African Journal of Mathematics and Computer Science Research Vol. 4 (6), pp. 217-220, June 2011
Available online at http://www.academicjournals.org/AJMCSR
ISSN 2006-9731 © 2011 Academic Journals

Full Length Research Paper

The formulation of complex queries and interactions using SMS text messages

Stephen Agyepong, Michael Franklin Bosu* and James Ben Hayfron-Acquah

Department of Computer Science, Kwame Nkrumah University of Science and Technology, Ghana.

Accepted 16 May, 2011

There is no question that short message service (SMS) has achieved huge success in the wireless world. Billions of SMS text messages are sent every day worldwide. SMS is now a major revenue generator for wireless carriers like expresso and MTN of Ghana. A lot of innovative applications are now built on top of the SMS technology and more are being developed. One of the emerging applications is the use of text messages by users to search for things. Currently these applications allow for just simple interactions and queries. Currently SMS text messages are restricted to at most 140 bytes (1120 bits) of data, so one SMS message can contain up to 160 characters if 7-bit character encoding is used, 70 characters if 16-bit Unicode UCS2 character encoding is used. In this paper, we propose a way to create complex queries and interaction between the user and the external short message entity (ESME) which maintains database of information and responds to queries, and also maintain a dialog with the user when search result is longer than 160 characters.

Key words: Unicode, external short message entity, short message service, wireless application protocol, language analyzer.

INTRODUCTION

There are enormous amount of information on the World Wide Web and the amount continue to grow on a daily basis. According to internet usage statistics for Africa published by Internet World Stats (2010), Ghana has seen over 2000 percentage point increase in the number of internet users since year 2000. Sadly the number of internet users as a percentage of the population is a mere 5.3%. On the other hand, mobile phone penetration in Ghana was estimated to hit 50% by the end of 2010. According to Peprah (2008) Ghana had over 8 million mobile phone subscribers, and has already seen 36% penetration since the end of 2007. The trend suggests that more and more people are going to have cell phone than access to the internet. The cell phone has the capability to make most, if not all, internet content available to cell phone users.

One of the emerging internet applications to make web content available to cell phone users is the use of SMS text messages to search for things. Currently these applications allow for just simple interactions and queries. For example, a text message may be sent to register for a service like this: Register <service> or search for things like hotels: hotel <city>. These simple messages, while they are good, are not useful to the business traveler that needs to stay at Adabraka, Accra because of business or other reasons or the cash strapped business owner or employee who is looking for the cheapest hotel in town or on a budget and therefore looking for a hotel within a certain price range. Ideally, the user should be able to search for a hotel, for example, by specifying the city, suburb, hotel rating or grade and price. An ideal query would look something like this: hotel <city> <suburb> <grade> <price>. An actual query would be presented this way: hotel accra adabraka “guest house” 5,20 which translates to the following: I am looking for a hotel, that is of type guest house, in Adabraka, Accra which has any type of room for between 5 and 20 Ghana cedis. Note that case is not sensitive so for example, accra is the same as Accra, same as ACCRA as far as our language analyzer is concerned.

A language grammar or syntax that defines <city>, <suburb>, <grade>, and <price> as simple identifiers or symbols of the vocabulary of the language will not do as

*Corresponding author. E-mail: mfbosu@yahoo.com. Tel: 00233-246-372687.

we will see later in the paper. This paper is proposing a key-value pair vocabulary which will resolve any language ambiguities that may arise as we seek to use cell phones to create complex queries. The previous grammar will be changed to something like this: hotel <city=identifier> <suburb=identifier> <grade=identifier> <price=identifier>. For example, the previous query that is looking for a hotel in Adabraka will be rewritten as: hotel city=accra suburb=adabraka grade="guest house" price=5,20. There are also several issues that may have to be resolved in our attempt to create complex queries using our cell phones. We discuss these issues and justify why we are using SMS to begin with in the ensuing paragraphs.

There are other alternatives to SMS but we chose SMS (Metsker, 2001) for several reasons and subsequently, we discuss the advantages to using SMS as opposed to, for example, using Wireless Application Protocol (WAP) or Java Platform, Micro Edition (Java ME).

