



УДК 621.391

**КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ МНОГОМОДАЛЬНОГО ИНТЕРФЕЙСА
АБОНЕНТСКОГО ТЕРМИНАЛА****A CONCEPTUAL MODEL MULTIMODAL USER INTERFACE OF THE SUBSCRIBER
TERMINAL****Д.А. Струев, Н. В. Бондарева, В.Ю. Будков, О.О. Басов, А.Л. Ронжин
D.A. Struev, N.V. Bondareva, V.U. Budkov, O.O. Basov, A.L. Ronzhin**

*Федеральное государственное казённое военное образовательное учреждение высшего образования
«Академия Федеральной службы охраны Российской Федерации»,
Россия, 302034, Орёл, ул. Приборостроительная, 35*

*Federal state military educational institution of higher professional education
"Academy of the Federal security service of the Russian Federation",
35 Priborostroitel'naya St, Orel, 302034, Russia*

*Санкт-Петербургский институт информатики и автоматизации Российской академии наук,
Россия, 199178, Санкт-Петербург, 14 линия, 39*

*St. Petersburg Institute for Informatics and Automation of Russian academy of sciences,
39, 14 Line, St. Petersburg, 199178, Russia*

e-mail: dima_orel86@mail.ru, natalibond.ru@gmail.com, budkov@iias.spb.su, oobasov@mail.ru

Аннотация. В работе представлена частично формализованная система взглядов на построение многомодальных интерфейсов человека-машинного взаимодействия. Концептуальная модель представлена когнитивными моделями реализации информационных и телекоммуникационных услуг. Первая модель отражает взаимодействие объектов и процессов, сопровождающих сбор, обработку, хранение и передачу многомодальной информации. Модель реализации информационных услуг отражает взаимодействие объектов и процессов, сопровождающих получение смысла сообщаемых сообщений, оценивание их истинности, установление личности абонента, его настроение, эмоциональное, психофизиологическое и ольфакторное состояния и другие варианты использования многомодальной информации. Разработанная концептуальная модель служит для выбора рациональных системотехнических решений по построению многомодальных интерфейсов человека-машинного взаимодействия.

Resume. The paper presents a partially formalized paradigm for the construction of multimodal interfaces for human-machine interaction. The conceptual model is represented by cognitive models of information and telecommunication services. The first model reflects the interaction of objects and processes involved in the collection, processing, storage and transmission of multimodal information. The realization model of information services reflects the interaction of objects and processes accompanying the processes of receiving the meaning of reported messages, estimation of message validity, identifying subscribers, along with their mood, emotional, physiological and olfactory state, as well as other uses of multimodal information. The developed conceptual model is designed for finding effective solutions to the construction of multimodal interfaces for human-machine interaction.

Ключевые слова: многомодальный интерфейс, человек-машинное взаимодействие, когнитивная модель, базисный фактор.

Keywords: multimodal interface, human-computer interaction, cognitive model, basic factor.

В настоящее время наблюдаются тенденции по слиянию или взаимопроникновению традиционных услуг связи, появлению новых информационных технологий, позволяющих пользователям расширять функциональность абонентских терминалов инфокоммуникационных систем (ИКС). Однако эти приложения внедряются в узкоспециализированных информационных системах за счет аддитивного наращивания затрат ресурсов пропускной способности и вычислительных мощностей процессоров.

Для максимальной открытости интерфейсов человека-машинного взаимодействия происходит обезличивание отдельных инфокоммуникационных услуг, т. е. пользователь все чаще запрашивает у сети один вид сервиса под названием «соединение с сетью», который подразумевает возможность получения им доступного или наиболее удобного способа взаимодействия с дру-



гим субъектом или информационными ресурсами. Однако в условиях наблюдаемого сегодня роста мощности и интенсивности информационных потоков, а также объемов информации, циркулирующей в информационном пространстве, существующие технические средства ИКС оказываются функционально ограниченными и не обеспечивают пользователя надежно и своевременно предоставляемой, полной, достоверной и конфиденциальной информацией.

Одним из путей преодоления указанного противоречия является создание новых методов обработки и представления информации, передачи все большей доли информационной деятельности информационным системам, создаваемым на базе новых информационных и коммуникационных технологий и аппаратно-программных средств. Перспективным в этом смысле является отказ от традиционных принципов разделения передаваемой информации на услуги связи (информатизации) в пользу ее многомодального представления. Здесь под модальностью следует понимать принадлежность отражаемого раздражителя к определенной сенсорной системе – части нервной системы человека, ответственной за восприятие определенных сигналов из окружающей или внутренней среды [Островский М., Шевелев И. 2003].

