

**РАСПОЗНАВАНИЕ РЕЧЕВЫХ СООБЩЕНИЙ РАДИООБМЕНА
В АВИАЦИИ НА БАЗЕ КОРРЕЛЯЦИОННОГО АНАЛИЗА**

© 2021 Н.А. Андриянов

Финансовый университет при Правительстве Российской Федерации, г. Москва, Россия

Статья поступила в редакцию 12.12.2020

В работе рассматривается задача распознавания речевых сообщений фразеологического радиообмена в гражданской авиации. Во введении обосновывается актуальность данной задачи. Далее представлены методы исследования, базирующиеся на корреляционном анализе. Наконец, приводятся описание эксперимента и результаты работы алгоритмов распознавания на базе корреляционного анализа. Были записаны различные варианты звуковых сигналов для пяти речевых сообщений и построены спектральные представления таких сигналов. Спектральное преобразование может быть получено либо с использованием специализированного программного обеспечения, либо на основе преобразования Фурье сигнала во временной области. Для получения более универсального эталонного сигнала и устранения влияния помех было произведено усреднение спектральных составляющих одного и того же речевого сообщения, записанного несколько раз. В действительности, использовались три спектра одного и того же речевого сообщения для усреднения. Такое усреднение спектра по трем обучающим составляющим обеспечило получение эталонного образца для каждой фразы, позволив при этом сократить влияние аддитивных белых гауссовых шумов на эталонную запись. В дальнейшем на основе корреляционного анализа были посчитаны связи между тестовыми фразами и всеми эталонами. На базе этих связей построена корреляционная матрица эталонных фраз. Исследования показали, что фразы, произнесенные одним человеком, оказались сильно-коррелированными. Анализ показал, что выбор класса (содержания речевого сообщения) при решении задачи распознавания, соответствующего наиболее близкому к единице значению коэффициента корреляции, обеспечивает свыше 90% правильных распознаваний на тестовой выборке, содержащей в общей сложности 100 фраз, по 20 на каждую фразу. При этом следует отметить, что при записи тестовых сообщений в качестве фона дополнительно присутствовала аддитивная белая гауссова помеха, воспроизводимая другим аудиоустройством. В случае же анализа информации без искусственно генерируемого шума, вероятность правильных распознаваний для тестовой выборки из 100 фраз, по 20 на каждую фразу, составила 100%.

Ключевые слова: фразеология радиообмена, гражданская авиация, статистический анализ, корреляционный анализ, распознавание образцов, спектр сигнала.

DOI: 10.37313/1990-5378-2021-23-1-91-96

ВВЕДЕНИЕ

Одной из важных систем обеспечения качественного управления взлетом, полетом и посадкой воздушного судна является фразеологический радиообмен. При этом, в отличие от обычной разговорной речи, связь и обмен информацией во время фразеологического радиообмена осуществляются с использованием стандартных типовых фраз, в гражданской авиации на территории Российской Федерации определенных Федеральными авиационными правилами [1]. При этом часто условия приема на борту могут быть затруднены в силу различных факторов: напряженного состояния пилота (или члена экипажа), присутствия дополнительных шумов в принимаемом сигнале и т.д. Попытки уточнить информацию приводят к тому, что осуществляется повторная передача одного

и того же сообщения, что сокращает скорость канала передачи информации и в ряде критических ситуаций может быть неприемлемо.

Транспортные системы в настоящее время развиваются все более и более стремительно. В их числе воздушный транспорт, наземный транспорт и водный транспорт. Если на наземном транспорте уже протестировано достаточно большое число различных современных технологий, включая различные автоматические системы мониторинга состояния водителя [2], то в системах воздушного транспорта распространены беспилотные летательные средства, для которых, однако, зачастую требуется согласование плана полета. На беспилотном наземном транспорте активно используются технологии искусственного интеллекта, в том числе различные системы распознавания. При этом важной становится визуальная информация. Для систем воздушного транспорта важна информация, помогающая осуществлять полет и поступающая от диспетчеров пунктов управления воздушным

Андриянов Никита Андреевич, кандидат технических наук, доцент Департамента анализа данных и машинного обучения. E-mail: nikita-and-nov@mail.ru

