

Сравнительный анализ и перспективы развития

На основе проведенного анализа можно составить сравнительную таблицу применимости технологий в интеллектуальных сетях (таблица 1):

Таблица 1. Характеристики модели

QoS			
Характеристика	IntServ	DiffServ	MPLS + RSVP-TE
Масштабируемость	Низкая	Высокая	Высокая
Гарантии сервиса	Жесткие (Hard QoS)	Относительные (Soft QoS)	Жесткие для магистралей
Сложность внедрения	Высокая	Низкая	Средняя/Высокая
Область применения	Малые сети, видеостудии	Корпоративные сети, Edge	Ядро сети (Core), провайдеры

Для современных интеллектуальных сетей Узбекистана, где наблюдается активное внедрение технологий NGN и 5G, наиболее перспективным является гибридный подход. На уровне доступа (Edge) целесообразно использовать **DiffServ** для классификации абонентского трафика, а в магистральном ядре (Core) — **MPLS** для обеспечения скоростной коммутации и отказоустойчивости.

Заключение. Анализ технологий обеспечения QoS показал, что на современном этапе развития интеллектуальных сетей связи не существует единственной универсальной технологии. Оптимальное решение заключается в интеграции модели дифференцированного обслуживания с механизмами управления трафиком на базе MPLS. Использование связки **MPLS + RSVP-TE** позволяет достичь необходимого баланса между гибкостью управления ресурсами и масштабируемостью системы, что является критически важным для обеспечения высокого качества мультимедийных услуг в условиях постоянно растущей нагрузки на сети связи Ташкента и республики в целом.

Список литературы

- Северов Д. NGN сегодня и завтра: взгляд интегратора // ИнформКурьер-Связь. — 2004.
- Олифер В. Г., Олифер Н. А. Компьютерные сети. Принципы, технологии, протоколы. — СПб.: Питер, 2020.
- Б.М. Умирзаков, Р.Х. Джураев Анализ качества обслуживания мультимедийного трафика в сетях IP. Цифровая экономика и информационные технологии. №1. 2025 г. Ташкент
- Kurose J. F. and Ross K. W. Computer Networking: A Top-Down Approach, 7th Edition / Hoboken, New Jersey: Pearson, [2017]
- Шварцман В. О. Интеграция в электросвязи. — М.: Агентство ИРИАС, 2001.

NUTQQA INTELLEKTUAL ISHLOV BERISH TIZIMLARIDA PLDA MODEL ASOSIDA SO'ZLOVCHILARNI TANIB OLISH

Shukurov Kamoliddin Elbobo o'g'li

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston, keshukurov@gmail.com

Xasanov Umidjon Komiljon o'g'li

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti, Toshkent, O'zbekiston, umidjon0923@gmail.com

Annotatsiya: Mazkur maqolada nutqqa intellektual ishlov berish tizimlarida so'zlovchilarni tanib olish masalalari ko'rib chiqiladi. Zamonaviy nutqni qayta ishlash usullari asosida so'zlovchining individual xususiyatlarini ajratib olish va ularni solishtirish jarayoni tahlil qilinadi. Tadqiqotda PLDA modelidan foydalanish orqali so'zlovchilarni aniqlashning samarali usullari yoritilib, modelni turli sharoitlarda qo'llash imkoniyatlari hamda hisoblash resurslarini tejash bo'yicha afzalliklari ko'rsatib berilgan.

Kalit so'zlar: nutq signallari, intellektual ishlov berish, so'zlovchini tanib olish, PLDA, embedding, sun'iy intellekt, ECAPA-TDNN.

SPEAKER RECOGNITION BASED ON THE PLDA MODEL IN INTELLIGENT SPEECH PROCESSING SYSTEMS

Shukurov Kamoliddin

Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent,
Uzbekistan

keshukurov@gmail.com

Khasanov Umidjon

Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent,
Uzbekistan

umidjon0923@gmail.com

Abstract. This article examines the issues of speaker recognition within intelligent speech processing systems. It analyzes the process of extracting and comparing individual speaker characteristics based on modern speech processing techniques. The study highlights effective methods for speaker recognition using the PLDA model, demonstrating its applicability under various conditions and its advantages in conserving computational resources.

