



МЕТОДИКА ПОЗТАПНОГО ВНЕДРЕНИЯ ПОЛИМОДАЛЬНЫХ ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ

О. О. БАСОВ¹
А. Л. РОНЖИН²

¹⁾ *Академия Федеральной службы охраны Российской Федерации*

²⁾ *Санкт-Петербургский институт информатики и автоматизации Российской академии наук*

e-mail:
oobasov@mail.ru

В работе предложено использовать методику Wizard of OZ как наиболее быстрый и продуктивный подход к оптимизации и настройке пользовательских приложений в реальных абонентских терминалах полимодальных инфокоммуникационных систем. Данный подход позволяет вести наблюдение за абонентом, работающим с системой, недостающие службы которой восполняются скрытыми операторами. Для снижения когнитивной нагрузки предложена структура с несколькими операторами. Сложность ее реализации заключается в необходимости обеспечения их совместной работы, особенно в условиях отклонения психофизиологического состояния операторов от нормы. Для определения последнего предложено использовать сигналы многомодального входного интерфейса абонентского терминала. Показан выигрыш в эффективности распределения функций между операторами при учете их психофизиологического состояния.

Ключевые слова: полимодальная инфокоммуникационная система, абонентский терминал, многомодальный интерфейс, оператор, психофизиологическое состояние.

В сложной природе многомодального взаимодействия когнитивная наука приобретает особую важность при разработке полимодальных инфокоммуникационных систем (ПИКС). Под ПИКС следует понимать взаимоувязанную совокупность систем обработки и хранения информации, телекоммуникационных сетей, их объединяющих, функционирующих под единым управлением с целью сбора, обработки, хранения, защиты, передачи и распределения, отображения и использования многомодальной информации, учитывающей смысл сообщаемых сообщений, личность абонентов (пользователей), их настройку, физиологическое и психоэмоциональное состояния [1].

На стадии первичного тестирования апробировать и исследовать эффективность синтезированной ПИКС позволяет методика Wizard of OZ (WOZ) [2, 3]. Данный подход позволяет вести наблюдение за пользователем (абонентом), работающим с системой, недостающие службы которой восполняются скрытыми операторами (рис. 1). Абонент не знает о присутствии оператора и уверен, что система работает в автоматическом режиме. Оператор наблюдает за абонентом, используя специализированную компьютерную систему, соединенную с оконечными терминалами через сеть.

Существует два направления в технике моделирования и тестирования. В первом случае в основу закладываются теоретические предположения о природе и способах взаимодействия, во втором – информация, полученная в ходе предыдущих экспериментов. Прогнозируемые модели не требуют какой-либо системной реализации и не нуждаются в подготовленных пользователях. Примерами подобных моделей являются GOMS [4] и родственные ему (например, SST [5]), а также модели, основанные на теории (ICS [6], KRI [7]) и когнитивный метод критического анализа [8]. Их главное преимущество заключается в том, что они позволяют провести оценивание пользовательского интерфейса на ранней стадии проектирования. Однако подобные модели основаны на теоретических гипотезах, а не на реальных данных, поэтому им может не хватать точности или они могут быть ограничены в возможностях. Кроме того, регулирование и интерпретация таких систем затруднительны и требуют существенных временных затрат.

Напротив, экспериментальная техника оценивания имеет дело с реальными данными, полученными в ходе наблюдений за реальными пользователями, исполняющими реальные задания и оперирующими с физическими предметами. В качестве предметов могут выступать план действий, написанный на бумаге, экспериментальная модель, опытный объект, исследуемый на компьютере. Когда

пользователь вызывает функцию, которая не доступна в разрабатываемой системе, оператор имитирует ее действие, используя специализированное программное обеспечение. Таким образом, разработчики могут изучить потребности пользователя, а затем настроить и оценить тот особый интерфейс, который следует использовать при выполнении заданий данного класса.