движением (УВД). Как правило, это голосовые сообщения. Потенциально применение систем автоматического распознавания позволяет снизить загруженность с «человека за рулём» вне зависимости от типа транспорта. Подобного рода прикладные задачи связаны с цифровой обработкой сигналов и обеспечением качественной связи. Обработка сигналов выполняется системами искусственного интеллекта и технического зрения, направленными на анализ изображений, выделение ориентиров, извлечение из них информации о местоположении, одновременное позиционирование и построение карт [3-5]. Однако такие системы в большей степени направлены на развитие автономных транспортных средств. Сегодня такие аппараты используются лишь для решения узкоспециализированных задач и, как правило, выполняются полностью в автономном режиме, т.е. при отсутствии человека внутри. В связи с этим к таким транспортным средствам априори предъявляются меньшие требования по безопасности.

Вместе с тем широкое применение пассажирского транспорта связано с определенными рисками. Немаловажным фактором здесь является физическая и эмоциональная нагрузка на «человека за рулем». Применительно к авиации речь идет о нагрузке на пилотов. Радиообмен является неотъемлемой частью работы пилота гражданской авиации (ГА), с которой он сталкивается постоянно. Фразеологический радиообмен подразумевает обмен короткими типовыми сообщениями, например, между экипажем воздушного судна (ВС) и диспетчерским пунктом для постоянного контроля обстановки на земле и в воздухе.

На данный момент радиообмен не стабилен, присутствует множество помех, ухудшающих восприятие сообщения и его правильное понимание. Для повышения безопасности полета обмен данными должен быть максимально качественным, пилот должен мгновенно реагировать на принятую от наземных служб информацию, а для этого необходимо свести к минимуму запросы на повтор сообщения.

Несмотря на то, что Федеральными авиационными правилами (ФАП) [1] установлены ограничения на участках по связи с Органами управления воздушным движением (УВД), радиообмен является неотъемлемой частью любого полета и оказывает определенное воздействие на пилотов. Действительно, правильный прием и трактовка сообщений в ходе фразеологического радиообмена являются важным элементом спокойствия как пилота, так и воздушной обстановки. При этом чаще всего полученные сообщения могут приниматься в условиях сильных помех, что также требует систем подавления, работающих в режиме реального времени.

Базой работы таких систем служат различные алгоритмы, разработка и исследование которых могут повысить эффективность распознавания и фильтрации речевых сообщений (РС).

Поэтому важно иметь на борту вспомогательные системы, которые могли бы определять фразы из сообщения в соответствии с некоторыми заранее известными классами. В связи с этим существует задача распознавания речевых сообщений радиообмена. Несмотря на то, что в настоящее время для решения задач распознавания речи активно применяются нейросетевые технологии [6], при ограниченном числе эталонов наивысшую точность может обеспечивать корреляционный алгоритм [7]. В данной работе исследуется распознавание на базе спектров эталонных фраз, произнесенных одним диктором.

Целью настоящей работы является исследование эффективности распознавания в различных условиях: при наличии шума в сигнале, при отсутствии (минимизации) шума в сигнале. При этом к сигналам не применяются алгоритмы предварительной обработки, например, фильтрации, позволяющие потенциально повысить качество постобработки.

МЕТОДЫ ИССЛЕДОВАНИЙ

Поскольку речевые сообщения можно представить как временные ряды, то к ним можно применять методы цифровой обработки сигналов, в том числе на базе случайных процессов [8-10]. Запись речевого сообщения осуществлялась в условиях, имитирующих слабый уровень дополнительных помех (искусственно был сгенерирован шум, отношение среднего уровня сигнала к дисперсии шума приблизительно равно 3). Кроме того, были записаны также 20 тестовых сообщений без шума. Эталонные образцы, все 3 записи, происходили также в условиях без генерации искусственного белого шума.

Для записи РС использовалась бесплатная программа Audacity. Данное программное обеспечение позволяет на основе записанного звукового файла, построить его спектральные составляющие: 255 значений спектра. При этом приведение значений спектра [дБ] в [В] позволяет получить временное представление сигнала в абсолютных значениях.

Для построения автокорреляционных функций (АКФ) записанных сигналов можно воспользоваться следующей связью амплитудного спектра сигнала с его энергетическим спектром (спектральной плотностью мощности)

$$G(F) = |S(F)|^2, \quad (1)$$

где $G(F)$ – энергетический спектр сигнала, $|\cdot|$ – операция вычисления модуля, $S(F)$ – спектр сигнала.

