Metadata Harvesting Framework in P2P-Based Digital Libraries

Hao Ding and Ingeborg Solvberg
NTNU/IDI, Trondheim 7491, Norway
Tel. 47-73594168
Fax. 47-73594466
Mail: {hao.ding, ingeborg}@idi.ntnu.no

Abstract: This paper reviews the main efforts in the research of harvesting metadata records in distributed digital libraries environment. The advantages and remained issues are briefly presented in order to have a comparison in our new approaches. A general scenario of integrating metadata in Peer-to-Peer based environment is introduced with respect to three situations, such as sharing a common schema; different schema but in the same community; different schema and community. There are many issues in the framework, whilst in this paper we present our approaches in extending OAI metadata harvesting protocol to fulfill the requirements in P2P network and adopting popular inference engine for heterogeneous schema mappings. Additionally, a rough idea is also presented for locating peer in such a framework.

Keywords: digital library; metadata; Peer-to-Peer; services

1 Introduction

Digital Library (DL) now faces a dilemma. Library systems are usually based on various metadata formats and built by different persons for their special purposes. Instead of requiring the users to search individual collection separately, people agree upon that providing a unified interface for a relatively complete view for the users is essential. Such dilemma forces us to reconsider the development and integration of building blocks and services for the future DL architectures. Almost six years, some DL researchers consider that this vision is very difficult to realize, because it is not clear how to describe arbitrary functionality so that other components can inspect the description and 'decide' automatically that this functionality is appropriate for a given task and what all the parameters are intended to convey [2]. From the arguments we can find that the main problem lies in how to render different annotating information, such as metadata formats and classification schemes, interoperable within a distributed and heterogeneous environment.

The rest of the paper is organized as follows. Section 2 introduces related works in integrating heterogeneous information. Section 3 presents a general P2P architecture which is devoted for our research scenario. Section 4 and 5 present our approaches in detail, which covers the extension of the OAI-PMH protocol for the P2P-based metadata harvesting and how to use JENA inference engine for varied metadata schema mapping. Finally, we introduce our current work in locating relevant peers in such a large and hybrid network.

2 Related Work

Based on characteristics such as structure, scalability and capabilities in interoperating with other heterogeneous metadata, we group existing library systems into three categories, although some of them may coexist within one library services.

1) Bibliographic metadata based method: this method is used to integrate information resources into a library system by embedding the access information within the metadata information. For example, the element 'identifier' in the Dublin Core Metadata Set (DCMS) can contain a URL/URN that links to the full text access information. However, this method is constrained to specific metadata set and it is extremely hard to interoperate with other information resources built in different metadata standards.

2) Database browsing and navigation based method: this method is a widely used one. Similar to the OpenDirectory and Yahoo!, this

kind of library systems classify the collections/databases into different categories according the predefined subjects, media types, or even the alphabet. The users will find it easier to find relevant information if he wants to something within a special subject. But, meanwhile, the users will find the returned information is quite limited (That is, the recall is low), because generally, such systems have to take manual or semi-automatic efforts to organize collections. Furthermore, the search result or link does not return the full text of an article to the user but instead provides the access link to the particular collections.

3) Global as View (GaV) based Method: in order to integrate and search cross heterogeneous resources built in different types of metadata, one simple idea is to create a global schema as a view over local ones [9]. Figure 1 illustrates the approach. A practical example is the NSDL[1] library which has generated a global metadata set. This global set contains 9 different metadata formats and furthermore, each of them is mapped separately into qualified DC which is adopted to annotate the records returned to the users. Here, the query transformation can be heavily reduced to rule deduction. However, the global view has to be modified whenever a new metadata set (e.g., a heterogeneous resource) is added [7].

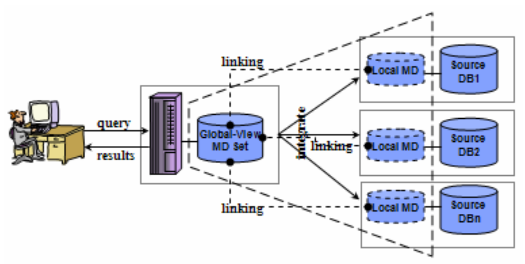

Fig. 1 Global as View Method

4) Local as View (LaV) based Method: Opposite to the GaV method, the Local View based method [9] does generate a union of various available metadata formats. Contrarily, it leaves the individual metadata to the data provider side, while it just distributes user queries to all digital collections, gathers results from each of them and delivers the list of search results back to the users' browser. Figure 2 illustrates the approach. Many examples can be found from the Web-based search engines, such as the MetaCrawler[2]. One argument for this method lies in the difficulties in generating a global metadata format. The advantages of a meta-search are that one search can highlight the strengths of many top search engines like Google, Yahoo! and AllTheWeb.

