CORRELATION OF CONTEXT INFORMATION FOR MOBILE SERVICES

Stephan Haslinger, Miguel Jiménez
Tisco GmbH, Tigergasse 21/13, 1080 Wien, Austria
Universidad Politecnica de Madrid, Spain

Schahram Dustdar
Distributed Systems Group, Institute of Information Systems
Vienna University of Technology, Argentinierstrasse 8, 1040 Wien, Austria

Keywords: Context Information, Context Awareness, Location Based Services, Data Correlation, Mobile Handsets, MyMobileWeb.

Abstract: Location Based Services are a key driver in today's telecom market, even if the power of Location Based Services is not nearly exhausted in nowadays telecom systems. To build intuitive Location Based Services for mobile handsets one success factor is to cover a broad range of mobile handsets available on the market and to make the services context aware. Within the EUREKA project MyMobileWeb we implement a framework to obtain contextual information from handsets using various capabilities of the mobiles. Contextual Information is every information we can obtain from the handset and that can be used for any kind of service. The most obvious information is location information. Within our framework we built an architecture that can obtain location information from various sources and is not bound to any special handset capability. Furthermore the architecture can be used to obtain various other context information, such as e.g. battery level. This information in addition is then used to offer special services to the customer. For this a correlation of the context information has to be done, which is based on a correlation engine for contextual information. This paper presents a framework that can handle and correlate contextual information in a very flexible way.

1 INTRODUCTION

In the last years, the technology world witnessed a very powerful new trend – more and more people started using their handsets when accessing the Internet. The main driving force behind this trend is the improved versatility, power and usability of mobile handsets, which in turn enabled the market to invest more in building mobile-phone-friendly web services. One new revenue stream for telecom operators in the next years is definitely Location Based Services on Mobile Handsets. Analysts at Gartner expect location to become a mainstream mobile application within two to five years. They see the market growing from 16m users in 2007, to 43.2m in 2008 and 300m by 2011 (Palmer, 2008). To launch successful Location Based Services there is a need for a proven development framework for those services. It has to be reliable, create a community of developers that are focusing on it and, of course, it should try to cover nearly all mobile handsets on the market. The current situation, with big suppliers (Google, Apple, Nokia, RIM, etc.) developing their own frameworks and ecosystems, does not encourage interoperability. Location Information is nothing else then contextual information that can be retrieved from the mobile handset and can be used to build special services for the user. Besides location there is a lot of other contextual information that is interesting to obtain, such as battery level of the handset, weather forecast for a special location, etc. All this information correlated can be used to build new mobile services for the handsets with a very high flexibility. This paper will present a contextual framework how context information from various input channels can be retrieved and correlated. For this a special data structure has to be used. Before we start...
working out the data structure we will give an example that should guide us through the paper. Imagine a mobile service you can access from your web browser to get tourist information for e.g. Barcelona. You are a tourist on a trip visiting Barcelona. On the airport you get an email from your hotel reservation system informing you that they had to cancel your booked hotel room. We assume that you installed the framework from MyMobileWeb on your handset, which has the possibility to retrieve context information about your environment. You browse to your hotel information site, which shows you all hotels in Barcelona, which are 4 and 5 star, as this is predefined in your private context information stored on your handset. You book the hotel. Afterwards the site leads you to another site offering you a plan how to use public transport to get to your hotels. Once you came to your hotel and enjoyed some hours of rest you are in the mood of doing some sightseeing and ask for interested places. The website offering you this information gives you a list of interesting places. As the service is aware that it is one of the rare rainy days in Barcelona, indoor places, such as museums, are ranked better than outdoor. You just click on the sights of interest and another service starts where a navigation system for walkers is coming up and guides you through Barcelona. Services like this exist rarely, and are always bound to a certain handset or to a certain technology your mobile handset has to support. We are proposing a framework that can assist developers to build such services with a very broad range of mobile handsets available on the market. To build such a framework it is crucial to build data sets of context information that is well known by the services on the backend side, as well as on the mobile side.

