

ОРДЕНА ЛЕНИНА
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
им. М.В.Келдыша РАН

С.Л. Головков, Н.Е. Каленов

Принципы построения системы удаленного доступа к
библиотечным ресурсам на базе сотовой связи

Москва
2003

С.Л. Головков, Е.Н. Каленов

Принципы построения системы удаленного доступа к библиотечным ресурсам на базе сотовой связи. ♦

Аннотация

Описываются принципы построения системы доступа к информационным ресурсам с использованием мобильных средств связи. Рассматриваются архитектура системы доступа, базирующейся на объединении интернет-технологий и сотовой связи, необходимые инструментальные средства, базовые функции и прототип интерфейса пользователя.

S.L. Golovkov, E.N. Kalenov

The Principles of the building of the remote access system to library resources on the base cellular network.

Abstract

The principles of the building of the access system to information resources with use the mobile communications is described. System architecture, basing on association internet technology and cellular network, necessary tools, base functions and prototype of the user interface are considered.

♦ Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных Исследований (проект N 03-07-90170)

Содержание

1. Введение	3
2. Постановка задачи	4
Промежуточное хранилище данных	6
Интерфейс	7
3. Прототип системы	7
Выбор инструментальных средств	8
Базовые функции	9
Прототип интерфейса пользователя	10
4. Заключение	13

Литература

1. Введение

В последние годы крайняя ограниченность финансирования российских библиотек вызвала необходимость решения проблемы организации оперативного доступа широкого научного сообщества к распределенному фонду отечественной и зарубежной научно-технической информации. Доступ к электронным библиотечным ресурсам обычно предполагает наличие стационарного рабочего места, оснащенного компьютером, подключенным к локальной сети библиотеки либо к сети Интернет.

Это обстоятельство ограничивает научную деятельность, поскольку такая возможность не вполне и не всегда доступна для массового пользователя, который к тому же не всегда уверенно применяет компьютер в своей работе. Предоставление пользователю мобильных средств доступа, базирующихся на объединении **интернет-технологий** и **сотовой связи**, позволяет снять эти ограничения. С одной стороны, существенно расширяется доступность электронных информационных ресурсов, с другой стороны, психологический барьер при использовании сотового телефона существенно ниже, чем при освоении компьютера.

Перспективность подобного подхода обусловлена тем, что цифровые мобильные средства связи совершили в последнее время революционный прорыв на массовый рынок, и нет сомнения в том, что они станут привычным инструментом современного ученого и специалиста. С теоретической точки зрения актуальность данного проекта заключается в его направленности на изучение новых форм и методов работы с информационными ресурсами, на исследование и решение проблем управления информацией в средах с мобильными средствами связи.

Проект, осуществляемый совместными усилиями Института прикладной математики им. М.В.Келдыша РАН и Библиотеки по естественным наукам

РАН, направлен на создание системы, обеспечивающей оперативный доступ к удаленным информационным ресурсам научной библиотеки (каталогам, справочным и реферативным базам данных, электронным коллекциям, библиографическим массивам и другим библиотечным ресурсам) с использованием современных средств беспроводной связи. Эти средства позволяют обеспечить существенное расширение возможностей доступа к информационному полю библиотек, используя разнообразные аппаратно-программные платформы, включая переносные или карманные компьютеры и сотовый телефон.

2. Постановка задачи

В общем виде модель, описывающая организацию мобильного удаленного доступа к библиотечным информационным ресурсам, представлена на рис.1