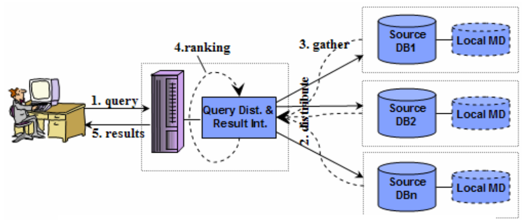

Fig. 2 Local as View Method

3 A Super-Peer Based Architecture for Metadata Harvesting over Heterogeneous Resources

Section 2 represents the mainstream in integrating digital collections. Each of them has its strong points and shortcomings. To meet the increasing requirements of library services in terms of scalability and interoperability, at least two key issues must be considered. One of them lies in the topology of the system infrastructure; the other focuses on the autonomous understanding of the complex semantics/metadata formats used in current library systems.

From the scalability perspective, the Peer-to-Peer (P2P) based infrastructure has received much attention because the decentralized nature of P2P computing infrastructure makes it ideal for economic environments that foster knowledge sharing and collaboration as well as cooperative and non-cooperative behavior sharing resources [13], such as a distributed digital library environment. Actually, the popularity of Napster (MP3), Gnutella, and eDonkey proves that P2P computing architecture has evolved as a new method for digital collections sharing. In contrast to the client/server computing model, each participating peer (e.g., a library system) can be the data provider and consumer in a P2P

---

[1] http://www.nsdl.org

[2] http://www.metacrawler.com/

network [3]. So, the compatibility with P2P architecture is expected since the library system itself needs to harvest relevant digital resources to enrich its repository.

A rendezvous *Super-Peer* based P2P architecture for metadata harvesting has been illustrated in Figure 3. The aim of such topology is to improve the scalability by reducing propagation traffic. [5] introduces the advantages for selecting super-peer networks are that they can strike a balance between the inherent efficiency of centralized search, and the autonomy, load balancing and robustness to attacks. Furthermore, some open source P2P frameworks are already available, such as JXTA, which can be reused for the implementation.

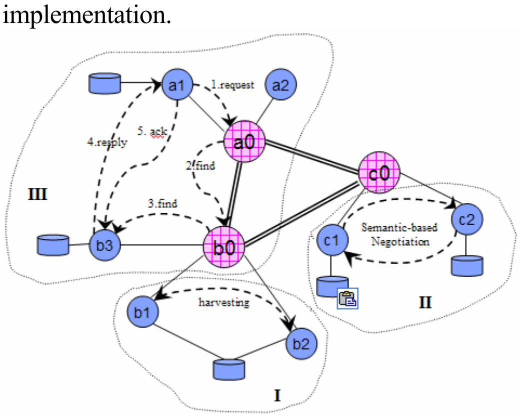

Fig. 3 A P2P architecture for Distributed DLs. Lattice nodes represent super-peers, solid nodes represents clients. Scenarios are categorized by the dashed lines.

In Figure 3, there are three specific scenarios. The first one describes two digital collections which share a common metadata format, e.g., Dublin Core. An application-independent interoperability framework based on metadata harvesting, the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH[3]), has been widely adopted in the library communities recently. However, it focuses on the client/server model and we can not 'import' it directly into the P2P computing environment. Section 4 introduces the work in extending the OAI-PMH protocol to meet such requirement.

Scenario II shows that two peers have different metadata formats but they locate in the same virtual community. In order to understand and harvest metadata information from each other, we introduce domain ontologies to support the mapping between heterogeneous metadata elements. We assume here that these digital collections are content-related or belong to the same domain, for example, two scientific paper collections built in different schema. Section 5 discusses this scenario in detail.

The most complicated scenario comes in scenario III where systems are created in different metadata formats and belong to different communities as well. An efficient mechanism is proposed in Section 6 to cluster relevant peers without regard to whether they are located in different community (fig. 3) or not. For example, if super-peer 'b0' has relevant digital collections to meet a1's queries through super-peer 'a0', 'b0' and 'a0' should be clustered together in the index space.

4 Extending OAI-PMH to P2P Network

Because of the low application barrier capability of OAI-PMH protocol, it has been widely adopted in quite a few institutes for harvesting metadata records from other digital collections. Although few digital resources are free for access, many of their metadata resources are or will be open for harvesting freely. However, few activities work on the extending OAI protocol with P2P concepts except Lagoze's vision in the OAI annual conference [6] and Edutella project [4].

Figure 4 shows the approach. In the original OAI-PMH model (fig.4.a), the 'service providers (SP)' act as the role of 'mediator' since they harvest metadata records from 'data providers (DP)' to local collections and then respond the users' queries. The problems in such topology are: the data in the SP's side are not up-to-date; the personal-level (small-sized) data providers may easily be overwhelmed by the large-scale data providers, that is, the service providers may never harvest them. In P2P network (fig.4.b), in order to solve the above problems, the SP functionalities are weakened. That is, users can directly request the DPs without dropping into SPs although users still maintain the capabilities to request SPs. Additionally, data are always up-to-date for the users and user queries can reach all available small sized DPs, Since DPs and SPs

[3]http://www.openarchives.org/

are loosely couple, they are allowed to 'join' and 'leave' freely.

