

УДК 004.934

АНАЛИЗ СУЩЕСТВУЮЩИХ СВОБОДНО РАСПРОСТРАНЯЕМЫХ СИСТЕМ РАЗДЕЛЕНИЯ ДИКТОРОВ НА ФОНОГРАММЕ**Рогов А.А., Петров Е.А.***Петрозаводский государственный университет, Петрозаводск, e-mail: johnp@petsu.ru*

В последнее время возрос интерес к задаче разделения дикторов на фонограмме – известной как «who spoke when». Решение данной задачи востребовано в таких областях, как распознавание речи и поиск дикторов на аудиозаписи. В статье представлен обзор свободно распространяемого программного обеспечения, применяемого для решения задач разделения дикторов на фонограмме. Рассмотрен критерий оценки эффективности работы систем разделения дикторов – Diarization Error Rate, предложенный национальным институтом стандартов и технологий США. Проведено тестирование систем разделения дикторов на нескольких фонограммах: шесть из корпуса NIST2008-ENG и одна сторонняя фонограмма, подготовлена в соответствии с рекомендациями NIST. Тестирование проводилось в условиях отсутствия информации о количестве дикторов на аудиозаписи и об их личности. Результаты тестирования представлены в статье. Оценка качества работы систем осуществлялась с помощью сценария, предоставляемого NIST. В ходе тестирования лучшие результаты показала система LIUM. Представлены результаты тестирования работы LIUM с применением различного количества акустических признаков MFCC от 13 до 19.

Ключевые слова: фонограмма, речь, сегмент, кластеризация, диктор, LIUM, AudioSeg, ELIS, DiarTK, DER**AN OVERVIEW OF OPEN SOURCE SOFTWARE USED TO SOLVE THE SPEAKER SEGMENTATION PROBLEM ON A PHONOGRAM****Rogov A.A., Petrov E.A.***Petrozavodsk State University, Petrozavodsk, e-mail: johnp@petsu.ru*

Recently there've been growing interest in the "speaker segmentation" problem on a phonogram, known as "who spoke when". The solution to the problem is much in demand in the field of speech recognition and speaker seeking on an audio record. The article presents an overview of the open source software being used to solve the "speaker segmentation" problem on a phonogram. The consideration has been given to the Diarization Error Rate – the criteria for the speaker segmentation systems efficiency evaluation, offered by the US National Institute of Standards and Technologies. The testing of the speaker segmentation systems has been made for several phonograms: six phonograms have been taken from the NIST2008-ENG case and an external one, which has been prepared as per NIST recommendations. The testing has been made in the absence of information of the amount of speakers on the audio record and of their identity. Test results are shown in the article. The system operational quality evaluation has been made with the script provided by NIST. LIUM system has shown the best results during the testing. Additionally the LIUM was tested with different amount of acoustic features MFCC from 13 to 19 and results was presented in the paper.

Keywords: phonogram, speech, segment, clustering, speaker, LIUM, AudioSeg, ELIS, DiarTK, DER

Среди множества различных задач по обработке речи можно выделить отдельное направление – задачу разделения дикторов на фонограмме. В зарубежных источниках она формулируется как «who spoke when» [5]. Она состоит в выделении речевых сегментов фонограммы и кластеризации выделенных сегментов по принадлежности к одному диктору. Решение данной задачи востребовано в различных областях человеческой деятельности. Например, в системах аннотирования, которые добавляют к речевым аудиофайлам различные метаданные, такие как временная разметка границ фраз, информация о говорящем и др. В системах автоматического распознавания речи сегментация речи дикторов используется для адаптации моделей к речи пользователя, что повышает точность распознавания речи.

Уже более десяти лет идет активное исследование в области разделения дикторов на фонограмме. Как показывают исследования, подходы и методы, используемые для решения данной задачи, сильно зависят от области применения, условий постановки и решения задачи. Одними из наиболее сложных условий для решения данной задачи являются частая смена дикторов, отсутствие какой-либо априорной информации о количестве дикторов на фонограмме и отсутствие информации о личности дикторов, образцах их голоса и даже пола. Существует большое количество публикаций в российских и зарубежных научных журналах, в которых представлены различные методы и алгоритмы применяемые для решения задач разделения дикторов на фонограмме в разных условиях. Однако в публикациях

отсутствует полная информация о представленных алгоритмах, а также не указываются ссылки на разработанные авторами программные продукты. В связи с этим зачастую невозможно быстро воспользоваться наработками, выполненными другими исследователями, для оценки работоспособности разработанных и представленных ими методов на собственных данных, а также для дальнейшего использования их в собственных исследованиях.

