Semantic-Based File Annotation on Mobile Device

Viresh. N. Vanarote, P. M. Kamde

Abstract— The mobile technology improved in the development of mobile operating system and storage capacity it also brings new challeges for user to find the files on the mobile device effectively because of large number of files are stored on mobile device. The file annotation and retrieval framework (FARM) proposed in the paper automatically annotate the files with their basic file attributes by extracting them from the underlying operating system of the device. A file is searched by matching the search query with the stored meta-data which means that any of the exact field from metadata is required to get the search successful.

Index Terms— File Annotation, J2ME, kXML Mobile Devices.

I. INTRODUCTION

Usage of mobile technologies cannot be neglected which plays a vital role to cope with our daily life. As technology advances, mobile phones are now commonly equipped with larger storage capacity and added features like navigation system, camera and additional communication interfaces. These stored files may any data types likes .pdf, .doc, ,png, ,jpeg.

When a user of particular device uses such kind of feature and application, usually it generates many files. It becomes a challenging issue to search efficiently and effectively for files on the device itself or in a networked environment. The meta-data analyses are also presented to show its significance while searching a required file. Meta-data is extracted from the underlying file system of the device and each file is decorated with its attributes. The associated meta-data is stored locally in Extensible Mark-up Language (XML) format, which means that the framework is not contingent to communication medium for neither storage nor retrieval of meta-data. The process of annotation automatically annotate the each file with three basic attributes however two optional tags can also added. A compact version of XML is used in order to parse the XML data in search module.

Meta-data is extracted from the underlying file system of the device and each file is decorated with its attributes. The meta-data is subsequently stored locally in Extensible Mark-up Language (XML) format, which means that the framework is not contingent to communication medium for neither storage nor retrieval of meta-data. Annotation process automatically annotates each file with three basic attributes however two optional tags can also added. A compact version of XML is used in order to parse the XML data in search module of the framework. File can be searched automatically through any of the attributes on device itself or even on other connected devices. The meta-data of all annotated files can also be viewed as a complete list. Performance evaluation shows that the parser does not degrade the search performance in the framework.

II. RELATED WORK

Information retrieval has been a challenging research issue in recent years because of the huge expansion in information resources and technological advances. The information retrieval system is one of the most successful method that is used to annotate the data for giving additional descriptions of the archived information. An element of metadata describe the information resource, or helps to provide access to an information resource and its purpose as Whether in the traditional context or in the Internet context, the important purpose of metadata is to provide and improve the retrieval of information.

Metadata was found to be valuable in the previous bibliographic retrieval systems, in which online information can be used to access via associated metadata. Its use was further extended in various methods for the management and retrieval of text, images, multimedia repositories web documents and file systems .Current researches shows some different techniques to handle file management and its efficient retrieval . These efforts are also much needed on handheld and resource-limited devices such as PDAs and mobile phones. Various techniques and research directions have been explored recently including semantic aware file system which is one of the most agreed alternatives to traditional hierarchical file system and extended in [6].by analyzing the meta data user can search efficiently. Meta-data is getting from the underlying file system of the device and each file is decorated with its attributes. The meta-data is also stored locally in XML format, which means that the framework is not dependent to communication medium for neither storage nor retrieval of meta-data.

Annotation process annotates each file with three basic attributes however two optional tags can also added. A compact version of XML is used in order to parse the XML data in search module of the framework. File can be searched automatically through any of the attributes on device itself or in network .Files on low-end devices are stored in hierarchical directory structures with application-specific default naming settings. For instance, a mobile phone camera generates an image file and usually names it as image001 . These are non-descriptive and very complicated for users to remember while searching the files. A file is searched by matching the search query with the stored meta-data which means that any of the exact field from metadata is required to get the search successful. The limited input and output result make worse for the user to interact with device. Such a challenging issue research is required to provide a realistic and effective approach to handle the following limitations such as Limited output resources, Limited computing resources ,no straight forward mechanism exists. The main purpose of this research is to find the practical approach retrieving files on low-end mobile and portable devices with minimum effort required by its users in the various aspects such as Compatibility ,User-friendliness, Intelligent and also the search on other connected devices.

