

3. Kersta L. G. Voiceprint Identification // Nature. 1962. Vol. 196, N 4861. P. 1253—1257.
4. Матвеев Ю. Н. Технологии биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н. Э. Баумана. Сер. „Приборостроение“. 2012. № 3(3). С. 46—61.
5. Harrison P. Variability of Formant Measurements. Submitted in partial fulfilment of the degree of MA. University of York, UK, 2004.
6. Enzinger E., Zhang C., Morrison G. S. Voice source features for forensic voice comparison — an evaluation of the Glottex software package // Proc. of Odyssey. Singapore, 2012. P. 78—85.
7. Hollien H. Forensic voice identification. NY, 2002.
8. Программный комплекс „ОТExpert“ [Электронный ресурс]: <[http://фоноскопическая.рф/forensic\\_audio-video/forensic\\_tools/otexpert/](http://фоноскопическая.рф/forensic_audio-video/forensic_tools/otexpert/)>.
9. Каганов А. Ш. Криминалистическая экспертиза звукозаписей. М.: Юрлитинформ, 2005.
10. Идентификация лиц по фонограммам русской речи на автоматизированной системе „Диалект“: Пособие для экспертов / Под ред. А. В. Фесенко. М., 1996.
11. Тимофеев И. Н. Применение автоматизированной системы „Диалект“ на базе компьютерной речевой лаборатории CLS (США) при решении задач идентификации дикторов: Метод. рекомендации. М.: ЭКЦ МВД России, 2000.
12. Матвеев Ю. Н. Оценка доверительного интервала общего решения ансамбля классификаторов // Изв. вузов. Приборостроение. 2013. Т. 56, № 2. С. 74—79.

**Сведения об авторах****Елена Владимировна Булгакова**— аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра речевых информационных систем;  
E-mail: bulgakova@speechpro.com**Екатерина Викторовна Краснова**— ООО „ЦРТ“, Санкт-Петербург; научный сотрудник;  
E-mail: krasnova@speechpro.comРекомендована кафедрой  
речевых информационных системПоступила в редакцию  
22.10.13 г.

УДК 004.93+57.087.1

Д. В. ДЫРМОВСКИЙ, С. Л. КОВАЛЬ, М. В. ХИТРОВ

**КОНЦЕПЦИЯ СИСТЕМЫ НАЦИОНАЛЬНОГО ФОНОУЧЕТА  
И ГОЛОСОВОГО БИОМЕТРИЧЕСКОГО ПОИСКА**

Представлена концепция системы национального фоноучета и голосового биометрического поиска.

**Ключевые слова:** фоноучет, голосовой биометрический поиск, инфраструктура системы.

**Введение.** Многие государства в настоящее время столкнулись с целым рядом задач, связанных с отражением угроз нелегальной эмиграции и контрабанды, предотвращением террористических атак и применением новых подходов при создании систем национальной и международной безопасности. Особую важность в связи с этим приобретает совершенствование систем учета, поиска, выявления и идентификации лиц, нарушивших закон. Решение подобных задач в глобальном масштабе требует применения передовых технологических решений [1].

К инфраструктуре системы национального фоноучета могут быть предъявлены те же требования, что и к инфраструктуре любой информационной системы национального масштаба: экономическая эффективность, с возможностью масштабирования и модернизации (в случае появления более новых технологий) при минимальных дополнительных инвестициях. Соответствовать таким требованиям может система с модульной структурой, отвечающая следующим критериям [2]:

- 1) достаточно общая и гибкая физическая инфраструктура для совершенствования по мере изменения технологий;
- 2) возможность интеграции в информационную систему как можно большего количества существующего в различных учреждениях и компаниях оборудования;
- 3) сетевой сервер размещается таким образом, чтобы было возможным его эффективное управление и обслуживание.
- 4) использование таких топологий соединений, которые будут как экономически эффективными, так и управляемыми.

