英単語学習を支援するモバイル用ＷＷＷテキストブラウザ

ワサカ・ヴィスーティヴィセット、並木 美太郎
東京農工大学大学院 電子情報工学専攻 情報工学講座

本研究では，持ち運びやすいという PDA の利点を活かすために，英語能力を高めるための英単語学習を支援するモバイル WWW テキストブラウザを設計した．主な機能は普通のホームページを閲覧する機能のほかに，英文ホームページにおける英単語の意味説明の表示機能，英単語復習機能の三つがある．英文ホームページを閲覧する場合，ユーザは分からない単語をクリックし，その単語意味説明を表示させる．単語の意味説明はユーザのサーバ上にある辞書からまたはインターネットから取り出す．このため，本ブラウザはプロキシーを利用し，英文ホームページに単語説明の情報を加える．また，ユーザが英文ホームページを閲覧した後，システムは自動的に復習問題を生成し，ユーザの学習レベルに合わせて学習を行なう．学習結果はユーザの履歴としてプロキシー上に保存される．また，ユーザの能力に合わせるために，次回の復習問題を生成するとき，この履歴を参考して問題を生成する．

WWW Text Browser for Mobile Computing with English Vocabulary Learning Facility

Vasaka Visoottiviseth and Mitaro Namiki
Tokyo University of Agriculture and Technology

In this paper, we try to make mobile device more useful by developing a mobile application on PDA, a kind of mobile devices. Not only users can browse home page from their PDA easily, our application also supports English vocabulary learning. We choose English home page on the Internet to be our learning materials, since it has many interesting areas of information. We dissolve the problem of poor performance of PDA by using a proxy server. This proxy server will prepare the learning materials by getting the translation of vocabulary from dictionary both on server and on the Internet. In addition, it will create an exercise and feedback the result of this exercise to make more appropriate exercise in the next time.

1. INTRODUCTION

Mobile computing is an interesting research topics in today-communication. In this paper, we focus on personal digital assistant (PDA), which is one of the small mobile devices. We chose a PDA called Palm Pilot as our target platform. We try to make most use of mobile device's characteristics by developing a mobile application with English vocabulary learning facility. While English is spoken widely in the world, it is essential for us, Japanese students, to know English well in order to communicate with people from other countries and can understand their cultures. In this paper, we designed a mobile WWW text-browser on Palm Pilot, a popular PDA for today.

Since vocabulary is the basic in learning language, we chose to help users to increase their English vocabulary efficiency. To help users learn English vocabulary, our browser provides some of the information of the vocabulary, such as its meaning both in English and in Japanese and example of sentences, when users view an English home page.

We get English vocabulary from free English-Japanese dictionary and Japanese-English dictionary, which we put it on the server. In the case that there are some difficult vocabulary that are not included in our dictionary, our system will search their meaning from on-line dictionary in the Internet.

If no information is found, we provide a facilities for user to add its meaning to the dictionary on the server. All of the vocabulary which all users added will be included and merged to our dictionary. In addition, we will provide exercises after users learned the vocabulary. The result of this exercise will be feedback to our program and be used in making the next exercise problems. The exercise and its results will be created in the form of home page,

so that users can view and exercise it wherever they want.

In this paper, we will describe our policy in design of our WWW text-browser that is appropriate to mobile computing.

2. MOBILE COMPUTING WITH PDA

In order to see what is the problem in using PDA in mobile computing, in this section, we will discuss on the characteristics of PDA and problems in presence mobile computing.

2.1 The Characteristics of PDA

PDA has advantage and disadvantage has the following:

(1) Can be used every time and everywhere

Because of its small size and its light weight, it can be used everywhere and every time users want. Many PDA today are so small that they can enter shirt pocket, and some of them weigh only 5 ounces.

(2) Not interactive

Because of its small size, its display is too small to view home page like on desktop machine and can display only little amount of data. Moreover, most of their displays are black/white that they can not have an interactive presentation.

(3) Poor computational performance

Poor CPU and can put only few capacity of RAM. PDA can not deal with a program that needs a lot of memory and can not maintained a large amount of data.

(4) Poor input pen interface

Many PDA use pen device to input their data. For many users, using a pen is slower and more difficult than using a keyboard.

