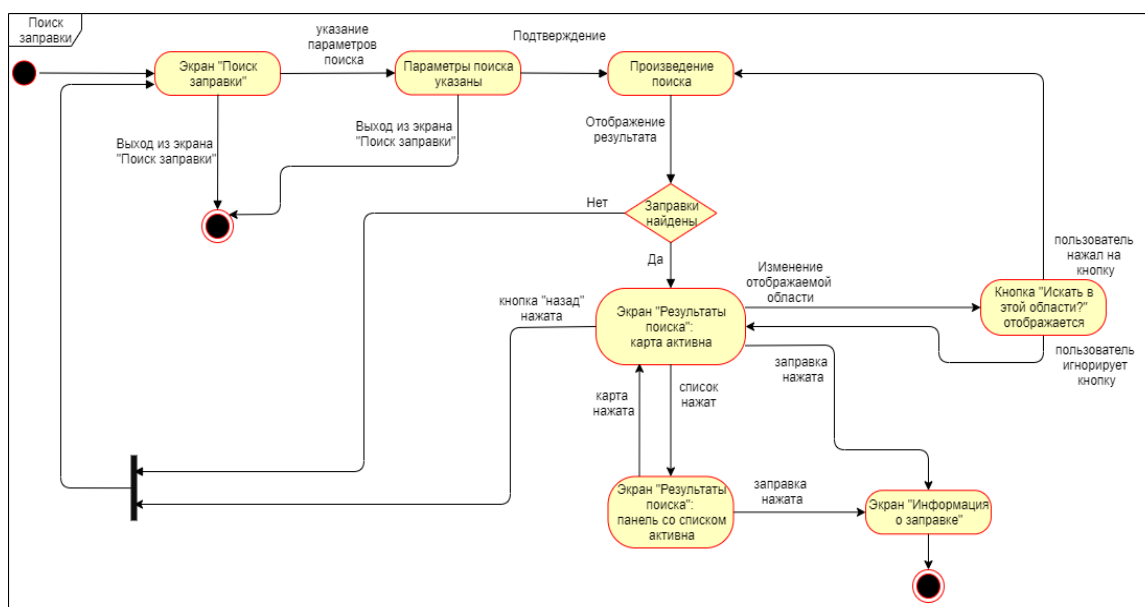


Рис. 3. Пример диаграммы активности. «Поиск заправки»

На стадии дизайна параллельно с составлением диаграмм активностей происходила не менее важная для этой стадии работа – моделирование макета пользовательского интерфейса. Такая визуализация позволяет объединить и попытожить всю проделанную до этого работу, а именно визуализировать все задокументированные и описанные элементы системы [10]. Это дает возможность наглядно продемонстрировать будущее программное обеспечение. Так же, как и с диаграммами активностей, каждый пользовательский интерфейс (он же «экран») должен соответствовать хотя бы одному сценарию использования для большей наглядности. Таким образом, образуется связь между требованиями (которые между собой также связаны), сценариями использования, диаграммами активностей и пользовательскими интерфейсами, что создает более полное представление разрабатываемом приложении. Для этой цели было смоделировано 18 пользовательских интерфейсов, которые в полной мере покрывают весь функционал приложения. Один из составленных пользовательских интерфейсов можно видеть на рис. 4.

Описанные выше работы можно классифицировать как пользовательские, так как их ход и результат будет виден конечному пользователю. Но на стадии дизайна также проходит работа, ход которой в приложении будет скрыт от пользовательских глаз. Речь идет о расчете маршрутов.

Для этого расчета планируется использовать поиск по «зависимым областям», площадь и расположение которых будет определяться общей протяженностью маршрута. Так, чем ближе друг к другу располагаются конечные точки маршрута, тем ближе стартовая точка маршрута будет располагаться к центру области поиска дорог. И наоборот – чем дальше точки расположены друг от друга, тем ближе точки маршрута будут смещаться к краям области. Наглядно это можно наблюдать на рис. 5 ниже.

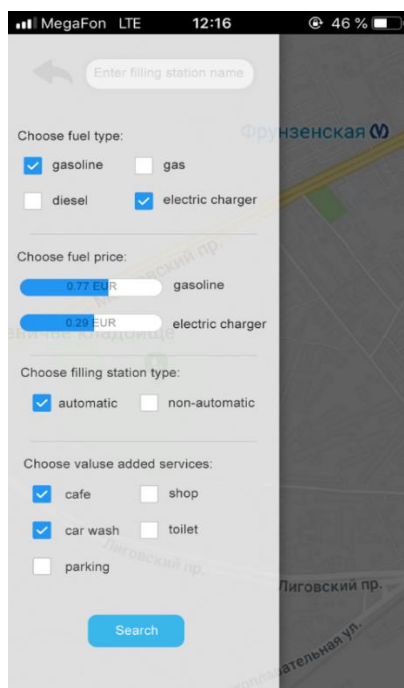


Рис. 4. Пример пользовательского интерфейса. Экран «Поиск заправки»

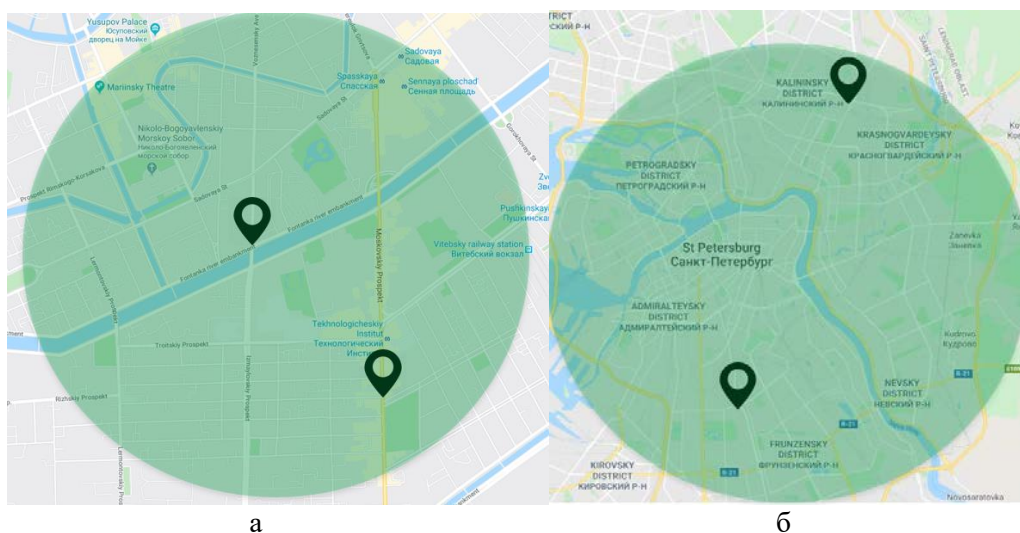


Рис. 5. Размер области поиска: на малом пути (а); на большом пути (б)

Стоит отметить, что даже когда область поиска будет определена, система должна будет производить дополнительный поиск за пределами обозначенной области, учитывая высокоскоростные и свободные дороги.