**Принципы построения систем.** Основными задачами систем национального фоноучета и голосового биометрического поиска являются:

- централизованный сбор и хранение образцов речи и голосовых биометрических признаков личностей, подлежащих учету;
- хранение сопутствующей информации — фотоизображений, установочных данных и любых других данных (метаданных), характеризующей личность или обстоятельства получения образца речи;
- поиск по биометрическим признакам и сопровождающей информации дикторов по всей базе данных или ее части;
- автоматическая идентификация/верификация с целью определения или подтверждения личности стоящего на учете лица.

Система национального фоноучета и голосового биометрического поиска, способная решить множество задач в области правоохранительной деятельности, должна быть построена на следующих принципах:

- 1) *территориальная распределенность* — возможность вести учет и осуществлять поиск по банкам системы из любого подразделения правоохранительных органов;
- 2) *открытость и консолидация данных* — возможность добавлять в систему банки и базы речевых данных, содержащие иную информацию, например, цифровые модели других биометрических признаков (лица, отпечатков пальцев, сетчатки глаза, ДНК) или учетные личные карточки подозреваемых;
- 3) *безопасность данных* — должен быть реализован контроль и разграничение доступа к данным системы на базе голосовой биометрии (возможно использование традиционных систем);
- 4) *автоматизация рабочих процессов* — в функционирование системы должен быть вовлечен каждый пользователь согласно его компетенциям через гибкие механизмы полей, групп и рабочих процессов.

**Территориальная распределенность системы.** Сбор биометрической информации у населения также должен производиться непосредственно на местах (муниципальные образования, организации).

Для решения этой задачи необходимо, чтобы:

- был обеспечен доступ к системе всех пользователей по защищенному протоколу HTTPS, сочетающему криптографическую стойкость и повсеместную доступность как в локальных и ведомственных сетях, так и в сети Интернет;
- система не имела технических ограничений ни по объему хранимой информации, ни по числу пользователей, которые получают доступ к ресурсам центрального хранилища и поисковой машины удаленно;

— обеспечивались web-интерфейс и поддержка большинства браузеров в совокупности с защищенными протоколами передачи данных для высокой мобильности специалиста, ведущего расследование. Благодаря этому можно будет образцы биометрических данных передавать в систему непосредственно с места совершения преступлений и там же получать результаты поиска.

На рис. 1 представлена схема территориального распределения пользователей системы.