2.2 Problems in Mobile Computing

The two big problems in mobile computing is caused by its characteristics of nomadic connection and low bandwidth network :

(1) Nomadic Connection

In mobile computing, the way always used to connect the network is by using cellular phone, or other PPP connection. Because of this cellular phone, the connection can be terminated easily during transferring of data

(2) Low-Bandwidth Network

By using low-bandwidth network likes PPP connection, it can not transmit a big file such as movie's data in real time. Transmitting multimedia data will in low-bandwidth network will cause a big lag time.

3. DESIGN GOAL

Our design goals are the four following issues:

3.1 Make use from the Characteristics of PDA

Because of the portability of PDA, users can bring it with them everywhere and every time they want. And because of its small display screen, it can not display a lot of information. By these characteristics, the appropriate software for PDA should be something like memory-training program, such as learning vocabulary by oneself.

In addition, most PDA use pen interface to input the data. Even it is more difficult for some users to input by pen more than by keyboard, pen is appropriate for learning vocabulary. Many users found that they can memorize vocabulary well after they wrote them many times.

3.2 Improve Mobile Computing that use PDA

As we described, even though PDA has an advantage that can be brought with everywhere and every time, it has a poor computational performance that can not deal with a large amount of data. And by using low bandwidth network likes PPP connection, it can not have an interactive performance.

In this paper, we attempt to improve the poor performance of PDA by using proxy server that will be a mediator between the Internet and PDA.

3.3 Increase English Ability

Most computers in the world speak by using English. And most information maintained in the world is English. English is not only for British or American people, but it became the international language which joins all of the people around the world. As we saw the necessary for students to get familiar with English, we decided to create an English-vocabulary learning software.

Our target of users are low-intermediate users to

low-advance user, who knows some English grammars but has poor vocabulary, such as high school students to graduate students.

3.4 Make use from Resources on the Internet

As we described that most information on the Internet nowadays is maintained in English, we tried to make use of them for our learning facility. The information on Internet does not like the information in ordinary English text book. There are many kinds and many interesting areas of information, such as news, technical reports, novels, advertising, music, etc. Learners always can concentrate on the topics they are interested. Each learner has different interest, but all of them can choose which topic they want to learn from the Internet. Accordingly, we will use these home pages to be our learning materials.

4. HOW TO INCREASE ENGLISH READING ABILITY

In this paper, we focus on learning English by reading. Many home pages on the Internet are written in English, and there are several topics for users to choose. In learning English, we have to read, comprehend its meaning, and increase our vocabulary in the same time. For English learners, we divide reading into two types as we will describe here.

4.1 Read for Comprehension

The best way to learn reading for comprehension is to guess the meaning of the new vocabulary. [1], [2] Even there are some vocabulary that learners do not know, in many cases, learners can understand the meaning of the contexts. The importance for comprehensive reading is learn to guess and grasp the meaning of the entire context. There are certain key words that a learner must understand the text, but skipping some other words will now prevent him/her from understanding a sentence or a passage reasonably well.

4.2 Read for More Vocabulary

A good way to increase vocabulary is to read. The more a user read, the larger his/her vocabulary

becomes.[3] Reading the same vocabulary in many contexts will help user get more familiar with that vocabulary and understand its meaning well. It is important for users to choose their interesting topics for reading. If users choose that they enjoy, they can concentrate on reading and can finish reading them.
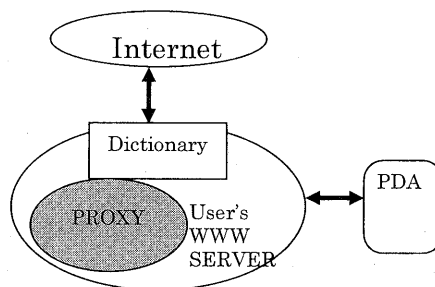
5. SYSTEM ARCHITECTURE

Fig. 1   Entire Systems and its Components

In order to success our goal, we designed a WWW text browser on Palm Pilot with English vocabulary learning facility. This browser can view home pages like the ordinary text browser. With its facility, it will help user improve their English reading skills by reading English home page. Our system support both reading for comprehension and reading for more vocabulary. We also prepare exercise after user has read a home page.