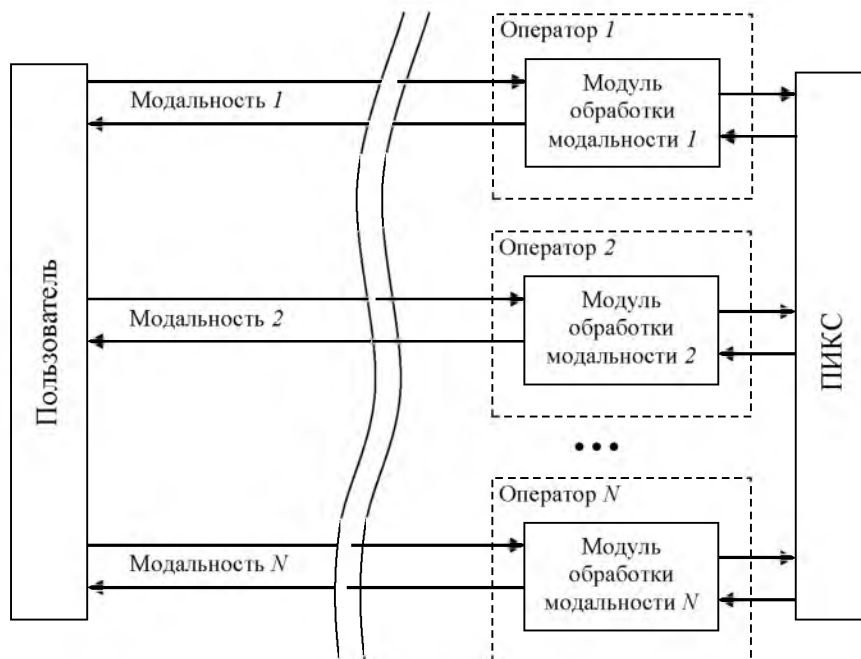


Рис. 1. Моделирование работы ПИКС в режиме скрытой экспертной поддержки

Большинство существующих систем, основанных на WOZ, первоначально были разработаны для изучения особенностей употребления естественного языка в области информационных услуг. Телефонные службы, такие как телефонные справочники, службы предоставления информации, требуют автоматизации операторских услуг и представляют собой широкое поле деятельности для проведения экспериментов [9]. Схема эксперимента довольно проста: абоненты ведут диалог с автоматической информационной системой, а оператор имитирует ее работу, отвечая на телефонные звонки. Для того чтобы у абонентов создавалась иллюзия, что они действительно разговаривают с компьютером, голос оператора пропускается через систему искажения (например, вокодер), имитирующую голос робота. Вопросы и ответы записываются для дальнейшего ручного транскрибирования и анализа. Главной целью данного эксперимента является накопление речевых баз данных (звуковых, текстовых), необходимых для обучения системы распознавания, менеджера диалога и других модулей обработки и ведения диалога в естественной форме. Например, системы, описанные в [2, 10], осуществляют наблюдение за движениями рук в сочетании с текстовыми сообщениями и речью. Несмотря на сложность реализации данного эксперимента с WOZ, все же был накоплен достаточный объем данных, позволяющий сделать некоторые выводы.

Замечено, что когда пользователи полагают, что общаются с компьютером, сложность языка гораздо ниже, чем при естественном общении человека с человеком [2]. Однако это сказывается на естественности общения и создает некоторый дискомфорт, поэтому одна из основных задач разработчиков направлена на повышение робастности системы к различным факторам, ухудшающим качество входных сигналов (неточности в движениях, речи пользователя, посторонние аудио-видео шумы, помехи в канале и приемно-передающих устройствах и т. д.) [10].

Другой интересный результат, полученный в ходе моделирования на основе WOZ, связан с особенностями поведения оператора. Для сохранения реализма автоматической



обработки и управления требуется, чтобы его действия были строго согласованы по смыслу, стилю и скорости. В частности:

- 1) в одинаковых ситуациях команда, поступившая от абонента, должна инициировать одинаковое поведение оператора;
- 2) время ответа должно соответствовать ожиданиям пользователя: если оператор реагирует слишком медленно, абонент может отменить свой запрос (имитируемую функцию) или подумать, что система перегружена.