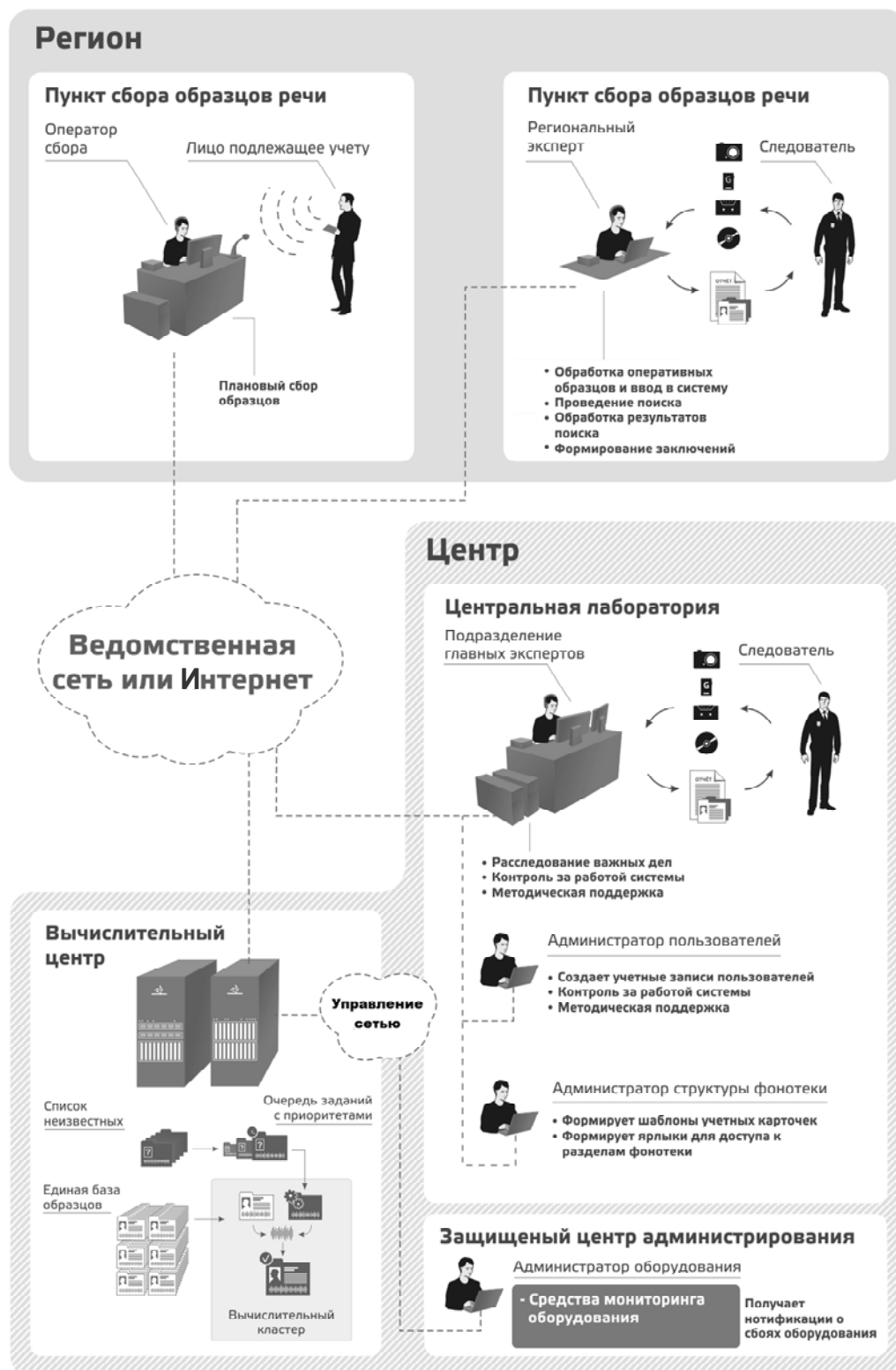


Рис. 1

**Функциональные уровни системы.** Система голосового биометрического поиска и национального криминалистического фоноучета включает три уровня.

1. *Сбор данных, подлежащих криминалистическому учету:* учетные карточки разыскиваемых лиц и образцы их речи, пригодные для идентификации.

2. *Хранение и обработка данных, осуществление вычислительных операций:* хранение банка биометрических данных, база данных учетных карточек, осуществление непосредственно идентификации и поиска по банку данных с помощью мощного вычислительного кластера.

3. *Операции пользователя:* постановка заданий на поиск, определение параметров поиска, вывод результатов в форме отчетов, оперативная нотификация сотрудников в случае обнаружения разыскиваемого лица в каналах связи или банках фонограмм, пополняемых в процессе оперативно-розыскных мероприятий, задания экспертам на обработку и анализ результатов поиска и фонограмм.

**Уровень сбора данных** базируется на существующих в правоохранительных органах технологиях сбора и учета криминалистических данных: результаты проведения оперативно-розыскных мероприятий (в том числе с помощью СОПМ, скрытых средств записи переговоров и т.д.); аудиозаписи приема вызовов дежурными частями территориальных подразделений; аудиозаписи дачи показаний; судебных разбирательств; данные экспертного фоноскопического учета.

Кроме того, в соответствии с законодательством возможна организация планового сбора образцов речи у населения: заключенных в тюрьмах; задержанных, находящихся в полицейских участках; сотрудников правоохранительных органов; населения в точках оформления удостоверений личности.