В рамках решения частной проблемы человеко-машинного взаимодействия разработаны многомодальные интерфейсы – системы, объединяющие различные средства ввода (передачи) информации в едином пользовательском интерфейсе человеко-машинного взаимодействия [Карпов А.А., Ронжин А.Л., Лобанов Б.М., Цирульник Л.И., Железны М. 2008; Ронжин А.Л., Ли И.В. 2007; Ронжин А.Л., Карпов А.А. 2008; Oviatt, 2003].

Соотнесение достижений когнитивной науки с результатами исследований в области информационных и телекоммуникационных технологий свидетельствует о необходимости и возможности перераспределения и согласования между пользователями информации от различных (традиционных и новых) сенсорных систем (анализаторов). Все это позволяет говорить о перспективности реализации инфокоммуникационных систем на основе многомодальных интерфейсов абонентских терминалов. Для количественного обоснования решений по их реализации разработана специальная система моделей. Моделью метауровня исследований является представленная ниже концептуальная модель.

Когнитивная модель многомодального интерфейса абонентского терминала

Концептуальная модель формируется проектировщиками системы и модернизируется специалистами организации-заказчика в ходе эксплуатации, а ее построение является предварительным этапом формулировки математической модели. Она относится к классу логико-лингвистических моделей [Хрусталев Е., Макаренко Д. 2008] и позволяет описать предметную область в обычных для пользователя многомодального интерфейса понятиях. В качестве концептуальных моделей в операционных исследованиях слабоструктурированных систем могут быть использованы когнитивные модели (карты).

Определение 1. Когнитивная карта многомодального интерфейса – структура (сеть) причинно-следственных связей между его элементами и средой функционирования, отображающая представление индивида (эксперта) или группы индивидов об устройстве и функционировании многомодального интерфейса человеко-машинного взаимодействия.

Определение 2. Ситуация – состояние многомодального интерфейса и корпоративной среды, характеризующееся структурой когнитивной карты, оценками динамики параметров базисных факторов и их значениями, зафиксированными в определенный момент времени.

Когнитивная карта представляет собой взвешенный ориентированный граф $G_{\text{КК}} = (X, C)$, вершинами X которого являются базисные факторы, а ребрами C – прямые влияния факторов друг на друга. С графом G ассоциируется матрица смежности графа $A_g = \{a_{ij}\}$, элементы которой характеризуют влияние фактора x_i на фактор x_j . Перечень базисных факторов когнитивной модели ПИКС представлен в таблице 1.

Для удобства восприятия и последующего анализа когнитивная (концептуальная) модель многомодального интерфейса человеко-машинного взаимодействия представлена моделями реализации телекоммуникационных (рис. 1) и информационных услуг (рис. 2). Структура указанных моделей и начальные значения параметров базисных факторов определены на основе методологии, описанной в работе [Хрусталев Е., Макаренко Д. 2008]. В качестве управляющих факторов выбраны факторы 1–10, относящиеся к многомодальному интерфейсу и среде функционирования, в качестве целевых (23–29, 39–48) – в наибольшей степени характеризующие состояние многомодального интерфейса и цели его функционирования.

Модель реализации телекоммуникационных услуг (рис. 1) отражает взаимодействие объектов и процессов, сопровождающих сбор, обработку, хранение и передачу многомодальной информации. На степень реализации телекоммуникационных услуг (факторы 23–29) оказывают влияние качество обработки сигналов различной модальности (факторы 11–22), а также ограничения:



- на способы человеко-машинного взаимодействия со стороны абонента – связанные с его навыками использования АТ и информационных технологий (фактор 1), личными предпочтениями и физическими ограничениями (причинно-следственные связи между факторами 1 и 2–7);
- среды человеко-машинного взаимодействия (причинно-следственные связи между факторами 2-7 и 11–22);
- на способы человеко-машинного взаимодействия со стороны многомодального интерфейса, связанные с его программно-аппаратными возможностями (фактор 9, причинно-следственные связи между факторами 9 и 11–22, 11–22 и 23–29);
- предоставляемых услуг, связанные с наличием доступа к инфокоммуникационным ресурсам (причинно-следственные связи между факторами 8 и 23–29) и их объемом (фактор 8).

Таблица 1
Table 1

Базисные факторы когнитивной модели многомодального интерфейса
Basic factors of a cognitive model of multimodal user interface