The rest of the Paper is organized as follows. Section 1 is focusing on the framework that is now built within MyMobileWeb to give the reader the possibility to understand the idea of the framework easily. Section 2 explains the architectural framework that implements the conceptual work from this paper. Section 3 focuses on how contextual data within the framework is structured. This is done with extending the W3C Draft of the Delivery Context Ontology (Lewis and Fonseca, 2008). Section 4 and section 5 deal with the concept of the Correlation Engine and patterns that are used for location and context data correlation. Section 6 covers the evaluation of the framework Section 7 provides the Related Work. Section 8 gives an outlook of our future work and summarizes the paper.

2 ARCHITECTURE OF THE FRAMEWORK

Figure 1 shows the architectural overview of the framework which is built within the project MyMobileWeb. Input Channels are channels that can retrieve contextual data from the handset. The details about how this data is structured will be further explained in Section 3. An Input Channel can be any sort of hardware capability of the handset that can deliver context data. In Figure 1 we assume as examples the GPS Sensor, QR Codes, Battery Level, etc. This contextual data alone would not allow a mobile service to give any valuable information to the user. Therefore this data needs to be correlated with suitable information to build mobile services described in the introduction. Location Based Services, just as any other context-aware service, are aimed at customizing their contents, appearance and
behavior to optimize their usefulness. Pure location information is central to such services, but it is not enough to make the most with respect to usability. Imagine a recommender system that offers you nearby restaurants on a map, not only will it have to adapt such map to your screen size, but also it will have to consider the resized result and know whether icons on the map are visible or not, seamlessly adapting to small devices, and it may filter vegetarian restaurants as you love meat. So there is some context information, not only the location that can enhance LBS. Furthermore, since application domains are unpredictable, establishing a boundary on what may be useful or not in a given domain will constraint the developers trying to offer innovative LBS. The presence of such, potentially huge, amount of factors and their influence in the usefulness of the offered service makes it necessary to have an automatic correlation process. Correlation in then crucial in LBSs as a mechanism for selecting context-dependent contents based on location and non-location information.

3 CONTEXTUAL DATA

Contextual Data is every data which can be used for mobile services to make them context aware. We believe that a classification of context is needed within mobile services.

In the next subsections we explain how contextual data has to be structured formally within our framework. For this we are extending the W3C “Delivery Context Ontology”, which semantically describes both internal and external context properties. It includes a thorough description of the device capabilities, but it also describes the environment (external dimension) and the user, though her goals or intentions are not yet defined. This paper focuses on location data which is defined within the environment.

3.1 Location Data

Context Aware Services on Mobile Handsets nowadays mainly work with location data. Nokia Maps at example uses GPS data to retrieve a user’s position. Additionally the user can also ask for restaurant and bars close to any location. Services like Nokia Maps usually are bounded to one certain technology to retrieve location data. In the case of Nokia Maps this is the GPS sensor. Besides this the service is proprietary and cannot be extended easily from anybody to reuse. One reason for this is that the contextual data the handset is handing over to the backend is not defined. Adhering to standards is crucial for using location data within any framework. The W3C Draft “Delivery Context Ontology” was our starting point for declaring location data for context aware services on mobile handsets. Within the Delivery Context Ontology (short: DCO) there are various classes defining the delivery context, which includes the characteristics of the device, the software used to access the service and the network providing the connection among others. When working with location data we are mainly focusing on the Delivery Context Environment Entity that has a location class to represent location. In the previous section we described the architecture of our framework that deals with various input channels to broaden the possibilities how a mobile handset can deliver contextual data. Such architecture cannot be fully covered by the current W3C Draft of the Delivery Context Ontology because it lacks from some support for incorporating location data from different sources. Extending the Delivery Context Ontology with following properties is a sufficient way to deal with this drawback. The ontology is formally specified in OWL. Describing OWL is out of the scope of this document, so we will focus on our extensions of OWL. Extensions are defined using an intuitive XML format. Datatype properties have types that are not themselves classes. Examples for this are “xsd:int” or “xsd:string”. Object properties have types that are classes. To clarify if an element within the following class description is a datatype property of an object property we use the namespaces “xsd” for datatype properties and “obj” for object properties.

