

Эдита К. Куулар, Андрей И. Труфанов, Алексей А. Тихомиров  
ДВУХКОМПОНЕНТНАЯ СЕТЕВАЯ МОДЕЛЬ В ТЕХНОЛОГИЯХ ГОЛОСОВОЙ  
ИДЕНТИФИКАЦИИ ЛИЧНОСТИ

---

Эдита К. Куулар<sup>1</sup>, Андрей И. Труфанов<sup>1</sup>, Алексей А. Тихомиров<sup>2</sup>  
<sup>1</sup>Иркутский Национальный Исследовательский Технический Университет,  
Лермонтова, 83, Иркутск, 6664074, Россия  
e-mail: kuular1991@mail.ru, <https://orcid.org/0000-0003-4832-3397>  
e-mail: troufan@gmail.com, <https://orcid.org/0000-0002-6967-3495>  
<sup>2</sup>Университет ИНХА,  
Инcheon, Республика Корея  
e-mail: alexeitikhomirovprof@gmail.com, <https://orcid.org/0000-0002-7317-5851>

ДВУХКОМПОНЕНТНАЯ СЕТЕВАЯ МОДЕЛЬ В ТЕХНОЛОГИЯХ ГОЛОСОВОЙ  
ИДЕНТИФИКАЦИИ ЛИЧНОСТИ  
DOI: <http://dx.doi.org/10.26583/bit.2018.1.08>

*Аннотация.* Среди важнейших параметров биометрических систем голосовой модальности, определяющих их эффективность, наряду с надежностью и помехоустойчивостью выделяют скорость идентификации и верификации личности. Данный параметр особенно чувствителен при обработке крупномасштабных баз голосовых данных в режиме реального времени. Многие исследовательские работы в данной области направлены на разработку новых и совершенствование имеющихся алгоритмов представления и обработки речевых записей, обеспечивающих высокую производительность голосовых биометрических систем. Здесь перспективным представляется современный подход к решению сложных объемных задач с большим числом элементов и учетом их взаимосвязей, использующий платформу комплексных сетей. Так, известны работы, в которых при решении задач анализа и распознавания лиц по фотографиям изображения трансформируются в комплексные сети для последующей их обработки стандартными приемами. Одним из первых применений комплексных сетей в задачах анализа звуковых рядов (музыкальных и речевых) являлось описание их частотных характеристик с помощью построения сетевых моделей – конвертации рядов в сети. На платформе сетевой онтологии в данной работе развит предложенный ранее метод представления аудиоинформации с целью ее высокопроизводительного автоматического анализа и снижения погрешности в распознавании личности – диктора. Для этого предполагается конвертирование звуковой информации в форму ассоциативной семантической (когнитивной) сетевой структуры, содержащей две компоненты: амплитудную и частотную. Были записаны два речевых примера с последующей их трансформацией в соответствующие сети и сравнением топологических метрик. Набор топологических метрик каждой из сетевых моделей (амплитудной и частотной) представляет собой вектор, а вместе – матрицу, как цифровой "сетевой" отпечаток голоса. Предлагаемый сетевой подход с его чувствительностью к состоянию индивида – физиологическому, психологическому, эмоциональному, может оказаться полезным не только для задач идентификации личности, но и для определения ее состояния.

*Ключевые слова:* биометрические модальности, звуковая информация, идентификация диктора, комплексные сети, амплитуда, частота.

*Для цитирования.* КУУЛАР, Эдита К.; ТРУФАНОВ, Андрей И.; ТИХОМИРОВ, Алексей А. ДВУХКОМПОНЕНТНАЯ СЕТЕВАЯ МОДЕЛЬ В ТЕХНОЛОГИЯХ ГОЛОСОВОЙ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ. *Безопасность информационных технологий*, [S.l.], v. 25, n. 1, p. 81-89, 2018. ISSN 2074-7136. Доступно на: <<https://bit.mephi.ru/index.php/bit/article/view/1096>>. Дата доступа: 15 feb. 2018. doi:<http://dx.doi.org/10.26583/bit.2018.1.08>.