**Уровень обработки данных** должен опираться на возможности современных технологий голосовой биометрической идентификации личностей и мировой опыт внедрения систем фоноучета различного масштаба — от экспертной лаборатории до национальной системы. Выделим основные характеристики, которыми должна обладать система голосового биометрического поиска и национального криминалистического фоноучета:

— высокая надежность поиска/идентификации по голосу (не менее 97 % правильной идентификации) с использованием языко- и текстонезависимых биометрических методов;

— высокая скорость поиска (поиск из 20 000 образцов за 3—10 мин при одновременной обработке не менее 100 запросов);

— надежная защита передачи и хранения информации (использование протокола HTTPS, распределение решения по слоям безопасности на аппаратном уровне);

— быстрота развертывания (возможность доступа по любым сетям с использованием защищенных протоколов, наличие web-интерфейса, автоматическое обновление клиентского программного обеспечения);

— работа в режиме высокой готовности 24×7 (автоматическое перераспределение нагрузки в случае выхода оборудования из строя, резервирование на всех уровнях, возможность обновления программного обеспечения без остановки работы, дистанционный мониторинг оборудования);

— эффективная работа с большим числом пользователей (автоматическая балансировка нагрузки, кластерные параллельные вычисления, очередь заданий с приоритетами);

— масштабируемость с возможностью расширения системы без остановки работы (от рабочей станции для нужд подразделения до вычислительного центра для национальной системы).

**Уровень операций пользователя** должен обеспечивать гибкую схему рабочих процессов и ролей пользователей. Каждый пользователь согласно своей роли и положению в иерархии групп, с одной стороны, получает доступ к необходимой для его работы информации, а с

другой — привлекается к решению тех задач, в которых он наиболее компетентен. Таким образом, высокая степень автоматизации и простота использования позволяют обучать пользователей непосредственно в процессе эксплуатации системы, а общие для всех пользователей средства (оповещения по e-mail, настраиваемые рабочие процессы) позволяют системе эффективно решать правоохранные задачи.

**Рабочие процессы системы.** Работа любой информационной системы национального масштаба невозможна без разделения ролей пользователей, автоматизации рабочих процессов и аудита, что обеспечивает такие свойства системы, как безопасность выполняемых пользователями операций и эффективность работы каждого пользователя вне зависимости от его квалификации.

Основными сценариями использования системы криминалистического учета являются помещение образцов в базу и поиск образцов по степени сходства.

Можно выделить различные логические типы поиска:

1) поиск „известного среди известных“:

— проверка постановки на учет уже имеющегося в базе диктора (исключение постановки на учет диктора под чужим именем);

2) поиск „неизвестного среди известных“ и наоборот:

— установление личности интересующего диктора;

3) поиск „неизвестного среди неизвестных“:

— проверка принадлежности образцов дикторов одному или нескольким людям;

— проверка причастности диктора к нескольким преступлениям.

Рабочие процессы системы должны обеспечивать максимальную эффективность и оперативность всех перечисленных типов поиска.

Существенной сложностью при внедрении систем фоноучета является подготовка персонала. При внедрении систем на уровне лабораторий все пользователи должны иметь квалификацию экспертов-фоноскопистов, подготовка которых является долгим и дорогим процессом. Внедрение системы национального масштаба с высоким уровнем автоматизации рабочих процессов позволяет преодолеть эту проблему за счет повышения квалификации оператора системы, от выполнения самых простых задач до уровня эксперта. Оператор обучается постепенно, выполняя практические задачи уже с первых дней работы.

Рекомендуется выделять следующие роли пользователей системы:

— оператор сбора образцов — производит плановый сбор образцов речи и отвечает только за качественную запись и соответствие учетных данных известной личности;

— оператор проверки образцов — анализирует качество записанных или полученных в результате оперативно-розыскной деятельности образцов и при необходимости сегментирует их;

— оператор поиска — выполняет несложные поисковые запросы, требующие немедленного ответа.

— эксперт — анализирует результаты поиска для предоставления их в суд.

Квалифицированные пользователи могут совмещать несколько ролей.

Схемы рабочих процессов разрабатываются в зависимости от возможностей и потребностей ведомства, внедряющего эту систему. Рабочие процессы не являются фиксированными во времени и могут эволюционировать вместе с системой. При внедрении может быть выбрана наиболее простая схема, а по мере вовлечения в работу с системой новых пользователей могут вводиться как новые роли, так и новые схемы рабочих процессов.