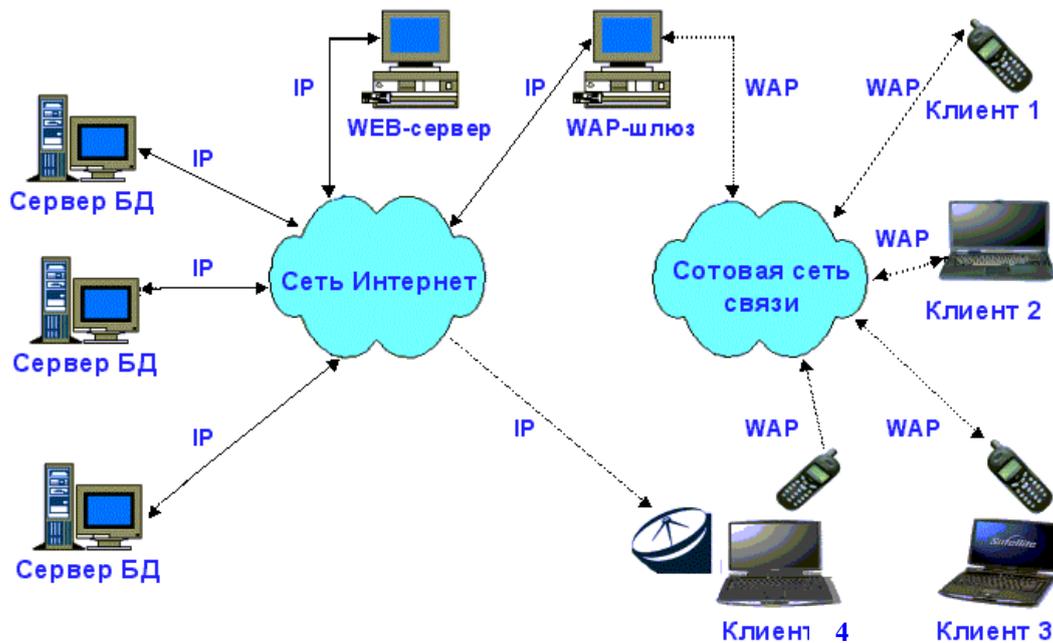


Рис.1. Архитектура системы доступа

WAP-шлюз обеспечивает сопряжение сотовой сети с сетью Интернет. Шлюз работает с DNS, кодирует и декодирует информацию, преобразует пакеты из WTP (Wireless Transport Protocol) в TCP. Шлюз также поддерживает безопасность передачи информации, используя стандарт WTLS (Wireless Transport Layer Security).

На WEB-сервере размещаются сервисы, которые, с одной стороны, генерируют WML- и HTML-страницы, через которые пользователь

осуществляет взаимодействие с системой, с другой стороны, обеспечивают доступ к серверам баз данных.

Функционирование системы происходит следующим образом:

- запросы пользователей поступают от WML-браузеров клиентов и проходят через сотовую сеть связи и WAP-шлюз;
- производится обработка полученной информации на WEB-сервере и преобразование ее в последовательность обращений к серверам БД через сеть Интернет;
- осуществляется собственно обращение к серверу БД за результатами;
- формируются блоки данных от серверов БД на WEB-сервере и производится вывод данных в виде WML-страниц на браузеры клиентов.

Для того, чтобы подчеркнуть то обстоятельство, что обслуживание клиентов осуществляется с использованием WAP-протокола (Wireless Application Protocol), будем повсеместно использовать префикс «WAP»: WAP-сервер, WAP-сервис, WAP-сайт и т.д.

Взаимодействовать с сервером библиотечных баз данных через сотовую сеть связи могут различные типы клиентов. На рис.1 изображены следующие типы клиентов:

- Клиент 1. Сотовый телефон, имеющий WAP-браузер, используется в качестве оконечного клиентского устройства. Транспортный протокол может обеспечивать как связь точка-точка, так и пакетную передачу на основе протокола GPRS.
- Клиент 2. Оконечным устройством является переносной или стационарный компьютер с сотовым модемом. В этом случае для клиента сохраняются все средства полноэкранного браузера, а сотовая сеть используется лишь в качестве транспортной среды.
- Клиент 3. Совместное использование компьютера (переносного или стационарного) и сотового телефона, имеющего средства синхронизации с компьютером - соединение с компьютером через инфракрасный порт, последовательный порт или USB-порт.
- Клиент 4. Если в качестве оконечного терминала используется переносной или стационарный компьютер, то его связь с сервером может быть ассиметричной: исходящий поток направляется через сотовую сеть, а входящий поступает через коммуникационный спутник - обычно это спутник цифрового телевидения. Такой способ подключения может оказаться удобным в тех случаях, когда ответ сервера содержит большой объем информации.

При построении системы доступа к библиотечным информационным ресурсам на базе сотовой связи возникает ряд проблем. На первом этапе мы выделили две из них.