№	Наименование базисного фактора	ОДФ
Управляющие факторы		
1	Навыки использования АТ и информационных технологий (ИТ)	+1
2	Наличие устройств традиционного ввода (клавиатура и «мышь»)	0
3	Наличие устройств рукописного ввода	0
4	Наличие микрофонов	0
5	Наличие видеокамер	0
6	Наличие тактильных датчиков	0
7	Наличие анализаторов запаха	0
8	Объем инфокоммуникационных ресурсов	-1
9	Объем вычислительных ресурсов АТ	-1
10	Объем требуемой памяти	-1
11	Выделение текстовых модальностей (традиционный и рукописный ввод)	+1
12	Выделение акустических модальностей (речевой и неречевые сигналы, паузы)	+1
13	Анализ артикуляции губ	+1
14	Выделение модальностей лица (строение, выражение и цвет кожи лица)	+1
15	Анализ движения рук	+1
16	Анализ телодвижений	+1
17	Анализ положения тела (положение тела и относительное положение частей тела)	+1
18	Выделение модальностей глаз (движение зрачков и моргания)	+1
19	Анализ внешнего вида	0
20	Выделение мезансцены	0
21	Выделение тактильных модальностей (влажность рук, давление, температура)	+1
22	Выделение ольфакторных модальностей (запахи)	+1
23	Передача текстовых сообщений	-1
24	Передача неподвижных изображений	-1
25	Передача аудио	-1
26	Передача видео	-1
27	Передача данных полимодальной услуги	-1
28	Тактильный вывод	-1
29	Передача запахов	-1
30	Определение семантики сообщений	+1
31	Распознавание речи	+1
32	Чтение по губам	+1
33	Распознавание лиц	+1
34	Распознавание мимики	+1
35	Определение направления взгляда	+1
36	Определение положения тела	+1
37	Распознавание жестов	+1
38	Распознавание запахов	+1
39	Стенографирование передаваемых текстовых и речевых сообщений	+1
40	Информационный поиск по передаваемым сообщениям	+1
41	Идентификация абонента	+1
42	Управление объектами (навигация)	+1
43	Мониторинг активности абонента	+1
44	Определение степени алкогольной интоксикации	+1
45	Определение ольфакторного состояния	+1
46	Определение эмоционального состояния	+1
47	Определение психофизиологического состояния	+1
48	Оценка истинности передаваемой информации	+1