Определив дисперсию сигнала, как значение АКФ в нулевой момент времени (при нулевом сдвиге), можно с помощью обратного преобразования Фурье рассчитать значения АКФ в соответствии с выражением на базе связи АКФ и энергетического спектра (1)

$$R(t) = \frac{\int_{-F_{\max}}^{F_{\max}} G(F)e^{j2\pi Ft} dF}{\sigma^2}, \quad (2)$$

где $\sigma^2 = \int_{-F_{\max}}^{F_{\max}} G(F)dF$ – значение АКФ в нулевой момент времени (при нулевом сдвиге).

На рис. 1 показаны АКФ первой (а) и второй (б) тестовых фраз, полученных в результате расчетов (2).

Следует отметить, что на рис. 1 по оси оУ представлены нормированные значения сигнала, а по оси оХ – дискретные номера интервалов времени с шагом 0.15 с. Из представленных графиков видно, например, что фраза из рис. 1б оказывается гораздо более коррелированной с собой (отсчеты фразы друг с другом), чем фраза из рис. 1а. Это может быть использовано при корреляционном анализе для различения гипотез о фразах.

Все исследуемые в данной работе фразы будут рассмотрены далее. По временным отсчетам речевого сигнала с помощью преобразования Фурье могут быть построены спектры для каждого сообщения, показывающие составляющие для 255 различных частот в диапазоне от 86 Гц

до 22 кГц. Характеристики эталонной фразы были построены путем усреднения трех речевых сигналов на каждой частотной составляющей. Тестовые фразы записывались одним и тем же диктором по 20 раз. Всего для исследования было выбрано пять фраз, представленных в табл. 1.

На рис. 2 представлена усредненная эталонная фраза во временной (а) и частотной области (б). По оси оУ – уровень сигнала в дБ, по оси оХ – время в с (а) и частота в Гц (б).

Затем для проверки корреляционных связей была рассчитана корреляционная матрица для всех пяти исследуемых фраз

$$R = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} & \rho_{25} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} & \rho_{35} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 & \rho_{45} \\ \rho_{51} & \rho_{52} & \rho_{53} & \rho_{54} & 1 \end{pmatrix}, \quad (3)$$

где $\rho_{ij} = \frac{M\{(\Phi_i - m_i)(\Phi_j - m_j)\}}{\sigma_i \sigma_j}$ – коэффициент

корреляции между i -й и j -й фразами. Здесь Φ_i – это выборка составляющих спектра для i -й фразы, m_i – среднее значение в выборке спектральных составляющих для i -й фразы, σ_i – среднее отклонение в выборке спектральных составляющих для i -й фразы. Аналогично и для j -й фразы.

Следует отметить, что в силу того, что анализируется спектральное представление голоса одного и того же диктора, все произнесенные диктором фразы имеют сильную корреляционную