Промежуточное хранилище данных

Сотовая связь является ненадежным и дорогостоящим (во-всяком случае, в настоящее время) каналом коммуникаций. Поэтому неразумно и неэффективно заставлять пользователя ждать ответа системы при сложных запросах к БД (при сложном поиске или при поиске в распределенной базе данных). Для таких случаев можно предложить пользователю коммуникацию, состоящую из двух разделенных по времени этапов: на первом этапе пользователь отправляет запрос на WAP-сервер, на втором этапе пользователь обращается к WAP-серверу за готовым результатом.

Для реализации такого варианта взаимодействия необходимо включить в состав системы доступа **промежуточное хранилище данных**, в которое WAP-сервер помещает готовые для отправки клиенту данные, полученные от сервера БД (для краткости будем называть это хранилище **Архивом**).

Логическое и физическое представление такого Архива является открытым вопросом.

Например, одним из решений является хранение в Архиве готовых данных с помощью XML-представления. Поскольку XML позиционируется как универсальное средство для обмена данными между приложениями и в последнее время становится распространенной практикой хранения готовых данных, такое представление может облегчить задачу обеспечения доступа к одним и тем же данным со стороны самых различных программно-аппаратных платформ.

Другим аспектом построения Архива является его логическая структура. Можно, например, считать, что он является «почтовым ящиком до востребования». Это означает, что в нем хранится заказанная, но еще не отправленная пользователю информация. Кроме того, можно хранить в Архиве готовые ответы для некоторого множества наиболее типичных запросов и даже периодически обновлять их, не дожидаясь запроса пользователя. При этом можно персонализировать Архив – формировать ответы на типичные запросы для каждого пользователя, реализовывая тем самым принятую в информационно-библиотечной технологии систему избирательного распространения информации (ИРИ). Возможность индивидуальной настройки на каждого отдельного пользователя особенно важна в условиях беспроводной связи, так как позволяет сократить объем передаваемой информации.

При выборе архитектуры Архива следует учитывать возможность использования особенности сотового оборудования, называемой "активным приложением" или "WAP-подталкиванием" (WAP Push).

Современные WAP-устройства могут работать в двух режимах: в режиме запроса и в режиме подталкивания. В первом случае пользователь посылает запрос на услугу, и сервер сразу выдает ему на браузер ответ. В режиме подталкивания (push) сервер выдает на аппарат пользователя по своей инициативе сигнал с некоторым уведомлением, в ответ на который пользователь по своему усмотрению может выполнить нужные ему действия.

Этот последний вариант хорошо подходит для случаев продолжительной реакции на запрос к БД. Пользователь заказывает некоторую услугу, выполнение которой может потребовать значительного времени, и отключается от сервера, переходя к другим делам. Сервер, после выполнения услуги, извещает об этом пользователя. Другой сервис на основе этой технологии может состоять в «активном» информировании пользователя о важных для него изменениях в библиотечном ресурсе, например, о наступлении срока возврата в библиотеку книги или журнала.

Интерфейс

Следующее направление работы – разработка концепции интерфейса пользователя. В этом направлении можно выделить следующие «азимуты».

- **Поиск метафоры.** Для мобильных устройств не существует общепринятой метафоры представления контента (как, например, метафоры «рабочего стола» или «бесконечной прокрутки» для иных типов устройств). Также нет общепринятой концепции навигации, представления карты WAP-сайта и т.п.
- **Единообразие презентации контента.** Весьма желательно достижение визуальной и навигационной корреляции между представлением информации на разных типах мобильных и стационарных устройств (характеризуемых разными размерами экрана, графическими возможностями и т.д.). Иначе говоря, речь идет о стандартах представления информации для разных типов клиентов - пользователей библиотек.
- **Способы вывода информации.** Интересно рассмотреть возможность реализации альтернативных (по сравнению с набором WML-страниц) способов вывода информации на мобильное устройство. Например, речевого вывода или отправки SMS-сообщений. В будущем может идти речь о передаче мультимедийных сообщений (Multimedia Message Service - MMS). Эта служба поддерживает передачу мультимедийных сообщений, содержащих картинки, музыку, графику и пр. Ожидается, что MMS будет одним из ключевых сервисов, которые определяют потребность WAP- и GPRS-технологий для широкого круга пользователей.
- **Модель взаимодействия «клиент–сервер».** Следует обратить внимание на модель "клиент-сервер", в которой претерпевает изменение парадигма взаимодействия клиента и сервера. Сотовый телефон - это самый «тонкий клиент» из всех известных прежде, обеспечивающих визуальное отображение. Рабочее место пользователя представляет собой простейшее универсальное устройство. Фактически, это терминал, снабженный только программой навигации. Вся потребляемая информация порождается на WAP-сервере.

