LOCATION BASED CELL PHONE ACCESS TO THE WIKIPEDIA ENCYCLOPEDIA FOR M-LEARNING

Andreas Bischoff
FernUniversität in Hagen
Faculty of Mathematics and Computer Science
Control Systems Engineering group (PRT)
Philipp-Reis-Building
Universitätsstraße 27
D-58097 Hagen
Germany
andreas.bischoff@fernuni-hagen.de

ABSTRACT

This paper describes a new approach to access the content of the free Wikipedia on line encyclopedia with ordinary cell phones. No connection to the mobile Internet is required. Mobile users are able to request Wikipedia articles via SMS or via touch tone input in a voice menu. An audio representation of the requested articles will be generated in real time by the help of computer synthesized speech (text-to-speech) technologies. A location based service which reads out the best fitting Wikipedia article depending on the users position is integrated.

KEYWORDS

Wikipedia, cell phone, Speech synthesis, VoIP, m-learning, location based services

1. INTRODUCTION

This paper will introduce the adaptation of a Web based service the 'Pediaphon', a tool for audio based m-learning, to cell phone usage. The service utilize existing text-to-speech technologies to render spoken article out of all German, French, Spanish and English Wikipedia articles. No special smartphone or mobile Internet access is required, every cell phone is suitable to use the service.

2. MOTIVATION

The ubiquitous availability of mobile communication devices which are connected to the Internet, makes it possible to use small amounts of spare time for mobile learning (m-learning). Travel and latency times can be used for the so called microlearning [11]. The term 'microlearning' describes a new e-learning paradigm with small or very small and short learning units, the so called 'microcontent' [9]. The main reasons which limit the usage of m-Learning services for the end-user are usability problems, mainly the limited screen size and input facilities of highly mobile devices like smartphones and PDA's. Communication costs and bandwidth limitations are also limiting factors for possible users of mobile learning services. Bandwidth limitations are solved today with UMTS enhancements like HSDPA and HSUPA (uplink). With the upcoming flat fees for UMTS and GPRS based Internet access, communication costs will be no longer an important issue. Nevertheless many users of mobile phones do not use mobile Internet access because they are afraid of the costs.

3. AUDIO BASED TEACHING MATERIAL

As an alternative to display large text documents on very small displays, audio based learning material can be a solution for hand held devices. The usage of audio based learning material in distance education is state of the art since the seventies. Audio based learning material can also be used by blind people without modification. The production of audio learning material is expensive and time-consuming. As an alternative approach, automatically generated audio material can replace time-consuming audio reproduction. Despite to the fact that the quality of text-to-speech generation is not perfect for m- and e-learning purposes, it is still usable for rapid prototyping of learning material. Especially to generate an audio representation of a text, dynamically text-to-speech conversion is the only solution.

4. SPOKEN WIKIPEDIA

The growing amount of high quality articles available via the on-line encyclopedia Wikipedia [22], [13] is very suitable as dynamic content for microlearning purposes. The established project 'Spoken Wikipedia' [20] implement an audio representation of chosen Wikipedia articles by the help of various contributors. But these solutions are lacking the features of the text based Wikipedia articles in some points. The underlying principle of Wikipedia is user changeable content. The content of Wikipedia articles is changing often. Manual recorded audio representations of articles will reflect only a state of an article at a certain moment. For the users of 'Spoken Wikipedia' it is difficult/impossible to correct an article directly. Since the audio recording of articles is time consumptive the 'Spoken Wikipedia' project only covers 727 articles of the 2,565,900 (2008-09-29, http://en.wikipedia.org) available articles. Audio and speakers quality differ. Since 'Spoken Wikipedia' is supported by volunteers due to varying quality of their audio equipment (microphones, sound cards) general audio quality is not standardized. The speakers are non professionals and sometimes non native speakers so their pronunciation differ widely. Since Wikipedia tries to establish an objective view of any topic, an emotional interpretation of an article may break this objectivity. Objectivity in the content of articles is often a topic of heavy discussion. The only solution to the first two points, actuality and completeness, are automatically generated audio articles by text to speech techniques.