Edita K. Kuular<sup>1</sup>, Andrey I. Trufanov<sup>1</sup>, Alexei A. Tikhomirov<sup>2</sup>

<sup>1</sup>*Irkutsk National Research Technical University,*

*Lermontov street, 83, Irkutsk, 664074, Russia*

*e-mail: kuular1991@mail.ru, <https://orcid.org/0000-0003-4832-3397>*

*e-mail: troufan@gmail.com, <https://orcid.org/0000-0002-6967-3495>*

<sup>2</sup>*INHA University,*

*Incheon, Republic of Korea*

*e-mail: alexeitikhomirovprof@gmail.com, <https://orcid.org/0000-0002-7317-5851>*

### **Two-component network model in voice identification technologies**

DOI: <http://dx.doi.org/10.26583/bit.2018.1.08>

*Abstract.* Among the most important parameters of biometric systems with voice modalities that determine their effectiveness, along with reliability and noise immunity, a speed of identification and verification of a person has been accentuated. This parameter is especially sensitive while processing large-scale voice databases in real time regime. Many research studies in this area are aimed at developing new and improving existing algorithms for presentation and processing voice records to ensure high performance of voice biometric systems. Here, it seems promising to apply a modern approach, which is based on complex network platform for solving complex massive problems with a large number of elements and taking into account their interrelationships. Thus, there are known some works which while solving problems of analysis and recognition of faces from photographs, transform images into complex networks for their subsequent processing by standard techniques. One of the first applications of complex networks to sound series (musical and speech) analysis are description of frequency characteristics by constructing network models - converting the series into networks. On the network ontology platform a previously proposed technique of audio information representation aimed on its automatic analysis and speaker recognition has been developed. This implies converting information into the form of associative semantic (cognitive) network structure with amplitude and frequency components both. Two speaker exemplars have been recorded and transformed into pertinent networks with consequent comparison of their topological metrics. The set of topological metrics for each of network models (amplitude and frequency one) is a vector, and together those combine a matrix, as a digital "network" voiceprint. The proposed network approach, with its sensitivity to personal conditions-physiological, psychological, emotional, might be useful not only for person identification, but also for determining his /her condition.

*Keywords:* *biometric modalities, sound information, speaker identification, complex net-works, amplitude, frequency.*

*For citation.* **KUULAR, Edita K.; TRUFANOV, Andrey I.; TIKHOMIROV, Alexei A.** *Two-component network model in voice identification technologies. IT Security (Russia), [S.l.], v. 25, n. 1, p. 81-89, 2018. ISSN 2074-7136. Available at: <<https://bit.mephi.ru/index.php/bit/article/view/1096>>. Date accessed: 15 feb. 2018. doi:<http://dx.doi.org/10.26583/bit.2018.1.08>.*

### **Введение.**

В последнее время биометрия – наука о распознавании личности по ее физическим и поведенческим признакам – стремительно развивалась, а ее достижения нашли применение в ряде областей, таких как правоохранительные, медицинские и финансовые службы [1]. Основная задача биометрической системы заключается в сравнении двух биометрических экземпляров и установлении принадлежности их к одной или разным личностям. Как в целом в проблеме управления доступом выделяют два класса задач – идентификации и аутентификации субъектов. Можно отметить, что субъект является носителем модуля биометрик – идентификаторов и одновременно признаков, позволяющих провести проверку – верификацию- и тем самым подтвердить или опровергнуть искомую аутентичность. Характерно, что для биометрии трудности задач идентификации в большей мере носят технологический характер, когда одна

регистрируемая биометрика пользователя сравнивается со всеми  $N$ , хранящимися в базе данных, (1: $N$ ). Так, Integrated Automated Fingerprint System (IAFIS), репозиторий в США, управляемый ФБР в 2016 году содержал сведения о более чем 100 млн. лиц, давая возможность доступа к базе данных правоохранительным органам в режиме 24x7x365 [2].

Биометрическая аутентификация не предполагает сравнения регистрируемой биометрики со всеми записями в репозитории, а только с единственной, (1:1), но для поддержки формального подтверждения подлинности необходимы сложные организационно-правовые решения, включая стандартизацию процессов, принятие протоколов.

Таким образом технологические требования к решениям (производительности и надежности) проявляются в большей степени в биометрических задачах идентификации.

Анализ голоса представляет собой одну из форм биометрической аутентификации, позволяющая установить личность индивида по набору его уникальных акустических характеристик.

Голосовая модальность, среди прочих [3], демонстрирует свою эффективность в задачах идентификации личности, что особенно важно в телекоммуникационных приложениях. [4] подчеркивалось, что это единственная биометрическая модальность, которая позволяет идентифицировать человека по телефону. Оценка точности идентификации, как правило основывается на характеристике, введенной Национальным институтом стандартов и технологий США (National Institute of Standards and Technology, NIST)- Equal Error Rate, EER, равной ошибке первого и второго рода:  $FRR=FAR$ . (Традиционно, под ошибкой первого рода подразумевается отказ в правах легальному пользователю, FRR; ошибка второго рода, – предоставление прав самозванцу, FAR, при этом с противоположным смыслом в NIST [5]). Среди систем распознавания личности по голосу, использующих базы данных, с голосовыми фразами сотен актеров -дикторов, наилучшие показатели EER составляют 3–5% [6,7]

В настоящее время для голосовой идентификации применяется широкий спектр моделей, методик, программных и технических средств. [8-15]

Главные усилия исследователей и разработчиков в данной предметной области направлены на снижение как погрешности в распознавании личности - диктора, так и сопутствующих технологических затрат в процессе распознавания. В работе [16] было предложено использовать сетевой подход, как позволяющий обойтись низко затратными вычислениями.