3. Прототип системы

Для решения поставленных задач было необходимо выбрать модельную задачу, на примере которой можно было бы опробовать основные

архитектурные концепции, выбрать и опробовать подходящий инструментарий, выявить проблемы и определить пути их решения.

В качестве модельной задачи была выбрана задача поиска описаний публикаций в каталоге. Задача формулируется следующим образом: пользователь формулирует запрос к базе данных, пользуясь словарями (фамилии, заглавия, фрагменты библиографического описания, года публикаций и т.п.), система отыскивает в удаленном электронном каталоге публикации, соответствующие запросу, и возвращает описания публикаций пользователю.

Выбор этой задачи обусловлен следующим:

- она является типовой задачей работы с библиотечными базами данных;
- задача охватывает основные технологические этапы доступа к удаленной библиотечной базе данных (работа с пользователем, работа с базой данных, коммуникации).

Решение задачи позволяет достичь следующие цели:

- выбрать необходимые инструментальные средства;
- разработать и реализовать базовые функции взаимодействия пользователя с библиотечными базами данных БЕН РАН (включающими сводные каталоги более 100 библиотек институтов РАН, БД публикаций сотрудников многих институтов и т.п.) через сотовый телефон;
- разработать прототип мобильного интерфейса.

Выбор инструментальных средств

При выборе инструментальных средств использовались следующие критерии:

- средства должны быть не только стандартные, но и широко распространенные и, следовательно, поддерживаемые научным сообществом;
- средства должны быть свободно распространяемыми.

В итоге остановились на следующей технологической цепочке инструментов:

- WML, WMLScript
Языки программирования мобильного клиента.
- Apache
Пакет, реализующий WEB- и WAP-серверы.
- Perl
Язык программирования CGI-скриптов, реализующих WAP-сервисы.
- DBI/DBD
Библиотеки, обеспечивающие интерфейс между языком Perl и СУБД MySQL. Библиотека DBI (DataBase Interface) обеспечивает интерфейс между языком Perl и абстрактной базой данных (иначе говоря, является посредником между программой и всеми драйверами, предназначенными

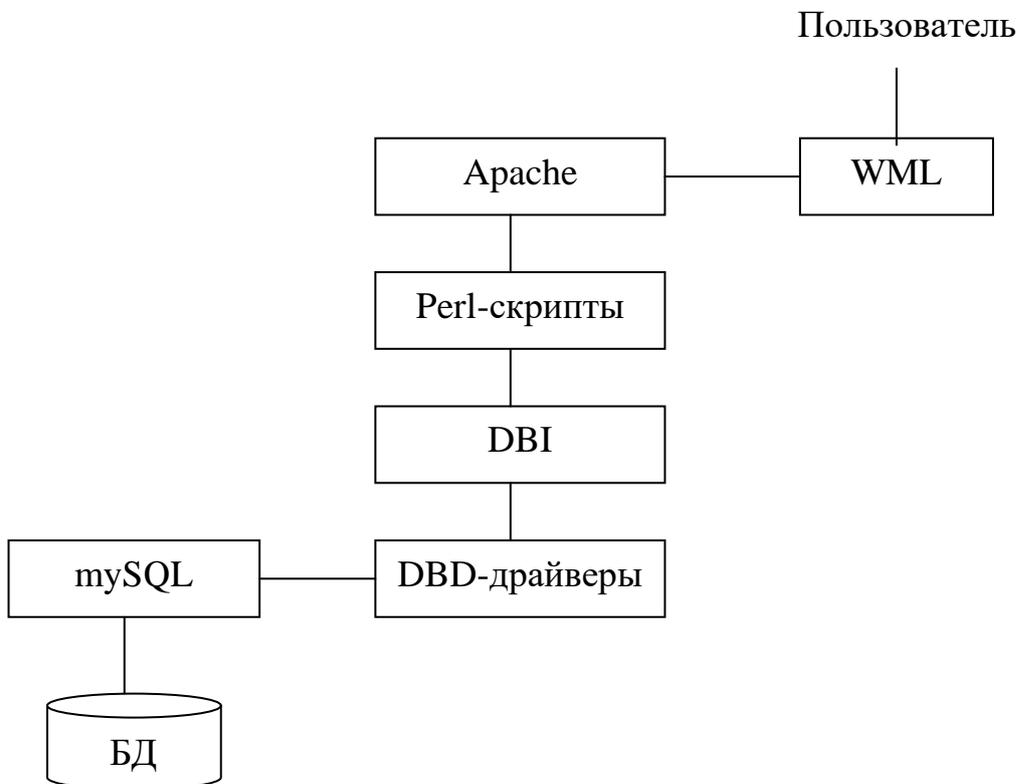
для работы с конкретными СУБД). Библиотека DBD (DataBase Driver) содержит драйверы для работы с конкретными СУБД. В частности, библиотека DBD::mysql обеспечивает работу с базой данных, управляемой СУБД MySQL.