В целом операторы не могут позволить себе импровизации. Чтобы добиться желаемого поведения, они должны обучаться на четко определенных заданиях и быть обеспечены мощными программно-аппаратными инструментами. Некоторые системы включают в себя ограниченные, но довольно удобные механизмы для формирования стандартных ответов или меню, содержащие предварительно заготовленные части ответов. Например, инструментальные средства SUEDE для разработки речевых пользовательских интерфейсов позволяют быстро создать интерфейс диалоговой системы на базе WOZ с помощью типовых запросов/ответов [11].

Для того чтобы снизить когнитивную нагрузку, была предложена структура с двумя операторами (по каждому каналу коммуникации), в которой один оператор специализируется на преобразовании входной/выходной модальности, а второй – выполняет обработку заданий и формирование ответа (рис. 2) [12, 13].

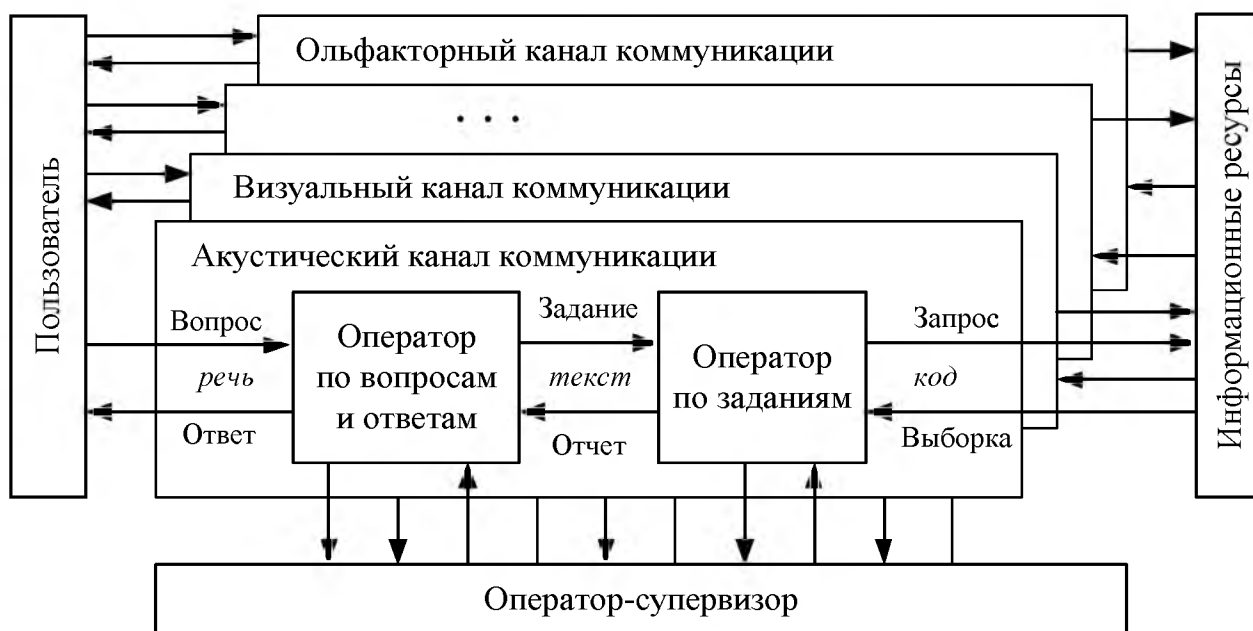


Рис. 2. Распределение функций операторов при реализации инфокоммуникационных услуг

Первый оператор получает запросы пользователя и выдает полученные ответы; оператор по заданиям интерпретирует запросы пользователя, переданные ему первым оператором, и генерирует ответы, которые будут изложены оператором по вопросам и ответам. Такое распределение функций обеспечивает слаженность и устойчивость работы информационной службы.