После того, как область будет определена, будет применяться один из алгоритмов для решения задачи о кратчайшем пути: это может быть алгоритм Беллмана-Форда [11] или Флойда-Уоршелла [12]. Здесь же было принято использовать для решения этой задачи алгоритм Дейкстры. Выбор именно этого алгоритма

обусловлен тем, что он широко применяется в самых разных областях и может быть адаптирован для текущей задачи [13].

Таким образом была произведена разработка нового навигационного решения, которое способно учитывать расположение автомобильных заправок при производстве расчета автомобильных маршрутов с помощью описанного алгоритма. В ходе работы применялось множество эффективных методов и способов документирования и описания программного обеспечения – выявление требований, детализация сценариев использования, визуализация состояний и поведений системы, – что позволило оптимизировать как время прodelываемой работы, так и ее качество.

Литература

1. Рейтинг российских городов-миллионников по обеспеченности автомобилями в 2019 году // Автостат: Аналитическое агентство. / [Электронный ресурс]. Режим доступа: URL: <https://www.autostat.ru/press-releases/41923/> (дата обращения: 01.03.2020).
2. Navigation devices & usage - Statistics & Facts // Statista. / [Электронный ресурс]. Режим доступа: URL: <https://www.statista.com/topics/2221/navigation-devices-and-usage/> (дата обращения: 01.03.2020).
3. Крачтен Ф. Введение в Rational Unified Process. Изд. 2-е: Пер. с англ //М.: Издательский дом “Вильямс. 2002.
4. Project Initiation // Project Management. / [Электронный ресурс]. <https://opentextbc.ca/projectmanagement/chapter/chapter-7-project-initiation-project-management/>. Дата обр. 10.12 2019.
5. Вигерс К. Разработка требований к программному обеспечению/Пер. с англ. М.: Издательско-торговый дом «Русская Редакция». 2004. 576с.: ил.
6. Madan M., Dave M., Tandon A. Need and Usage of Traceability Matrix for Managing Requirements //International Journal of Engineering Research. 2016. Т. 5. №. 8. С. 666-668.
7. Буч Г., Рамбо Д., Якобсон И. Язык UML. Руководство пользователя. 2-е изд.: Пер. с англ. Мухин Н. М.: ДМК Пресс. 2006. 496 с.: ил.
8. Баринов В.Р. Применение диаграмм сценариев использования при разработке программного обеспечения //Проблемы современной науки и образования. 2016. №. 32 (74).
9. Sulaiman N., Ahmad S.S.S., Ahmad S. Logical approach: Consistency rules between activity diagram and class diagram //International Journal on Advanced Science, Engineering and Information Technology. 2019. Т. 9. №. 2. С. 552.
10. Бубарева О.А., Вайцель Н.С. Подход к проектированию пользовательского интерфейса в системах реального времени на базе онтологий //Южно-Сибирский научный вестник. 2018. №. 1. С. 82-86.
11. Чертков А.А. Автоматизация выбора кратчайших маршрутов судов на основе модифицированного алгоритма Беллмана-Форда //Вестник государственного университета морского и речного флота им. адмирала СО Макарова. 2017. №. 5 (45).
12. Прихожий А.А., Карасик О.Н. Разнородный блочный алгоритм поиска кратчайших путей между всеми парами вершин графа //Системный анализ и прикладная информатика. 2017. №. 3.
13. Игнатюк В.А., Ничипоренко С., Ногаев В.М. Разработка модели сети дорог с параметрами для прокладки кратчайшего пути по алгоритму Дейкстры / Территория новых возможностей. Вестник Владивостокского государственного университета экономики и сервиса. 2009.



Соловьева Валентина Викторовна

Год рождения: 1984

Университет ИТМО,

Институт дизайна и урбанистики,

студент группы №7738,

направление подготовки: 05.13.10 – Управление

в социальных и экономических системах,

e-mail: valentina.soloveva@itmo.ru

УДК 311.213

АНАЛИЗ МЕТОДОВ ПРОВЕДЕНИЯ СОЦИОЛОГИЧЕСКИХ ОПРОСОВ С ЦЕЛЮ ИССЛЕДОВАНИЯ ТРАНСПОРТНОГО ПОВЕДЕНИЯ НАСЕЛЕНИЯ

В.В. Соловьева

Научный руководитель – к.т.н. С.А. Митягин¹

1 – Университет ИТМО

Аннотация

В работе рассмотрены методы проведения социологических опросов с целью выявления характеристик транспортного поведения населения. Рассмотрен российский и международный опыт, выявлены особенности проведения транспортных опросов в различных регионах. Выявлена проблематика проведения транспортных опросов и предложены пути решения выявленных проблем.

Ключевые слова

Социологический опрос, транспортное поведение, мобильность, дневниковый опрос, подвижность населения.

Необходимость проведения социологических опросов с целью выявления мобильности населения обусловлена задачей прогнозирования транспортной подвижности населения. Для прогнозирования мобильности и эффектов от мероприятий по развитию транспортной системы всё чаще используется транспортное моделирование. Решение задачи сбора данных для построения моделей транспортного спроса также обуславливает необходимость сбора данных о подвижности населения. Самым распространенным способом для сбора таких данных является социологический опрос.

Исследование транспортного поведения жителей исследуемого региона – города, агломерации, региона, страны – позволяет выявить такие показатели как статистика о наличии и количестве автомобилей, наличии водительских прав, данные о подвижности населения, предпочтения по видам транспорта, временные затраты на перемещения, способы перемещения, структура поездок, стоимость времени в пути и другие.

Транспортные опросы проводятся различными методами – личное интервью, телефонный опрос, онлайн-опрос. Опрашиваться могут отдельные респонденты или целые домохозяйства. Методы личного интервью и телефонного опроса позволяют набрать необходимую репрезентативную выборку и осуществить выборочную проверку, но являются трудоемкими и дорогостоящими. Интернет-опросы позволяют собрать большую базу респондентов, они менее ресурсозатратны, однако такие опросы

часто характеризуются слабой репрезентативностью, невозможностью проверки результата и смещенной выборкой.

Отдельный интерес в сборе данных о подвижности населения представляют дневниковые опросы. В таком опросе респондент воспроизводит все корреспонденции за вчерашний день: цель поездки, место старта, вид транспорта, пересадки, время ожидания транспорта/время поиска места для парковки, длина поездки, место финиша.

С помощью транспортных опросов можно выявить заявленные предпочтения: респонденту дается выбор между несколькими опциями с учётом его персональных поездок. Например, выбрал ли бы он другой вид транспорта (варьируется время ожидания, стоимость, время на поиск парковочного места и т. д.), выбрал ли бы он поездку с пересадкой вместо прямой при изменении времени в пути или других параметров. Таким образом, выявляется чувствительность респондентов к тем или иным параметрам, благодаря чему можно прогнозировать, при каких условиях люди изменят своё транспортное поведение, выявляется стоимость времени для различных групп населения.

