

ИЗВЛЕЧЕНИЕ ТЕМЫ ИЗ ГОЛОСОВЫХ АННОТАЦИЙ ДОРОЖНОЙ ОБСТАНОВКИ НА БАЗЕ СИСТЕМЫ РАСПОЗНАВАНИЯ РЕЧИ CMU SPHINX

Сосунова И.А.

Университет ИТМО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики», Санкт-Петербург, Россия (197101, г. Санкт-Петербург, Кронверкский проспект, д.49.), e-mail: inna_sosunova@corp.ifmo.ru

В статье предлагается подход, позволяющий получать информацию о дорожной обстановке и критических ситуациях на дороге напрямую от водителя посредством распознавания речи. Примером таких ситуаций являются: ямы, колдобины, обледенение дороги, открытые люки, аварии и аварийные ситуации, несоблюдение ПДД, неработающие или некорректно работающие светофоры и табло, некорректная разметка. Для формирования словаря ключевых слов, на основе которого производится распознавание речи, был собран и проанализирован корпус сообщений о дорожных проблемах. Из корпуса были выделены 11 подкорпусов, описывающих различные дорожные проблемы, в соответствии с которыми был составлен словарь ключевых слов для распознавания и реализован алгоритм выделения темы сообщения на основе списков ключевых слов, соответствующих различным дорожным проблемам. Итогом текущего этапа исследования стал прототип системы извлечения темы из голосовых аннотаций дорожной обстановки.

Ключевые слова: распознавание речи, голосовое аннотирование, извлечение информации, интеллектуальные транспортные системы, аннотирование дорожной обстановки, корпусная лингвистика.

EXTRACTING TOPICS FROM VOICE ANNOTATIONS OF ROAD CONDITIONS BASED ON SPEECH RECOGNITION SYSTEM CMU SPHINX

Sosunova I.A.

ITMO University « Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics», Saint-Petersburg, Russia (197101, Saint-Petersburg, 49 Kronverksky Pr.), e-mail: inna_sosunova@corp.ifmo.ru

The paper proposes an approach that allows to obtain information about road conditions and emergency situations on the road directly from the driver using speech recognition technology. Such situations are the pits, potholes, ice on the road, open hatches, accidents and emergency situations, traffic violations, broken or improperly working traffic lights and displays, incorrect road markings. The corpus of messages about traffic problems was collected and analyzed. On the basis of this corpus dictionary of keywords for speech recognition was composed. From the corpus 11 subcorpus describing various road problems were received. Based on this the dictionary of keywords for the speech recognition was compiled. Also, the algorithm for topic extraction based on lists of keywords corresponding to various traffic issues was implemented. The outcome of the current stage of the research is a prototype of the system based on speech recognition for extraction topics from drivers annotations of road conditions.

Keywords: speech recognition, voice annotation, information retrieval, intelligent transport systems, annotation of road conditions, corpus linguistics.

С ростом процента городского населения и количества мегаполисов, всё более остро встают проблемы городского трафика, дорожной безопасности и скорости передвижения в городской черте. Для решения проблем мобильности, безопасности и повышения уровня жизни городского населения была разработана концепция Smart City. Используя передовые цифровые технологии, “умный” город будущего сможет предоставлять и собирать всю необходимую его жителям информацию, управлять ресурсами, повышая уровень жизни и безопасность населения. Одной из важнейших составляющих Smart City является

интеллектуальная транспортная система (ИТС). Важнейшая функция ИТС - обеспечение безопасности на дорогах и своевременное реагирование на критические и потенциально опасные ситуации.

В статье предлагается подход, при котором информация о таких ситуациях получается напрямую от водителя. Примером таких ситуаций являются: ямы, колдобины, обледенение дороги, открытые люки, аварии и аварийные ситуации, несоблюдение ПДД, неработающие или некорректно работающие светофоры и табло, некорректная разметка. Согласно предлагаемому методу, увидев ситуацию, о которой необходимо сообщить, водитель (вместо того чтобы останавливаться, выяснять телефон соответствующей службы и дозваниваться, как это происходит на данный момент), сможет используя гарнитуру сообщить о проблеме (проаннотировать ситуацию) голосом, после чего это сообщение, и все необходимые для решения проблемы данные (фото, видео, аудио файл, расшифровка аудио файла) будут автоматически отправлены в соответствующую службу и, при необходимости, другим водителям.

Целью настоящего исследования является выявление ключевых слов, описывающих различные дорожные проблемы и разработка алгоритма, позволяющего извлекать тему голосовых сообщений водителей о дорожных проблемах на основе ключевых слов.

Материалы и методы. Исходный корпус состоит из 5806 сообщений, включает в себя 11 подкорпусов, описывающих основные дорожные проблемы. Для каждого подкорпуса на основе частотных словарей были составлены словари, в соответствии с которыми производится распознавание речи на базе CMU Sphinx [3]. Метод извлечения темы основан на количестве ключевых слов, содержащихся в сообщении, входящих в подкорпуса, описывающие различные дорожные проблемы.