Сложность построения систем с несколькими операторами заключается в том, что необходимо обеспечить совместную работу операторов и гарантировать их непротиворечивое поведение. В идеале, нагрузка должна быть одинаково распределена между операторами. Однако объем работы сложно оценить, так как он во многом зависит от поведения пользователя. При многомодальном взаимодействии могут использоваться несколько модальностей параллельно. В этих условиях необходимо, чтобы операторы динамично меняли свои роли. Кроме того, необходим специальный оператор, который будет выступать в качестве супервизора [14]. Оператор-супервизор не выполняет каких-



либо обычных заданий, а регулирует поведение операторов, отслеживает сеанс работы пользователя с системой и принимает соответствующие решения при неисправной работе.

При гибкой конфигурации системы (рис. 2) число операторов, активных приложений, пользователей и используемых модальностей может динамически изменяться. Работа оператора-супервизора состоит в оптимальном распределении человеко-машинных ресурсов во избежание перегрузок. Однако при этом не учитываются влияния, связанные с возникновением нервно-эмоционального напряжения, утомления, заболевания и других отклонений психофизиологического состояния (ПФС), характерных для операторов, что негативно сказывается на качестве их работы. Развитию методов оценки психофизиологического состояния посвящены работы [15–19]. Однако в них, как правило, используется либо косвенная оценка ПФС операторов, либо рассматриваются одномодальные (например, сигнал клавиатуры) входные интерфейсы автоматизированных рабочих мест операторов, что значительно снижает точность определения таких состояний.

Повысить точность определения ПФС операторов позволяет использование информации от многомодальных (речь, клавиатура, мышь) входных интерфейсов автоматизированных рабочих мест операторов. Научно-методический аппарат оценивания ПФС $f_i, i = \overline{1, M}$, операторов достаточно подробно изложен в [20–24].

Полученные оценки f_i могут быть положены в основу задачи оптимального распределения функций между операторами [25]:

$$E = \sum_{i=1}^M E_i = \sum_{i=1}^M f_i \sum_{j=1}^N e_{ij} x_{ij} \rightarrow \max \quad (1)$$

при условии закрепления достаточного числа операторов за каждой функцией

$$\sum_{i=1}^M x_{ij} = 1;$$

и ограничениях на их загрузку

$$\sum_{j=1}^N x_{ij} Q_j \leq q_i,$$

где $E_i = f_i \sum_{j=1}^N e_{ij} x_{ij}$ – производительность i -го оператора; $e_{ij} = w_j \sum_{k=1}^K (\min(P_{ik}, R_{jk})) / \sum_{k'=1}^K R_{jk'}$ – эффективность закрепления i -го оператора за j -й функцией; w_j – показатель значимости j -й функции $\left(\sum_{j=1}^N w_j = 1 \right)$; $\bar{R}_j = (R_{j1}, R_{j2}, \dots, R_{jk}, \dots, R_{jK})$ – нормативный профиль j -й функции, представляющий требуемые уровни компетенций оператора, необходимые для выполнения j -й функции $(j = \overline{1, N})$; $\bar{P}_i = (P_{i1}, P_{i2}, \dots, P_{ik}, \dots, P_{iK})$ – квалификационный профиль (профиль компетенций) i -го оператора $(i = \overline{1, M})$; P_{ik} – фактический уровень i -го оператора по k -ой компетенции $(k = \overline{1, K})$; $0 \leq x_{ij} \leq 1$ – доля j -й функции, выполняемая i -м оператором; Q_j – трудоемкость выполнения j -й функции; q_i – трудовой ресурс i -го оператора.



Решение задачи (1) симплекс-методом позволяет супервизору перераспределить функции между операторами:

$$q_i^{\text{ФАКТ}} = \sum_{j=1}^N x_{ij} Q_j,$$

временно исключить из процесса экспертной поддержки операторов, у которых наблюдается прогрессивное снижение работоспособности.