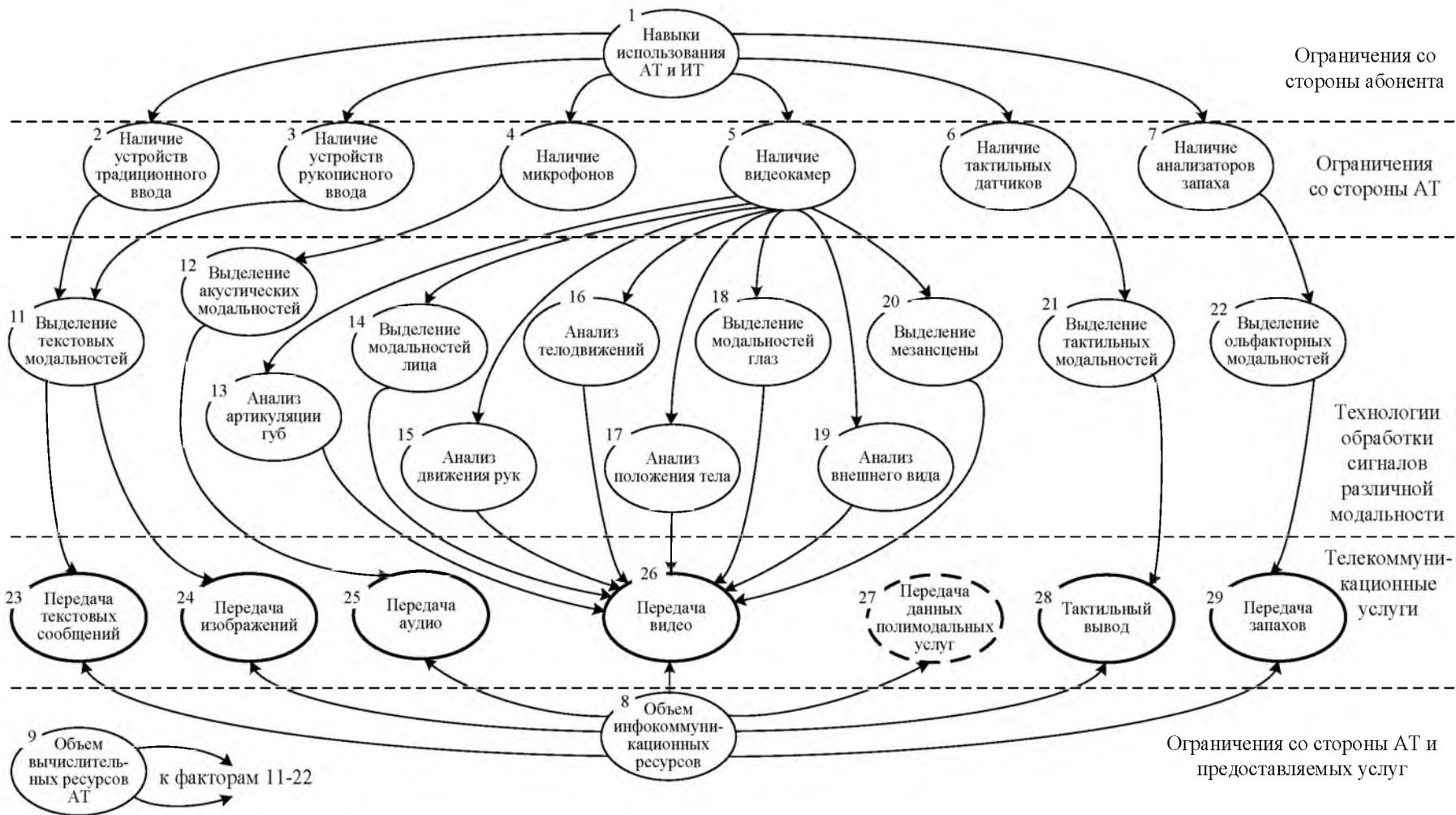


Рис. 1. Когнитивная модель реализации телекоммуникационных услуг
 Fig. 1. A cognitive model of the implementation of telecommunication services

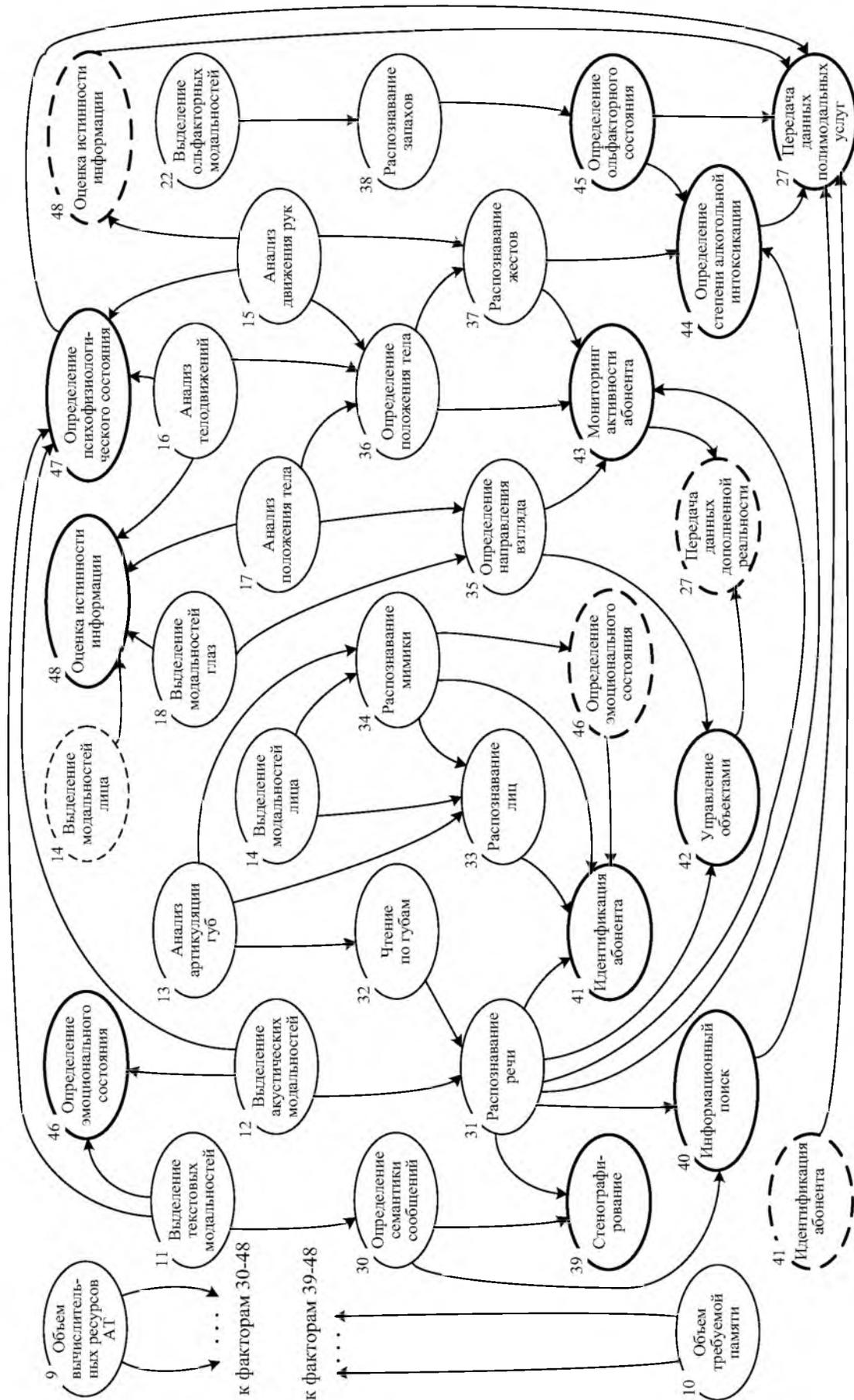


Рис. 2. Когнитивная модель реализации полимодальных услуг
Fig. 2. Cognitive model of realization of multimodal services



Модель реализации информационных услуг (рис. 2) отражает взаимодействие объектов и процессов, сопровождающих получение смысла сообщаемых сообщений, установление личности абонента, его настроение, эмоциональное, психофизиологическое и ольфакторное состояния и другие варианты использования многомодальной информации.