To improve the poor computational performance of PDA, we also designed a proxy server. This proxy server will be put on user's WWW server and will be a mediator between Palm Pilot and the Internet. We assume that user use cellular phone to connect proxy server and PDA. We will prepare English learning materials on this server, so that users can download English home pages they want to read immediately.

In preparing learning materials we need English-Japanese dictionary to tell us the meaning of vocabulary that included in English home page. Accordingly, we also put this dictionary on user's WWW server.

Our system is consist of components like in **Figure 1**.

6. WWW TEXT BROWSER FOR MOBILE COMPUTING

Since PDA has a low advance display screen, it can not display image file as beautiful as the desktop machine does. And most of them have so small screen that can not display all the information in one screen. Although image file may be important for understanding the details of a documents, browsing image file from such a small screen can not help user's comprehension much. Accordingly, in our attempt to develop this browser, we chose to display only text data.
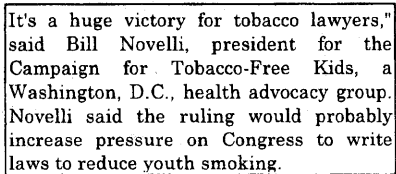
It have the following three ways or modes that users can choose to view a home page:

(1) Normal Mode

This mode is used when user want to view home page as they look from other browser. We do not distinguish between English and Japanese, and will not show any meaning of the vocabulary. On the other hand, user can click any links in a home page, that the other two modes can not do.
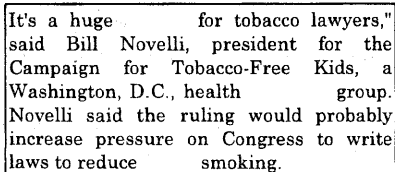
(2) Comprehension Mode

This mode will support reading for comprehension, when we found that the home page user wants to view is written in English. To practice reading for comprehension, a user has to learn to figure out the meaning of the entire context. When s/he meet unknown word, they should skip it and learn the meaning of the context from the rest of the sentences. Delete some difficult vocabulary user doesn't know and grasp only the main idea of the context.

It's a huge victory for tobacco lawyers," said Bill Novelli, president for the Campaign for Tobacco-Free Kids, a Washington, D.C., health advocacy group. Novelli said the ruling would probably increase pressure on Congress to write laws to reduce youth smoking.

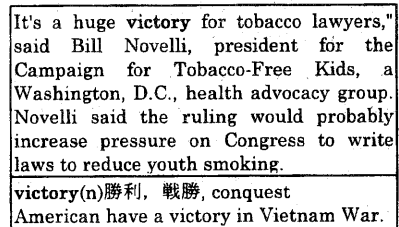Fig. 2   View in Normal Mode

It's a huge   for tobacco lawyers," said Bill Novelli, president for the Campaign for Tobacco-Free Kids, a Washington, D.C., health   group. Novelli said the ruling would probably increase pressure on Congress to write laws to reduce   smoking.

Fig. 3   View in Comprehension mode

When user click to the unknown word, we will hide that word in order that user can view and guess the meaning from the rest of context. If difference between normal mode and comprehension mode is shown in **Figure 2** and **Figure 3**.

(3) Vocabulary Mode

This mode will show the meaning of the vocabulary, when user click the unknown word. To display the explanations of word, we first thought of display it in pop-up menu, but because of the small screen of PDA, we finally designed to display it in the bottom part of browser. This mode will show user the meaning of the word both in Japanese and in English, and some examples of sentence (**Fig. 4**). User also can specify which word s/he want to exercise in. The list of words that users chose to exercise can be saved in text format and can be shown after users terminate our browser.

It's a huge **victory** for tobacco lawyers," said Bill Novelli, president for the Campaign for Tobacco-Free Kids, a Washington, D.C., health advocacy group. Novelli said the ruling would probably increase pressure on Congress to write laws to reduce youth smoking.

victory(n)勝利，戦勝, conquest
American have a victory in Vietnam War.

Fig. 4   View in Vocabulary mode

When user want to view a home page, our browser has to send a request of the home page to the proxy server and also tell which mode user want to user, in order that our proxy server can prepare the appropriate data to user.

7. PROXY SERVER

We have to change plain HTML text file of home page to adjust with our mobile browser. Our proxy server will search for the English vocabulary in plain HTML text file and search for its meaning from English-Japanese dictionary on server. These tasks needs many computational resources, and are not appropriate to be executed on low performance machine like PDA. For these reason, we decided to use proxy server in order to reduce PDA's tasks.