Повышение производительности операторов определяется превышением кривой 2 (рис. 3), соответствующей значениям показателя эффективности (1) с учетом ПФС f_i операторов, над кривой 1, рассчитанной без учета таких состояний. В представленном примере показано (верхний график) изменение ПФС трех операторов:

- 1) оператора по заданиям текстового канала коммуникации, работающего только с использованием «мыши» – f_1 ;
- 2) оператора по вопросам и ответам текстового канала коммуникации, работающего только с использованием, клавиатуры – f_5 ;
- 3) оператора по вопросам и ответам акустического канала коммуникации, использующего речь – f_7 .

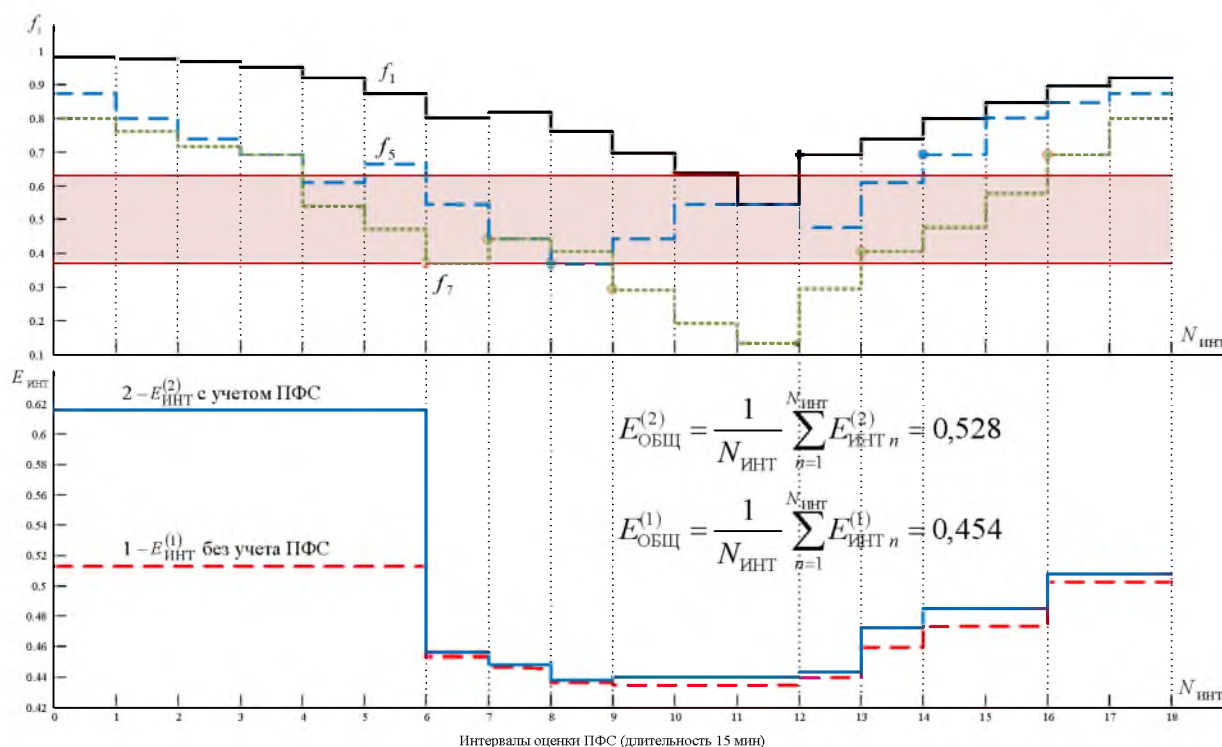


Рис. 3. Эффективность распределения функций между операторами

При этом эффективность распределения функций между операторами в течение сеанса взаимодействия с абонентом (18 интервалов оценки ПФС) в рассматриваемом примере повышается на 7,4 %. На основе полученных данных сделан вывод, что динамическое распределение функций на основе разработанного критерия (1) является адекватным задаче распределения функций между операторами (рис. 2) с учетом их ПФС, и позволяет автоматизировать соответствующий процесс [26].