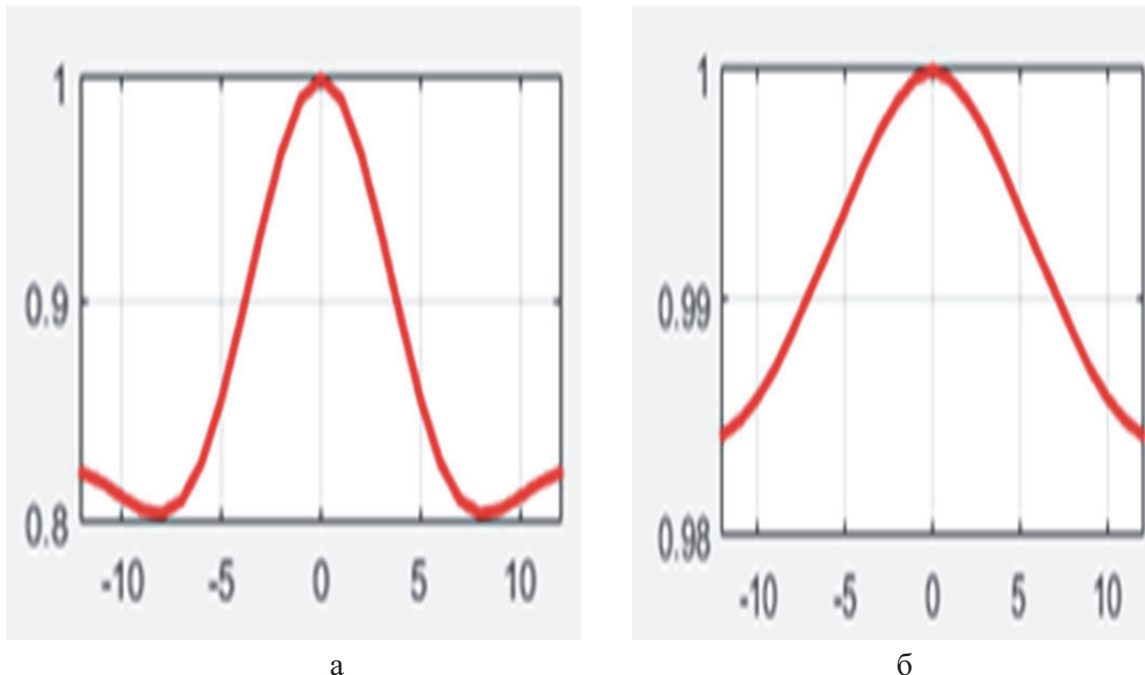


Рис. 1. Автокорреляционная функции фразы «Предполетная проверка связи» (а) и фразы «Принимаю вас на 5» (б)

Таблица 1. Распределение выборки

№ п/п	ФРАЗА	Диктор 1 Эталон	Диктор 1 Тест
1	Предполетная проверка связи	3	20
2	Принимаю вас на 5	3	20
3	Информация «Элфа» на борту	3	20
4	Разрешите запуск на стоянке	3	20
5	Запуск разрешили	3	20

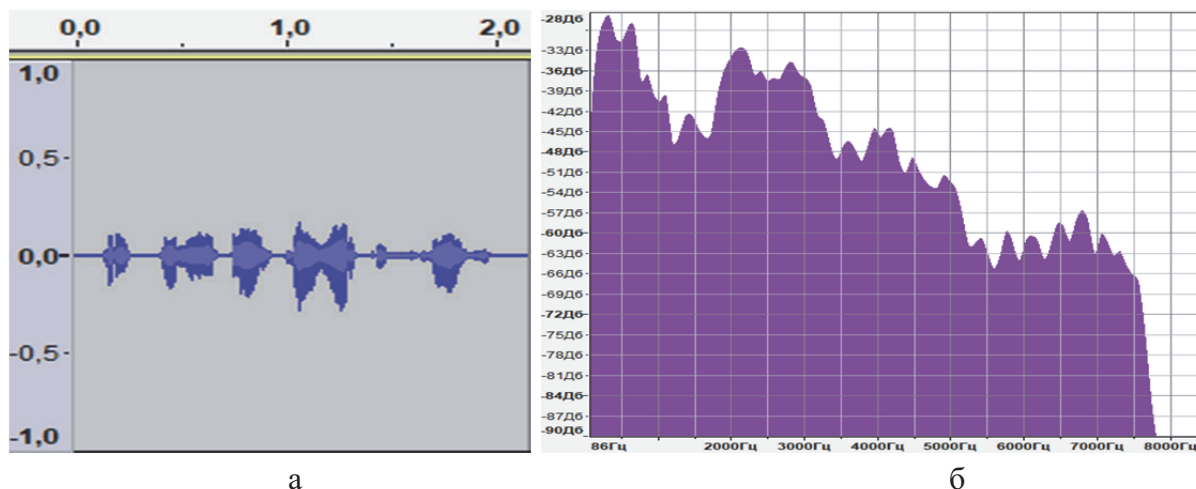


Рис. 2. Временное (а) и частотное (б) представление эталона для первой фразы.

связь. В ходе эксперимента, как видно из таблицы 1, анализировалось распознавание 20 образцов каждой фразы. При этом считалась корреляция отсчетов выборки спектра распознаваемой фразы и каждого эталона. Таким образом, при распознавании 5 фраз были посчитаны 5 коэффициентов корреляции для каждого тестового образца. Например, расчеты для 5 различных тестовых фраз представлены ниже. Поиск коэффициентов корреляции выполняется по формуле:

$$\rho_{ij} = \frac{M\{(\Phi_{zi} - m_{zi})(\Phi_{mj} - m_{mj})\}}{\sigma_{zi}\sigma_{mj}}, \quad i = 1,2,3,4,5, \quad (4)$$

где индексы «z» и «m» введены для указания на характеристику эталонной и тестовой фразы соответственно.