This proxy server has to be put on user's WWW server, and a user can use PPP connection to connect

our proxy server and his/her PDA. Because our proxy server is developed by language C, it is required C-compiler to be installed on this WWW server. On this WWW server, we also put user's task-history that we will describe in the next section.

Our proxy server will be the mediator between the Internet and PDA. User can log-in to this proxy by identifying user ID and password. When s/he wants to view a home page, s/he has to send a request to their server. This server is user's WWW server which maintains user's home page. Our program on server will request for the home page from the Internet. After receiving data of home page, it will rewrite data and create learning materials.

7.1 Create Learning materials

As we describe in former section, because of low performance of PDA, we designed to prepare learning materials on the proxy server. When users want to view a home page, our mobile browser will send a request to proxy server. Proxy server will send a request and receive data of home page from the appropriate server.

After proxy server received data of home page from the Internet, it will decide whether this data is English or not, and which view mode user has selected. If the data is not English, our system will deal with it like in normal view mode.

In normal mode, we will rewrite home page that is written in HTML to be the easier one that appropriate to our WWW text browser, and then send it to PDA that requested that home page.

In comprehension mode and vocabulary mode and the home page is in English, our mobile browser will search for the meaning and explanation of each word included in a home page by using shared dictionary on their WWW server. Then, it will rewrite data of the home page in order to be viewed with comprehension mode and vocabulary mode. We will describe this shared dictionary in the next section.

For learning materials, we recommend users to use the news home page, such as CNN, TIME, and asahi.com. Different from ordinary English text books, these news home page not only have many

different interesting areas that can match to most users. Users can read know a lot of vocabulary while they can know what is occurring in our world. Accordingly, we set these news home page to be default home page of our system.

On the other hand, reading the interesting area will help users keep reading English. User can also set other home pages to be their default page. If there are some home page that are set by users, our proxy server will pre-cache the data from the Internet and keep it in the hard disk. For example, if a user love to see CNN's home page, especially for business page, we will pre-catch the data from CNN business page and rewrite it in order that it can be used with vocabulary mode. When users log-in and get the data from the server in the next day. They have no use to wait for our server in preparing the data.

7.2 Exercise problems

After user viewed our home page, we will provide exercise problems for them. These exercise will be created by random function from the database of words a user has clicked to learn, and they will be created on proxy server. When user finished reading an English home page and wants to do the exercise, s/he should send a request to proxy server. After finishing creating exercise, proxy server will send the exercise to PDA that requested it.

Answers of exercise will be transfer to PDA with the exercise itself. After user finished doing the exercise, our program on PDA will check it, and remain its result. In the next time PDA is connected to proxy server, these result of exercise will be transfer from PDA to server.

8. NETWORK BETWEEN PDA AND PROXY SERVER

Although our system attempt to reduce network traffic by transferring only text data, in low-bandwidth and nomadic connection like PPP connection, the connection can be terminated during a transferring of data. Transferring data from the beginning will cost time and useless.

To prevent this, we designed to memorize the position (byte) of data in the file that is being

transferred from Proxy server to PDA. If the connection is terminated during our transfer, in our re-connecting, we will not transfer data from the beginning of file, but from the position that the data has been transferred.

9. USING DICTIONARY

We first put a free English-Japanese dictionary on user's WWW server. This dictionary is so small that it can not tell us all of the meaning of the word. But when there is a word that not included in this dictionary, our system will search for its meaning from on-line dictionary. The list of on-line dictionary is first set by us, but user can add other server to this list later.

In this section will be discuss on problems in using dictionary and how to resolve it.

9.1 Dynamic of the information

One of the big problems in the Internet is the dynamic of the information. Many home pages are rewritten, some are moved, and some are deleted. This also will be a problem in our attempt to search for unknown word from dictionary. We can not guarantee that the on-line dictionary server will be in used forever, and we can get their information forever too.

To solve this problem, we designed to add each searching result to our dictionary on user's WWW server. This method will help us can keep some information even the original on-line dictionary is deleted or moved.