Сетевое описание являет собой современную платформу исследований самых разных данных, в т.ч. временных рядов [17] что может представлять интерес для биометрического анализа. Здесь важным представляется разработка схем и алгоритмов трансформации изображений и звуковых записей [18-20] в комплексные сети для последующего их сопровождения надежными эффективными методами теории графов, теории вероятностей и линейной алгебры.

### **Модель.**

*Звуковой сигнал* можно представить, как совокупность различных синусоидальных составляющих. *Высота звука* - определяется частотой звуковой волны (или, периодом волны). Чем выше частота, тем выше звучание. *Громкость звука* определяется амплитудой сигнала. Чем выше амплитуда звуковой волны, тем громче сигнал.

Известно, что процедура распознавания личности по голосовому биометрическому показателю заключается в сравнении записей голоса, сделанных с помощью различных аудио устройств. Но сравнивать предъявляемый голос с огромным объемом аудио данных с целью выявления совпадения голоса диктора довольно сложно и технологически затратно. В отличие от традиционных методов в [16] предложен метод сравнения аудиоданных путем построения сетевой модели и получения матрицы сетевых метрик. Обеспечивая высокое быстродействие, данный метод, однако, характеризуется

недостаточной надежностью и годится лишь для предварительной (грубой) оценки звуковых записей из-за близости сетевых метрик различных голосовых рядов и, тем самым значительной вероятности ошибок второго рода.

С использованием общего подхода [16] в данной работе ставилась задача сетевого представления аудиоинформации как с целью ее высокопроизводительного автоматического анализа, так и снижения погрешности в распознавании личности – диктора. Решение предполагает конвертирование звуковой информации в форму ассоциативной семантической (когнитивной) сетевой структуры с двумя компонентами: амплитудной и частотной.

Для сравнительного анализа были записаны 2 речевых файла двух разных дикторов в формате WAV. Записи сделаны микрофоном мобильного телефона. Условно речевые актеры- дикторы обозначены как С1 и Э1. С1- мужского рода, 27 лет. Э1- женского рода, 26 лет. В качестве речевой информации, воспроизводимой дикторами выбрано одно предложение с одинаковым количеством слов. Аудио файлы предварительно обрабатывались, переводились из MP3 формата в формат данных WAV.

С помощью специализированной программы Spectral Plus (программа для работы с аудио –файлами [21]) определены временные точки с пиковыми амплитудными и частотными данными. Данные представлены на рис. 1-2.



Рис. 1. Диаграмма временных точек с пиковыми амплитудными и частотными показателями (речевая запись С1)

(Fig. 1. Chart of time points with peak amplitude and frequency indicators (speech record C1))