III. PROGRAMERS DESIGN

Nowadays Mobile devices have larger memory capacity to store large amount files and data. Suppose user want to find a particular file from mobile memory location then user should do the lot of manual work and its time consuming to find the file .On mobile device user store any type of file because of improvement in mobile technology. To overcome these problem file annotation process is help to user to find the file from memory location. In an annotation process mobile operating system interact with the attributes of the corresponding file. All attributes of the file parsed with kXML parser and stored in XML format. The annotation process using this XML meta data for search the file. For searching file user uses the input in the form of attributes by using these input given to the annotation process to annot the particular file.

Where,

FA= File attributes

Ud= User data

Xd=Xml data

Output can be shown as O={F,Xu,Xm}

Where,

F=Searched File

Xu= Xml data updated

Xm=Xml meta data

Success state : If file is present on mobile device and user found the file through this annotation process then the process is successful.

3.1. Mathematical Model

The system S for above description can be given as,

S={I,O,SO,Xm,Xu,Bpd}

Where

I=Set of input

O=Set of output

SO=Search option

Xm=Xml data

Xu=Xml update data

Bpd=Bluetooth paired device

The file read option can be shown as

FRo={ Ann,Bs}

Where

Ann=Annotation process

Bs=Bluetooth system

The annotation process can be shown as follows

Ann={Af,Um,As,Xm,}

Where

Af=Annotate file

Um=Update meta data

As=Advanced search

Xm=xml meta data

The Annotation process :

In annotation process mobile device uses the operating system to create the file attributes. User can give the input in the different form for annotation. File is searched by using default attributes like File name ,File size ,File creation date and optional keyword and distribution which users have to type in for each file. By using the xml parser parse theses all attributes and stored in xml format. when user want to search the file it checks in xml data file if file is available it gives the output to the user.
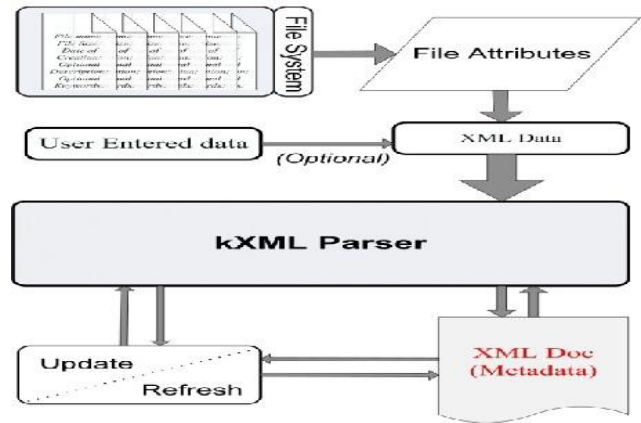

Figure 1: Annotation process

Search Module:

Search module plays very important role in searching the files and data. user search the file by using the file attributes. Mobile file system create by default three attributes. while two additional tags can be appended through Optional meta-data part. These two optional tags are Keyword and Description, which users have to type in for each file. As mentioned in files are annotated with three automatic and two optional tags, search can be performed with different options using available attributes as shown in Fig.2 However the two optional tags can only be used as search option if that file is annotated with the optional tags.

In figure 3 shows the search option by using File name, File size, File creation , Keyword , Description. Select any potion for file searching and find the particular file. First three are the default attributes and remaining two are optional but when user search the file bu usung keyword and distribution then actual file searching more faster so that these are very important.

Bluetooth module:

Bluetooth is a short range, low power and low cost communication and its mostly available in new mobile devices so by using Bluetooth user connect with other mobile device and paired with that device and find the file on paired device also. Bluetooth protocol is standardized which gives good control for stack initialization, device management, device discovery, service discovery, and communication. Service Discovery Protocol (SDP) is used. This protocol is to discover the nearest devices. And files shared by other users in the network. It also called (PAN) personal area network.