На рис. 2, 3 приведены схемы жизненного цикла учетной карточки в системе криминалистического учета, отражающие автоматизацию рабочих процессов. В данных примерах учетные карточки известных и неизвестных личностей обрабатываются по разным схемам рабочих процессов.

Система обеспечивает высочайшую степень автоматизации. Таким образом, любая учетная карточка, введенная в систему, будет обработана (с привлечением экспертов, если это требуется) и использована для получения новой информации, полезной при расследовании или в оперативно-розыскных мероприятиях.

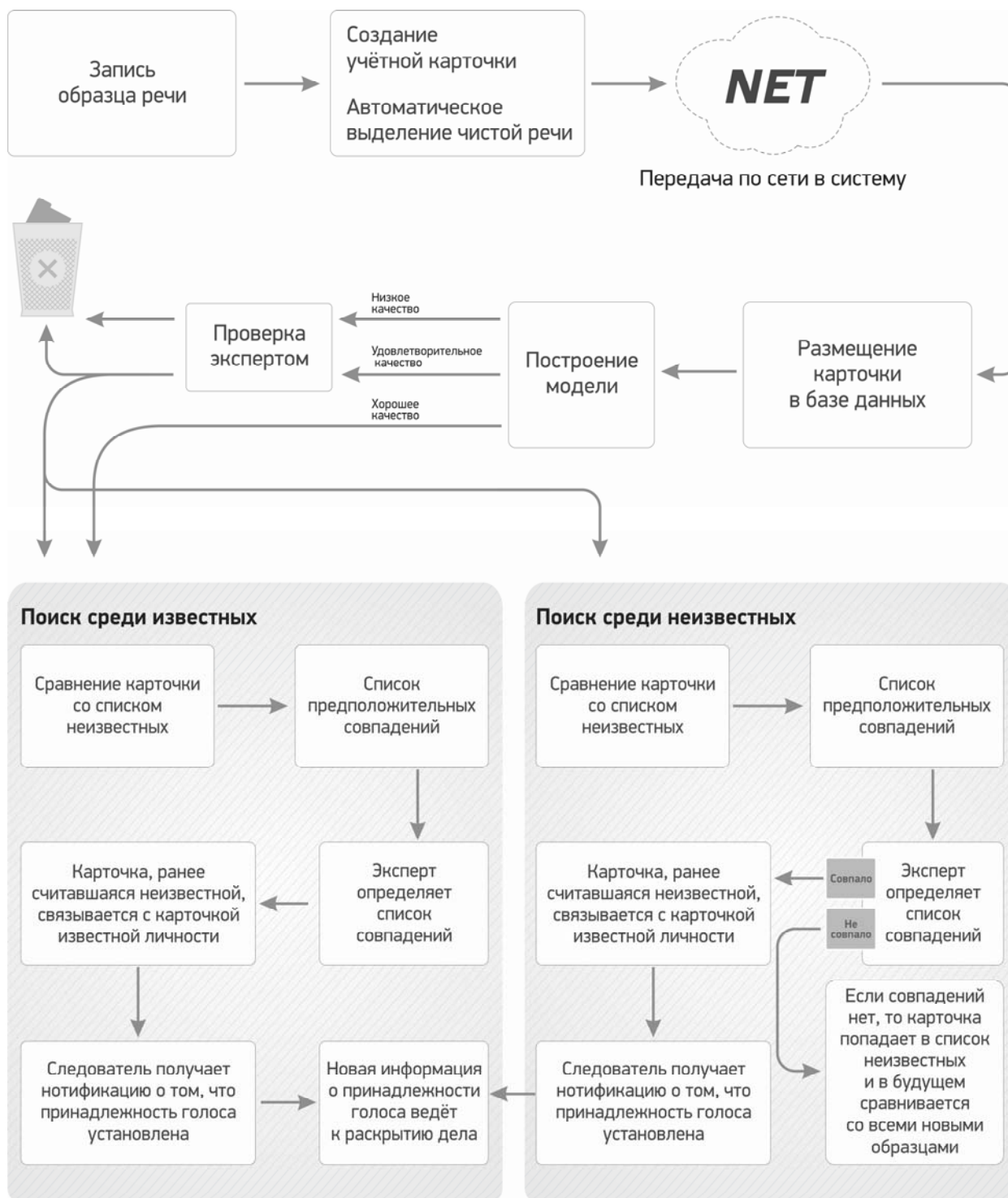


Рис. 2

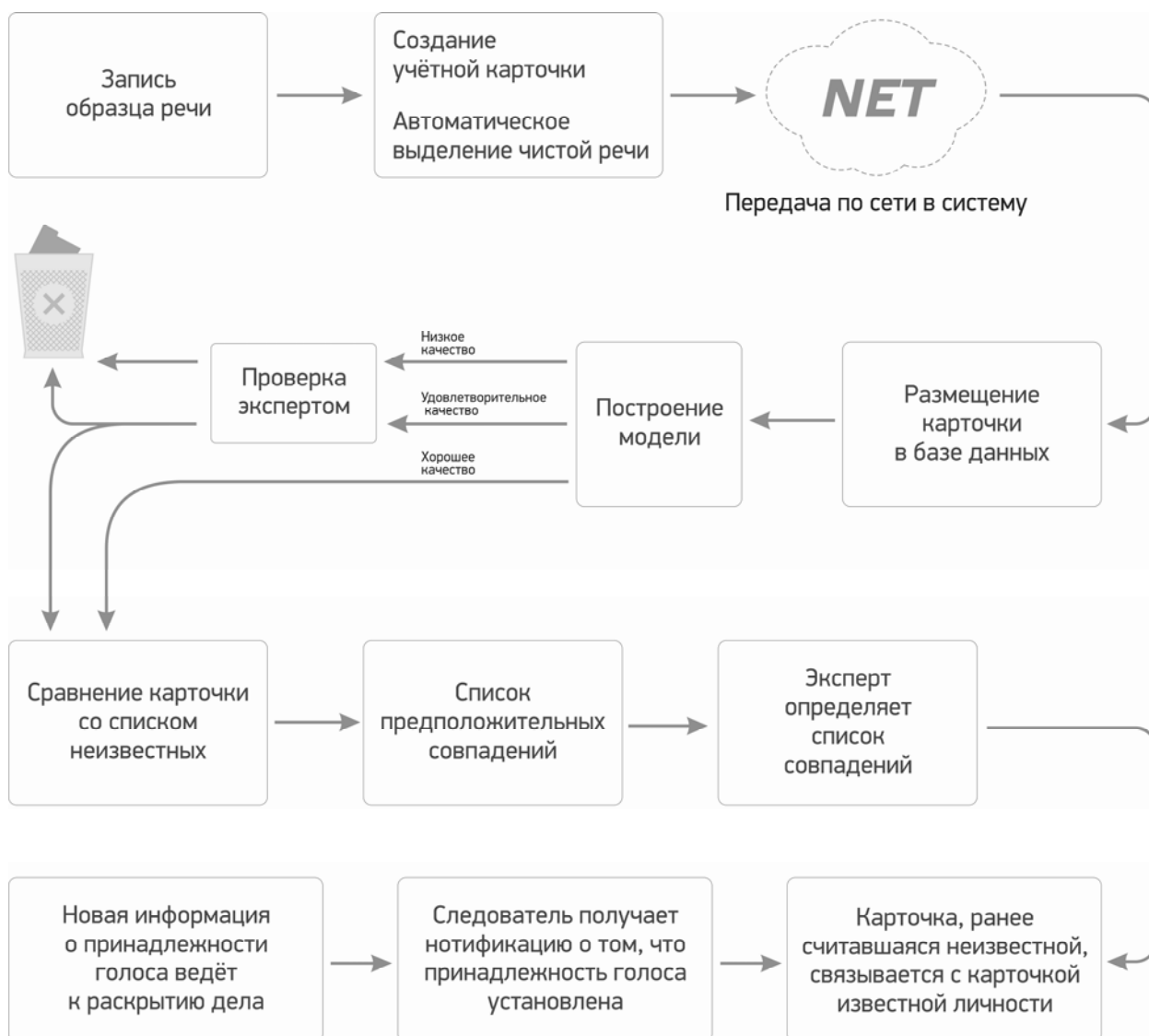


Рис. 3

Автоматизация рабочих процессов обеспечивает максимально оперативную обработку речевого материала, являющегося уликой.