Результаты исследования. В статье приведены результаты первого этапа работы: собранный корпус текстов; выделенные в нём подкорпуса, соответствующие различным дорожным проблемам; ключевые слова, описывающие эти проблемы; алгоритм позволяющий отнести сообщение к тому или иному типу дорожных ситуаций на основе распознанных ключевых слов.

Очевидно, что перечисленные дорожные проблемы можно зафиксировать различными другими методами.

Система мониторинга качества дорожного покрытия посредством мобильных приложений. Мобильные приложения, такие как Street Bump (США) [5], позволяют считывать данные о колебаниях подвески автомобиля и после их анализа, делают вывод о качестве дороги. Система может различать данные, поступающие от машины и от пешехода, что минимизирует погрешности и позволяет оперировать реальными данными. Также

мобильные приложения позволяют: оценивать качество дорог, отправлять снимки дефектов дорожного покрытия, просматривать статистику и карту состояния дорог. Подобные приложения частично решают проблемы фиксирования ям и колдобин, однако, чтобы сенсор сработал, в яму необходимо въехать.

Распознавание видео в режиме реального времени. Очевидно, что распознавание видео в реальном времени является ценным источником информации о ситуации на дороге, хотя и только некоторые проблемы на дороге можно выявить таким образом. К таким проблемам можно отнести: распознавание предварительно заданных объектов или классов объектов и идентификация - распознавание индивидуального экземпляра объекта. На данный момент в транспортной сфере применяются системы распознавания номерных знаков и системы распознавания лиц. Однако подобные системы требуют больших системных ресурсов и памяти.

Трехмерные датчики. Могут использоваться для распознавания ям, колдобин, препятствий проезду, предотвращения столкновений. Проблемой тут является высокая стоимость и ограниченная зона действия.

Предлагаемый в статье подход, сочетая в себе преимущества перечисленных методов, позволяет водителям сообщать о потенциально опасных ситуациях в режиме реального времени в соответствующие дорожные и муниципальные службы. Помимо этого, применение распознавания речи значительно повышает безопасность водителя, позволяя сообщать о дорожной ситуации, не отрывая взгляд от дороги и держа руки на руле [1, 2].

Этапы работы:

- Составление корпуса аннотаций водителями дорожных ситуаций
- Выделение в корпусе сообщений подкорпусов, соответствующих различным типам дорожных ситуаций
- Проведение предварительной обработки корпуса
- Составление частотных словарей для подкорпусов
- Составление списков ключевых слов, описывающих дорожную ситуацию
- Выбор системы распознавания речи
- Создание словаря ключевых слов для распознавания
- Описание и реализация алгоритма извлечения темы из голосовых аннотаций дорожной обстановки.

Корпус был составлен на основе сообщения водителей, размещённых на русскоязычных ресурсах в сети Интернет, таких как твиттер Яндекс Пробки и сайты электронного правительства, позволяющие гражданам оставить заявки о нарушениях правил парковки, состоянии дорог, неработающих светофорах и прочих дорожных проблемах. Язык корпуса –

русский, объём – 5806 сообщений. После предварительной обработки с использованием алгоритмов, описанных в [4] (удаления идентификаторов, заголовков, временных отметок; приведения текстов к нижнему регистру; удаления гиперссылок; удаления эмодиконов (смайлов); удаления стоп-слов и лемматизации) корпус был разбит на 11 подкорпусов, содержащих 400-600 сообщений, описывающих основные дорожные проблемы.

- Пробки;
- Аварийно-опасные участки дороги;
- Погодные условия;
- Состояние трассы (обледенение, неубранный снег, грязь);
- Дорожное покрытие (качество асфальта, некачественно выполненный ремонт дорог);
- Разметка проезжей части;
- Дорожные знаки, табло, ограждения;
- Нарушение ПДД;
- Создание аварийно-опасной ситуации;
- Нарушение правил парковки;
- Ямы и выбоины.

Для каждого из 11 подкорпусов в полуавтоматическом режиме был составлен частотный словарь, словарь устойчивых выражений (биграмм) и список ключевых слов. Приведём пример частотного словаря для подкорпуса Аварийно-опасные участки дороги (Таблица).

Частотный словарь подкорпуса Аварийно-опасные участки дороги

дорога	123	проезд, яма	58	встречный, находится	37	ограждение, отсутствовать, пешеходный, устранение	30
участок	118	работа	55	аварийный	36	обвалиться, переулок	21
дорожный	115	разметка	51	проезжий, направление	34	переход, вынужденный, пдд	22
дом	111	покрытие	45	день	33	проблема, утечка, газ	20
улица	64	асфальт, автомобиль	41	метро, полоса, проводится	32	здание, тротуар	18
знак	61	адрес	40	принять, бетонный, администрация	31	транспорт, проехать, опасно,	17

После составления частотных словарей для каждого подкорпуса был сформирован список ключевых слов – слов, встречающихся в подкорпусе с наибольшей частотой и не

входящих в первые 40 пунктов частотных словарей других корпусов. Приведём примеры списков ключевых слов для нескольких подкорпусов.