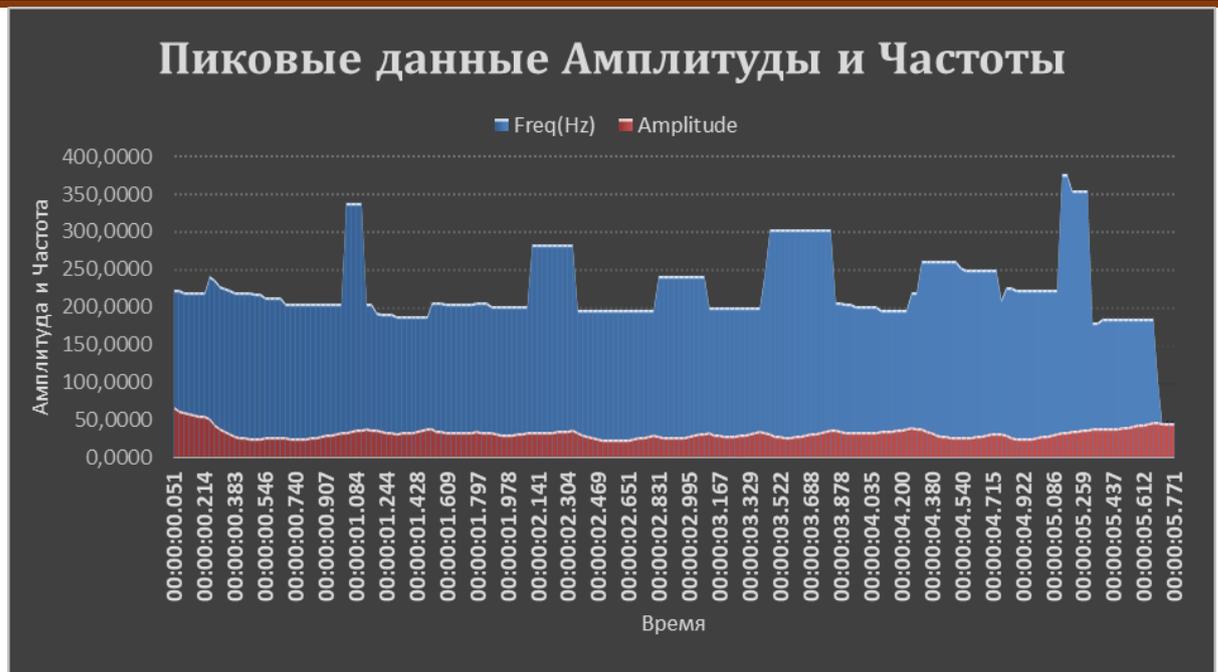


Рис. 2. Диаграмма временных точек с пиковыми амплитудными и частотными показателями (речевая запись Э1)

(Fig. 2. Chart of time points with peak amplitude and frequency indicators (speech record E1))

Поскольку наблюдения проводятся в некотором временном интервале, то для каждой временной точки следует определить значение амплитуды (рис. 3.) и частоты (рис. 4.) сигнала.

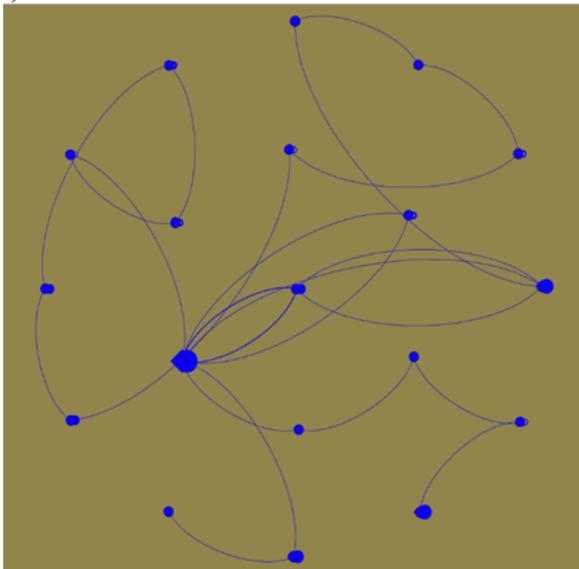


Рис. 3. Граф по частотным показателям (речевая запись C1)  
(Fig. 3. Graph built on frequency indicators (speech record C1))

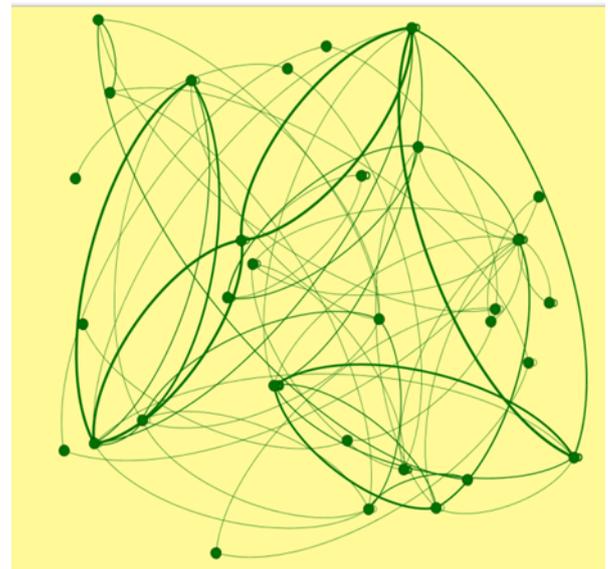


Рис. 4. Граф по амплитудным показателям (речевая запись C1)  
(Fig. 4. Graph built on the amplitude indicators (speech record C1))

В нашем случае при конвертировании данных в комплексную сеть узлом для построения сетевой модели является либо амплитуда (рис. 5.), либо частота (рис. 6.) в определённой временной точке. Подобно [16] за счет применения схемы «модели соседства» [18] устанавливается последовательная связь между узлами – частотами - соседняя