В 2015 году НИУ ВШЭ было проведено исследование транспортного поведения домохозяйств РФ. Исследование охватывало множество населённых пунктов по всей России. Репрезентативная выборка составила 8028 человек (а также 1041 человек приняли участие в дневниковом исследовании). Опрос проводился методом автоматизированного компьютеризированного телефонного интервью (САТИ). Выборка для дневникового исследования была рекрутирована из телефонной выборки (21,5%), а 78,5% – за счёт квотированной интернет-панели. 57 вопросов составили 8 блоков: вопросы об общей характеристике подвижности и наиболее характерных видах транспорта, о специфике занятости и типичных для респондента трудовых корреспонденциях, о составе домохозяйства, о личном транспорте, об общественном транспорте, о вчерашних перемещениях (этот раздел призван определить общую структуру подвижности респондентов в случае отказа от заполнения дневника), о дальних поездках и о социально-демографических характеристиках. Опрашивались респонденты с 18 лет.

В результате опроса, а также дневникового исследования были получены усреднённые характеристики и проведён анализ таких параметров как общая подвижность населения, предпочтения по видам транспорта, временные затраты на перемещения, способы перемещения, структура поездок [1].

Результаты исследования, проведённого НИУ ВШЭ характеризуют существующую мобильность населения России, однако не предполагают анализ чувствительности к тем или иным изменениям параметров поездки, что не позволяет прогнозировать изменение транспортного поведения населения в зависимости от изменения тех или иных характеристик транспортной системы.

Исследование мобильности населения в Германии проводится каждые несколько лет. Опрос населения проводится круглогодично, при этом учитываются передвижения и в будние, и в выходные дни. Дети до 10 – 13 лет опрашиваются в присутствии взрослых членов семьи. Опрашиваются как отдельные респонденты, так и домохозяйства целиком. Выборка составляет примерно 25000 (в 2002 году) – 30000 (в 2016 году) домохозяйств. Опрос проводится методами личного интервью, по телефону, по обычной и электронной почте.

Опрос выявляет влияние на повседневную мобильность таких индивидуальных характеристик, как размер домохозяйства, возраст и пол респондентов, наличие и доступность индивидуального автомобиля, наличие водительского удостоверения и т. д.

Благодаря тому, что опросы населения проводятся регулярно, результаты представляют собой не только данные о текущей мобильности населения Германии, но также данные о динамике изменения параметров транспортного поведения во времени,

что позволяет использовать их для прогнозирования ситуации. Отличительной чертой опроса, проводимого в Германии, является открытость собранных данных – не только результаты анализа, но и сырые исходные данные находятся в открытом доступе [2, 3].

В Шотландии опрос транспортного поведения населения проводится в рамках опроса домохозяйств по целому ряду направлений. Опрос проводится каждые два года, круглогодично, при этом опрашиваются респонденты старше 16 лет. Опрашиваются два члена домохозяйства – член с самым высоким доходом и ещё один произвольный взрослый член домохозяйства.

Анкета содержит 3 части и большое количество подробных вопросов о характеристиках домохозяйства, респондента и его повседневной мобильности, третья часть представляет «дневник путешествий», который заполняется за предыдущий день [4-6].

В отличие от рассмотренных выше примеров, опрос в Португалии проводился только в одном городе – Лиссабоне. Пилотный опрос был проведён в 2008 году на фокус-группе 150 человек. Опрос проводился в интернете и состоял из 5 частей.

В первой части кроме социально-экономических характеристик, респондент давал описание своих поездок за день. Далее с учётом определённых характеристик выбирается одна корреспонденция респондента и задаются подробные вопросы относительно этой поездки. Это вопросы, касающиеся стоимости парковки, платного проезда, количества людей в машине, времени задержек (для поездок на личном автомобиле); стоимости проезда, времени поездки, времени задержек; возможности воспользоваться другим видом транспорта, возможности регулировать время выезда/прибытия (для поездок на общественном транспорте) и другие. Далее респонденту предлагались на выбор несколько гипотетических сценариев этой поездки при различном выборе вида транспорта в зависимости от таких параметров, как время поездки, время ожидания, стоимость, время пешеходного пути, количество пересадок; далее – выбор наилучшего сценария из выбранных. Ещё одна часть опроса касалась информационных сервисов. В конце респонденту было предложено ответить на вопросы, направленные на то, сделал ли он осознанный выбор на предыдущих этапах, а также на что он ориентировался при выборе (например, время, комфорт, стоимость поездки). Также респонденты отвечали, согласны ли они с предложенными утверждениями, например, относительно влияния транспорта на окружающую среду. Таким образом был выявлен вес каждого из параметров.

После проведения пилотного исследования опрос был скорректирован и распространён на выборку из всего города. Опрос был проведён в мае–сентябре 2009 года [7,8].

Опрос, проведённый в Лиссабоне, характерен тем, что вопросы, задаваемые респонденту, зависели от его предыдущих ответов о совершённых поездках, то есть участнику на выбор предлагается индивидуальный набор сценариев конкретно для его корреспонденций.

Транспортный опрос в Лондоне охватывает целые домохозяйства, анкета содержит большое количество вопросов, дополнительно к этому заполняется лист поездок за вчерашний день. В Лондоне практикуется вознаграждение респондентам. Так же, как и в Германии публикуются методология, агрегированные данные и результаты исследования [9].

Опрос в Дании проводится по телефону и с помощью онлайн-форм. Если респондент не заполнил опрос самостоятельно, то через 2 дня ему позвонит оператор и проведет тот же опрос по телефону [10, 11]. В интернете отвечает около 15% получивших письма, еще 60% отвечают после телефонного звонка. Все адреса, которые вводятся респондентом или интервьюером автоматически преобразуются в координаты еще при заполнении формы, что значительно позволяет снизить затрачиваемые

ресурсы на обработку и цифровизацию данных.

В результате проведённого анализа были выявлены следующие достоинства и недостатки транспортных опросов. К достоинствам можно отнести регулярность опросов (Германия, Шотландия, Лондон) – таким образом можно анализировать динамику мобильности; информирование жителей (Шотландия, Германия) – это повышает долю людей, готовых пройти опрос; публикацию результатов исследования и собранных данных (Германия, Лондон) – повышает уровень доверия населения; использование дневника поездок (Лондон, Германия, Шотландия, Лиссабон) – позволяет использовать данные для создания спроса в транспортных моделях; возможность настраивать анкету в зависимости от предыдущих ответов респондента (Лиссабон) – позволяет выявить чувствительность респондента к тем или иным параметрам; простота и скорость проведения (Дания) – снижает временные и финансовые затраты. К недостаткам можно отнести необходимость ручной обработки и оцифровки анкет и высокий показатель затрачиваемого времени на проведение опросов (все рассмотренные примеры, кроме Дании).