Keywords: speech signals, intelligent processing, speaker recognition, PLDA, embedding, artificial intelligence, ECAPA-TDNN.

РАСПОЗНАВАНИЕ ДИКТОРОВ НА ОСНОВЕ МОДЕЛИ PLDA В СИСТЕМАХ ИНТЕЛЛЕКТУАЛЬНОЙ ОБРАБОТКИ РЕЧИ

Шукуров Камолиддин

Ташкентский университет информационных технологий имени Мухаммада аль-Хорезми,
Ташкент,
Узбекистан

keshukurov@gmail.com

Хасанов Умиджон

Ташкентский университет информационных технологий имени Мухаммада аль-Хорезми,
Ташкент,
Узбекистан

umidjon0923@gmail.com

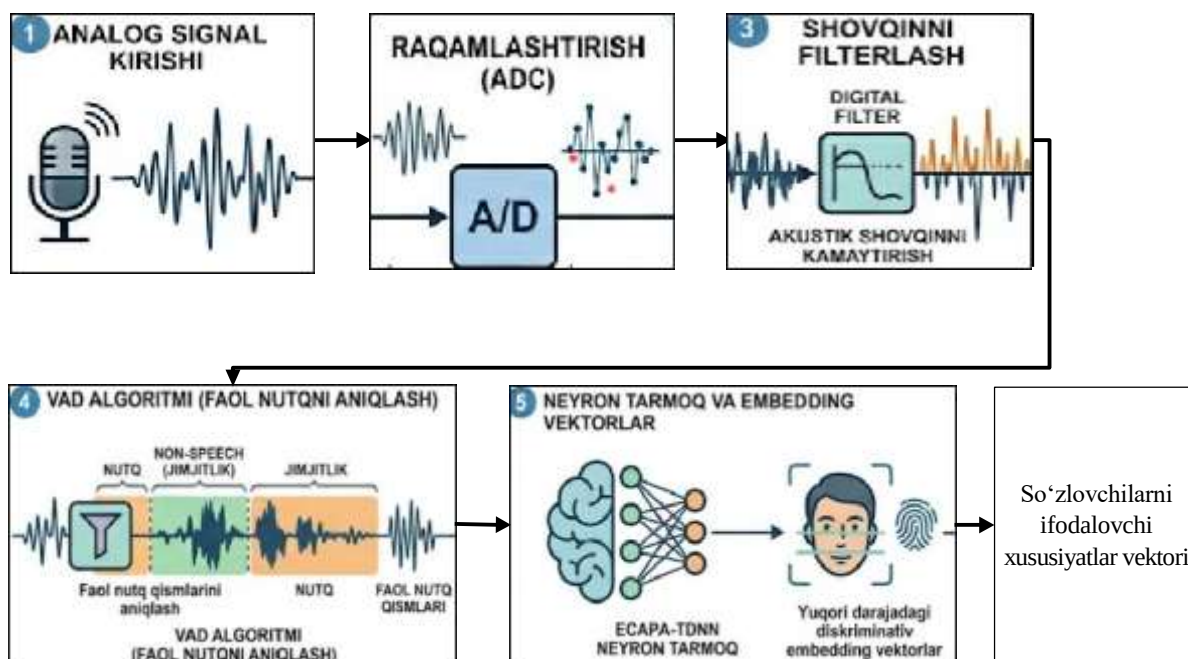
Аннотация: В данной статье рассматриваются вопросы распознавания дикторов в системах интеллектуальной обработки речи. Анализируется процесс извлечения и сравнения индивидуальных характеристик диктора на основе современных методов обработки речи. В исследовании освещаются эффективные методы распознавания дикторов с использованием модели PLDA, демонстрируется возможность ее применения в различных условиях, а также преимущества с точки зрения экономии вычислительных ресурсов.

Ключевые слова: речевые сигналы, интеллектуальная обработка, распознавание диктора, PLDA, эмбединг, искусственный интеллект, ECAPA-TDNN.