На степень реализации информационных услуг (факторы 39–48) оказывают влияние качество обработки многомодальных сигналов (факторы 30–38), а также ограничения:

- на способы человеко-машинного взаимодействия со стороны многомодального интерфейса, связанные с его программно-аппаратными возможностями (причинно-следственные связи между факторами 9 и 30–38; 10 и 39–48);
- предоставляемых услуг, связанные с предметной областью (причинно-следственные связи между факторами 39–48).

Для описания взаимовлияния базисных факторов использованы лингвистические переменные [Хрусталева Е., Макаренко Д. 2008], на их основе формализована обобщенная матрица взаимовлияний базовых факторов.

Общая методика формирования множества допустимых системотехнических решений на основе концептуального моделирования

Разработанная концептуальная модель служит для выбора рациональных системотехнических решений по построению многомодальных интерфейсов абонентских терминалов.

Множество допустимых системотехнических решений определяется:

- требованиями пользователей к номенклатуре и качеству предоставляемых на основе многомодальных интерфейсов услуг (факторы 23–29 когнитивной модели реализации телекоммуникационных услуг и факторы 39–48 когнитивной модели реализации информационных услуг);
- доступными устройствами ввода/вывода сигналов различных модальностей (факторы 2–7 когнитивной модели);
- объемом и типом предоставляемых инфокоммуникационных услуг (факторы 8–10 когнитивной модели);
- математическими моделями, методами и алгоритмами, реализующими соответствующие технологии обработки сигналов различной модальности (факторы 11–22 когнитивной модели) и многомодальных сигналов (факторы 30–38 когнитивной модели);
- инцидентными им потоками естественных и искусственных сигналов,
- обеспечивающими требуемый коэффициент целедостижения и максимум эффективности управленческого решения.

Для выбора рациональных системотехнических решений по построению многомодальных интерфейсов человеко-машинного взаимодействия на основе когнитивного моделирования осуществляется следующая последовательность процедур.

Шаг 1. Определение требований пользователей к номенклатуре и качеству предоставляемых услуг. На данном шаге задается вектор начальных тенденций целевых базисных факторов $X_t = (x_1(t), x_2(t), \dots, x_n(t)); t = 0$. Для «неиспользуемых» услуг начальные тенденции соответствующих факторов приравниваются к нулю.

Шаг 2. Структурный анализ когнитивной карты, включающий в себя анализ целей функционирования на непротиворечивость, управлений (влияния управляющих факторов) на согласованность с целями и оценивание эффективности интегрального воздействия управляющих факторов на целевые на основе научно-методического инструментария, представленного в работе [Хрусталева Е., Макаренко Д. 2008]. Основу такого анализа составляют оценки динамики фактора (ОДФ) x_i , представленные для рассматриваемой концептуальной модели в таблице 1.

Структурный анализ разработанной концептуальной модели показал, что вектор целей (факторы 23–29, 39–48) не противоречив, вектор управлений (факторы 1–10) согласован с вектором целей, а наиболее сильным влиянием на целевые факторы обладают факторы 8–10.

Шаг 3. Сценарное моделирование развития ситуации заключается в моделировании возможных вариантов структур когнитивной карты, оценок динамики параметров базисных факторов и их значений с целью выявления совокупности факторов, способствующих достижению целей такого развития.

На данном шаге целесообразным является применение управляемого развития, предполагающего целенаправленное воздействие на один или несколько управляющих факторов 1–10 (табл. 1).

Шаг 4. Оценка эффективности проектных решений. Формирование множества допустимых системотехнических решений.

Для оценки эффективности принятых управленческих решений предлагается использовать коэффициент эффективности управленческого решения, характеризующий отношение степени



достижения целей к объему и ценности ресурсов, необходимых для реализации соответствующего управленческого решения:

$$e_s = \frac{f_s}{r_s},$$

где f_s – коэффициент благоприятности ситуации для сценария s ($s = 1, \dots, S$):

$$f_s = \frac{\sum_{k=1}^m g_k^l i_k}{\sum_{k=1}^m i_k};$$

$g_k, k=1, \dots, 10$ – коэффициент целедостижения:

$$g_k = \left| \frac{y_k^* - y_k^n}{y_k^* - y_k^h} \right|;$$

y_k^*, y_k^n, y_k^h – текущее, прогнозируемое и начальное значения целевого фактора соответственно; i_k – важность цели; S – количество вариантов решений (сценариев); r_s – ресурсоемкость управленческого решения для сценария s :

$$r_s = \frac{\sum_{i=1}^p u_i v_i}{\sum_{i=1}^p v_i},$$

где u_i – управляющее воздействие; v_i – ценность ресурса [7].