ADVANTAGES OF SHORT MESSAGE SERVICE (SMS)

SMS is widely available. Any cell phone can send and receive SMS text message. The same cannot be said about the other means of communication such as instant messaging (IM) and WAP or technologies such as Java ME. A limited number of cell phones support WAP. We want our solution to be widely available and see SMS as the only way to go.

Messages are instantly recorded. Unlike a phone call, an SMS message is automatically stored where it can be re-read. This proves particularly useful to our solution in the case where a query returns fairly detailed information that might otherwise be forgotten.

SMS is personal. Unlike an e-mail, SMS is much more likely to be read by one person at any point in time at any time during the day, since the majority of people have their mobile phones at arms reach Twenty four (24) hours a day.

SMS messaging is relatively Spam free. Unlike e-mail, SMS as it currently stands is relatively Spam free. Although this may change over time, at present it is the ideal communication channel to cut through the clutter. As a result, marketing departments worldwide are climbing aboard to try and target their customers in this one-on-one manner that has SMS text messaging soaring way above any other medium of communication.

SMS is discreet. Unlike a phone call you do not have to run out of the restaurant where you are eating to receive your SMS message, you know when an SMS has arrived. The discreet nature of text messaging ensures you stay in touch with minimal disturbance.

SMS leads to smaller phone bills. An SMS is far cheaper than a phone call yet in most instances you will convey just as much information as you would if you had called.

SMS is a store-and-forward service, meaning that when you send a text message to a friend, the message does not go directly to your friend's cell phone. The advantage of this method is that your friend's cell phone does not have to be active or in range for you to send a message. The message is stored in the Short Message Service Center (SMSC) (for days if necessary) until your friend turns his/her cell phone on or moves into range, at which point the message is delivered. The message will remain stored on your friend's SIM card until he/she deletes it.

In addition to person-to-person messages, SMS can be used to send a message to a large number of people at a time, either from a list of contacts or to all the users within a particular area. This service is called broadcasting and is used by companies to contact groups of employees or by online services to distribute news and other information to subscribers.

CREATING AMBIGUITIES USING SIMPLE SYNTAX OR GRAMMAR

Every language displays a structure called its grammar or syntax. The syntax decomposes sentences into their constituent parts. For example, a correct English language sentence always consists of a subject followed by a predicate, correct here means well formed. This fact can be described by the following formula:

Sentence = Subject + Predicate

It follows that every sentence must have a single structure in order to be unambiguous. Here, we show how treating the vocabulary of our language as simple identifiers or symbols can lead to ambiguities. Take for example, the following sentences:

Hotel Kumasi budget (1)
Hotel Kumasi dichemso (2)

In (1), the language analyzer sees a literal (hotel), followed by a word (kumasi), and followed by another word (budget). For a human we can see the difference but to the language analyzer both (1) and (2) have the same structure (literal word word) and therefore ambiguous. Let us look at another example:

Hotel Kumasi dichemso 2 (3)
Hotel Kumasi dichemso 5 (4)

Even for a human it is very difficult to say exactly what the sender really wants – a 2 or 5 star hotel in kumasi dichemso or any hotel room that costs 2 or 5 Ghana cedi in kumasi dichemso? The language analyzer sees a literal, word, word, and a number for both (3) and (4). From the previous examples we can see how easy it is to create ambiguity in our language treating our vocabulary as simple variables.

RESOLVING LANGUAGE AMBIGUITIES USING KEY VALUE PAIRS

The ambiguities can be resolved by preceding each variable with exactly what the variable following it means. For example, the four sentences above can be resolved by re-writing them as follows:

Hotel city = Kumasi grade = budget (5)
Hotel city = Kumasi suburb = dichemso (6)
Hotel city = Kumasi suburb =dichemso grade = 2 (7)
Hotel city = Kumasi suburb = dichemso price = 5 (8)

Or as follows:

Hotel city is Kumasi grade is budget (9)
Hotel city is Kumasi suburb is dichemso (10)
Hotel city is Kumasi suburb is dichemso grade is 2 (11)
Hotel city is Kumasi suburb is dichemso price is 5 (12)