Рассчитанная по выражению (3) корреляционная матрица эталонов выглядит следующим образом (табл. 2).

Таблица 2. Корреляция эталонов

	Фраза 1 (Э)	Фраза 2 (Э)	Фраза 3 (Э)	Фраза 4 (Э)	Фраза 5 (Э)
Фраза 1 (Э)	1	0,97694	0,982227589	0,978056	0,978333
Фраза 2 (Э)	0,976939782	1	0,985586624	0,991763	0,991827
Фраза 3 (Э)	0,982227589	0,985587	1	0,978174	0,989996
Фраза 4 (Э)	0,978055947	0,991763	0,978173871	1	0,983958
Фраза 5 (Э)	0,978332623	0,991827	0,989996099	0,983958	1

После расчета пяти параметров по выражению (4) решение об отнесении тестового образца к одному из классов производится на основе максимума корреляции

$$i_m = i_{\text{Max}\{\rho_{ij}\}}, \quad i = 1,2,3,4,5 \quad (5)$$

Таким образом, выражение (5) позволяет определить класс наиболее вероятного сообщения.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

Рассмотренный выше алгоритм построения корреляционной матрицы эталонов может быть применен и к тестовым сообщениям. В частности, для пяти разных тестовых фраз получены следующие характеристики (табл. 3).

Исследование для записанных сигналов без

Таблица 3. Корреляция тестовых фраз и эталонов

	Фраза 1 (Э)	Фраза 2 (Э)	Фраза 3 (Э)	Фраза 4 (Э)	Фраза 5 (Э)
Фраза 1 (Т)	0,992250837	0,987923	0,987936849	0,984963	0,988123
Фраза 2 (Т)	0,976362706	0,994272	0,980751845	0,988217	0,985951
Фраза 3 (Т)	0,980980353	0,983718	0,997973913	0,973522	0,988553
Фраза 4 (Т)	0,982253031	0,991562	0,978709376	0,995522	0,985845
Фраза 5 (Т)	0,975854959	0,990852	0,991329316	0,983671	0,997937

Таблица 4. Процент правильных распознаваний в отсутствии помех

№ п/п	ФРАЗА	Эталон 1	Эталон 2	Эталон 3	Эталон 4	Эталон 5	Доля
1	Предполетная проверка связи	20	0	0	0	0	100%
2	Принимаю вас на 5	0	20	0	0	0	100%
3	Информация «Элфа» на борту	0	0	20	0	0	100%
4	Разрешите запуск на стоянке	0	0	0	20	0	100%
5	Запуск разрешили	0	0	0	0	20	100%
	Общая:						100%

Таблица 5. Процент правильных распознаваний при шуме

№ п/п	ФРАЗА	Эталон 1	Эталон 2	Эталон 3	Эталон 4	Эталон 5	Доля
1	Предполетная проверка связи	18	1	0	0	1	90%
2	Принимаю вас на 5	0	19	0	0	0	95%
3	Информация «Элфа» на борту	1	0	18	1	0	90%
4	Разрешите запуск на стоянке	0	0	2	18	0	90%
5	Запуск разрешили	1	0	0	1	18	90%
	Общая:						91%

шума показали абсолютную точность корреляционного анализа. Результаты представлены в табл. 4.

Анализ представленных результатов показывает, что, несмотря на сильную корреляцию между различными фразами, максимальные значения коэффициентов корреляции позволяют правильно распознать фразу. В табл. 5 представлены характеристики распознавания для всего тестового набора фраз.

Как видно из табл. 5, доля правильных распознаваний составляет не менее 90%, несмотря на то, что здесь корреляционные алгоритмы применялись для зашумленных речевых сигналов. Однако следует отметить,

что была достаточно маленькая тестовая выборка. Общая доля правильных распознаваний считалась как отношение числа верных распознаваний на всей тестовой выборке (вне зависимости от класса) к общему числу тестовых фраз и составила 91%. При этом не наблюдается большого разброса для разных фраз. Вместе с тем, следует учитывать присутствие белого шума при записи тестовых образцов (отношение сигнал/шум порядка 3).

ЗАКЛЮЧЕНИЕ

Таким образом, в работе исследованы характеристики распознавания речевых сообщений

фразеологического радиообмена гражданской авиации на русском языке на базе корреляционного анализа спектрального представления сигналов. Получена достаточно высокая точность для тестовой выборки – 91% (9% неправильных распознаваний на все фразы).