При изменении ситуации в метасистеме (ИКС) производится корректировка концептуальной модели и повторение шагов 2–4.

Пример реализации общей методики формирования множества допустимых системотехнических решений на основе концептуального моделирования

Предположим, что наряду с обменом сообщениями пользователей при коммуникативном взаимодействии посредством ИКС с многомодальными интерфейсами интересуют их истинность (ложность). Будем считать, что при осуществлении сеанса видеотелефонии пользователи способны полностью определить смысл передаваемой информации, но не имеют специальной подготовки для определения истинности передаваемой информации. Тогда для определения истинности в таких системах необходимо применение дополнительных аппаратно-программных средств (например, полиграфа). Использование последних неизбежно приведет к увеличению времени обработки входных сигналов и/или увеличению затрат на их реализацию.

При многомодальном представлении информации возможна оценка ложности сообщаемой информации по динамике параметров невербального поведения пользователя, представляющей собой одно из наиболее перспективных направлений в области теоретической и прикладной психологии. В работе [Basov et al., 2015] изложена сущность услуги «честная видеотелефония» с применением методов оценивания истинности передаваемых многомодальных потоков данных [Басов О.О., Носов М.В., Шалагинов В.А. 2014; Будков В.Ю., Савельев А.И., Вольф Д.А. 2016; Budkov V., Vatamaniuk I., Basov V., Volf D. 2016.].

Результаты применения общей методики формирования множества допустимых системотехнических решений (рис. 3) на основе когнитивного (концептуального) моделирования соответствующего многомодального интерфейса абонентского терминала взаимодействия показали, что по силе влияния управляющих факторов на целевые (см. табл. 1) их можно расположить в следующем порядке (от наиболее сильно влияющего к менее влияющему): 26–15–17–48–14–16–25–27–12–13–18–19–20 (рис. 4).

На основе указанных факторов определены основные задачи, реализуемые моделируемого многомодального интерфейса абонентского терминала и сформированы множества соответствующих алгоритмов их реализации [Басов О.О. 2015].

Заключение

Задача выбора рациональных системотехнических решений по построению многомодальных интерфейсов абонентских терминалов может быть решена с использованием разработанной концептуальной модели. Для устойчивого поиска допустимых элементов множества системотех-

нических решений применен анализ когнитивных карт, основанный на линейной алгебре. Ограничение области поиска и соответствующие этому снижение общей размерности задачи достигается за счет контроля из всей совокупности свойств системы только тех характеристик, которые максимально влияют на результаты функционирования (в соответствии с методом главных компонент) многомодальных интерфейсов. Результаты анализа концептуальной модели и моделирования управляемого развития ситуации служат начальными условиями для решения задач оптимизации структуры программно-математического и информационного обеспечения абонентских терминалов.