Выявлены основные проблемы, которые могут возникнуть при проведении опроса о транспортном поведении в российских городах. В первую очередь, это практически полное отсутствие опыта проведения таких опросов, а также связанное с этим отсутствие культуры опросов в России – большинство жителей не знакомы с практикой опросов и не понимает их значимости, что приведет к большому проценту отказов от участия. Кроме того, транспортные опросы характеризуются высокой длительностью, что приводит к низкой доле завершённых опросов), а устаревшие методы проведения (бумажные бланки и последующая ручная оцифровка) приводит к огромным затратам времени и трудовых ресурсов.

Для решения этих проблем предлагаются следующие возможные решения.

Во-первых, это использование цифровых технологий для проведения опросов, что снизит вероятность ошибок из-за ручного ввода и трудозатраты на оцифровку анкет. Это может быть использование специальных приложений для ввода ответов, использование автоматической геопривязки адресов, для онлайн-опросов возможно применение геймификации для того, чтобы респонденты с большей охотой доходили до конца анкеты. Также использование цифровой формы опросника облегчит возможность задавать вопросы, основанные на предыдущих ответах респондента, как это делается при выявлении заявленных предпочтений.

Во-вторых, это информирование жителей о проведении опроса, его целях и методике для повышения уровня доверия потенциальных респондентов к проведению опроса. В странах, где такие опросы проводятся регулярно, жители с большей готовностью принимают участие в анкетировании. Также для повышения доверия жителей к опросам важно публиковать в открытых источниках и методику, и результаты проведённого опроса.

Литература

1. Мулеев Е.Ю. Транспортное поведение населения России: краткий отчёт об исследовании. Москва. 2015.
2. Uwe Kunert R.F. Methodological Advances in National Travel Surveys: Mobility in Germany 2002 // *Transport Reviews*. Vol. 25. No. 4. Jul 2005. Pp. 415–431.
3. Angelika Schulz C.N. European transport conference // *German national travel survey MID 2016 – Mobility in Germany: New challenges – new approaches*. 2016. Pp. 1-12.
4. Scottish government statistician group. *Statistical Bulletin Transport Series Scottish Household Survey: Travel Diary 2009/2010*, Scottish Government, Edinburgh, 978-1-908181-17-6. 2011.

5. // Scottish Government: [сайт]. [2017]. URL: <http://www.gov.scot/Topics/Statistics/16002/SurveyDetails> (дата обращения: 27.09.2019).
6. Scottish Household Survey Questionnaire. 2017.
7. Lang Yang CFC&MBA. Stated Preference Survey for New Smart, Massachusetts Institute of Technology Transport Modes and Services: Design, Pilot Study and New Revision, ITS-SCUSSE-09-02. 2009.
8. Lang Yang CFCMBA. Evaluating innovative travel modes and services in Lisbon: a stated preference approach, Massachusetts Institute of Technology, ITS-SCUSSE-09-02. 2009.
9. TfL Strategic Analysis. London Travel Demand Survey – Survey Design and Sample Design. May 2017
10. Hjalmar Christiansen. Documentation of the Danish National Travel Survey. August 2012, DTU Transport.
11. // TRIMIS (<https://trimis.ec.europa.eu/project/documentation-danish-national-travel-survey>). Documentation of the Danish National Travel Survey: [сайт]. [2015]. URL: <https://trimis.ec.europa.eu/project/documentation-danish-national-travel-survey> (дата обращения: 10.03.2020).



Умаров Миrabбосбек Абдимуталиб угли

Год рождения: 1995

Университет ИТМО,

факультет инфокоммуникационных технологий,

магистрант группы № К4240

направление подготовки: 09.04.02 – Информационные системы и технологии,

e-mail: mirabbosbek.umarov@gmail.com



Крастев Йордан Георгиев

Год рождения: 1992

Университет ИТМО, факультет инфокоммуникационных технологий,

магистрант группы № К4240,

направление подготовки: 09.04.02 – Информационные системы и технологии

e-mail: danny@krastev.org



Грудинин Владимир Алексеевич

Год рождения: 1973

Университет ИТМО,

факультет инфокоммуникационных технологий,

к.т.н, доцент

e-mail: grudinin@itmo.ru

УДК 004.855.5

**ИССЛЕДОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ ОПЛАТЫ
ПАРКОВОК И ПЛАТНЫХ ДОРОГ С ПОМОЩЬЮ МОБИЛЬНОГО
СЕРВИСА И СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ**

М.А. Умаров, Й.Г. Крастев

Научный руководитель – к.т.н., доцент В.А. Грудинин

Аннотация

В работе представлен проект «Plate Vision», который состоит из веб-приложения и мобильного приложения, основанное на Ruby on Rails и React Native. Данный проект служит платформой оплаты парковочных пространств и платных дорог с помощью мобильного приложения. Также платформа включает в себя систему автоматического распознавания автомобильных номеров ANPR.

Ключевые слова

Ruby, rails, react native, автоматическое распознавание номеров, выделение области номерного знака, оптическое распознавание символов (OCR), ANPR.

В результате бесконечного увеличения объема транспортных средств в нашей окружающей среде, автоматическое распознавание номерных знаков (ANPR) стало развивающимся решением для управления и мониторинга транспортных средств во всем мире с целью обеспечения соблюдения правил и предотвращения преступной деятельности, такой как нарушение правил стоянки, проезд на красный свет, превышение скорости и угон автомобиля. Несмотря на то, что уже существует множество общедоступных и частных методов и библиотек, которые были

разработаны, и используются для автоматического распознавания номеров автомобильных номеров по всему миру, все-таки не уделялось большого внимания продвижению к кроссплатформенному решению ANPR, которое поддерживает все автомобильные номера по всему миру (рис. 1).