3.1.1 Point

The current class description of the class Point holds coordinates representing a point on the Earth surface. When dealing with location input channels that are designed to deliver a specific location with absolute coordinates this representation is sufficient. In our framework we are also working with input channels that cannot deliver such information. An example for this is the NFC Tag, or a QR Code Tag, which might deliver only the ID that has to be correlated in the background to deliver an appropriate location information. Imagine a railway station with a NFC Tag. By touching the NFC Tag a mobile service could deliver exact information about the current timetable of the railway station. Such a mobile service can just work with location information that is more precise then absolute coordinates on the Earth surface.
To cover this approach the point class has to be extended with following object property.

```xml
<locationid type="obj:LocationID">
  <LocationID>
    <id type="xsd:string"/>
    <namespace type="xsd:string"/>
    <channelName type="obj:Channels"/>
    <timestamp type="xsd:dateTime"/>
    <isdirectinfo type="xsd:Boolean"/>
    <scoreinfo type="obj:NeighbourInfo" minOccurs="0"/>
  </LocationID>
  <Channels>
  <Bluetooth type="xsd:string">...
  </Bluetooth>
  <NFC type="xsd:string">...
  </NFC>
  <QRCode type="xsd:string">...
  </QRCode>
  <ZigBee type="xsd:string">...
  </ZigBee>
  <WLAN type="xsd:string">...
  </WLAN>
  </choice>
  <CellID type="xsd:string">...
  </CellID>
  </Channels>
  <NeighbourInfo>
    <score type="xsd:int"/>
    <id type="xsd:string"/>
  </NeighbourInfo>
</LocationID>
</Channels>
</Point>
```

The `LocationMethod` class has to be extended with the new location method:

**ID_Correlation_Method**

The datatype property is in the class `LocationID` nothing else than a unique identifier that can be used to stamp location data, and even other contextual data, by the correlation engine. To make the `locationid` unique across multiple systems working with the same technology the class `LocationID` holds the datatype property namespace. With this approach the `locationid` is bounded to a special domain. Applying to our timetable example, the namespace takes care that a user on the railway station in Barcelona really gets the timetable he needs and not the one from Sevilla, just because the NFC Tags in Sevilla and Barcelona share the same ID. It is obvious that within one domain a `locationid` must be unique.

The object property `channelName` holds the name of the input channel. This has an informative character and has to be seen in conjunction with the object property `ScoreInfo`. `ScoreInfo` is a class that for the time being will just appear with the channel `CellID` or `WLAN`. When using `CellID` or `WLAN` as location input channels it is possible to calculate location data with triangulation of a list of neighbours together with the information how strong the information from the neighbours can be measured. How strong the signal of a neighbour can be measured is referred with `score`, which is a number ranging from 0 to 99. How the score together with the list of neighbours leads to a location is part of the correlation engine. The datatype property `timestamp` gives the exact time when the `locationid` was measured. This is a very important information and very interesting when retrieving information from another handset by calling a function via Bluetooth. Within such a mesh network location data can be retrieved by asking a neighbour about location data he shares. In this case the datatype property `isdirectinfo` would be set to false, as it indicates that information was retrieved by another handset. Within such an environment, the `timestamp` is needed since location data might be outdated. The question of when the location data is indicated as outdated is not part of this paper. It is also not part of this paper to discuss whether an information from a partner is trusted or not. Another extension or change of the point class is that the coordinates cannot be bound to appear exactly once.

Therefore a change of the point class would lead to following revised form:

```xml
<Point>
  <coordinates minOccurs="1" maxOccurs="unbounded"/>
</Point>
```

Revising the `Point` class is needed as a location can be the correlation of multiple `locationids`. To apply this behaviour it would be also possible to change the `Environment` class and set `Place` to be unbounded. Extending the `point` class with these parts helps to deal with input channels that are delivering unique IDs and where content of these IDs has to be generated from various data sources within a correlation step. It is worth pointing out that with this extensions the W3C Draft might get ambiguous. Within the current draft of the DCO there is a location method called `Cell Id Method` that is used when retrieving the location via a `Cell Id`, but does not cover the possibility of correlating data with a triangulation algorithm in the backend.