9.2 Expanding and Sharing Dictionary

In the case that many users have their directory on WWW server, they can share a dictionary. When they share a dictionary together, its size will grow bigger because of they will add unknown vocabulary that have been searched from on-line dictionary to this shared dictionary. This will cause creating learning materials be faster than not sharing the dictionary, because some vocabulary may be already added by other user. This assumption will be approved when all of the users have the same interesting area.

On the other hand, if each user, who shares the same dictionary, has different interest topics, home page that they select will be different, and also the word used in each interest area will be different too. In this case, if the dictionary becomes bigger, the probability that the word a user want to know is added to the shared dictionary by other user may be smaller. In a word, the bigger the dictionary becomes, the more tasks proxy server need to search for vocabulary meaning. As the result, it will delay response from proxy server.
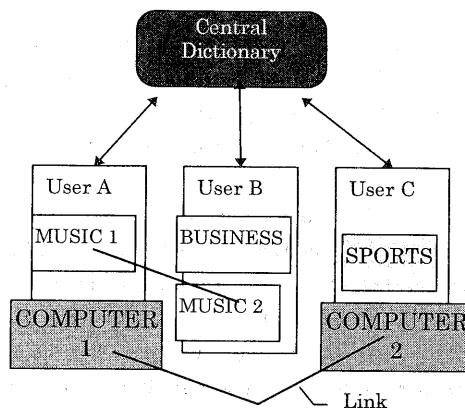

Fig. 5 Sharing dictionaries with other users (Physical Image)

As the result, even sharing dictionary will be helpful for users in the future, user has to choose who has to same interesting and want to share the dictionary. If a user has interesting area that other users have, s/he should keep this dictionary on their local directory and their only use. Some users who have many interesting area can share some of dictionaries with other users who has the same interesting area, but should keep some dictionaries that is in the different interesting to others (**Fig. 5**).

9.3 Latency on searching on-line dictionary

Most on-line dictionary use CGI to retrieve a word. It may cause such a lag time that will delay the response from on-line dictionary server. And it will cause more lag time, if we use only one on-line dictionary to search for all vocabulary.

To prevent this, we have to distributed our request to several server. For example, we have to search for "millennium", "casualty", "Thermometer", and "dysfunction". We will search for the meaning of

"millennium" on on-line dictionary server named "A", "casualty on on-line dictionary server named "B", "Thermometer" on on-line dictionary server named "C", and so on.

If there is no response from an on-line dictionary server after a limited time, our proxy server will give up and try to search on another on-line dictionary server.

In addition, some dictionary may be out of use when they are crashed, deleted moved to another server. We have to checked and get rid them from our list of on-line dictionary server.

9.4 Aware of Network Traffics

To search the meaning of word from on-line dictionary may cause a lot of time. An on-line dictionary may be on high speed in this time zone, but it may be on low speed in the other time zone. The thing that we should do is to keep watch at the condition of traffics all the time, and decide when is the best time for searching on that on-line dictionary server.

To know exactly which on-line dictionary server will be best on what time and day may take us a long time, for example, we have to keep watch on their response times and in one or two months. These response times will be taken in statistics and will be used to decide which on-line dictionary server can response in the highest speed on what time and day. The information of the highest-speed server on exact time and day will be used when there is no response from an on-line server.

On-line Dictionary   The highest speed on Monday, 10:00-11:00 am.

(1) search for a word; (2) no response
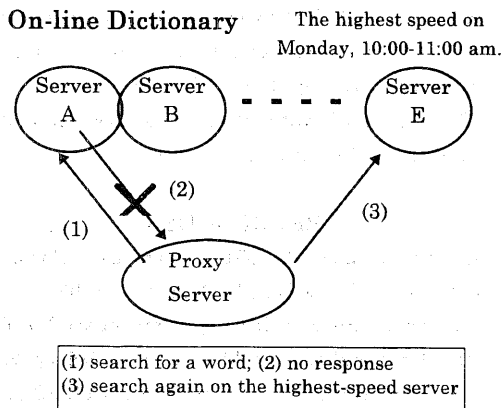(3) search again on the highest-speed server

Fig. 6   Searching on on-line dictionary

For example, we try to search on the on-line dictionary server named "A", but there is no response from server "A" until a limited time. The search of word on that server will be given up. Then, if the highest-speed server on that time and day is server "E", we will try to search the same word on server "E". If searching on server "E" failed, we will try other high-speed server on that time and day (Fig. 6).