5. SPEECH INTERFACE TO THE WIKIPEDIA

The overall approach of the introduced service is to fetch a Wikipedia article on every user request directly from the Wikipedia web servers. After some preprocessing (removal of links images and huge tables) the HTML documents will be rendered to plain text. After speech synthesis (see 6.) the generated audio representation of an article will be provided to the users in different ways, depending of the capabilities of the target cell phone devices.

6. TEXT-TO-SPEECH PROCESSING

Text to speech audio generation is available since the eighties in differing quality for different languages but it is still not a trivial task. A rule set for preprocessing of text files is necessary to cover special cases like spoken numbers, abbreviations and text formatting. Spoken language consists of a set of phonemes and the generation of these phonemes out of text files vary largely on this between different languages. Especially if the pronunciation of a word is depending on its meaning the phoneme generation will fail. A free digitally available pronunciation encyclopedia for the target language will simplify this task. The open source tool 'txt2pho' [25] is used to get a text based representation of the phonemes for a given German language text file. The English language variant of 'Pediaphon' takes advantage of a similar tool for British English ('freephone') [15]. The French version of the Pediaphon service uses the 'LIA PHON' [4], the new Spanish version the TTS [19] text to phoneme translator. After identifying all phonemes each single phoneme must be synthesized as digital audio output. The free 'mbrola' [10] speech synthesizer was adapted for this task.

'Mbrola' is a universal solution for voice files from different languages. A web-based interface of the Pediaphon service for PC usage has been introduced in [31].

7. WAP INTERFACE

For mobile phone access we have established an WAP interface to the Pediaphon [6] service, but the communication costs are still a limiting factor for mobile users. The WAP interface to the Pediaphon service provides a WML 1.1 input mask as well as a WML server response. The response includes a link to the corresponding article at the Wapedia [2] WAP service (a WAP representation of all Wikipedia articles) as well as a link to the generated MP3 audio file. Downloading the usually huge MP3 files can be very expensive for users without a GPRS or UMTS flat fee contract. To type in type URL's on mobile phones is usually painful. Common approaches to assist the users for typing, like the T9 method, are not useful for URL's. Especially long URL's with special characters require much effort for potential users.

8. PHONE INTERFACE

The main disadvantage of the WAP solution is the large amount of data which will be generated by each Pediaphon query. A MP3 file generated from a typical Wikipedia article will be about 12 MB of data size. For mobile Internet users without a flat fee the download may be very costly. The downloaded file has to be stored locally on the device. Some mobile phones have very limited resources to store files. To avoid those communication costs and the limitation to users of the mobile internet we have created a pure SMS and GSM based service, too. No (probably expensive) GPRS or UMTS mobile Internet access is required. The service works with all mobile phones today without the need of additional configuration or expert knowledge. The costs are transparent for the users, only communications fees for an SMS request and a land line call have to be paid. Since the users are paying directly to the communications service provider, the Pediaphon service can be established without any costs for the university. The Pediaphon phone interface was realized by help of voice over IP technology (VoIP) and the Asterisk [14] open source PBX (private branch exchange) . Since user independent voice recognition technology today is not able to detect more than 30-50 different words without a training phase a suitable input method for the Wikipedia search terms is required.

SMS input approach

The German language Pediaphon SMS interface could be accessed by sending a SMS with the text: 'pedia [search word]' to +49-151-59111661. After two minutes (to be sure that the SMS arrives and the text to speech (TTS) processing is done) the users have to call (with the same caller ID) the land line number +49231-1774088 (International, Germany). The announcement of the generated speech will start automatically. The communication costs are transparent to the users. This SMS based service is available for the English and the German language version. The SMS request for the English language version must include the tokens 'pedia en [search word]'.