- Аварийно-опасные участки дороги: участок, проезд, яма, покрытие, асфальт, аварийный, бетонный, ограждение, отсутствовать, обвалиться, утечка, газ.
- Нарушение правил парковки: тротуар, парковка пешеходный, систематический, газон, парковаться, остановка, подъезд, двор, припарковать, парковочный, вынужденный, проезд, напротив, пройти.
- Пробки: пробка, час, центр, стоять, пик, шоссе, движение, перекрыть, кольцо, загруженность, ожидать, глухой, фура, развязка, объезжать, ряд.

Технология распознавания речи значительно повышает безопасность водителя, позволяя выполнять такие операции как навигация, климат контроль, переключение радиостанций, ответ на телефонный звонок, не задействуя при этом руки, и минимизирует необходимость отводить глаза от дороги. В данном исследовании распознавание речи является наиболее действенным способом получения информации от водителя.

Существует два подхода к механизму распознавания речи: распознавание речи на сервере (Google ASR , Яндекс SpeechKit) и решение на базе систем распознавания речи с открытым исходным кодом. Готовые решения предоставляют распознавание речи на основе большой статистической модели языка, что позволяет распознавать речь в режиме диктовки. Наиболее хорошо этот способ подходит для поисковых запросов и диктовки текста. Для исследования был выбран вариант полностью локального решения. К преимуществам выбранного метода относятся: скорость и независимость от сервера, на котором производится обработка, настраиваемость и гибкость. Помимо прочего, распознавание на данном этапе исследования производится по отобраным ключевым словам, словарь которых можно внедрить в систему с открытым исходным кодом. То есть, отказавшись от готовых решений, мы повышаем процент распознавания и снижаем вероятность ошибки.

Для разработки прототипа была выбрана система распознавания речи Sphinx [4] – дикторнезависимый распознаватель непрерывной речи, который использует Скрытую Марковскую модель и n-граммную статистическую языковую модель.

На основе списков ключевых слов был сформирован словарь ключевых слов, по которому ведётся распознавание речи. Используя акустическую модель, система распознавания речи CMU Sphinx получает на вход данные с микрофона, на выходе, основываясь на словаре ключевых слов, возвращает гипотезы о том, что говорит пользователь в виде текста и цифрового эквивалента точности. В полученном текстовом файле производится поиск по спискам ключевых слов подкорпусов, после чего сообщение относится к той дорожной ситуации, ключевых слов, соответствующих которой было

найденно больше. На выходе мы получаем текстовый файл с темой сообщения и списком ключевых слов (Рисунок). Например, из сообщения “Перед Администрацией города после утечки газа или чего-то подобного, полностью разрушено дорожное покрытие” будут выделены ключевые слова “утечка”, “газ”, “покрытие”, соответствующие подкорпусу Аварийно-опасные участки дорог и, следовательно, сообщение будет отнесено к дорожной ситуации Аварийно-опасные участки дорог.



Алгоритм определения темы сообщения

Выводы

На первом этапе исследования был собран и обработан корпус сообщений, описывающих различные дорожные проблемы, выделено 11 подкорпусов и для каждого из них составлены частотные словари и списки ключевых слов. Была выбрана система распознавания речи, составлен словарь ключевых слов для распознавания и реализован алгоритм выделения темы сообщения на основе списков ключевых слов, соответствующим различным дорожным проблемам. Итогом первого этапа исследования стал прототип системы голосового аннотирования дорожных ситуаций водителями. На основе полученных материалов в дальнейшем предполагается разработка системы голосового аннотирования, применимая в реальных дорожных условиях.

Список литературы

1. Cameron, H. Ghent, Belgium, Speech at the Interface., 2000 r., In Workshop on 'Voice Operated Telecom Services', стр. 1–7.
2. Chen F., Ing-Marie Jonsson, Jessica Villing, and Staffan Larsson, Application of Speech Technology in Vehicles, Speech Technology, Springer US, 2010.
3. CMU Sphinx [Электронный ресурс] // Информационный сайт проекта CMU Sphinx. – Режим доступа: www.cmusphinx.sourceforge.net/wiki/start.
4. Huang X., Acero A., Hon H.-W. Spoken Language Processing. Guide to Algorithms and System Development, 2001.
5. Streetbump [Электронный ресурс] // Сайт проекта Streetbump. – Режим доступа: www.cityofboston.gov/DoIT/apps/streetbump.asp.

Рецензенты:

Григорьев В.А., д.т.н., профессор, заведующий кафедрой беспроводных телекоммуникаций
Университета ИТМО, г. Санкт-Петербург;

Парфенов В.Г., д.т.н., профессор, декан Факультета информационных технологий и
программирования Университета ИТМО, г. Санкт-Петербург.