Целью данной статьи является создание обзора свободно распространяемых программных продуктов и библиотек, предназначенных для организации систем разделения дикторов, а также проведение апробирования рассмотренных систем на реальных данных.

LIUM

Первым рассмотрим продукт LIUM_SpkDiarization [13]. Он разработан на основе более раннего проекта mClust, написанного на языке C++ в 2005 году [3]. Первая публичная версия LIUM была представлена в 2010 году на конференции 2010 Sphinx Workshop [7]. В отличие от своего предшественника проект LIUM был перенесен на JAVA платформу для минимизации различных проблем при переходе от одной операционной системы к другой и для большей совместимости с различными библиотеками. LIUM_SpkDiarization распространяется как единый JAR архив, который может быть запущен без необходимости использования сторонних приложений [13]. Он включает в себя полный набор инструментов для быстрого создания простых систем разделения дикторов на основе аудиофайлов. На официальном сайте проекта [13] говорится, что данный набор утилит позволяет производить расчет MFCC признаков, содержит детекторы речевой активности и различные методы сегментации и кластеризации речи.

Для создания более сложных систем разделения дикторов можно воспользоваться исходными кодами системы LIUM, которые можно скачать с официального сайта проекта [12] и доработать их для решения поставленной задачи. LIUM является свободно распространяемым программным продуктом и распространяется по лицензии GPL. Единственное условие, которое выдвигают разработчики, – это указывать в публикациях ссылку на проект LIUM, если в работе использовались библиотеки или исходные коды LIUM. С официального сайта проекта [12] мож-

но скачать BASH сценарий, который реализует простую систему разделения дикторов на фонограмме, созданную на основе LIUM. Данная система диаризации дикторов позволяет обрабатывать файлы в формате Wave или Sphere (16 kHz/16 bit PCM моно).

AudioSeg

AudioSeg [10] – набор утилит, распространяемых по лицензии GPL, разработанный в 2005 году, последняя версия пакета вышла в 2012 году [2]. Он включает в себя инструменты для реализации различных компонентов системы разделения диктора на фонограмме. Данный набор утилит написан на языке C и распространяется в виде исходных кодов. Программное обеспечение можно скачать с сайта разработчиков [10], после чего его необходимо скомпилировать на Вашем компьютере. Кроме исходных кодов программ, разработчики предоставляют программные библиотеки для языка C. На основе данных библиотек можно реализовать компоненты системы разделения дикторов, однако функционал этих библиотек ограничен. Например, такие функции, предоставляемые AudioSeg, как обнаружение тишины и Витерби декодирование, не включены в C библиотеки.

При работе с аудиофайлами система использует акустические признаки MFCC, однако в набор утилит AudioSeg не входят инструменты для построения акустических признаков на основе аудиофайла. Для этого используется сторонняя свободно распространяемая утилита Spro версии 4.0 или выше, которую необходимо скачать с сайта разработчиков [11] и скомпилировать. Более подробную информацию о процессе установки системы AudioSeg можно найти в официальной документации проекта [2]. Утилиты, входящие в пакет AudioSeg, позволяют обрабатывать аудиофайлы в формате Wave (16 kHz/16 bit PCM моно) а также (A-Law или Mu-Law 8 bit моно).

ELIS

ELIST [8] – утилита, которая позволяет производить сегментацию дикторов на аудиозаписях WAVE формата 16 kHz/16 bit моно или стерео. Данное программное обеспечение разработано в университете Гента в Бельгии. Оно распространяется в виде уже скомпилированной утилиты под 64-битную платформу Windows и Linux.

Для того чтобы получить утилиту, необходимо написать письмо разработчикам, в ответ вам вышлют файл с лицензионным соглашением который требуется заполнить, распечатать, подписать и отправить скан-копию подписанного файла разработчикам, после чего вам вышлют архив с программой. По условиям лицензии в случае использования утилиты необходимо сослаться на публикацию разработчиков [4]. Утилита настраивается через конфигурационный файл, который содержит очень ограниченный набор настроек. Например, в него входят такие параметры, как минимальная длительность речевого сегмента и минимальная длительность неречевого сегмента.