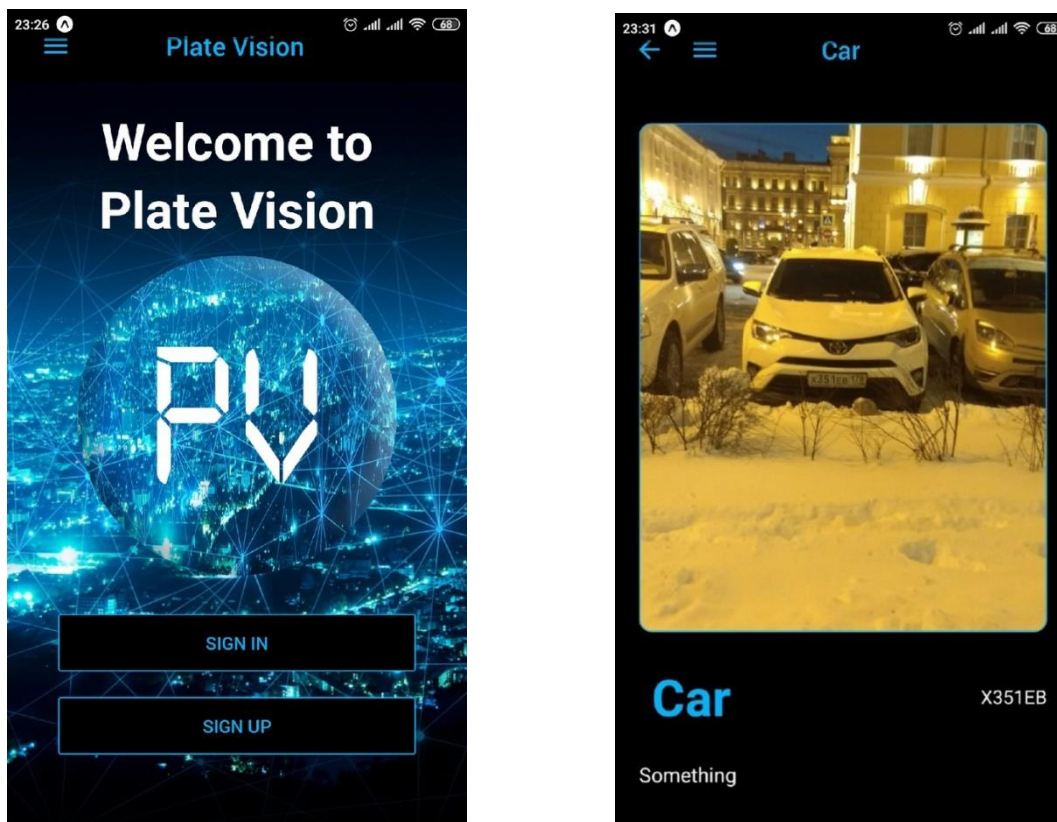


Рис. 1. Интерфейс мобильного приложения

Поскольку основные функциональные возможности и характеристики автоматического распознавания номерных знаков (ANPR) основаны на многолетних исследованиях, в которых используются различные комбинации методов оптического распознавания символов (OCR) на изображениях для поиска и считывания номерных знаков на транспортных средствах, а также различных решений уже существует, нет необходимости проходить постороннюю разработку и обучение, необходимые для полного воссоздания самой функциональности OCR.

Целью данной работы – представить пример того, как объединить и оптимизировать существующие решения OCR и ANPR для разработки «Plate Vision», уникальной платформы ANPR, которая может работать на любом устройстве, используя любую камеру, транспортное средство и местоположение.

Существует многие корпоративные технологические гиганты, такие как Google, Amazon, Microsoft и IBM, предоставляют свои собственные облачные API-интерфейсы распознавания, которые позволяют анализировать изображения и маркировать определенные объекты и элементы с использованием алгоритмов. Однако, на разработку и обучению этих алгоритмов у исследователей ушли годы, из-за этого они являются платными решениями, которые могут быстро стать дорогостоящими при попытке масштабировать или даже продемонстрировать платформу, такую как ANPR, для большого количества пользователей. Поэтому использование платформ и библиотек с открытым исходным кодом является не только лучшим экономическим

решением, но и тем, которое обеспечивает необходимую способность модифицировать и контролировать алгоритмы распознавания ANPR.

Из всех существующих на данный момент различных решений OpenALPR является единственным, имеющим открытый исходный код и позволяющим модифицировать и контролировать свои алгоритмы распознавания ANPR для конкретных стран, поэтому он был выбран для использования в «Plate Vision».

OpenALPR (Open Source Automatic License Plate Recognition) – это библиотека с открытым исходным кодом, применяющаяся для распознавания автомобильных номерных знаков, как из потока видеокamеры, так и из локального видеофайла, и из файлов изображений. Данная библиотека была написана на языке программирования C++, и имеет интеграцию для таких языков программирования как C/C#, Java, Node.js, Python [6-8].

OpenALPR также имеет 2 основных зависимостей, OpenCV и Tesseract OCR, которые служат основой для выполнения ANPR. Лучше всего сохранить минимальное количество зависимостей, которые есть у платформы, для лучшей производительности, однако в любом проекте, в котором используется OCR, обязательно должно быть несколько больших зависимостей [1].

Платформа

Платформа «Plate Vision» построена с использованием двух фреймворков: Ruby on Rails и React Native, и состоит из 2 отдельных приложений, мобильного и веб приложения. Сервер единственный для обеих приложений и включает REST API для мобильного приложения.

Ruby on Rails фреймворк, написанный на языке программирования **Ruby**, реализует архитектурный шаблон Model-View-Controller для веб-приложений, а также обеспечивает их интеграцию с веб-сервером и сервером баз данных. Является открытым программным обеспечением и распространяется под лицензией MIT [4].

React Native – это JS-фреймворк для создания нативно отображаемых iOS- и Android-приложений. В его основе лежит разработанная в Facebook JS-библиотека React, предназначенная для создания пользовательских интерфейсов [5] (рис. 2).

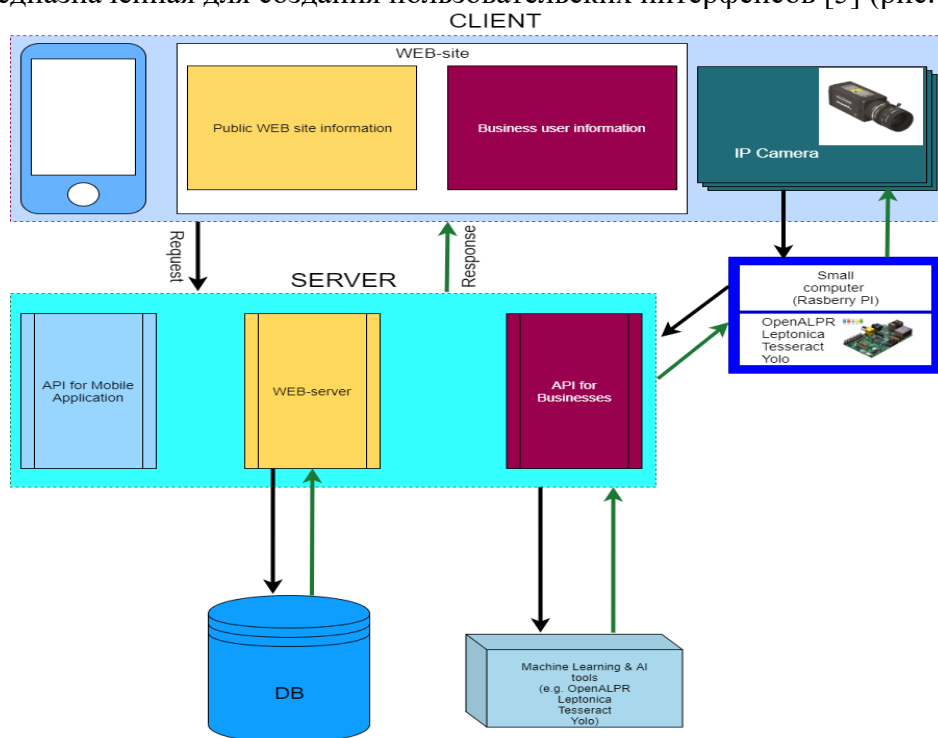


Рис. 2. Архитектура системы

Серверная часть (Backend)

Серверная часть платформы состоит из различных библиотек (gem) и контроллеров, которые обрабатывают запросы с веб-сайта и мобильного приложения (рис. 3).