Таким образом, представленные результаты свидетельствуют о возможности использования методики Wizard of OZ на стадии первичного тестирования ПИКС, в том числе в условиях повышенной когнитивной и психологической нагрузки на операторов.



Литература

1. Басов О.О. Качество функционирования и эффективность полимодальных инфокоммуникационных систем / О.О. Басов, И.А. Саитов // Труды СПИИРАН. 2014. Вып. 1 (32). – С. 152–170.
2. Dahlback N. Wizard of Oz Studies – Why and How / N. Dahlback, A. Jonsson, L. Ahrenberg // Knowledge-Based Systems. 1993. Vol. 6. No. 4. – P. 258–266.
3. Salber J.C. Applying the Wizard of Oz Technique to the Study of Multimodal Systems / J.C. Salber // Proc. EWHCI'93. Berlin: Springer-Verlag, 1993. – P. 219–230.
4. Ронжин А.Л. Применение техники «Гудвин» для моделирования человеко-машинного взаимодействия / А.Л. Ронжин, А.Б. Леонтьева // Известия вузов. Приборостроение. 2006. Т. 49. № 11. – С. 70–75.
5. Card S.K. The Psychology of Human-Computer Interaction / S.K. Card, T.P. Moran, A. Newell // Lawrence Erlbaum Associates, 1983.
6. Kieras D. An Approach to the Formal Analysis of User Complexity / D. Kieras, P.G. Polson // International Journal of Man-Machine Studies. 1985. 22. – P. 365–394.
7. Barnard, P. J. Cognitive Resources and the Learning of Human-Computer Dialogues / P.J. Barnard // Interfacing Thought. Cognitive Aspects of Human-Computer Interaction / J.M. Carroll (ed.). MIT Press Publ., 1987. – P. 112–158.
8. Löwgren J. Knowledge-Based Tool for User Interface Evaluation and its Integration in a UIMS / J. Löwgren, T.A. Nordqvist // Human-Computer Interaction INTERACT '90. – P. 395–400.
9. Lewis C. Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces / C. Lewis, P. Polson, C. Wharton et al. // Proc. CHI'90. – P. 235–241.
10. Maulsby D. Prototyping an Intelligent Agent through Wizard of Oz / D. Maulsby, S. Greenberg, R. Mander // Proc. InterCHI'93. – P. 277–284.
11. Geppener V. V. Applications of empirical mode decomposition for processing nonstationary signals / V.V. Geppener, D.M. Klionski, N.I. Oreshko et al. // Pattern Recognition and Image Analysis. 2008. Vol. 18. No. 3. – P. 390–399.
12. Klemmer S.R. SUEDE: A Wizard of Oz Prototyping Tool for Speech User Interfaces / S.R. Klemmer et al. // Proc. UIST 2000. CHI Letters. 2000. 2(2). – P. 1–10.
13. Moran D. Multimodal user interfaces in the Open Agent Architecture / D. Moran, A. Cheyer, L. Julia et al. // Proc. of IUI-97. Orlando, 1997. – P. 61–68.
14. Лосев Ю.И. Модель технологии информационного обеспечения решения задач управления / Ю.И. Лосев, С.И. Шматков, К.М. Руккас Мохаммед Саламе Абрахим Арабиат // Научные ведомости Белгородского государственного университета. Серия «История. Политология. Экономика. Информатика». – 2013. – № 22 (165). Выпуск 28/1. – С. 195–203.
15. Абашин В.Г. К вопросу принятия решения о текущем психофизическом состоянии оперативного персонала по клавиатурному почерку / В.Г. Абашин // Известия ОрелГТУ. Серия «Фундаментальные и прикладные проблемы техники и технологии», 2007. №4/268 (535). – С.250–251.
16. Абашин В.Г. Исходные данные клавиатурного почерка для определения работоспособности человека / В.Г. Абашин // Известия Орел-ГТУ. Серия «Фундаментальные и прикладные проблемы техники и технологии», 2007. – №4-21268(525). – С. 136–139.
17. Савченко В.В. Анализ эмоционального состояния диктора по голосу на основе фонетического детектора лжи / В.В. Савченко, Р.А. Васильев // Научные ведомости Белгородского государственного университета. Серия «История. Политология. Экономика. Информатика». – 2014. – № 21 (192). Выпуск 32/1. – С. 186–195.
18. Смирнов С.В. Идентификация параметров загрузки оператора по требуемому уровню надежности / С.В. Смирнов // Информационные системы и технологии. 2012. № 4 (72). – С. 78–85.
19. Сизов А.Г. Оценка качества голоса лиц, страдающих дисфонией / А.Г. Сизов, С.Д. Тиунов, Р.В. Мещеряков // Труды СПИИРАН. 2012. Вып. 20. – С. 138–152.
20. Носов М.В. Математические модели и алгоритмы формирования джиттера сигналов текстового канала взаимодействия технических средств и оператора АРМ / М.В. Носов, Е.А. Васечкин, О.О. Басов // Материалы VI Международной научно-технической конференции «Информационные технологии в науке, образовании и производстве» (22-23 мая 2014 года). Орел: ГУ-УНПК, 2014. http://youconf.ru/files/itnor2014/Носов_Васечкин_Басов.pdf.
21. Носов М.В. Методика разделения джиттера сигналов различных каналов взаимодействия технических средств и оператора АРМ и оценки характеристик его компонент / М.В. Носов // Информационные системы и технологии. – Орел: Госуниверситет-УНПК. 2014. № 3(83). – С. 63–72.