### 3.2 Mobile Handset Capabilities

Contextual Data within a mobile environment is close coupled with the the capabilities of the mobile handset. It is obvious that data about the environment can be just used if the handset has the
capability of measuring such information. For context aware mobile services increasing capabilities of mobile phones increase the feature possibilities of the services. To cover our approach we have to add the following extension to the DCO. We believe that the extensions are self explainable. The architecture we explained in section 1 can be extended with plugins. Whenever a handset with a new capability approaches the market developers can implement a plugin for the framework to capture this new form of data. An example could be a light-intensity sensor in the handset used to set the background lighting of the handset appropriately. If the developer wants to retrieve the information from the light sensor he would have to build a plugin which handles the API of the light sensor and should extend the DCO. Therefore every new class we will add for this framework will have a datatype property <pluginname>_plugin_installed of type boolean. Another property which is needed for a lot classes is <pluginname>_enabled, which indicates if the mobile handset itself has the channel turned on. This is of course just needed if the channel can be turned on or off by the user, which is true for NFC, ZigBee and WLAN. For the plugin NFC an example would look like:

```xml
<NFC_reader>
  <nfc_enabled type="xsd:boolean"/>
  <nfc_plugin_installed type="xsd:boolean"/>
</NFC_reader>
```

We believe it is intuitive to follow this guideline and will therefore focus more on extensions, which might be not as obvious.

```xml
<WLAN>
  <wlan_enabled type="xsd:boolean"/>
  <wlan_plugin_installed type="xsd:boolean"/>
  <encryption_protocol_provided type="obj:WLAN_Encryption_Protocol" maxOccurs="unbounded"/>
  <encryption_protocol_used type="obj:WLAN_Encryption_Protocol" minOccurs="0"/>
  <channel_number_used type="xsd:int"/>
</WLAN>
```

The object property encryption_protocol_provided holds a list of possible encryption methods the mobile handset can handle. The object property encryption_protocol_used holds the current encryption used by the handset or just does not appear if no encryption is used.

For covering the Bluetooth mesh network the class BluetoothSupport should be extended with the following data properties.

```xml
<Bluetooth_plugin_installed type="xsd:boolean"/>
<data_sharing_enabled type="xsd:boolean"/>
```

The data property data_sharing_enabled indicates if the user of the mobile handset wants his location data to be shared within the Bluetooth mesh network. Currently we assume that every location data can be shared. We will work on a trust class to define which location data can be shared and when it can be shared. We believe that with these extensions the DCO can fully cover the ideas from the framework, but have to point out that the list might not be exhaustive and extended during our continuing work. For the current DCO classes CellID and GPS we did not add special properties as this can be covered with the current DCO.

### 4 CORRELATION ENGINE DESCRIPTION

The Correlation Engine created in the MyMobileWeb project is aimed at correlating any kind of semantically described content taking into account all of the contextual information available. Its development has emphasized the domain independence, so it can be used to exploit any ontology-based semantic description of domain elements for arranging, selection or discovery purposes.

The strategy used to get this requirement has been to separate the expert knowledge of the domain from the correlation engine. This way, such knowledge is external to the engine, and can be defined as needed or tuned to each application's needs. This has been achieved by creating a custom high-level rule language that is designed to be used by non-technical users, as the domain experts usually are. This language is composed by correlation rules that modify the score of each element that takes part in it, named candidates, according to its characteristics and the context. Such rules have a set of condition expressions and their
corresponding score updating expression. Both expressions can reference any semantic description of the elements or the context, so the domain expert can express with these rules any criterion he wants to apply to the correlation.