- **mysql**

Широко распространенная реляционная СУБД, обладающая следующими важными для решения поставленной задачи свойствами:

- многопоточность: поддержка нескольких одновременных запросов;
- записи фиксированной и переменной длины;
- гибкая система привилегий и паролей;
- интерфейс с языками C и perl.

Указанные выше инструментальные средства можно условно выстроить в следующую цепочку от пользователя до базы данных:



Следует также отметить, что для отладки WAP-сервисов использовались имитаторы мобильных телефонов, в частности, Deck-It и M3Gate.

Базовые функции

Выбор базовых функций во многом определяется структурой базы данных, хранящей библиографическую информацию. Поэтому потребовалось создание «эталонной» (или «тестовой») базы данных – функционального эквивалента «реальной» библиотечной базы данных, отличающегося от последней сравнительно небольшим объемом и полной автономностью.

Созданная «эталонная база данных» (доступная по адресу <http://testbooks.benran.ru>) представляет собой выборку из «реального» электронного каталога книг, содержащую описание 201 наименования книг (на русском и английском языках).

Для создания «эталонной базы данных» пришлось реализовать конвертер – программу, выполняющую выборку содержательной информации из «реальной» библиотечной базы данных, преобразование информации и загрузку ее в «эталонную» базу данных.

Конвертер реализован в виде набора независимых программ, каждая из которых решает частную задачу конвертации. Такой подход позволяет контролировать каждый этап конвертации и «собирать» любую требуемую конфигурацию конвертера.

Конвертация компонент базы данных проводится в два этапа, с промежуточным представлением компонент в виде текстовых файлов (xml-документов).

Выбор базовых функций основывался на типовом сценарии работы пользователя при доступе к «эталонной» библиотечной базе данных (в рамках выбранной нами типовой задачи). В общем случае такой сценарий может состоять из следующих шагов:

1. Выбор словаря.
2. Просмотр словаря и, возможно, пометка (выбор) каких-то элементов словаря.
3. Возможно, повторение действий 1-2 с другими словарями.
4. Просмотр отобранных элементов словарей и построение из них «формулы поиска», т.е. запроса к базе данных.
5. Запуск поиска.
6. Просмотр результата. Возможно, возврат в пункт 4.

Декомпозиция каждого пункта сценария на элементарные типовые действия позволяет определить базовый набор функций. Например: «считывание фрагмента базы данных в Архив», «пометка элементов Архива» и т.п.

Прототип интерфейса пользователя

При проектировании интерфейса пользователя, в первую очередь, внимание уделялось следующему важному обстоятельству.