DiarTK

DiarTK [6] набор утилит, позволяющий производить сегментацию речи дикторов на аудиозаписях, последняя версия вышла в 2012 году. Набор утилит DiarTK написан на языке C++ и распространяется по лицензии GPL, его можно скачать с сайта разработчиков [16]. Он распространяется в виде архива с исходным кодом, который необходимо скомпилировать на вашем компьютере. DiarTK использует функции из пакета утилит MathLab, поэтому необходимо чтобы MathLab был установлен на компьютере. Вместе с программным обеспечением пользователю предоставляются три примера готовых систем разделения дикторов, созданных на основе набора утилит DiarTk.

При работе с аудиофайлами система использует акустические признаки, MFCC, а также может дополнительно использовать такие признаки как Time Delay of Arrivals(TDOA) и Frequency Domain Linear Prediction (FDLP)/Modulation Spectrum(MS). Однако в пакете утилит DiarTk отсутствуют инструменты для генерации этих признаков. Система предполагает использование готовых файлов с признаками в стандарте НТК. Для генерации файлов НТК стандарта можно воспользоваться свободно распространяемым набором утилит [9]. В работе [6] представлено более детальное описание алгоритмов и методов решения задачи разделения дикторов используемых DiarTk.

Критерий оценки систем разделения дикторов на фонограмме

Для оценки эффективности работы систем разделения дикторов на фонограмме существует несколько методик. Одна из существующих методик разработана национальным институтом стандартов и технологий США (National Institute of Standards and Technology, NIST). Она описывается в проекте «Rich Transcription Evaluation Project» [15], одной из задач которого является задача «Metadata Extraction Speaker Diarization Task».

В соответствии с этой методикой мерой оценки эффективности систем разделения дикторов на фонограмме выступает величина DER (Diarization Error Rate) (1), которая рассчитывается по формуле

$$DER = \frac{\sum (T(seg) \cdot \max(N_{ref}(seg), N_{sys}(seg)) - N_{correct}(seg))}{\sum (T(seg) \cdot N_{ref}(seg))} \cdot 100, \quad (1)$$

где $T(seg)$ – длительность речевого сегмента seg ; $N_{ref}(seg)$ – количество дикторов, голос которых присутствует на речевом сегменте seg в соответствии с эталонной разметкой; $N_{sys}(seg)$ – количество дикторов, голос которых присутствует на речевом сегменте seg в соответствии с результатом работы оцениваемой системы; $N_{correct}(seg)$ – количество верно отнесенных к речевому сегменту seg дикторов.

$$E_{\overline{FA}} = \frac{\sum T(seg) \cdot (N_{sys}(seg) - N_{ref}(seg))}{\sum T(seg) \cdot N_{ref}(seg)}, \quad (2)$$

$$E_{miss} = \frac{\sum T(seg) \cdot (N_{ref}(seg) - N_{sys}(seg))}{\sum T(seg) \cdot N_{ref}(seg)}, \quad (3)$$

$$E_{spkr} = \frac{\sum (T(seg) \cdot \min(N_{ref}(seg), N_{sys}(seg)) - N_{correct}(seg))}{\sum T(seg) \cdot N_{ref}(seg)}. \quad (4)$$

Величина DER является суммой трех ошибок: ошибка ложного детектирования речи E_{FA} (2), ошибка ложного пропуска речи E_{miss} (3) и ошибка разделения дикторов E_{spkr} (4). Автор работы [1] отмечает, что первые два вида ошибок относятся к оценке качества работы систем детектирования речевой активности, третья составляющая используется именно для сравнения систем разделения дикторов. Следует отметить, что, кроме представленной выше методики существуют и другие способы оценки систем diarизации, эффективность которых еще исследуется.