Kirish. Hozirgi kunda hisoblash mashinalari va sun'iy intellekt texnologiyalarining jadal rivojlanishi natijasida nutqqa intellektual ishlov berish tizimlarining matematik modellari va dasturiy ta'minotini takomillashtirish dolzarb masalaga aylanmoqda. Bunday dasturiy majmualar va algoritmlar yordamida inson nutqini tahlil qilish, raqamli signallardan semantik hamda biometrik ma'lumotlarni ajratib olish va avtomatlashtirilgan qaror qabul qilish imkoniyati yaratilmoqda[1]. So'zlovchilarni tanib olish tizimlari axborot xavfsizligi, biometrik autentifikatsiya hamda inson-kompyuter muloqoti (*ing: Human-Computer Interaction*) jarayonlarini tabiiy hamda uzluksiz ta'minlashda muhim o'rin tutadi[1-3]. Nutqqa asoslangan dasturiy tizimlarning asosiy afzalligi murakkab va qimmatbaho qo'shimcha apparat ta'minotini talab qilmasdan, standart mikrofon hamda hisoblash qurilmalarining o'zida optimal algoritmlar yordamida yuqori samaradorlik bilan ishlash imkoniyatining mavjudligidir.

Nutq signallariga ishlov berishning algoritmik modeli

Nutq signallarini dasturiy muhitda qayta ishlash bir necha algoritmik va hisoblash bosqichlaridan iborat. Dastlab, analog signal raqamlashtirilib, raqamli filtrlar yordamida turli xil akustik shovqinlardan tozalanadi. So'ngra VAD (*ing: Voice Activity Detection*) algoritmi orqali faol nutq qismlari aniqlanib, keraksiz jimjitlik qismlari (non-speech) qirqib olinadi[2-6]. Ushbu bosqichda tizimning hisoblash resurslarini optimallashtirish maqsadida chuqur neyron tarmoqlar, xususan, ECAPA-TDNN (*ing: Emphasized Channel Attention, Propagation and Aggregation in TDNN*) arxitekturasiga asoslangan dasturiy modullar qo'llaniladi. Mazkur neyron tarmoq signaldan yuqori darajadagi diskriminativ embedding vektorlarni shakllantirib, so'zlovchining biometrik xususiyatlarini tizim tushunadigan ixcham matematik ko'rinishga o'tkazadi[3,4].



1-rasm. Nutq signallaridan biometrik profil shakllantirish jarayoni

Ajratib olingan n -o'lchamli embedding vektorlarni solishtirish va qaror qabul qilish jarayonida PLDA (Probabilistic Linear Discriminant Analysis) matematik modelidan foydalaniladi. PLDA noma'lum vektorlarning o'xshashligini ehtimollik prinsiplari asosida baholaydi. Ushbu yondashuvning dasturiy afzalligi shundaki, u matritsali hisoblash murakkabligini keskin kamaytiradi va katta hajmli ma'lumotlar bazalarida tezkor qidiruv hamda solishtirish imkonini beradi. Matematik jihatdan model so'zlovchi variatsiyasi va sharoit variatsiyasi ta'sirini alohida qismlarga ajratib tahlil qiladi. Bu esa apparat qismida mikrofon sifatining o'zgarishi yoki tarmoqdagi shovqinlar inson-kompyuter muloqoti tizimlarining ishonchligiga salbiy ta'sir ko'rsatishining oldini oladi[7-9]. PLDA modelining asosi quyidagi generativ tenglamaga tayanadi:

$$x = \mu + \Phi y + \Gamma z + \epsilon$$

Bu yerda x – ajratib olingan n – o'lchamli so'zlovchini ifodalovchi xususiyat vektori, μ – barcha

ma'lumotlar uchun umumiy o'rtacha qiymat, Φy – so'zlovchining o'ziga xos xususiyatlarini ifodalovchi qism, Γz – yozib olish sharoiti va kanal xususiyatlarini ifodalovchi qism, ϵ – qoldiq xatolik yoki gauss shovqini.