Мобильный телефон (как целевое клиентское устройство) отличается небольшим размером экрана и довольно ограниченными возможностями управления. Вследствие этого желательно обеспечить:

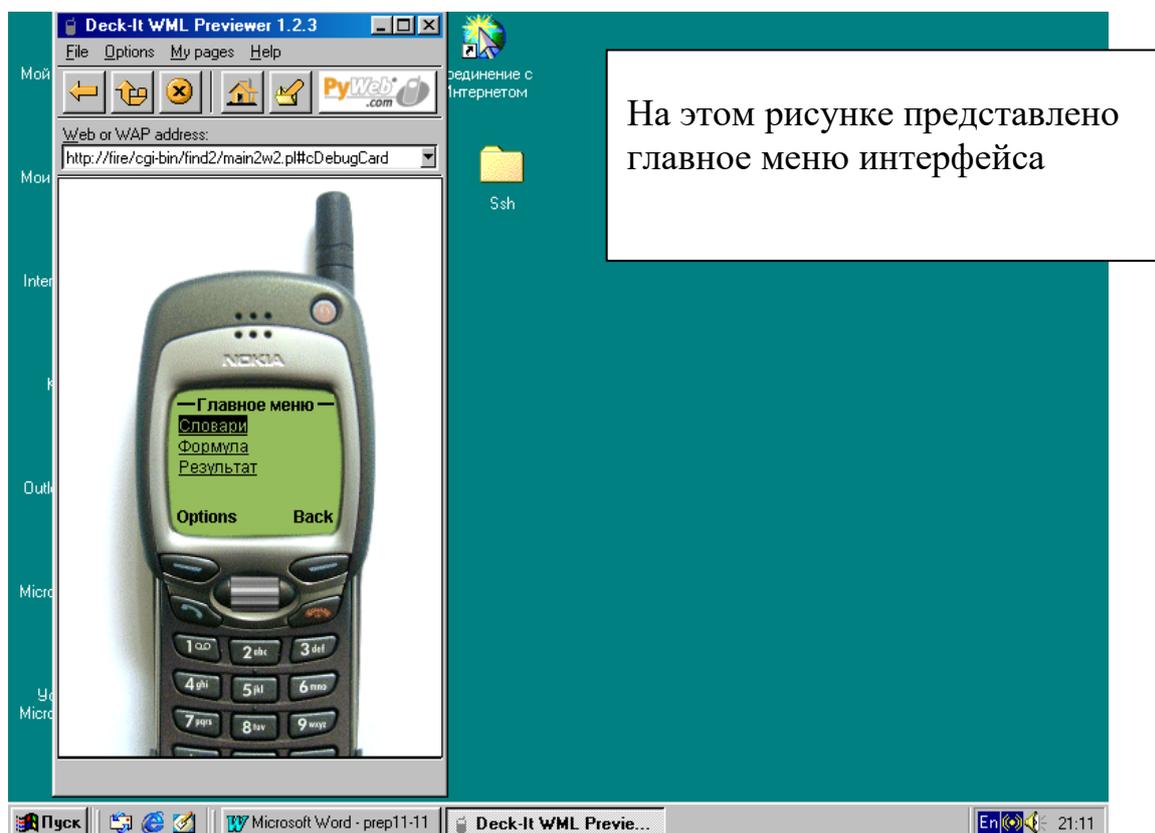
- возможность продолжить прерванный сеанс связи с того же места, на котором он был прерван, чтобы не повторять уже проделанную работу;
- предельно простую навигацию по экранным формам, идеально – каждый раз иметь только два направления дальнейшего «движения»;
- минимизацию объема информации, передаваемой между клиентом и сервером.

Учитывая эти особенности, решено хранить все характеристики текущего состояния интерфейса (всех карт интерфейса) каждого пользователя в Архиве на WAP-сервере.