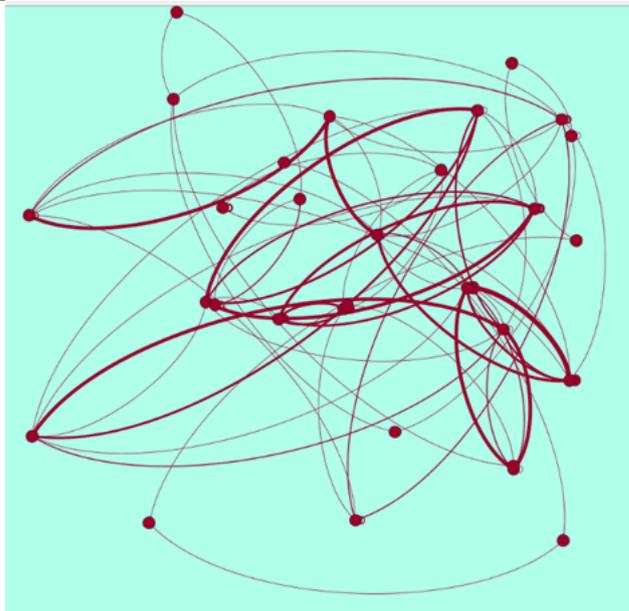


Рис. 5. Граф по амплитудным показателям (речевая запись Э1)  
 (Fig. 5. Graph built on the amplitude indicators (speech record E1))

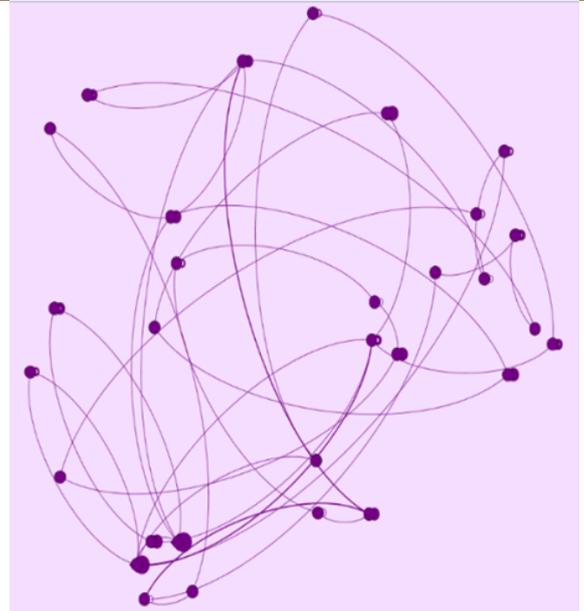


Рис. 6 - Граф по частотным показателям (речевая запись Э1)  
 (Fig. 6. Graph built on frequency indicators (speech record E1))

присоединяется к соседней (по значению, не по времени) и, аналогично, между амплитудами. Повторение узлов определено увеличением их весовых значений. Для получения матрицы метрик строятся по отдельности 2 графа по частотным и амплитудным индикаторам. Естественно, что объединить два графа ни в мультиплекс [22], ни в общую стволую структуру [23] не представляется возможным, хотя изначально данные соответствуют одной временной точке. Отметим, что при установлении связи амплитуда и частота будут определены их первыми значениями. Повторяющиеся амплитуды и частоты просто приведут к увеличению весовых значений узлов.

#### Результаты исследований.

Выбор ключевых сетевых метрик, характеризующих процессы в исследуемой системе является вторым важнейшим этапом после строительства самой сетевой модели.

Таблица 1 Матрица метрик (Речевая запись С1)

| метрика                           | Статистика амплитуды | Статистика частоты |
|-----------------------------------|----------------------|--------------------|
| узлы                              | 32                   | 19                 |
| ребра                             | 82                   | 37                 |
| Средняя степень                   | 2,562                | 1,947              |
| Средняя взвешенная степень        | 4,844                | 3,158              |
| Средний коэффициент кластеризации | 0,368                | 0,32               |
| Средняя длина пути                | 6,215                | 5,522              |

Таблица 2 Матрица метрик (Речевая запись Э1)

| метрика | Статистика амплитуды | Статистика частоты |
|---------|----------------------|--------------------|
| узлы    | 30                   | 30                 |
| ребра   | 80                   | 62                 |

Эдита К. Куулар, Андрей И. Труфанов, Алексей А. Тихомиров  
ДВУХКОМПОНЕНТНАЯ СЕТЕВАЯ МОДЕЛЬ В ТЕХНОЛОГИЯХ ГОЛОСОВОЙ  
ИДЕНТИФИКАЦИИ ЛИЧНОСТИ

|                                   |       |       |
|-----------------------------------|-------|-------|
| Средняя степень                   | 2,667 | 2,067 |
| Средняя взвешенная степень        | 6,6   | 6,6   |
| Средний коэффициент кластеризации | 0,359 | 0,272 |
| Средняя длина пути                | 5,562 | 5,969 |