Подготовка и проведение экспериментов

Суть экспериментов заключается в тестировании свободно распространяемых систем разделения дикторов на реальных данных и оценке качества их работы с помощью критерия DER . В качестве тестовых данных для экспериментов использовались шесть аудиофайлов из корпуса NIST2008-ENG: ENG_fabrl.wav, ENG_fadxu.wav, ENG_fafgm.wav, ENG_fafkr.wav, ENG_fagru.wav, ENG_faicw.wav. Формат файлов Wav 16 бит, частота дискретизации 11,025 кГц, для каждого файла имеется текстовый файл с эталонной разметкой в формате RTTM. В связи с тем, что часть систем разделения дикторов требуют использовать файлы с частотой дискретизации 16 кГц, с помощью утилиты ffmpeg исходные аудиофайлы были перекодированы в файлы с частотой дискретизации 16 кГц. В список файлов для тестирования был добавлен еще один аудиофайл формата Wav, для которого в соответствии с рекомендациями NIST [14] был создан специальный файл ключевой разметки.

Для оценки качества работы систем разделения дикторов использовался сценарий на языке Perl md-eval-v21.pl [17], который предоставляет NIST. Он позволяет рассчитывать значения DER на основе двух файлов в формате RTTM, файла эталонной разметки и файла разметки, которую формирует система разделения дикторов. Каждая из рассматриваемых систем разделения дикторов сохраняет результаты своей работы в текстовых файлах своего собственного формата, поэтому были написаны Python сценарии перевода файлов в файлы, соответствующие формату RTTM.

На основе алгоритмов, предложенных автором диссертационной работы [1], разработано платное программное обеспечение, осуществляющее разделение дикторов на фонограмме. После того как мы обратились к автору работы [1], он предоставил нам результаты тестирования своей системы на 6

аудиофайлах корпуса NIST2008-ENG, для сравнения результатов работы свободно распространяемых систем.

Нами было проведено тестирование систем разделения дикторов на файлах из корпуса NIST2008-ENG. Система разделения дикторов на основе LIUM по неустановленным причинам не смогла обработать два аудиофайла: ENG_fadxu.wav, ENG_fagru.wav. В ходе выполнения процесс разделения дикторов зависал и переставал отвечать. Система ELIS некорректно обработала файл, взятый не из корпуса. Произвести тестирование системы DiarTK не удалось. При запуске тестирования она зависала или завершалась с ошибкой. Было написано письмо разработчикам системы с просьбой более подробно описать алгоритм запуска системы и алгоритм подготовки исходных данных для тестирования. Авторы не ответили.

Результаты тестирования представлены в табл. 1. В ней отображены средние значения ошибок для соответствующей системы. Отдельно в табл. 2 вынесены результаты тестирования систем на аудиофайле, взятом не из корпуса.

Таблица 1
Результаты тестирования систем на файлах из корпуса NIST2008-ENG

Система	E_{FA}	E_{miss}	E_{spkr}	DER
LIUM	40	0,05	11,8	51,85
AudioSeg	47,51	0,32	41,1	88,93
ELIS	29,78	2,37	36,5	68,65
система [1][6]	8,56	39,1	2,17	49,83

Таблица 2
Результаты тестирования систем на файле, взятом не из корпуса

Система	E_{FA}	E_{miss}	E_{spkr}	DER
LIUM	18,1	0	30,6	48,7
AudioSeg	18,1	0	82	100,1

Из табл. 1 и 2 видно, что наименьшую ошибку из протестированных систем дала система разделения дикторов на основе LIUM. По умолчанию система разделения дикторов на основе LIUM использует в качестве акустических признаков 13 MFCC, были проведены опыты с применением различного количества признаков от 13 до 19. В табл. 3 представлены результаты работы системы со стандартным количеством признаков MFCC и результаты, давшие минимальную ошибку.

Таблица 3

Результаты тестирования LIUM с различным количеством MFCC

Файл	MFCC	E_{FA}	E_{miss}	E_{spkr}	DER
ENG_fafrk	13	27,7	0	9,0	36,7
ENG_fafrk	16	27,7	0	11,6	39,3
ENG_fabrl	13	46,3	0	23,0	69,3
ENG_fabrl	14	46,3	0	23,5	69,8
ENG_fafgm	13	33,1	0	8,7	41,8
ENG_fafgm	16	33,1	0	6,3	39,4
ENG_faicw	13	52,9	0	6,5	59,4
ENG_faicw	17,18,19	52,9	0	3,8	56,7
Файл не из корпуса	13	18,1	0	30,6	48,7
Файл не из корпуса	19	18,1	0	19,9	38

Выводы

Рассмотрены и представлены свободные распространяемые программные продукты для организации систем разделения дикторов. В ходе тестирования наилучшие результаты показали система разделения дикторов, созданная на основе алгоритмов [1], а также система разделения дикторов на основе LIUM. Однако, учитывая, что тестирование систем проводилось с использованием настроек по умолчанию, однозначно выделить наиболее удачные системы разделения дикторов нельзя. Заметим, что суммарная длительность аудиоданных, на которых проводилось тестирование, была недостаточной для окончательных выводов о работоспособности систем – всего 30 минут. Обычно тесты систем проводятся на аудиоданных суммарной длительностью более нескольких часов.

