

Научно-производственное республиканское
унитарное предприятие
«Научно-исследовательский институт технической



О практических аспектах применения технологий автоматического распознавания речи в системах технических средств для обеспечения оперативно-розыскных мероприятий

М.А. Мазилкин
государственное предприятие «НИИ ТЗИ»



Содержание

- ❖ Введение
- ❖ Актуальность применения технологий автоматического распознавания речи
- ❖ Пример практической реализации



- ❖ Оперативно-розыскная деятельность – один из видов правоохранительной деятельности любого государства
- ❖ Основная особенность ОРД – преимущественно негласный характер

Контроль на сетях электросвязи – законный или санкционированный перехват

Lawful interception



- ❖ большое количество речевой информации – тысячи часов
- ❖ оперативность обработки ?????
- ❖ охват обработки ?????



Комплекс распознавания речевой информации «Речь»

- ❖ разработан в рамках государственной программы
- ❖ основан на недорогих серверах (2 Intel Xeon 2680, 64 Гб ОЗУ, 4 сервера)
- ❖ объем информации – не менее 4000 часов (не менее 50 тысяч файлов) в сутки
- ❖ средний процентный показатель распознанных слов в сутки для голосового трафика сетей сотовой подвижной электросвязи – не менее 70 %



Комплекс распознавания речевой информации «Речь»

- ❖ масштабируемость
- ❖ расширение используемых языков



Комплекс распознавания речевой информации «Речь»

- ❖ автоматическая классификация текстов по заданным категориям с дообучением модели на основании ручной классификации операторами (Word2vec)
- ❖ использование пополняемых пользовательских словарей фамилий, географических наименований, наименований штатных должностей и т.д
- ❖ последовательное цветовое выделение распознаваемого текста синхронно с воспроизведением фонограммы



Комплекс распознавания речевой информации «Речь»

- ❖ информирование оператора-аналитика о наличии в фонограмме ключевых слов из перечня с учетом весовых коэффициентов и других критериев
- ❖ полнотекстовый поиск по всему массиву текстов с учетом морфологии и с подсветом найденных слов (словосочетаний) в текстах фонограмм



Спасибо за внимание

Научно-производственное республиканское унитарное предприятие «Научно-исследовательский институт технической защиты информации»

✉ г.Минск, ул. Первомайская 26/2

☎ (017)-294-01-71

✉ info@niitzi.by

💻 www.niitzi.by