Ряд топологических метрик (2-й и 3-й столбцы в таблицах 1-2) соответственно амплитудной и частотной сетевых моделей речевого текста представляет собой вектор, а вместе - матрицу, как цифровой "сетевой" отпечаток голоса. Этот цифровой отпечаток, в нашем случае – матрица сетевых метрик, должен наилучшим образом отражать индивидуальные голосовые особенности актора-диктора, и при этом в соответствии с требованиями к биометрическим системам безопасности [24] оставаться стабильным во времени, легко измеряемым и не поддаваться имитации. Анализ матриц (Табл.1-2), характеризующих метрики, указывает на их различие, которое обеспечивает возможность надежной идентификации диктора. Для подтверждения чувствительности матрицы метрик к голосу актора-диктора, дальнейшие исследования будут включать сравнение записей одного диктора с аутентичными и неаутентичными записями.

Область применения системы идентификации по голосовым биометрическим данным достаточно обширна: верификация клиентов при доступе к данным, тем более что остальные биометрические данные человека не так-то просто получить через системы удаленного доступа. Обработка речевых баз данных, сравнительный анализ на соответствие данных. Довольно часто применяется в криминалистике.

Данный метод можно применять при верификации, пользователь предъявляет в том или ином виде свой идентификатор, и система распознавания должна сравнить полученную матрицу метрик с имеющейся матрицей метрик соответствующего идентификатору личности предъявителя, и система распознавания должна подтвердить или отвергнуть этот идентификатор. Поскольку пользователь заинтересован в подтверждении его идентификатора, то старается не вносить в речевой пароль вариаций, которые отсутствовали в период записи его голосовых биометрических данных.

### **Выводы.**

Подготовлена двухкомпонентная сетевая модель голосовой биометрической модальности, включающей амплитудную и частотную составляющие, к которым можно применить известные алгоритмы топологического сопоставления. Сделан вывод о том, что матрица ключевых топологических параметров сетевых компонентов модели может служить сетевым цифровым отпечатком, позволяющим эффективно сравнивать голоса предъявителя и имеющиеся в базе данных, в предположении большей чувствительности метода и значительно меньших технологических затрат, чем того требуют традиционные системы.

Планируется исследовать особенности использования сетевых моделей в текст зависимых и текст независимых системах голосовой идентификации.

Важным и необходимым представляется дальнейшее изучение влияния на результаты голосового распознавания личности с помощью сетевых метрик таких факторов, как помехи и шумы, ошибочность произношения, состояние диктора, различие устройств голосовой записи. С другой стороны, если предлагаемый сетевой подход продемонстрирует чувствительность к состоянию индивида – физиологическому, психологическому, эмоциональному, он может оказаться полезным не только для задач идентификации личности, но и для определения ее состояния. Концепцию комплексных сетей, как исследовательскую платформу разумно оценить для применения также к иным биометрическим модальностям.

СПИСОК ЛИТЕРАТУРЫ:

- 1 Talreja V., Ferrett M. T., Valenti C., Ross A. Biometrics-as-a-Service: A Framework to Promote Innovative Biometric Recognition in the Cloud. arXiv:1710.09183v1. 2017. -6 p. URL: <https://arxiv.org/pdf/1710.09183.pdf> (дата обращения: 25.11.17)
- 2 Goode A. Biometric Identification or Biometric Authentication? URL: <https://www.veridiumid.com/blog/biometric-identification-and-biometric-authentication/> (дата обращения: 18.02.16)
- 3 Ibrahim D.R., Tamimi A.A., Abdalla A.M. Performance Analysis of Biometric Recognition Modalities 8th International Conference on Information Technology (ICIT). Amman, Jordan. 2017. pp. 980-984
- 4 Матвеев Ю.Н. Технологии биометрической идентификации личности по голосу и другим модальностям. Вестник МГТУ им. Н.Э. Баумана. Сер. “Приборостроение”. 2012, С. 46-61
- 5 Wu J. C., Martin A. F., Greenberg C. S., Kacker R. N. The Impact of Data Dependence on Speaker Recognition Evaluation IEEE/ACM Trans Audio Speech Lang Process. 2017. vol. 25, no.1, pp. 5–18
- 6 Сорокин В.Н. Распознавание личности по голосу: аналитический обзор. В.Н. Сорокин, В.В. Вьюгин, А.А. Тананыкин. Информационные процессы. 2012. Т. 12, вып. 1. С. 1–30
- 7 Poddar A., Sahidullah M., Saha G. Performance comparison of speaker recognition systems in presence of duration variability. India Conference (INDICON), 2015 Annual IEEE, 2015 -6 p.
- 8 Sahidullah M., Saha G. A novel windowing technique for efficient computation of MFCC for speaker recognition. IEEE signal processing letters. 2013. vol. 20. no. 2. pp. 149–152.
- 9 Bahmaninezhad F., Hansen J.H.L. I-vector/PLDA speaker recognition using support vectors with discriminant analysis. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2017. pp. 5410–5414
- 10 Yao T., Meng C., Liang H., Jia L. Speaker recognition system based on deep neural networks and bottleneck features. Journal of Tsinghua University (Science and Technology). 2016. vol. 56, no. 11, pp. 1143–1148
- 11 Shulipa A., Novoselov S., Melnikov A. Approaches for Out-of-Domain Adaptation to Improve Speaker Recognition Performance. A. Ronzhin et al. (Eds.): SPECOM 2016, LNAI 9811.2016. pp. 124–130
- 12 Reynolds D.A. Gaussian mixture models. Encyclopedia of biometric recognition. Heidelberg: Springer 2015–pp. 827–832
- 13 Козлов А.В., Кудашев О.Ю., Матвеев Ю.Н., Пеховский Т.С., Симончик К.К., Шулипа А.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2012. Труды СПИИРАН. 2013. Вып. 25. С. 350-370
- 14 Рахманенко И.А. Программный комплекс для идентификации диктора по голосу с применением параллельных вычислений на центральном и графическом процессорах. Доклады ТУСУРа, Т. 20, № 1, 2017. С. 70-74
- 15 Идентификация человека по голосу URL: [http://infoprotect.net/varia/identifikaciya\\_po\\_golosu](http://infoprotect.net/varia/identifikaciya_po_golosu) (дата обращения: 06.11.17)
- 16 Куулар Э.К., Тихомиров А.А., Труфанов А.И. Идентификация источников звуковой информации методом сетевого анализа. Безопасность информационных технологий 2016. №2 С. 43-48 URL: <https://bit.mephi.ru/index.php/bit/article/view/78/85>
- 17 Yang D., Li X. Bridge time series and complex networks with a frequency-degree mapping algorithm. Proc. IEEE International Symposium on Circuits and Systems (ISCAS). 2012. pp.910-913
- 18 Bezsudnov I.V., Snarskii A.A. From the time series to the complex networks: the parametric (dynamical) natural visibility graph. Physica. 2014 no. 414 pp. 53-60 arXiv:1208.6365 URL: <https://arxiv.org/ftp/arxiv/papers/1208/1208.6365.pdf>
- 19 Zinoviev D. Networks of Music Groups as Success Predictors. arXiv:1709.08995v1. 2017 -6 p. URL: <https://arxiv.org/abs/1709.08995> (дата обращения: 04.08.17)
- 20 Ferretti S. On the Complex Network Structure of Musical Pieces: Analysis of Some Use Cases from Different Music Genres. arXiv:1709.09708v1. 2017 -32 p. URL: <https://arxiv.org/abs/1709.09708> (дата обращения: 13.09.17)
- 21 SpectraPLUS ®. Pioneer Hill Software. [Электронный ресурс] URL: <http://www.spectraplus.com/>
- 22 Boccaletti S., Bianconi G., Criado R., Del Genio C.I., Gómez-Gardeñes J., Romance M., Sendiña-Nadal I., Wang Z., Zanin M. The structure and dynamics of multilayer networks. Physics Reports, 2014, no. 544 (1) pp. 1-122
- 23 Тихомиров А. А., Труфанов А. И., Россодивита А. Модель взаимодействующих стволых сетей в решении задач топологической устойчивости сложных систем. Безопасность информационных технологий. 2013, №1, с.125-126
- 24 Ворона В. А., Костенко В. О. Биометрические технологии идентификации в системах контроля и управления доступом. Computational nanotechnology. 2016 вып. 3, с. 224–241

REFERENCES:

- [1] Talreja V., Ferrett M. T., Valenti C., Ross A. Biometrics-as-a-Service: A Framework to Promote Innovative Biometric Recognition in the Cloud. arXiv:1710.09183v1. 2017. -6 p. URL: <https://arxiv.org/pdf/1710.09183.pdf> (дата обращения: 25.11.17)
- [2] Goode A. Biometric Identification or Biometric Authentication? URL: <https://www.veridiumid.com/blog/biometric-identification-and-biometric-authentication/> (дата обращения: 18.02.16)
- [3] Ibrahim D.R., Tamimi A.A., Abdalla A.M. Performance Analysis of Biometric Recognition Modalities 8th International Conference on Information Technology (ICIT). Amman, Jordan. 2017. pp. 980-984
- [4] Matveev Yu.N. Technologies of biometric identification of a person by voice and other modalities. Vestnik of MSTU. N.E. Bauman. Ser. Instrument making. 2012 P. 46-61. (in Russian).
- [5] Wu J. C., Martin A. F., Greenberg C. S., Kacker R. N. The Impact of Data Dependence on Speaker Recognition Evaluation IEEE/ACM Trans Audio Speech Lang Process. 2017. vol. 25, no.1, pp. 5–18
- [6] Sorokin V.N. Personality recognition by voice: analytical review. V.N. Sorokin, V.V. Vjugin, A.A. Tananykin. Information Processes. 2012. T. 12, vip. 1. - P. 1-30. (in Russian).
- [7] Poddar A., Sahidullah M., Saha G. Performance comparison of speaker recognition systems in presence of duration variability. India Conference (INDICON), 2015 Annual IEEE, 2015 -6 p.
- [8] Sahidullah M., Saha G. A novel windowing technique for efficient computation of MFCC for speaker recognition. IEEE signal processing letters. 2013. vol. 20. no. 2. pp. 149–152.
- [9] Bahmaninezhad F., Hansen J.H.L. I-vector/PLDA speaker recognition using support vectors with discriminant analysis. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2017. pp. 5410-5414
- [10] Yao T., Meng C., Liang H., Jia L. Speaker recognition system based on deep neural networks and bottleneck features. Journal of Tsinghua University (Science and Technology). 2016. vol. 56, no. 11, pp. 1143–1148
- [11] Shulipa A., Novoselov S., Melnikov A. Approaches for Out-of-Domain Adaptation to Improve Speaker Recognition Performance. A. Ronzhin et al. (Eds.): SPECOM 2016, LNAI 9811.2016. pp. 124–130
- [12] Reynolds D.A. Gaussian mixture models. Encyclopedia of biometric recognition. Heidelberg: Springer 2015–pp. 827–832
- [13] Kozlov A.V., Kudashev O.Yu., Matveev Yu.N., Pehovskiy TS, Simonchik K.K., Shulipa A.K. The system of voice announcers identification for the NIST SRE 2012 competition. Proceedings of SPIIRAS. 2013. vol. 25. P. 350-370. (in Russian).
- [14] Rakhmanenko I.A. A software package for voice speaker identification using parallel computations on the central and graphics processors. Reports of TUSUR, vol. 20, no. 1, 2017. pp. 70-74. (in Russian).
- [15] Identification of a person by voice URL: [http://infoprotect.net/varia/identifikaciya\\_po\\_golosu](http://infoprotect.net/varia/identifikaciya_po_golosu) (reference date: 06.11.17). (in Russian).
- [16] Kuular E.K., Tikhomirov A.A., Trufanov A.I. Identification of sources of sound information using the method of network analysis. Information Technology Security. 2016 No. 2 URL: <https://bit.mephi.ru/index.php/bit/article/view/78/85>.(in Russian).
- [17] Yang D., Li X. Bridge time series and complex networks with a frequency-degree mapping algorithm. Proc. IEEE International Symposium on Circuits and Systems (ISCAS). 2012. pp.910-913
- [18] Bezsudnov I.V., Snarskii A.A. From the time series to the complex networks: the parametric (dynamical) natural visibility graph. Physica. 2014 no. 414 pp. 53-60 arXiv:1208.6365 URL: <https://arxiv.org/ftp/arxiv/papers/1208/1208.6365.pdf>
- [19] Zinoviev D. Networks of Music Groups as Success Predictors. arXiv:1709.08995v1. 2017 -6 p. URL: <https://arxiv.org/abs/1709.08995> (дата обращения: 04.08.17)
- [20] Ferretti S. On the Complex Network Structure of Musical Pieces: Analysis of Some Use Cases from Different Music Genres. arXiv:1709.09708v1. 2017 -32 p. URL: <https://arxiv.org/abs/1709.09708> (дата обращения: 13.09.17)
- [21] SpectraPLUS ®. Pioneer Hill Software. [Электронный ресурс] URL: <http://www.spectraplus.com/>
- [22] Boccaletti S., Bianconi G., Criado R., Del Genio C.I., Gómez-Gardeñes J., Romance M., Sendiña-Nadal I., Wang Z., Zanin M. The structure and dynamics of multilayer networks. Physics Reports, 2014, no. 544 (1) pp. 1-122
- [23] Tikhomirov AA, Trufanov AI, Rossodivita A. A model of interacting stem networks in solving problems of topological stability of complex systems. Security of information technology. 2013, №1, p.125-126. (in Russian).
- [24] Vorona VA, Kostenko V. O. Biometric identification technologies in access control systems. Computational nanotechnology. 2016 vol. 3, p. 224-241. (in Russian).

*Поступила в редакцию – 28 декабря 2017 г. Окончательный вариант – 06 февраля 2018 г.  
Received – December 28, 2017. The final version – February 06, 2018.*