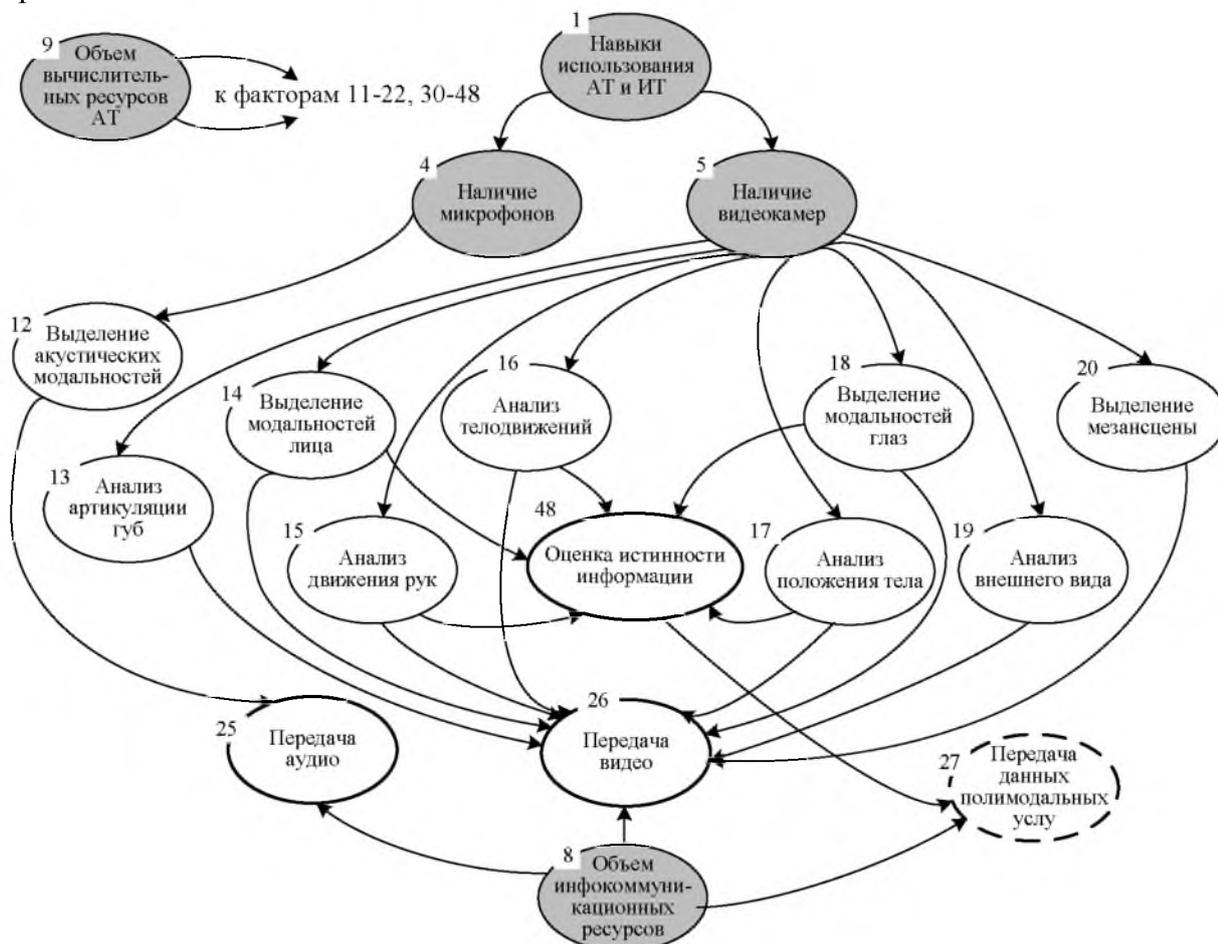


Рис. 3. Структура когнитивной модели реализации услуги «честная видеотелефония»
 Fig. 3. The structure of the cognitive model of realization "fair videotelephony"

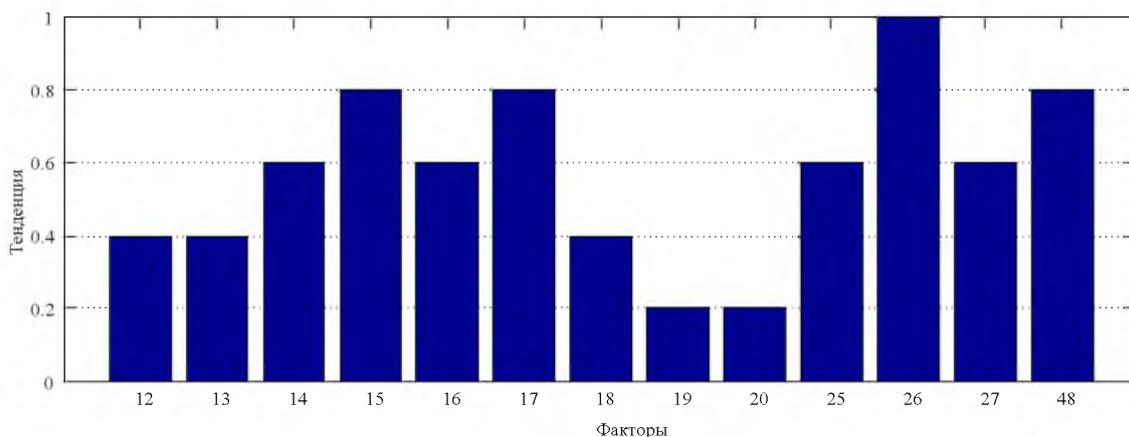


Рис. 4. Анализ эффективности интегрального воздействия управляющих факторов на целевые
 Fig. 4. Analysis of the effectiveness of the integral effect of the control factors on the target

Исследование выполнено при частичной поддержке гранта РФФИ (16-37-60085) и гранта Президента Российской Федерации (МК-7925.2016.9).



This work is partially supported by the Russian Foundation for Basic Research (project No. 16-37-60085) and by the Council for Grants of the President of Russia (project No. МК-7925.2016.9).