СПИСОК ЛИТЕРАТУРЫ

1. Приказ Минтранса РФ от 26 сентября 2012 г. № 362 «Об утверждении Федеральных авиационных правил «Порядок осуществления радиосвязи в воздушном пространстве Российской Федерации»
2. Козловский А.И., Порватов И.Н., Подольский М.С. Обзор автомобильных систем оперативного контроля состояния водителя. Результаты собственных исследований // Науковедение. 2013. № 6, С. 1-12.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс, М.: Техносфера 2012. – 1104 с.
4. Nemoto, Takuma and Onodera, Keichi and Mohan, Rajesh Elara and Iwase, Masami and Wood, Kristin. An Application of the Simultaneous Localization and Mapping (SLAM) Method Based on the Unscented Kalman Filter (UKF) to a Reconfigurable Quadrupe Robot with Crawling Locomotion // 2018 International Conference on Reconfigurable Mechanisms and Robots - Pp. 1-8
5. Альгирдас Руйбис. Беспилотные летательные аппараты: новейшие разработки // ВКС. 2016. №3-4 (88-89) - С. 94-96.
6. Казачкин, А. Е. Методы распознавания речи, современные речевые технологии / А. Е. Казачкин. — Текст : непосредственный // Молодой ученый. — 2019. — № 39 (277). — С. 6-8.
7. Красовская И.К., Смирнов М.Н., Смирнова М.А. Корреляционный метод в задаче распознавания речи // Процессы управления и устойчивость. 2016. №1. Т. 3. С. 409-413.
8. Андриянов Н.А., Гаврилина Ю.Н. Обнаружение и различение сигналов, моделируемых авторегрессиями с изменяющимися свойствами // В сборнике: 28-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии» (КрыМиКо'2018). Материалы конференции. 2018. С. 356-360.
9. Андриянов Н.А. Обзор методов прогнозирования временных рядов // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2018. № 11. С. 147-151.
10. Андриянов Н.А., Булдаковский А.И., Рисухин К.В. Анализ корреляционных функций речевых сообщений фразеологического радиообмена на русском языке // В книге: Актуальные проблемы физической и функциональной электроники. Материалы 22-й Всероссийской молодежной научной школы-семинара. 2019. С. 79-82.

RECOGNITION OF RADIO EXCHANGE VOICE MESSAGES IN AVIATION BASED ON CORRELATION ANALYSIS

© 2021 N.A. Andriyanov

Financial University under the Government of the Russian Federation, Moscow, Russia

The paper considers the problem of speech messages recognition in phraseological radio exchange for tasks of civil aviation. The introduction substantiates the relevance of this problem. The following are research methods based on correlation analysis. Finally, a description of the experiment and the results of the recognition algorithms based on correlation analysis are given. Various variants were recorded for five speech messages and spectral representations of such signals were constructed. Spectral transform can be obtained either using specialized software or based on the Fourier transform of the signal in the time domain. To obtain a more universal reference signal and eliminate the influence of interference, the spectral components of the same speech message recorded several times were averaged. In fact, three spectra of the same speech message were used for averaging. This spectrum averaging over three training components provided a reference sample of phrases or patterns for each phrase, and reduced the influence of additive white Gaussian noise in the reference. Later, on the basis of correlation analysis, the connections between test phrases and all patterns were calculated. On the basis of these connections, a correlation matrix of reference phrases is built. Research has shown that phrases spoken by one person were highly correlated. The analysis showed that the choice of the class (the content of the speech message) when solving the recognition problem corresponding to the value of the correlation coefficient closest to one provides over 90% of correct recognitions on a test sample containing a total of 100 phrases, 20 for each phrase. It should be noted that, when recording test messages, an additive white Gaussian noise was additionally present as a background, reproduced by another audio device. In the case of information analysis without artificially generated noise, the probability of correct recognition for a test sample of 100 phrases, 20 for each phrase, is 100% when using correlation analysis.

Keywords: radio phraseology, civil aviation, statistical analysis, correlation analysis, pattern recognition, signal spectrum

DOI: 10.37313/1990-5378-2021-23-1-91-96

*Nikita Andriyanov, Candidate of Technics, Assistant Professor
at the Data Analysis and Machine Learning Department.
E-mail: nikita-and-nov@mail.ru*