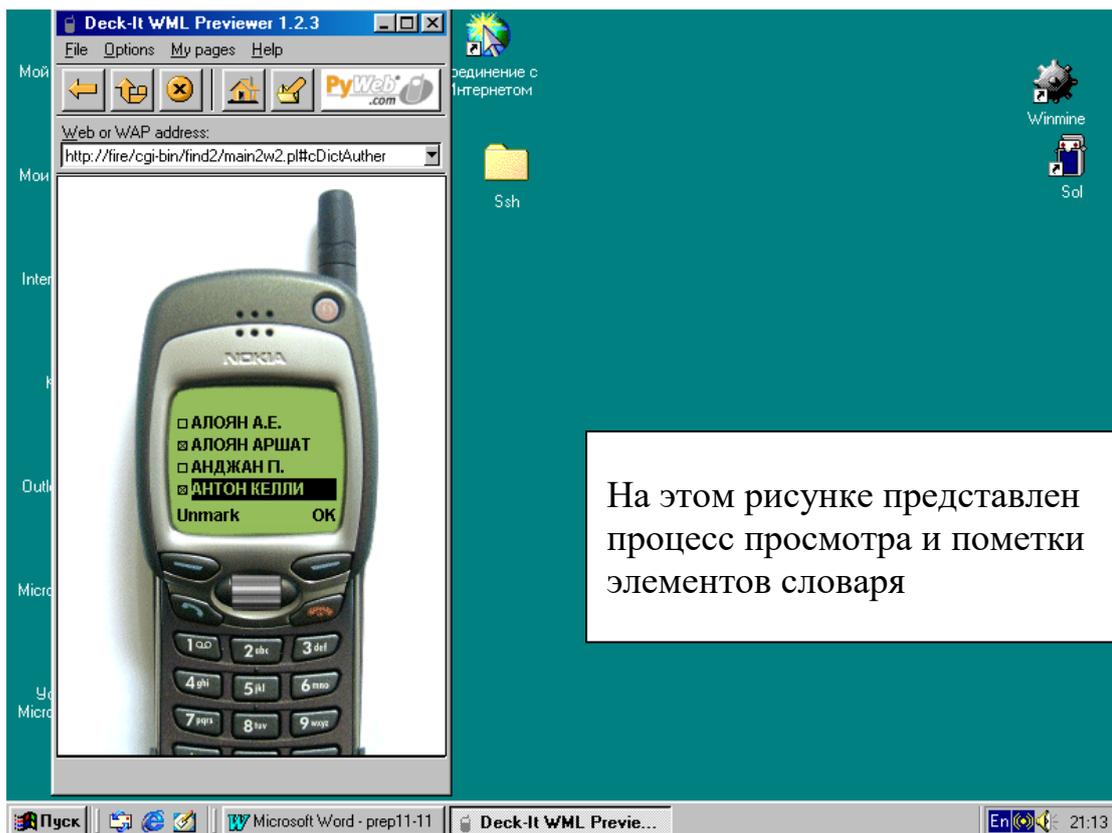
При таком решении структура CGI-скриптов, обслуживающих клиента, становится логически очень простой:

- считывание из Архива состояния интерфейса, соответствующего последней предъявленной данному пользователю карте;
- модификация состояния интерфейса в соответствии с присланной пользователем командой;
- сохранение нового состояния интерфейса в Архиве;
- предъявление пользователю очередной карты, соответствующей новому состоянию интерфейса.