Работа выполнена при финансовой поддержке Программы стратегического развития ПетрГУ в рамках реализации комплекса мероприятий по развитию научно-исследовательской деятельности.

Список литературы

1. Кудашев О.Ю. Система разделения дикторов на основе вероятностного линейного дискриминантного анализа. дис. ... канд. техн. наук. – СПб.: Санкт-Петербургский нац. исслед. ун-т. Информационных технологий механики и оптики, 2014. – 158 с.
2. Gravier G., Betsler M., Ben M. audiosseg: Audio Segmentation Toolkit, release 1.2., IRISA, january 2010.
3. Rouvier M., Dupuy G., Gay P., Khoury E., Merlin T., Meignier S. An Open-source State-of-the-art Toolbox for Broadcast News Diarization, Interspeech, Lyon (France), 25-29 Aug. 2013.

4. Vandecatseye A., Martens J.P. A fast, accurate and stream-based speaker segmentation and clustering algorithm, Proceeding Eurospeech (Geneva), 941–944.

5. Vijayasenan D., Valente F., Bourlard H. An Information Theoretic Combination of MFCC and TDOA Features for Speaker Diarization, in: IEEE Transactions on Audio Speech and Language Processing, 19(2), 2011.

6. Vijayasenan D., Valente F. DiarTk: An open source toolkit for research in multistream speaker diarization and its application to meetings recordings, in Proceedings of Interspeech, Portland, Oregon (USA), 2012.

7. Meignier S., Merlin T. LIUM SpkDiarization: an open source toolkit for diarization, in CMU SPUD Workshop, Dallas, Texas (USA), March 2010.

8. EAST: The ELIS Audio Segmentation Tool | DSSP – ELIS – UGent [Электронный ресурс]. – Режим доступа: <http://dssp.elis.ugent.be/download/audio-segmentation-software> (дата обращения: 05.04.2015).

9. HTK Speech Recognition Toolkit. [Электронный ресурс]. – Режим доступа: <http://htk.eng.cam.ac.uk/download.shtml> (дата обращения: 05.04.2015).

10. InriaForge: AudioSeg: Project Home. [Электронный ресурс]. – Режим доступа: <https://gforge.inria.fr/projects/audiosseg/> (дата обращения: 05.04.2015).

11. InriaForge: SPro: Project Home. [Электронный ресурс]. – Режим доступа: <http://gforge.inria.fr/projects/spro> (дата обращения: 05.04.2015).

12. LIUM Speaker Diarization Wiki. Download. [Электронный ресурс]. – Режим доступа: <http://www-lium.univ-lemans.fr/diarization/doku.php/download> (дата обращения: 05.04.2015).

13. LIUM Speaker Diarization Wiki. [Электронный ресурс]. – Режим доступа: <http://www-lium.univ-lemans.fr/diarization/doku.php> (дата обращения: 05.04.2015).

14. NIST, «The 2009 (RT-09) Rich Transcription Meeting Recognition Evaluation Plan» [Электронный ресурс]. – Режим доступа: <http://www.itl.nist.gov/iad/mig/tests/rt/2009/docs/rt09-meeting-eval-plan-v2.pdf> (дата обращения: 05.04.2015).

15. NIST Rich Transcription Evaluation Project. [Электронный ресурс]. – Режим доступа: <http://www.itl.nist.gov/iad/mig/tests/rt/> (дата обращения: 05.04.2015).

16. Speaker Diarization Toolkit. Idiap Research Institute. [Электронный ресурс]. – Режим доступа: <http://www.idiap>

ch/scientific-research/resources/speaker-diarization-toolkit (дата обращения:05.04.2015).

17. md-eval-v21.pl [Электронный ресурс]. – Режим доступа: <http://www.itl.nist.gov/iad/mig/tests/rt/2006-spring/code/md-eval-v21.pl> (дата обращения:05.04.2015).