22. Басов О.О. Исследование характеристик джиттера периода основного тона речевого сигнала / О.О. Басов, М.В. Носов, В.А. Шалагинов // Труды СПИИРАН. 2014. Вып. 1(32). – С. 27–44.

23. Басов О.О. Программа определения параметров текстовых модальностей / О.О. Басов, М.В. Носов, В.В. Никитин, Д.А. Гуляйкин // Свидетельство о государственной регистрации программы для ЭВМ № 2014613478 от 27.03.2014.

24. Басов О.О. Программа формирования характеристик случайного джиттера сигналов текстовых и речевого каналов коммуникации / О.О. Басов, М.В. Носов, В.В. Никитин, Д.А. Гуляйкин // Свидетельство о государственной регистрации программы для ЭВМ № 2014615750 от 02.06.2014.

25. Носов М.В. Критерий эффективности управления человеческими ресурсами / М.В. Носов, О.О. Басов // Сборник материалов VII Международной молодежной научно-практической конференции СКФ МТУСИ «ИНФОКОМ-2014». Ростов-на-Дону: СКФ МТУСИ. 2014. – С. 64–65.

26. Басов О.О. Программа динамического распределения производственно-технологических функций при изменении психофизиологических состояний исполнителей-операторов АРМ / О.О. Басов, М.В. Носов, П.А. Сысоев, Д.А. Гуляйкин // Свидетельство о государственной регистрации программы для ЭВМ № 2014616058 от 10.06.2014.

TECHNIQUE OF PHASED IMPLEMENTATION OF POLYMODAL COMMUNICATION SYSTEMS

O. O. BASOV¹
A. L. RONZHIN²

¹⁾ *Academy of Federal Agency of protection of Russian Federation*

²⁾ *St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences*

e-mail:
oobasov@mail.ru

In this paper we suggest using the Wizard of OZ methodology as the fastest and most productive approach to optimizing and tuning user applications in the actual subscriber terminals of multimodal communication systems. This approach allows us to monitor the subscriber, operating the system, the missing services of which are replenished by hidden operators. To reduce the cognitive load the structure with several operators is proposed. The complexity of its implementation is in the necessity of making them work together, especially if psycho-physiological states of operators deviate from the norm. To determine this fact the use of signals of the multimodal input interface of the subscriber terminal is proposed. The efficiency gain obtained by the allocation of functions between operators, taking into account their psycho-physiological states, is presented.

Key words: polymodal communication system, subscriber terminal, multimodal interface, operator, psycho-physiological state.