Figure 2: Main Display
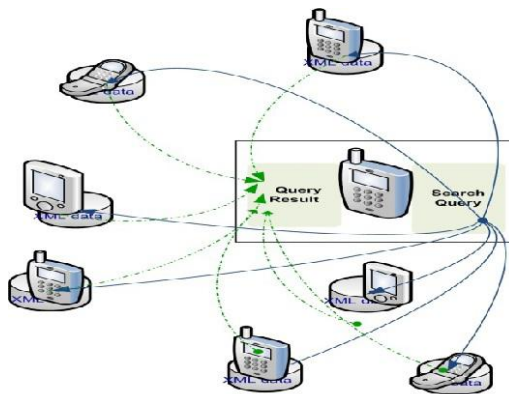
Figure 3: Search option

Figure 4: Bluetooth Environment

In figure 4 shown file searching by using Bluetooth paired devices.

XML format :

Why XML? XML (Extended markup language) is platform independent language it run on any mobile device with very low memory capacity and Xml have parser to parse the data (kXML). Universally supported specification for exchanging document and data across applications and platforms kXML parser for processing XML document in order to search for information about the required file. kXML is a lighter and compact version of XML parser which is specially designed for low-end devices and exclusively used on JME platform.

```
<? xml version="1.0" ?>
- <start>
  - <File>
      <FileName>nature.PNG</FileName>
      <FileSize>36031</FileSize>
      <FileCDate>Sat July 04 22:35:57 GMT 2009</FileCDate>
      <KeyWord>trees, green ,picture</KeyWord>
      <Description>Visit to a nice place on my 30th
          Birthday with my family and friends </Description>
  </File>
  - <File>
      <FileName>classnote.doc </FileName>
      <FileSize>278</FileSize>
      <FileCDate>Mon Jul 06 11:06:15 GMT 2009 </FileCDate>
      <KeyWord>J2ME, WNCC, notes </KeyWord>
      <Description>Lecture notes</Description>
  </File>
</start>
```

Figure 5: XML tags and Data format

J2ME :

J2ME is the powerful platform for developing mobile phone application[11]. J2ME develop standards for Java technology and they defined various configurations and profiles to fulfill diverse requirements of application developers. J2ME uses a compact version Kilo Virtual Machine (KVM). A complete working model of the proposed framework is implemented in J2ME and consists of several modules to perform different functions. It is compatible with all Java enabled devices which runs Java Virtual Machine. Samsung, Asus, Motorola, Panasonic, iball, nokia Nextel and many more have Java enabled mobile phones. Different hardware configuration on these small computing devices was a challenging task for Java Community Process (JCP) which is a mechanism to develop standard for Java technology, however the challenge was successfully overcome by defining Configurations and Profiles.

Data Flow architecture: Mobile system are comes with improved operating system then the file searching is some difficult so by using the file annotation technology user find file in les time. In file annotation process there are many process are done for finding the file.
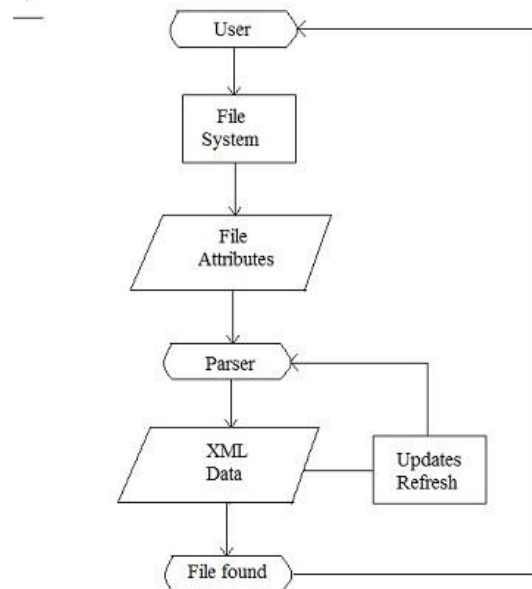

Figure 6: Data flow Diagram

IV. RESULTS AND ANLAYSIS

To evaluate the significance of meta-data in our proposed framework, we formulated two set of tests comprising of 40 search queries. For the first set, all files were automatically annotated and the optional tags were also added, however for the second test, files were only annotated automatically and additional tags were left empty. Results indicate that success rate is 84% for the set which was full annotated and 72% for the second set.