Список литературы References

- Басов О.О., Носов М.В., Шалагинов В.А. 2014. Исследование характеристик джиттера периода основного тона речевого сигнала. Труды СПИИРАН, 32: 27-44.
- Basov O.O., Nosov M.V., Shalaginov V.A. 2014. Issledovanie harakteristik dzhittera perioda osnovnogo tona rechevogo signala. Trudy SPIIRAN, 32: 27-44.
- Басов О.О. 2015. Принципы построения полимодальных инфокоммуникационных систем на основе многомодальных архитектур абонентских терминалов. Труды СПИИРАН, 39: 109-122.
- Basov O.O. 2015. Principy postroenija polimodal'nyh infokommunikacionnyh sistem na osnove mnogomodal'nyh arhitektur abonentskih terminalov. Trudy SPIIRAN, 39: 109-122.
- Будков В.Ю., Савельев А.И., Вольф Д.А. 2016. Методика исследования параметров речевого сигнала, отражающих истинность передаваемой информации. Доклады ТУСУРа, 19(2): 56-60.
- Budkov V.Ju., Savel'ev A.I., Vol'f D.A. 2016. Metodika issledovanija parametrov rechevogo signala, otrazhajushhih istinnost' peredavaemoj informacii. Doklady TUSURa, 19(2): 56-60.
- Островский М.А., Шевелев И.А. 2003. Физиология человека: учебник; в 2 т.; под ред. Покровского В.М., Коротко Г.Ф. М., Т. 2. 201-259.
- Ostrovskii M.A., Shevelev I.A. Fiziologiya cheloveka: uchebnik; v 2 t.; pod red. Pokrovskogo V.M., Korotko G.F. Moscow, T. 2. 201-259.
- Карпов А.А., Ронжин А.Л., Лобанов Б.М., Цирульник Л.И., Железны М. 2008. Разработка бимодальной системы аудиовизуального распознавания русской речи. Информационно-измерительные и управляющие системы. 10(6): 58-61.
- Karpov A.A., Ronzhin A.L., Lobanov B.M., Cirul'nik L.I., Zhelezny M. 2008. Razrabotka bimodal'noj sistemy audiovizual'nogo raspoznavanija russkoj rechi. Informacionno-izmeritel'nye i upravljajushhie sistemy. 10(6): 58-61.
- Ронжин А.Л., Карпов А.А. 2008. Сравнение методов локализации пользователя многомодальной системы по его речи. Известия высших учебных заведений. Приборостроение. 51(11): 41-47.
- Ronzhin A.L., Karpov A.A. 2008. Cravnenie metodov lokalizacii pol'zovatelja mnogomodal'noj sistemy po ego rechi. Izvestija vysshih uchebnyh zavedenij. Priborostroenie. 51(11): 41-47.
- Ронжин А.Л., Ли И.В., 2007. Автоматическое распознавание русской речи. Вестник Российской академии наук. 77(2): 133-138.
- Ronzhin A.L., Li I.V., 2007. Avtomaticheskoe raspoznavanie russkoj rechi. Vestnik Rossijskoj akademii nauk. 77(2): 133-138.
- Савченко В.В., Васильев Р.А. 2014. Анализ эмоционального состояния диктора по голосу на основе фонетического детектора лжи. Научные ведомости БелГУ. Сер. История. Политология. Экономика. Информатика. 21(192): 186-195.
- Savchenko V.V., Vasiliev R.A. 2014. Analiz jemocional'nogo sostojanija diktora po golosu na osnove foneticheskogo detektora lzhi. Nauchnye vedomosti BelGU. Ser. Istorija. Politologija. Jekonomika. Informatika [The analysis of the emotional condition of the announcer on the voice on the basis of the phonetic lie detector] 21(192): 186-195.
- Хрусталева Е.Ю., Макаренко Д.И. 2008. Концептуальное моделирование военной безопасности государства. М.: Наука: 303.
- Hrustalev E.Yu., Makarenko D.I. 2008. Konceptualnoe modelirovanie voennoi bezopasnosti gosudarstva. M.: Nauka: 303.
- Basov O.O., Ronzhin A.L., Budkov V.Yu., Saitov I.A. 2015. Method of Defining Multimodel Information Falsity for Smart Telecommunication Systems. 8th Conference ruSMART 2015. Springer International Publishing Switzerland. S. Balandin et al. (Eds.): ruSMART 2015, LNCS 9247. 163-176.
- Budkov V., Vatamaniuk I., Basov V., Volf D. 2016. Investigation of Speech Signal Parameters Reflecting the Truth of Transmitted Information. Speech and Computer, 18 International Conference SPECOM 2016, Budapest, Hungary, August 23-27, 2016, Springer International Publishing, LNAI 9811. 419-426.
- Oviatt S.L. Multimodal interfaces. 2003. The Human-Computer Interaction Handbook: Fundamentals. Evolving Technologies and Emerging Applications. Mahwah, 1(14): 286-304.