References

1. Kudashev O.Ju. Sistema razdelenija diktorov na osnove verojatnostnogo linejnogo diskriminantnogo analiza. dis. ... kand. tehn. nauk. SPb.: Sankt-Peterburgskij nac. issled. un-t. Informacionnyh tehnologij mehaniki i optiki, 2014. 158 s.

2. Gravier G., Betsler M., Ben M. audioseg: Audio Segmentation Toolkit, release 1.2., IRISA, January 2010.

3. Rouvier M., Dupuy G., Gay P., Khoury E., Merlin T., Meignier S. An Open-source State-of-the-art Toolbox for Broadcast News Diarization, Interspeech, Lyon (France), 25-29 Aug. 2013.

4. Vandecatseye A., Martens J.P. A fast, accurate and stream-based speaker segmentation and clustering algorithm, Proceeding Eurospeech (Geneva), 941–944.

5. Vijayasenan D., Valente F., Bourlard H. An Information Theoretic Combination of MFCC and TDOA Features for Speaker Diarization, in: IEEE Transactions on Audio Speech and Language Processing, 19(2), 2011.

6. Vijayasenan D., Valente F. DiarTk: An open source toolkit for research in multistream speaker diarization and its application to meetings recordings, in Proceedings of Interspeech, Portland, Oregon (USA), 2012.

7. Meignier S., Merlin T. LIUM SpkDiarization: an open source toolkit for diarization, in CMU SPUD Workshop, Dallas, Texas (USA), March 2010.

8. EAST: The ELIS Audio Segmentation Tool | DSSP ELIS UGent [Elektronnyj resurs]. Rezhim dostupa: <http://dssp.elis.ugent.be/download/audio-segmentation-software> (data obrashhenija:05.04.2015).

9. HTK Speech Recognition Toolkit. [Elektronnyj resurs]. Rezhim dostupa: <http://htk.eng.cam.ac.uk/download.shtml> (data obrashhenija:05.04.2015).

10. InriaForge: AudioSeg: Project Home.[Elektronnyj resurs]. Rezhim dostupa: <https://gforge.inria.fr/projects/audioseg/> (data obrashhenija:05.04.2015).

11. InriaForge: SPro: Project Home. [Elektronnyj resurs]. Rezhim dostupa: <http://gforge.inria.fr/projects/spro> (data obrashhenija:05.04.2015).

12. LIUM Speaker Diarization Wiki. Download. [Elektronnyj resurs]. Rezhim dostupa: <http://www-lium.univ-lemans.fr/diarization/doku.php/download> (data obrashhenija:05.04.2015).

13. LIUM Speaker Diarization Wiki. [Elektronnyj resurs]. Rezhim dostupa: <http://www-lium.univ-lemans.fr/diarization/doku.php>. (data obrashhenija:05.04.2015).

14. NIST, «The 2009 (RT-09) Rich Transcription Meeting Recognition Evaluation Plan» [Elektronnyj resurs]. Rezhim dostupa: <http://www.itl.nist.gov/iad/mig/tests/rt/2009/docs/rt09-meeting-eval-plan-v2.pdf> (data obrashhenija:05.04.2015).

15. NIST Rich Transcription Evaluation Project. [Elektronnyj resurs]. Rezhim dostupa: <http://www.itl.nist.gov/iad/mig/tests/rt/>. (data obrashhenija:05.04.2015).

16. Speaker Diarization Toolkit. Idiap Research Institute. [Elektronnyj resurs]. Rezhim dostupa: <http://www.idiap.ch/scientific-research/resources/speaker-diarization-toolkit> (data obrashhenija:05.04.2015).

17. md-eval-v21.pl [Elektronnyj resurs]. Rezhim dostupa: <http://www.itl.nist.gov/iad/mig/tests/rt/2006-spring/code/md-eval-v21.pl> (data obrashhenija:05.04.2015).

Рецензенты:

Колесников Г.Н., д.т.н., профессор, зав. кафедрой общетехнических дисциплин, Институт лесных, инженерных и строительных наук, Петрозаводский государственный университет, г. Петрозаводск;

Печников А.А., д.т.н., доцент, ведущий научный сотрудник, Институт прикладных математических исследований, Карельский научный центр РАН, г. Петрозаводск.