Tizimga kelib tushgan ikkita noma'lum vektorning (x_1 va x_2) bitta so'zlovchiga tegishli ekanligini baholashda ehtimollik prinsiplariga asoslangan Log-Likelihood Ratio (LLR) funksiyasi qo'llaniladi:

$$score(x_1, x_2) = \log \frac{P(x_1, x_2 | H_{same})}{P(x_1, x_2 | H_{diff})}$$

Ushbu yondashuvning dasturiy afzalligi shundaki, ehtimolliklarni hisoblashda oldindan o'qitilgan matritsalar qo'llanilishi sababli, u hisoblash murakkabligini keskin kamaytiradi va katta hajmli ma'lumotlar

bazalarida tezkor qidiruv hamda solishtirish imkonini beradi. Shuningdek, Γz vektorini baholash va ta'sirini kamaytirish hisobiga, apparat qismida mikrofon sifatining o'zgarishi yoki tarmoqdagi shovqinlar inson-kompyuter muloqoti tizimlarining ishonchliligiga salbiy ta'sir ko'rsatishining oldini oladi [10, 11].

Tajribalar. Tajriba davomida chuqur neyron tarmoq orqali olingan so'zlovchilarni ifodalovchi xususiyatlar vektorlar kosinus kosinus masofasi hamda ehtimollik asoslangan PLDA modellari yordamida qiyosiy tahlil qilindi. Sinov jarayonida tizimning barqarorligini tekshirish uchun VoxCeleb ochiq ma'lumotlar bazasidan olingan turli shovqin darajasiga ega nutq signallaridan hamda maxsus yig'ilgan nutq to'plamidan foydalanildi.

1-jadval. So'zlovchilarni tanib olish aniqligini qiyosiy tahlil qilish natijalari

Baholash usuli (Scoring method)	Qabul qilingan xususiyat (Embedding)	EER (%)	MinDCF (0.01)	Aniqlik darajasi (Accuracy, %)
Kosinus o'xshashligi (Cosine Similarity)	ECAPA-TDNN (192 o'lchamli)	1.92	0.185	88.07
Standart LDA (Linear Discriminant Analysis)	ECAPA-TDNN (192 o'lchamli)	1.45	0.142	90.55
PLDA modeli (Taklif etilgan usul)	ECAPA-TDNN (192 o'lchamli)	0.87	0.092	94.35

Biometrik tizimlarning dasturiy va algoritmik samaradorligini baholashda tizimning begona shaxsni adashib qabul qilish (*ing: False Acceptance Rate - FAR*) va haqiqiy foydalanuvchini inkor etish (*ing: False Rejection Rate - FRR*) xatoliklari o'zaro kesishgan nuqta, ya'ni EER (*ing: Equal Error Rate — teng xatolik ko'rsatkichi*) asosiy mezon hisoblanadi. Bu ko'rsatkich qancha past bo'lsa, dasturiy majmuaning ishonchlilik darajasi shuncha yuqori bo'ladi. 1-jadval ma'lumotlaridan ko'rinib turibdiki, taklif etilayotgan PLDA modeli an'anaviy Kosinus o'xshashligi (1.92%) va standart LDA (1.45%) modellariga nisbatan EER ko'rsatkichini 0.87% gacha qisqartirishga erishdi.

Shu bilan birga, real amaliyotda begona odamni tizimga kiritib yuborish xavfi haqiqiy foydalanuvchini tanimay qolishdan ko'ra qimmatroq oqibatlariga olib kelishi mumkin. Shu sababli, algoritmlar tahlilida qo'shimcha ravishda MinDCF (*ing: Minimum Detection Cost Function — aniqlash xarajatining minimal funksiyasi*) ham hisoblab chiqildi. Ushbu mezon biometrik tizimning ishchi muhitdagi "xatolik narxini" baholaydi. Natijalar shuni ko'rsatadiki, chuqur neyron tarmoqlar (ECAPA-TDNN) va ehtimollikka asoslangan PLDA modelining integratsiyasi MinDCF ko'rsatkichini minimal darajaga (0.092) tushirib, tizimning axborot xavfsizligi talablariga to'liq javob berishini matematik jihatdan isbotladi.

Xulosa. Mazkur maqolada nutqqa intellektual ishlov berish va so'zlovchilarni tanib olish tizimlarining algoritmik va matematik ta'minoti tahlil qilindi. Tadqiqot natijalari shuni ko'rsatadiki, chuqur neyron tarmoqlar va PLDA modelining birgalikdagi integratsiyasi matritsali hisoblash murakkabligini kamaytirib, yozib olish sharoiti variatsiyalarining salbiy ta'sirini minimallashtiradi. Ishlab chiqilgan ushbu dasturiy-algoritmik yechimlar kelgusida ovozli yordamchilar, biometrik xavfsizlik tizimlari va avtomatlashtirilgan inson-kompyuter muloqoti tizimlarini takomillashtirish uchun mustahkam poydevor bo'lib xizmat qiladi. Ushbu maqolada olingan natijalar va ishlab chiqilgan matematik modellar kelgusida

yanada murakkab va qamrovli bo'lgan speaker diarization (nutq signalidan so'zlovchilarni avtomatik ajratish va tanib olish) masalalarini hal qilish uchun mustahkam poydevor bo'lib xizmat qiladi.