Рис. 3. Архитектура классов (Серверная часть)

В следующем списке перечислены и описаны контроллеры, которые служат в качестве полноценного JSON API и обрабатывают запросы (React Native) мобильного приложения:

1. Alpr controller – содержит методы анализа изображения для того, чтобы определить номерной знак.
2. Devices controller – содержит методы для сохранения информации об устройстве пользователя в БД.
3. Locations controller – сохраняет данные о местоположении устройства пользователя в БД.
4. Users controller – обрабатывает запросы для авторизации/аутентификации пользователя.

5. Vehicles controller – содержит методы для создания, редактирования, удаления ТС и сохранения данных в БД.

Мобильное приложение (React Native)

В индустрии мобильных приложений правят два гиганта: Android и iOS. Для компаний и разработчиков очень важно предоставить продукт большинству пользователей, поэтому необходимо адаптировать обе платформы. Платформы имеют свой собственный способ разработки приложений и лишь небольшое сходство между собой.

Кроссплатформенные платформы, которые преодолевают этот пробел, приходят и уходят, поскольку им не удастся создавать приложения с тем же визуальным или функциональным стандартом, который предоставляют собственные платформы. Тем не менее, React Native обещает обеспечить полностью нативную работу с использованием только одной кодовой базы. React Native требует и работает на платформе NodeJS. NodeJS – это кроссплатформенная среда (выполнения JavaScript сценариев) с открытым исходным кодом [4].

Чтобы понять приложение React Native, необходимо знать о некоторых основных понятиях React, таких как JSX, состояние (state), реквизиты (props) и компоненты (components). JSX – это расширение XML-подобия синтаксиса ECMAScript без какой-либо определенной семантики. Компоненты – это строительные блоки любого нативного приложения React, и у типичного нативного приложения React их будет много. Проще говоря, компонент – это класс или функция JavaScript, которое по выбору принимает входные данные, т. е. свойства (props), и возвращает элемент React, который описывает, как должен выглядеть раздел пользовательского интерфейса.

Существует два типа данных, которые контролируют компонент:

1. Props (свойства) являются неизменяемыми и устанавливаются родителем, и они фиксируются в течение всего срока службы компонента [2].
2. State (состояние) изменчиво. Это означает, что состояние может быть обновлено в будущем, а свойства – нет. Состояние может быть инициализировано в конструкторе, а затем вызвано (setState), когда необходимо его изменить [3].

Тестирование и Хостинг

Хотя первоначальные тесты и исследования для этой платформы основаны на российских номерных знаках, конечной целью платформы Plate Vision является возможность считывания номерных знаков со всего мира. Поскольку OpenALPR уже прошел обучение для популярных стран, таких как США и Европа, первоначально предполагалось, что он также будет хорошо работать с российскими номерами, поскольку некоторые европейские номера похожи на российские.

Тем не менее, первоначальные тесты с использованием стандартных конфигураций OpenAlpr в Европе на российских номерных знаках показывают, что хотя OpenAlpr был способен идентифицировать и распознавать некоторое количество российских номерных знаков, для получения надежных и согласованных результатов была необходима дальнейшая настройка OpenAlpr. Чтобы улучшить анализ российских номеров без какого-либо дополнительного обучения с использованием OpenAlpr, было создано несколько пользовательских конфигураций, полученных и оптимизированных из европейских конфигураций с конкретными модификациями, основанными на различных тестовых изображениях, которые содержат автомобили с российскими номерными знаками (рис. 4).

Несмотря на то, что платформа была оптимизирована для работы с лучшими российскими номерными знаками, она уже поддерживает множество изображений

автомобилей и номерных знаков со всего мира, поскольку она запускает несколько модифицированных конфигураций библиотеки OpenAlpr, что приводит к множеству возможных результатов и в конечном итоге определяет лучший результат, основанный на сопоставлении шаблонов с номерами региона мира и точности считывания.

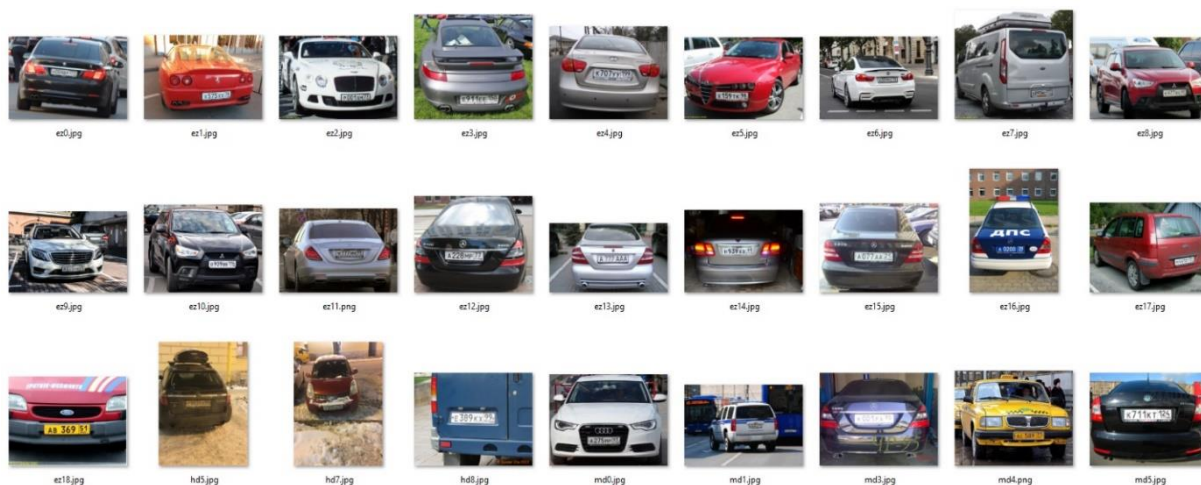


Рис. 4. Примеры фотографии для обучения алгоритма

Вся платформа размещается с использованием облачных сервисов Heroku и обновляется, и поддерживается через систему контроля версий git. Однако важно отметить, что настройка определенных зависимостей, таких как OpenAlpr, для работы в облачной серверной среде, такой как Heroku, была сложной из-за ограничений и проблем, связанных с облачной серверной средой Heroku. Поэтому службы Docker также использовались в сочетании с Heroku, чтобы преодолеть трудности настройки и обеспечения того, чтобы облако OpenAlpr и все его зависимости функционировали так же, как в локальной среде разработки [1].