The significance of user-typed additional tags can be measured by the difference between two sets of results, which is 12% in this case. We can use probabilistic evaluation to generalize this implication. A file search query may have two possible outcomes either success or failure, therefore, based on value of p and q, a generalized model is presented to quantify the efficiency of search mechanism based on Binomial Distribution $b(x; n, p)$.

The graph shows that the probability of getting a successful result is higher for P which is 70% and above as compared to Pa. It is more likely to get a successful result if files are annotated with the optional tags along with automated meta-data tags.
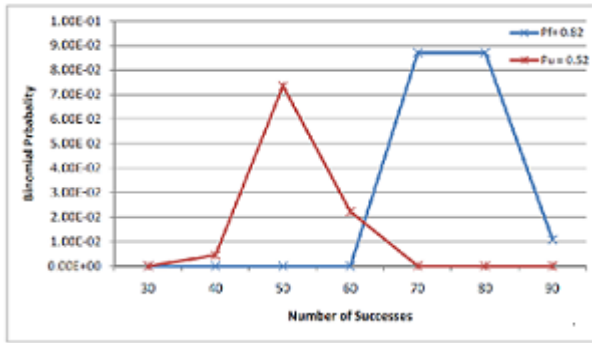

Fig. 7 Probability Comparison

V. CONCIUSION

To annotate and share files on resource limited handheld devices. A file is searched by matching the search query with the stored meta-data which means that any of the exact field from metadata is required to get the search successful. Various techniques will also be analyzed to create and update ontology and its role in file search. the emerging issue of managing large number of files on low end mobile devices and proposed practical frameworks for facilitating the file retrieval. The main reason for creating metadata in any information retrieval system is to facilitate the discovery of the required information and its use has proven to be very effectual. The FARM framework presented in this thesis automatically annotates stored files with their corresponding attributes. These attributes are considered as meta-data of files and parsed through a tiny parser to assemble them in XML structure. The annotation and file retrieval mechanisms proposed in FARM were   cautiously designed keeping the resource limitations of the devices in mind. Related graphs of the relative comparison of algorithmic efficiency parameters concluding the result of the research and future scope in enhancing the data parameters and technologies.

REFERENCES

1. W. Cathro, "Metadata: An Overview," in Standards Australia Seminar: Matching Discovery and Recovery,http://www.nla.gov.au/nla/cathro3.html,last accessed May (2011).
2. A. Sen, "Metadata management: past, present and future", Decision Support Systems 37 (1), pp. 151-173. (2004).
3. S. Handschuh, St. Staab, F. Ciravegna, "S-CREAMsemi-automatic creation of metadata", in: A. Gomez-Perez (Ed.), The 13th International Conference on Knowledge Engineering and Management (EKAW-2002), Springer Verlag, (2002).
4. M. Tuffield, S. Harris, D. Duplaw, A. Chakravarthy, C. Brewster, N. Gibbins, K. O'Hara,F. Ciravegna,D. Sleeman, N. Snadbolt, Y. Wilks, "Image Annotation with Photocopain" in the Proceedings of the fifteenth world wide web conference (www06), Edinburgh, May (2006).
5. J. Tang, X.-S. Hua, G.-J. Qi, Y. Song, and X. Wu,"Video Annotation Based on Kernel Linear Neighborhood Propagation," IEEE Transactions on Multimedia, vol. 10, no. 4, pp.620-628, (2008)
6. V. Lanfranchi, F. Ciravegna, D. Petrelli, "Semantic Webbased Document: Editing and Browsing in AktiveDoc",Proceedings of the 2nd European Semantic Web Conference,Heraklion, Greece, May 29-June 1,(2005)
7. F. Amardeilh, "Semantic Annotation and Ontology Population". Semantic Web Engineering in the Knowledge Society, ISI Global, (2008).
8. J. Vincent, "Emotional Attachment and Mobile Phones", Knowledge, Technology, Policy, Vol. 19,No. 1, pp. 39-44, (2006).
9. Extensible Markup Language, http://www.w3.org/XML(2011).
10. Parsing XML in J2ME, by Jonathan Knudsen.