Фойдаланилган адабиётлар

1. Hansen, John & Hasan, Taufiq. (2015). Speaker Recognition by Machines and Humans: A tutorial review. Signal Processing Magazine, IEEE. 32. 74-99. 10.1109/MSP.2015.2462851.
2. Rabiner L. R., Schafer R. W. Theory and Applications of Digital Speech Processing. – Pearson, 2010. – 1056 p.
3. Shukurov, K., Ochilov, M., Khasanov, U., Kholdorov, S., Rakhmanova, M. (2026). Advanced Speaker Diarization Techniques for Analyzing Uzbek Speech Signals. In: Bradford, P.G., Koul, S.K., Gadsden, S.A., Ghatak, K.P. (eds) Proceedings of IEMTRONICS 2025. IEMTRONICS 2025. Lecture Notes in Electrical Engineering, vol 1468. Springer, Singapore. https://doi.org/10.1007/978-981-95-0433-6_42
4. Shukurov K., Khasanov U., Turaev B., Kakhkharov A. (2023). The Effectiveness of the Implementation of Speech Command Recognition Algorithms in Embedded Systems. The Eurasia Proceedings of Science, Technology, Engineering and Mathematics, 23, 220-224. <https://doi.org/10.55549/epstem.1365795>
5. K. Shukurov, U. Khasanov, S. Kholdorov, M. Karimova and L. Murodjonov, "Emotion-Aware Speaker Diarization Based on Prosodic and Deep Embedding Integration," 2025 4th International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2025, pp. 988-992, doi: 10.1109/ICAAIC64647.2025.11330488.
6. Mukhamadiyev, A., Mukhiddinov, M., Khujayarov, I., Ochilov, M., & Cho, J. (2023). Development of Language Models for Continuous Uzbek Speech Recognition System. Sensors, 23(3), 1145. <https://doi.org/10.3390/s23031145>
7. Musaev, M., Abdullaeva, M., & Ochilov, M. (2022, August). Advanced feature extraction method for speaker identification using a classification algorithm. In AIP Conference Proceedings (Vol. 2656, No. 1, p. 020022). AIP Publishing LLC.
8. Prince S. J. D., Elder J. H. Probabilistic Linear Discriminant Analysis for Inferences About Identity // IEEE 11th International Conference on Computer Vision. – 2007. – P. 1-8.
9. Ioffe S. Probabilistic Linear Discriminant Analysis // Computer Vision – ECCV. – Springer, 2006. – P. 531-542.
10. Desplanques B., Thienpondt J., Demuynck P. ECAPA-TDNN: Emphasized Channel Attention, propagation and aggregation in TDNN based speaker verification // Interspeech. – 2020. – P. 3830-3834.
11. Snyder D., Garcia-Romero D., Sell G., Povey D., Khudanpur S. X-vectors: Robust DNN Embeddings for Speaker Recognition // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2018. – P. 5329-5333.

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОГНОЗИРОВАНИИ БИОЛОГИЧЕСКОЙ АКТИВНОСТИ ХИМИЧЕСКИХ СОЕДИНЕНИЙ

Шакирова А.П.

Старший преподаватель кафедры медицинских наук, филиала КФУ в городе Джизак
nasyapavlovna87@gmail.com

Юсуфова С.Г.

Старший преподаватель кафедры медицинских наук, филиала КФУ в городе Джизак
sabinayusufova1991@gmail.com

Аннотация: В данной статье рассматривается, как современные технологии искусственного интеллекта помогают предсказывать биологическую активность химических соединений. Особое внимание уделено методам машинного и глубокого обучения, применяемым в QSAR- моделировании, виртуальном скрининге и молекулярном дизайне. Показано, что