Touch tone input approach

As an alternative input method a touch tone based interface was created. Like the input of a SMS the numerical keys must be pressed more than once to access the corresponding letter. Numerical keypads are usually already labeled in that way. The chosen letter will be repeated after each input. A voice menu assist the user during input. After successful processing of the search keyword the playback of the requested article starts. During audio playback the users are able to navigate in the announcement by pressing predefined keypads. At any time users can stop the playback and redefine their search.

World wide access

The phone service is fully based on VoIP SIP (session initialization protocol). Therefore an international dial in with local communication rates can easily be established. In addition to the German dial in number as 'free world dial up' (FWD, www.freeworlddialup.com) SIP account was used to realize local international dial in numbers. By the help of the free Sipbroker service (www.sipbroker.com) local phone numbers in 27 countries distributed over 4 continents are available to the reach the service inexpensively.

9. PEDIAPHON AS A LOCATION BASED SERVICE

To extend the features of the introduced service to a location based server at first some information about the users position is required. The proposed approach assumes that the user is equipped with a GPS receiver, a cell phone (Some smartphones today are already GPS or AGPS (Assisted GPS) capable) or knows at least some address data of his location. In case the user only knows some address data this information can easily be converted to GPS coordinates by the help of the Google maps API Web service[16]. Even if only a PO.Box or city name is known, the Web service returns, relative to the users real position, inaccurate but near coordinates. In case of the user has a cell phone, base station cell ID (each base stations cell in a mobile phone network has a unique ID) can be used to estimate the users position (the so called cell tracking[26]). Most of the smartphones today are able to display the cell ID of the connected base station. Since most of the cell phone provides sell access to these position information them self, this data is often not available to the public. The German O2 network provider sends location information of each base station in Gauss Krueger notation [5] as a free cell broadcast service. This location information is originally provided to detect a special discount, the so called 'home zone', an area with reduced communication fees. As an alternative some free Web bases services exists, which collects the GPS coordinates of each detected base station by help of volunteers. For instance 'Patricks GSM Pages' [30] provides a Web-based access to database of GPS coordinates of the O2 mobile phone network base stations in Germany. This service is not a Web service, but can be used with some additional parsing of the HTTP response automatically in a similar manner. Depending on the size of the cell the users position is known in a range between 200 m - 10 km. In case of the users has a GPS receiver his position is known very precisely (~ 10m). In all the three cases the location of the users is known in different precision. With the help of another Web service at geonames.org [29] it is possible to get nearby geocoded Wikipedia articles (most of the Wikipedia articles today, whicht describe a place, a city or any physical existing unique object are enriched with its GPS coordinates) back to a given position. This approach is the so called reverse geocoding. Depending on the request the Web service returns one or more nearby articles or nearby article in a given range. Even information about the distance to the given position will be returned for each fitting article. The Pediaphon location based service converts the given article to a spoken MP3 audio file on the fly for each position request. The cell phone based Pediaphon service provides also a location based service. Additional to the keywords described in 8 a user is able to submit his position as address, as GPS or Gauss Krueger coordinates via SMS. The best fitting Wikipedia article will be played back during the following call to the service. In case of a user equipped with GPS receiver or cell phone (smartphone) this approach can easily be extended to an automatically play back of the nearest located Wikipedia article in case of a detection of a better fitting article in respect of user movements. This is the functionality of an automatically talking guide to information or attractions in an unknown environment (e.g. a tourist guide) [12].

10. CONCLUSION

The service itself is quite usable for short Wikipedia articles. It has been reported by some users that listening to a synthetic voice for a long time is inconvenient. The Pediaphon service is more usable with short articles just to refresh knowledge or to recover facts and issues. Especially the mobile phone version can be useful to act like a 'Hitchhiker's Guide to the Galaxy' [1] for mobile users in their private everyday life [8]. The control systems engineering group of the University uses the same technique to generate audio teaching material (and a Podcast) out of text based material.