Of course the user will type more but then his queries will be unambiguous. We can minimize the amount of typing by using abbreviations for the key. We can even eliminate the equal (=) symbol. We can re-write examples 5 through 8 as follows:

Hotel city Kumasi grade budget (13)
Hotel city Kumasi suburb dichemso (14)
Hotel city Kumasi suburb dichemso grade 2 (15)
Hotel city Kumasi suburb dichemso price 5 (16)

Or as follows:

Hotel c Kumasi g budget (17)
Hotel c Kumasi s dichemso (18)
Hotel c Kumasi s dichemso g 2 (19)
Hotel c Kumasi s dichemso p 5 (20)

A benefit of this approach is that it brings us closer to the natural language as examples 9 to 12 show. The approach also gives the user several options in which to formulate a query. Order is not important when formulating queries using our approach.

A restriction of our approach is that a parameter that has a space in them has to be put in quotes in order to avoid any ambiguities that may arise. For example, we will write:

Hotel c Accra s “ada foah” p 5 (21)

SPECIFYING RANGE

A cash strapped business owner or employee may want to know the cheapest hotel in town or hotel within his/her budget by specifying a price range. A user may specify a range in one of the following ways:

Hotel c Kumasi s dichemso p 30,50 (22)
Hotel c Kumasi s=dichemso p=30,50 (23)
Hotel c Kumasi s is dichemso p bet 30 and 50 (24)

UPDATING INFORMATION USING KEY VALUE PAIRS

A useful side effect of our proposal is that it can be extended to update information. When the rates for a single (s), double (d), suite (st), etc. changes a message can be sent to the ESME to update their records. For example, the following message could be sent:

Hotel update user = user1 password = pwd12 s = 20 d = 25 st = 30 (25)

We authenticate the user before we perform any update operations to ensure the integrity of information. Using key-value pairs we can update just one field or several fields at once. We can also use the position of a field in a record to update it. If the position of double is 1, single is 2, and suite is 3 the previous sentence can be written as:

hotel update user=user1 password=pwd12 2=20 1=25 3=30 (26)

PACKAGING USER MESSAGE(S) AND DEALING WITH USER TRANSACTION

A major limitation of SMS text messages is the 140 byte limit per message. This means that we will have to break up messages longer than 140 bytes and ask the user to let us know when they need the next message. The problem that arises when we break up messages is how do we know when the user is done asking for more messages? So, we break up the messages meant for the user into 10, and send the first message. The user asks for the next message and did not send us any next message for another five minutes. How do we know that the user is having problems sending us the next message? The user is done because the previous two messages gave them the answer they were looking for.

Each message to be sent to the user is stored on the ESME’s system and therefore is cost to the ESME. The more of such messages that are kept the longer it takes to respond to requests to send additional messages since we have a longer list to search through. We propose a timer service that runs periodically deleting messages that have been waiting for a specified time period.

CONCLUSION

An SMS language was presented that allow the user to craft complex queries and get useful results for the queries. We have been able to eliminate ambiguities in

our queries using key value pairs thus, make it possible for our users to accurately formulate queries. Another benefit of our approach is that, it could potentially bring us closer to natural language making it easy for the user to formulate complex and useful queries. Our system was able to overcome the traditional 140 bytes of data limitation by using a paging technique that allows an unlimited amount of data to be sent to the user.

ACKNOWLEDGEMENTS

The authors sincerely thank all who contributed their suggestions and criticisms to this research which materialized over the last year. They are also grateful to Kwadwo Osman, Kwaku Duah Boateng, Sophia Agyepong and Osei Bobie of Sabonay Technologies Ltd who contributed in no small way to the realisation of this research.

REFERENCES

Internet World Stats (2010). [WWW document] http://www.internetworldstats.com/stats1.htm(accessed 23rd February, 2011).

Metsker SJ (2001). Building Parsers with Java, Addison Wesley, pp. 10-56.

Peprah S (2008). Cell Phone Population to Hit 50%, [WWW document] http://www.ghanaweb.com/GhanaHomePage/NewsArchive/artikel.php?ID=144001(accessed 23rd February, 2011).