Every element is initialized with the same score, and the computed result of the updating expression whose condition has been satisfied is multiplied by its score. Those updating factors are bounded to a 0 to 1 interval, so the system applies a negative updating. However it can be easily turned into positive updating by means of the default score, which is applied when no rule condition is satisfied.

Updating expressions can be defined as computed or fixed values. Computed values depend on some numerical values of both the elements descriptions and the context. It can be shown in the following rule example, which penalizes the distance between each candidate and the device location:

```xml
<rule name="distance">
  <action>
    <when>true</when>
    <score_updt>Max(Exp(-λ*GeoDistance(candidate$vcard:latitude, candidate$vcard:longitude, currentPosition$coordinates:latitude, currentPosition$coordinates:longitude)), 0.2)</score_updt>
  </action>
  <def_score_updt>0.2</def_score_updt>
</rule>
```

Different criteria used in the algorithm, in the form of rules, may not have the same importance, or simply the domain expert may not want to assign the same relevance to every rule. This can be accomplished by weighted rules, implemented by raising the lower bound of the updating score factor. Thus, a lower bound of 0.3 has twice the ability to lower the score of elements than a lower bound of 0.6.

Although the language supports general-use operators, both logical and mathematical ones, they can be inappropriate for difficult calculations. An extension mechanism has been defined by creating user-defined functions encapsulated in Java classes. The “GeoDistance” function used in the depicted rule is a good example of a user-defined function.

This approach has some advantages compared to traditional matchmaking approaches, such as (Cali et al, 2004) and (Li and Horrocks, 2003). The most noticeable one is that this correlation engine does not impose any specific representation but the language, OWL/RDF, so it is easily deployed on any semantics-enabled project. It is also highly scalable since it only needs to have the precise information that will be used by the rules in memory instead of the whole database.

5 LOCATION AND CONTEXT CORRELATION

Correlation in Location Based Services is a complex task since the available location information may come from different sources, especially in the mobile Web, where the diversity of devices and the strict restrictions do not allow for an easy homogenization of the location data. Moreover, selecting or discovering the most suitable services or contents based on location data is not enough in a so changing environment as the mobile Web is.

Location information is only a small part of the context, which can be described as any information that can be used to characterize the situation of an entity and is considered relevant to the interaction between a user and an application (Dey, 2001). Since the context and the environment influence mobile browsing at a higher degree than desktop browsing, LBSs, aimed at enhancing the usefulness of services and improving user experience, should broaden the spectrum of contextual information used in correlations. These should take into account every piece of information that can affect or influence the consumption of a service or the usefulness of the content.

Integrating the whole context in the correlation process allows for the definition of both location-dependant and location independent criteria which are needed. Location-dependant criteria represent the guidance of the traditional correlation in LBS. They comprise all the ideas of correlation with location data such as distance, presence in the same city, building or room, to be within or out of sight, etc. Location-independent criteria are all the criteria that are not related to location information. Some of the location-independent criteria can still be used in LBSs correlation, such as the battery level in case of services which consume a large amount of energy, or the information about the device's capabilities that are needed to access a service, such as bluetooth. Nevertheless, correlation should not be constrained to any set of information. The actual data and criteria used come from the domain and depend on what is going to be correlated, so no restrictions should be laid.

One or more rules are deduced from every criterion. Different conditions of use generate

}
different condition-action pairs. As the criteria used are orthogonal, each criterion is independent from the rest and defined on its own. The rules are also independently defined. However, the integration of different rules in a single correlation implies the use of the weights mechanism to ponder their combined effect. As a result, completely different ideas of what are the correct elements to offer the user in a given context, are combined giving a single result that has been decided by all of them.

6 EVALUATION

The evaluation of the framework will be based on an implementation, which will be fully integrated within the MyMobileWeb platform. Once the framework is implemented a first mobile service offering tourist information will be launched. At the current state we are in discussions with partners to specify and design what are the minimum features for such a tourist information system. Once the ideas from our partners are evaluated we will then define the first minimum setup for the system. We expect to have a running system in the second quarter of 2009. In the beginning it is likely that not all location input channels will be fully functioning and the correlation engine will only just handle simple correlation. As the MyMobileWeb project is an open source project we believe to build a community of developers working on more on more plugins. The tourist information scenario gives us great possibility to test and proof the concept, as a lot of people with different background will use and test the system. We believe that usability issues, besides technical issues with very uncommon mobile handsets, will be the biggest points we will have to face in the test setup. With the feedback it will be possible for us to see where the system might need some redesign.