Заключение

Целью данной работы состояло в том, чтобы представить платформу распознавания «Plate Vision», а также описать ее 3 основных функциональных компонента и то, как она будет оптимизировать ANPR, при этом поддерживая любое устройство, камеру, транспортное средство и местоположение. Хотя текущая версия платформы только реализует и использует OpenAlpr с несколькими конфигурациями, чтобы улучшить автоматическое распознавание номерных знаков, все еще существуют разные способы улучшения ANPR, и с дальнейшими исследованиями, тестированием и экспериментами платформа «Plate Vision» надеется, что объединит уникальные методики и методы, и библиотеки OCR для создания действительно оптимального международного решения ANPR.

Литература

1. Nicholas J. Butko, Javier R. Movellan «Optimal scanning for faster object detection»[Электронный ресурс]. <https://ieeexplore.ieee.org/abstract/document/5206540/>. Дата обр. 10.12 2019.
2. Ivan Culjak, David Abram, Tomislav Pribanic «A brief introduction to OpenCV» [Электронный ресурс].<https://ieeexplore.ieee.org/abstract/document/6240859>. Дата обр. 10.12 2019.
3. OpenALPR Documentation [Электронный ресурс]. <http://doc.openalpr.com/>. Дата обр. 10.12 2019.

4. Ruby on Rails Documentation [Электронный ресурс]. <https://guides.rubyonrails.org/>. Дата обр. 10.12 2019.
5. React Native Documentation [Электронный ресурс]. <https://facebook.github.io/react-native/docs/getting-started.html>. Дата обр. 10.12 2019.
6. Axelsson Oscar, Carlstrom Fredrik [Evaluation Targeting React Native in Comparison to Native Mobile Development]. <https://lup.lub.lu.se/student-papers/search/publication/8886469>. Дата обр. 10.12 2019.
7. Hansson Niclas, Vidhall Tomas [Effects on performance and usability for cross-platform application development using React Native]. <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A946127&dswid=8626>. Дата обр. 10.12 2019.
8. OCR Documentation [Электронный ресурс]. <https://cloud.google.com/vision/docs/ocr>. Дата обр. 10.12 2019.

СОДЕРЖАНИЕ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ПРОГРАММИРОВАНИЕ; ИНФОКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ; ДИЗАЙН И УРБАНИСТИКА.....	4
Азина Л.В. ИССЛЕДОВАНИЕ ВЛИЯНИЯ ОСВЕЩЕНИЯ ДВОРОВЫХ ПРОСТРАНСТВ МНОГОКВАРТИРНЫХ ДОМОВ НА ПОЛЬЗОВАТЕЛЬСКУЮ АКТИВНОСТЬ НА ПРИМЕРЕ СПБ.....	5
Аксёнов А.А., Рюмина Е.В. АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ ДЕТЕКТИРОВАНИЯ ЛИЦ.....	12
Акулов А.В. РЕШЕНИЕ ЗАДАЧИ МУЛЬТИЛЕЙБЛИНГА.....	20
Астапов С.С., Лаврентьев А.В., Кабаров В.И. ИСПОЛЬЗОВАНИЕ ПРОСТРАНСТВЕННОЙ АКУСТИЧЕСКОЙ ИНФОРМАЦИИ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ ИДЕНТИФИКАЦИИ ДИКТОРОВ НА ДАЛЬНОМ МИКРОФОНЕ.....	25
Барсуков А.В. ИССЛЕДОВАНИЕ ПОПУЛЯРНЫХ АЛГОРИТМОВ ПРОЦЕДУРНОЙ ГЕНЕРАЦИИ ЛАНДШАФТОВ НА ВОЗМОЖНОСТЬ УСОВЕРШЕНСТВОВАНИЯ.....	30
Беген П.Н., Чугунов А.В. АНАЛИЗ РЕЗУЛЬТАТИВНОСТИ РАБОТЫ ИНТЕЛЛЕКТУАЛЬНОГО КЛАССИФИКАТОРА СООБЩЕНИЙ ГРАЖДАН НА ПОРТАЛЕ «НАШ САНКТ-ПЕТЕРБУРГ».....	36
Богорадникова Д.А. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ИЗВЛЕЧЕНИЯ ПРИЗНАКОВ ИЗ ТЕКСТА.....	42
Васильева В.О. ИССЛЕДОВАНИЕ ПОВЕДЕНЧЕСКОГО ОЖИДАНИЯ ПОЛЬЗОВАТЕЛЕЙ ОТ СРЕДЫ НА ОСНОВЕ АНАЛИЗА ПЕШЕХОДНЫХ УЛИЦ В КРОНШТАДТЕ.....	48
Газизуллина А.Р. УЛУЧШЕНИЕ КАЧЕСТВА ТЕКСТОНЕЗАВИСИМОГО РАСПОЗНАВАНИЯ ДИКТОРА.....	55

Галактионова А.А. ОБЗОР СУЩЕСТВУЮЩЕЙ ПРАКТИКИ СОВМЕСТНОГО КАРТИРОВАНИЯ ПРИ ПРОЕКТИРОВАНИИ ОБЩЕСТВЕННЫХ ПРОСТРАНСТВ.....	62
Гараев Н.Р. ОБЗОР МЕТОДОВ РАСШИРЕНИЯ НАБОРОВ ДАННЫХ ДЛЯ ЗАДАЧ КЛАССИФИКАЦИИ.....	69
Двойникова А.А., Верхоляк О.В., Карпов А.А. СЕНТИМЕНТ-АНАЛИЗ РАЗГОВОРНОЙ РЕЧИ ПРИ ПОМОЩИ МЕТОДА, ОСНОВАННОГО НА ТОНАЛЬНЫХ СЛОВАРЯХ.....	75
Дьякова В.А., Матросова Е.В. АНАЛИЗ ПРАКТИК И СЕРВИСОВ ДИСТАНЦИОННОГО МОНИТОРИНГА ПОКАЗАТЕЛЕЙ ЗДОРОВЬЯ.....	81
Евсеева Е.С. ДЕТЕКТИРОВАНИЕ АКУСТИЧЕСКИХ СОБЫТИЙ В РАЗЛИЧНЫХ УСЛОВИЯХ ОКРУЖАЮЩЕЙ СРЕДЫ.....	86
Захарова А.А. КЛАСТЕРИЗАЦИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ С УЧЕТОМ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ.....	90
Ильина Е.Р. ОБЗОР МЕТОДИК ОПРЕДЕЛЕНИЯ ДОСТУПНОСТИ СОЦИАЛЬНЫХ ОБЪЕКТОВ И ОБЕСПЕЧЕННОСТИ НАСЕЛЕНИЯ СОЦИАЛЬНЫМИ ОБЪЕКТАМИ.....	94
Карлов Б.А., Ананченко И.В. ИСПОЛЬЗОВАНИЕ ВЛАСКНАТ-ИНСТРУМЕНТА ДЛЯ ЗАДАЧ ИНВЕНТАРИЗАЦИИ СЕТЕВОЙ ИНФРАСТРУКТУРЫ.....	97
Коробова П.И., Мамаев Н.К. МЕТОДЫ ИЗВЛЕЧЕНИЯ ТЕРМИНОВ ИЗ ТЕКСТОВ НА РУССКОМ ЯЗЫКЕ.....	104
Котов Д.О. ПРИМЕНЕНИЕ ТЕХНОЛОГИИ РЕЛЕЙБЛИНГА К РАЗМЕТКЕ БАЗ ДАННЫХ.....	108
Куан Чонг Тхе ИССЛЕДОВАНИЕ АЛГОРИТМА MVDR ДЛЯ ДВУХЭЛЕМЕНТНЫХ МИКРОФОННЫХ РЕШЕТОК С ПРОДОЛЬНОЙ И ПОПЕРЕЧНОЙ АРХИТЕКТУРОЙ В УСЛОВИЯХ ПОМЕЩЕНИЯ.....	112
Лымарь В.В., Краснова О.А. МЕТОДЫ ОБРАБОТКИ РАЗНОРОДНЫХ ГОРОДСКИХ ДАННЫХ, ПРИМЕНИМЫЕ К ИНДУСТРИАЛЬНЫМ ТЕРРИТОРИЯМ.....	116