10. HISTORY OF USER'S TASK

History of tasks that a user execute will be maintained on user's WWW server. This history is created for two reasons; for memorizing tasks during learning and for future feedback. We will describe these design issues in the following:

10.1 History during Learning Time

Many users may view home page randomly. They may click to see the next page and then come back to continue reading the first page. This operation may be occurred now and then when users view news home page. View many home pages in one time may need a lot of cache. But for PDA like Palm Pilot that has a limited resources will be suffered that we can not cache a large size of data in PDA.

In present progress, we designed to cache only one home page, its exercise and its answers that can be used for only one time. On the other hand, we will keep almost all learning materials on user's WWW server. This capacity of cache on this server will be limited by a number of file and its size in kilo byte that a user has set.

When a user wants to see other home page when s/he did not finish reading the present page, our browser will send a request for the new home page to proxy server and also send the information of the present page to server. The information of the present page will be consist of what page that a user viewed, how long or where that s/he read that home page. A user can specify the position in the home page that s/he has read, but if s/he did not specify anything, we will figure out it from the last word s/he clicked to view its meaning.

This information will be saved on proxy server, and when a user requested the home page that s/he has not finished reading it again, we will send data

of that home page and the information of where s/he has read with it.

10.2 History for Future Feedback

Each user has an individual personality and different ability. In order to know their ability in reading English, we try to keep track on their learning progress. This learning progress can be notify from the result of exercise that are always transferred from PDA to proxy server and are kept on user's WWW server.

On user's WWW server, we maintain the address of home pages that a user has read, which vocabulary that s/he has learned, and s/he has learned. And from the result of exercise, we can know which word a user already remember and which word s/he can not remember, etc.

We will feedback to this information when we have to create learning materials and its exercise in the next time. For example, if user can not remember the word "venerate", we will find the home page including "venerate" for learning material in the next time.

Some words like "computer" may be very simple to some learners, but for others it may be difficult. We can not know the learner's level at first time s/he use our program. So, at first we will find for all of the words including in a home page. In vocabulary mode, if user didn't click a word, that word may be something s/he already knows its meaning. But only one time can not adjust this. We decided to keep look on user's behavior for at least five times. After five times, user still didn't click on that word, we will say it "Already known word". This information will be kept on user's WWW browser and be feedback in the next time of learning.

We divided vocabulary in four types:
(1) Already known word

At first, we prepare a word list that contains about 40 words, which we think are very simple and any one who learned English more than six years should know, such as "a", "dog", and "cat". If there is any word that that user never click it after five times of learning, we will include it to a list of already know word.
(2) May be known word

The word that a user has never click, but it appeared less than five times. We have to keep track on it.
(3) Already learned word

The word that a user did not know, but s/he learned and do its exercise correctly more than five times.
(4) May be learned word.

The word that user did not know and do its exercise correctly but less than five times. We have to keep track on it.

These lists of word will be kept in user's WWW server in order to be feed back in the future. User can view this history by using WWW browser from any computer. In this case, s/he has to identify her/himself by user ID and password before viewing the history.

11. SUMMARY

In this paper, we have described the design issue of our WWW text browser and it proxy server. This system is now on develop. We wish that with this system users can use their leisure time to improve their English reading ability by using our mobile WWW text browser.

Some users may be glad to read an English home page and learn the meaning of each word by view it from our browser, but they may not want to do the exercise by using PDA. In the future, we plan to develop our system in order that it can be used both on PDA and on desktop machine.

Moreover, we propose how to improve the mobile computing with PDA by giving an example of using proxy server and an example of how we use on-line dictionary. We hoped that these methods can be efficiently used with PDA in a low-bandwidth network like PPP connection.

REFERENCES

1) Eli Hinkel: TOEFL Strategies, Barron's Educational Series, Inc., 1994, pp. 94, pp. 100.

2) Pamela J. Sharpe: How to prepare for the TOEFL test, Barron's Educational Series, Inc., 1996, pp. 265

3) Michael A. Pyle, Mary Ellen Munoz: CLIFFS Test of English as a Foreign Language Preparation Guide, John Wiley & Sons (SEA) Pte. Ltd., pp. 21