**Заключение.** В работе предложена концепция построения системы национального фоноучета и голосового биометрического поиска, положенная в основу реальных систем, развернутых ООО „ЦРТ“ в различных странах (Мексика, Эквадор) и ведомствах (МВД, ФСКН и др.).

Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

#### СПИСОК ЛИТЕРАТУРЫ

1. Настасенко М. В., Дырмовский Д. В. Эффективное использование речевой информации и биометрических технологий в силовых структурах — научно-технические решения // Сб. тр. XX Междунар. науч. конф. „Информатизация и информационная безопасность правоохранительных органов“. М., 2011. С. 356—359.
2. Nair R. R. Strategic plans for designing information systems under local government // Proc. 18th National IASLIC Seminar. Thrissur, Kerala: Agricultural University, 1998. P. 11.

3. Настасенко М. В., Дырмовский Д. В. Эффективное использование речевой информации и биометрических технологий в силовых структурах // Вестн. МГТУ им. Н. Э. Баумана. Сер. „Приборостроение“. 2011. Вып. № 3. С. 18—25.
4. Матвеев Ю. Н. Технологии биометрической идентификации личности по голосу и другим модальностям // Вестн. МГТУ им. Н. Э. Баумана. Сер. „Приборостроение“. 2012. № 3 (3). С. 46—61.
5. Дырмовский Д. В., Коваль С. Л. Особенности человеко-машинного интерфейса современных систем биометрической идентификации // Изв. вузов. Приборостроение. 2013. Т. 56, № 2. С. 66—74.

#### *Сведения об авторах*

- Дмитрий Викторович Дырмовский** — ООО „ЦРТ“, Санкт-Петербург; директор филиала; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра речевых информационных систем; соискатель; E-mail: ddv@speechpro.com
- Сергей Львович Коваль** — канд. техн. наук, доцент; ООО „ЦРТ“, Санкт-Петербург; главный эксперт; E-mail: koval@speechpro.com
- Михаил Васильевич Хитров** — канд. техн. наук; ООО „ЦРТ“, Санкт-Петербург; генеральный директор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра речевых информационных систем; заведующий кафедрой; E-mail: khitrov@speechpro.com

Рекомендована кафедрой  
речевых информационных систем

Поступила в редакцию  
22.10.13 г.

УДК 004.93+57.087.1

Ю. Н. МАТВЕЕВ, А. К. ШУЛИПА

## АНАЛИЗ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ НА ОСНОВЕ МНОГООБРАЗИЙ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ДИКТОРОВ

Исследованы особенности применения методов обучения на основе многообразий, широко используемых в приложениях по распознаванию изображений, для решения задач распознавания личностей по голосу (дикторов). Проанализированы результаты экспериментов по использованию таких методов.

**Ключевые слова:** обучение на основе многообразий, распознавание диктора.

**Введение.** Алгоритмы машинного обучения на основе многообразий [1] пока мало применяются в системах распознавания дикторов. Для текстонезависимого распознавания дикторов, как правило, используются методы, базирующиеся на моделировании статистических распределений речевых признаков на основе смесей гауссовых распределений, GMM [2]. Согласно оценке Национального института стандартов и технологий США (NIST), компании, занимающие лидирующие позиции в распознавании дикторов, реализуют и совершенствуют свои алгоритмы в рамках подходов на основе гауссовых смесей [3, 4]. Тем не менее в некоторых работах [5, 6] предпринимались попытки использования алгоритмов машинного обучения на основе многообразий для решения задач распознавания дикторов.

В настоящей работе рассмотрены такие алгоритмы, описываются результаты их использования и делается заключение о возможности их применения для решения задач распознавания дикторов.

**Метод диффузных карт.** В работе [5] решалась задача текстонезависимой идентификации дикторов. Для отображения статистических моделей речевых признаков на низкоразмер-