7 RELATED WORK

Context-awareness has been the focus of many research efforts. Most of the available toolkits focus on gathering, aggregating and providing context information.

(Biegel and Cahill, 2004) present a framework for developing mobile, context-aware applications. Within their framework Communication, happens only in one direction. In our approach we have the possibility to build a two way communication between the backend and the mobile handset, therefore offering a flexible way building new mobile services.

(Costa et al., 2004) designed a platform for mobile context-aware applications. Context information is shared by subscribing to this platform using the WASP Subscription Language. Their framework as such offers a very big flexibility in how to build and setup mobile services, but lacks of information how to correlate contextual data. In our approach correlation of contextual data is a built in concept.

The Solar middleware by (Chen and Kotz, 2002) provides a platform for context aware mobile applications consisting of one star and several planet nodes. Client applications need not collect, aggregate, or process context themselves but subscribe to context changes at the central star. The approach of subscribing to context changes at any provider, which in our case will be the backend of the MyMobileWeb environment, is an idea we might introduce as well in our framework.

There are numerous other concepts, just as from Sørensen et al. and Hinze et al. that also introduce the idea of subscriptions.

Traditional approaches to correlation were defined as semantic matchmaking algorithms. All of them are based on a certain semantic description of the items to be matched, so they are not easily applied to a desired context. One of the most relevant ones is (Cali et al, 2004) which was defined for matchmaking advertisements with requests and describes four levels of matching: exact, request being a sub-concept of advertisement, advertisement being a sub-concept of the request, not-null intersection and disjointness. (Colucci et al, 2003) is focused on Web Service discovery and also defines similar levels of match, but introduces a ranking function that can state differences between the elements classified in the same group. Finally, (Li and Horrocks, 2003) define a mechanism based on concept abduction and contraction as another ranking function for personnel selection purposes.

There are a lot of concepts and ideas how to build context aware mobile services. To the best our knowledge none of the concepts is focusing mainly on an architecture to build context aware mobile services that can be used with various handsets and that introduce the power of correlating contextual data in an extent we do.

8 FUTURE WORK AND CONCLUSIONS

Building context aware mobile services is a massive upcoming market. Location Based Services as an
example of context aware mobile services are a key driver of today's telecom business and is growing rapidly. Nevertheless, building context aware mobile services nowadays is hard, as the underlying mobile handset technology varies from one model to another. Within this paper we presented a framework that tries to solve this problem. The framework will be implemented within the project MyMobileWeb. It presents a solution how mobile services can be build to cover a wide range of handsets and therefore offering the possibility to build context aware mobile services for the mass market. The framework introduced the concept of various input channels, which means that any contextual data a mobile handset offers can be retrieved. For this the architecture introduced the concept of plugins. The description of the delivery context used in this framework is based on the W3C Delivery Context, which we extended to our needs. The Correlation Engine created in the MyMobileWeb project is aimed at correlating any kind of semantically described content taking into account all of the contextual information available. Its development has emphasized the domain independence, so it can be used to exploit any ontology-based semantic description of domain elements for arranging, selection or discovery purposes. This approach increases the correlation possibilities for contextual data for mobile services.

We are now working on the reference implementation of the framework. Within 2009 the first location based services will be built and used with this framework. Besides the implementation we will work on conceptual ideas how to extend the framework.

We believe that our framework offers a very attractive way for anyone who is interested in building context aware mobile services for the mass market.

ACKNOWLEDGEMENTS

This work is supported in part by the European Social Fund and Madrid's Regional Government under their Research Personnel Training programs. Furthermore this work is supported by the Austrian Funding Authorities under their Basic Founding Program.

REFERENCES