11. FUTURE WORK

Future work on 'Pediaphon' will cover the implementation of new target languages like Russian and will improve the mobile phone based user interface (the voice menu). A voice menu for currently only web-based available Frensh and Spanish language versions have to be created. An evaluation of typical usage will be done.

12. REFERENCES

[1] D. Adams. The Hitchhiker's Guide to the Galaxy. 1979.

[2] F. Amrhein. Wapedia. http://www.wapedia.org/de/, 2007.

[4] F. Bechet. Liaphon - un systeme complet de phonetisation de textes. In Revue Traitement Automatique des Langues (T.A.L.), volume 42. edition Hermes, 2001.

[5] R. Bill and D. Fritsch. Grundlagen der Geo-Informationssysteme. Herbert Wichmann Verlag, Heidelberg, 1996.

[6] A. Bischoff. Pediaphon. http://www.pediaphon.de, 2006.

[7] A. Bischoff. Podcast based m-learning with pediaphon – a web based text-to-speech interface for the free wikipedia encyclopedia. In 7th International Conference 'Virtual University' VU'06, pages 173-176, Bratislava, Slovak Republic, DEC 2006.

[8] O. Bohl, S. Manouchehri, and U. Winand. Mobile information systems for the private everyday life. Mobile information systems, 3(3-4):135 – 152, 2007.

[9] A. Dash. Introducing the microcontent client. Technology and Entrepreneurship, NOV 2002.

[10] T. Dutoit, V. Pagel, N. Pierret, F. Bataille, and O. van der Vreken. The mbrola project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In ICSLP'96, pages 1393 –1396, Philadelphia, 1996.

[11] R. Fischer. Microlearning with mobile weblogs. In Microlearning Conference 2005. http://microlearning.org/micropapers/MLproc_2005_fischer.pdf, 2005.

[12] S. Gulliver, G. Ghinea, M. Patel, and T. Serif. A context aware tour guide: User implications. Mobile Information Systems, 3(2):71 – 88, 2007.

[13] N. T. Korfiatis, M. Poulos, and G. Bokos. Evaluating authoritative sources using social networks: an insight from wikipedia. Online Information Review, 30(3):252 – 262, 2006.

[14] N.N. Asterisk pbx. http://www.asterisk.org, 2007.

[15] S. Isard, A. Conkie, Freephone. http://tcts.fpms.ac.be/synthesis/mbrola.html, 2007.

[17] N.N. Google maps api. http://www.google.com/apis/maps/, 2007.

[19] A. Conkie, TTS, http://tcts.fpms.ac.be/synthesis/mbrola/tts/Spanish/perl_tts.zip, 2008.

[20] N.N. Spoken wikipedia project. http://en.wikipedia.org/wiki/Wikipedia:Spoken_Wikipedia, 2007.

[22] N.N. Wikipedia. http://en.wikipedia.org/wiki/Main_Page, 2007.

[24] M. Patocka. Links. the www text browser. http://links.sourceforge.net/, 2007.

[25] T. Portele, B. Steffan, R. Preuss, W. F. Sendlmeier, and W. Hess. Hadifix - a speech synthesis system for german. In International Conference on Spoken Language Processing, pages 1227–1230, Banff, Alberta, 1992.

[26] M. Pura. Linking perceived value and loyalty in location-based mobile services. Managing Service Quality, 15(6):509 – 538, 2005.

[29] M. Wick. Geonames. http://www.geonames.org, 2007.

[30] P. Wilczek. Patricks gsm-seiten. http://gsm.yz.to/, 2007.

[31] A. Bischoff. CONTEXT-AWARE MOBILE LEARNING WITH PEDIAPHON - A TEXT-TO-SPEECH INTERFACE TO THE FREE WIKIPEDIA ENCYCLOPEDIA FOR CELL PHONES AND MP3-PLAYERS. In Inmaculada Arnedillo Sánchez, editor, IADIS International Conference Mobile Learning 2007, pages 228-232, Lisbon, Portugal,, July 2007. IADIS, IADIS Press.