Мамаев Н.К. АНАЛИЗ НОВЫХ ПОДХОДОВ К РАЗРАБОТКЕ ДИАЛОГОВЫХ СИСТЕМ НА ОСНОВЕ ПОИСКА.....	120
Маркитантов М.В. АНАЛИТИЧЕСКИЙ ОБЗОР АУДИОВИЗУАЛЬНЫХ КОРПУСОВ РЕЧИ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ВОЗРАСТА ДИКТОРА.....	124
Михайлова И.А., Архипов А.А. ПРОЕКТИРОВАНИЕ СЕРВИСА О СОСТОЯНИИ ЗДОРОВЬЯ ПАЦИЕНТА ДЛЯ ЕДИНОЙ ГОСУДАРСТВЕННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ В СФЕРЕ ЗДРАВООХРАНЕНИЯ САНКТ ПЕТЕРБУРГА.....	129
Мурадян А.Х. КЛАССИФИКАЦИЯ СПОСОБОВ ДЕТЕКТИРОВАНИЯ И БЛОКИРОВАНИЯ АВТОМАТИЗИРОВАННОГО СБОРА ДАННЫХ С ВЕБ-РЕСУРСОВ.....	134
Петров О.Е., Кабаров В.И. ИСПОЛЬЗОВАНИЕ ВЕКТОРНЫХ ПРОСТРАНСТВ ДЛЯ РЕАЛИЗАЦИИ ТЕКСТОВОГО ПОИСКА ПО БОЛЬШИМ ОБЪЕМАМ МЕДИА ДАННЫХ.....	139
Пилясова Д.Д. АНАЛИЗ ДАННЫХ И МОДЕЛИРОВАНИЕ ЦИФРОВЫХ РЕШЕНИЙ ГОРОДА В СФЕРЕ ОБРАЗОВАНИЯ И КУЛЬТУРЫ.....	144
Рюмина Е.В. МЕТОД ИЗВЛЕЧЕНИЯ ИНФОРМАТИВНЫХ ВИДЕОПРИЗНАКОВ ДЛЯ РАСПОЗНАВАНИЯ ЭМОЦИЙ.....	151
Тимофеева А.О., Матросова Е.В., Кононова О.В. АНАЛИЗ БИЗНЕС И ИНФОРМАЦИОННЫХ ПРОЦЕССОВ В РАМКАХ ТЕЛЕМЕДИЦИНСКИХ КОНСУЛЬТАЦИЙ «ВРАЧ-ПАЦИЕНТ».....	156
Тимофеева Е.П. ДЕТЕКТИРОВАНИЕ АКУСТИЧЕСКИХ СОБЫТИЙ В БЫТОВЫХ УСЛОВИЯХ.....	162
Тропников А.С., Беген П.Н., Низомутдинов Б.А., Углова А.Б. ПРИМЕНЕНИЕ ЦИФРОВЫХ ОТПЕЧАТКОВ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ИХ ИНФОРМАЦИОННОГО ОБРАЗА.....	165
Фоменко Б.М. ВАЖНОСТЬ АНАЛИЗА МЕТОДОВ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ В УПРАВЛЯЮЩИХ СИСТЕМАХ.....	171

Хватова А.В. ОРГАНИЗАЦИЯ СВЕТОВОЙ СРЕДЫ В ЗАВИСИМОСТИ ОТ ФУНКЦИОНАЛЬНОГО НАЗНАЧЕНИЯ ПРОСТРАНСТВА.....	174
Чунькова Е.В. ВЛИЯНИЕ ДАЛЬНОСТИ ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ НА ВОСПРИЯТИЕ СВЕТОВОЙ СРЕДЫ.....	181
Белый В.А. К ВОПРОСУ О РАЗВИТИИ ЭЛЕКТРОННОГО УЧАСТИЯ ГРАЖДАН В КОНТЕКСТЕ УПРАВЛЕНИЯ ПОЛИТИЧЕСКИМИ ПРОЦЕССАМИ В СОВРЕМЕННОЙ РОССИИ.....	188
Каштанов К.М., Иванов С.Е. ТЕХНОЛОГИИ РАЗРАБОТКИ НАВИГАЦИОННОГО ПРИЛОЖЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ МАРШРУТА ТРАНСПОРТА С УЧЁТОМ АВТОЗАПРАВОК.....	192
Соловьева В.В. АНАЛИЗ МЕТОДОВ ПРОВЕДЕНИЯ СОЦИОЛОГИЧЕСКИХ ОПРОСОВ С ЦЕЛЬЮ ИССЛЕДОВАНИЯ ТРАНСПОРТНОГО ПОВЕДЕНИЯ НАСЕЛЕНИЯ.....	203
Умаров М.А., Крастев Й.Г. ИССЛЕДОВАНИЕ И РАЗРАБОТКА СИСТЕМЫ ОПЛАТЫ ПАРКОВОК И ПЛАТНЫХ ДОРОГ С ПОМОЩЬЮ МОБИЛЬНОГО СЕРВИСА И СИСТЕМ ВИДЕОНАБЛЮДЕНИЯ.....	208

АЛЬМАНАХ НАУЧНЫХ РАБОТ МОЛОДЫХ УЧЁНЫХ УНИВЕРСИТЕТА ИТМО

Том 3

Редакционно-издательский отдел Университета ИТМО

Зав. РИО

Н.Ф. Гусарова

Дизайн обложки

Н.А. Потехина

Вёрстка

Я.Я. Платунова

Подписано к печати 25.11.2020

Заказ № 4371 от 25.11.2020

Тираж 100 экз.

Печатается в авторской редакции