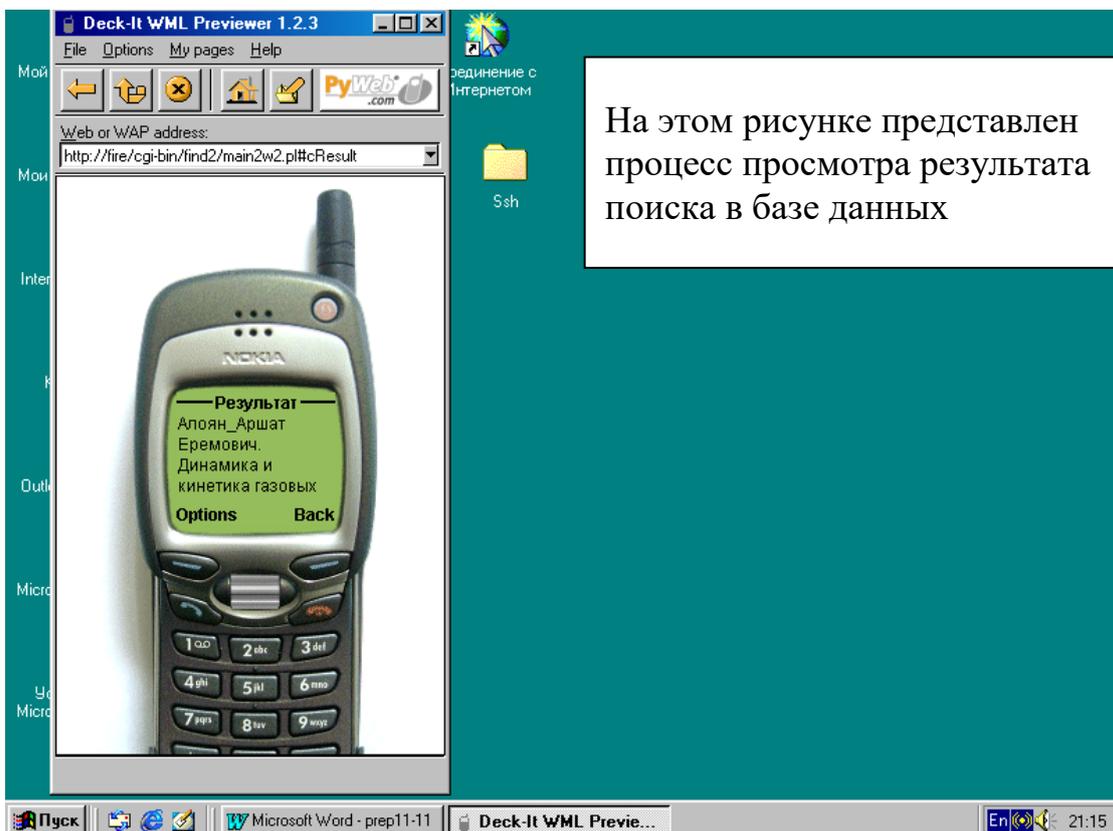
Ниже приведены несколько примеров, иллюстрирующих визуальное представление карт интерфейса.



Представим, что пользователь только что соединился с библиотечной базой данных. Однако состояние его интерфейса, считанное из Архива, соответствует моменту завершения предыдущего сеанса: в словарях сохранились пометки, можно посмотреть результат последнего поиска и т.д.



Отобранные элементы словаря попадают в Архив. С помощью пункта меню «Формула» из этих отобранных элементов можно строить запрос к базе данных. Результат поиска представлен на нижнем рисунке.



4. Заключение

Технологии мобильного доступа к информационным ресурсам сравнительно молоды. В частности, только к середине 2001 года на российском рынке впервые появились телефонные аппараты с протоколом цифровой передачи данных - GPRS, который представляется наиболее эффективным и удобным для использования при организации мобильного взаимодействия с удаленной базой данных. Только в 2002 году завершено создание инфраструктуры для GPRS-телефонии, а мобильная телефонная связь с протоколом WAP доступна лишь с 2001 года.

С другой стороны, как в мире, так и в России быстро расширяется рынок услуг по предоставлению доступа к удаленным базам данных с помощью мобильных средств связи. Объединяются две наиболее быстро растущие и развивающиеся отрасли - сотовая связь и Интернет. Все это определяет значительную актуальность работ в указанной сфере.

Поэтому в настоящем препринте авторы хотели не только изложить свое представление о том, как должна или может быть построена система мобильного доступа к научно-технической информации в базах данных БЕН РАН, но и привлечь внимание к этой перспективной области исследований.

Литература

1. Головков С.Л., Рубин А.Г., Смирнов В.К.
Мобильные средства поддержки научных исследований в среде Интернет. Всероссийская научная конференция "Научный сервис в сети Интернет", Тезисы докладов, Новороссийск, Изд-во Московского университета, 2002.
2. Власова С.А., Каленов Н.Е., Соловьева Т.Н.
Информационные ресурсы библиотеки по естественным наукам РАН в сети Интернет // 275 лет на службе науке: библиотеки и институты информации в системе РАН: Сб. науч. тр. - М., 2000. - с. 80 - 85.
3. Каленов Н.Е.
Формирование и использование электронных ресурсов в системе БЕН РАН // Новые технологии в информационном обеспечении науки. Сб. статей. - М.: Биоинформсервис, 2001. - с. 7-12