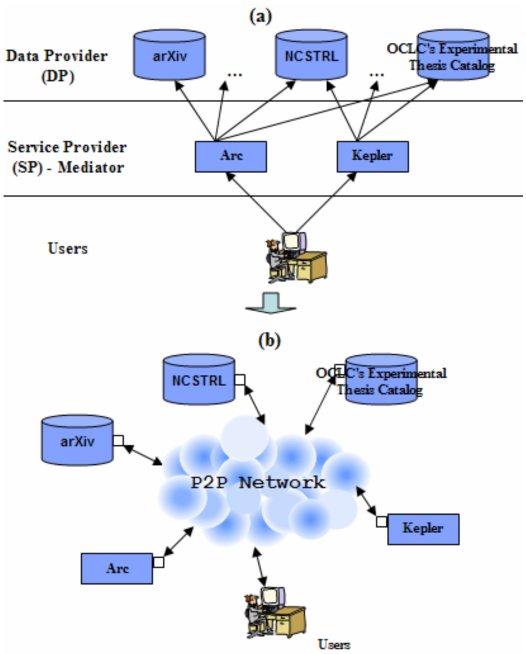

Fig. 4 (a). Topology of the typical OAI-PMH protocol; (b). Harvesting scenario in P2P network

Nejdl [4] describes two kinds of design considerations to enable an OAI DP to become an OAI-P2P peer. The first one is to wrap the DP with a peer which replicates the data to a RDF repository. The second one is to answer queries directly from the DP's database. Both of the two methods are appropriate to support the harvesting when two peers share a common metadata schema. However, because few DPs in the real-world development are willing to add an extra layer to their softwares in terms of the security, reliability issues, there remains the main drawback in implementing DP wrappers. But, considering the other side of the story, the user may not feel bothered to install a wrapper on his side if such 'transparent' component is able to help him to find more information. Figure 5 illustrates an example for the communication among three roles of user, user wrapper, and DP. In figure 5, it is assumed that both of the wrapper and DP support DCMS. The user sends queries as that in the traditional way while the wrapper transform the query to a format understandable by the specific DP. Consequently, the user can always retrieve the 'fresh' data from DPs. Furthermore, the wrapper can also be deployed to harvest data from other DPs into the local repository or cache which acts as the functionalities of SPs.

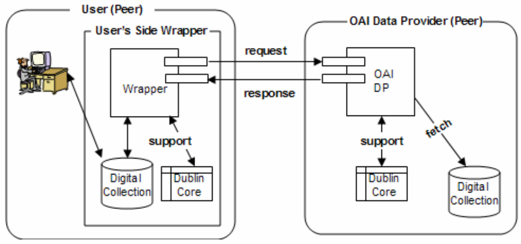

Fig. 5 P2P-based Harvesting with user's side wrapper

We introduce OWL/RDF [11,12] (cf. response example) is to express relationships among metadata records which have been harvested into the metadata repository. In a monotonous repository, it is necessary to figure out whether there are some equivalent records because there do have duplicated records floating around different DPs. In order to enhance the system reliability or find the data provenance, it is necessary to set up the relationships among relevant metadata records.

Additionally, there may also have metadata records which are harvested by other protocols besides OAI. It is unavoidable that the repository may contain metadata records in various formats. In order to alleviate the heterogeneity and improve the searching precision, it is necessary to set up a crude level relationship among those relevant metadata records. Section 4.3 will come to it.

5 Mapping between Heterogeneous Schema in P2P Network

Different collections may be built upon various metadata schemas. In the library community, even one specific collection may have several alternative metadata formats to choose from. Actually, having different schemas brings us many conveniences in describing specific collections for specific purposes. For example, if we want to annotate a set of records for scientific publishing purpose, we can adopt several alternatives, such as, Dublin Core, MARC and EAD, etc. However, problems may arise in combining underlying metadata schemas. It mainly comes from that objectives of some repositories differ from each other heavily, for instance, while combining terms used for medical diagnosis (the art or act of recognizing the presence of disease from its sign or symptoms) and health care (to help average people in identifying or preventing or treating illness or disability). Some may argue that mapping among multiple resources is possible if we can merge the terminologies with as much as possible. And one on-going work [6] uses Fedora [4] as a way of re-presenting the relationships between digital objects, and then synchronizing the relationships store. We agree with it if such directly mapping technology is adopted in a constrained domain-specific application, but for a large and distributed environment, such as P2P network, it is worthwhile to re-consider the infrastructure because for large terminologies the mapping efforts are also large [23]. Problems lie partially in that it is difficult for the layman to differentiate the tiny varieties between two similar terms resided in different schema. That is, the mapping has not come out of the blending of syntax and semantics. For example, should we know exactly the meanings of *AllowableQualifier* or Abbreviation in MeSH? How could we efficiently map *RecordOriginatorsList* and *DateRevised* into the conventional terms, like Contributor and Date not make users confused?

Dublin Core receives much attention at this time. It was designed to intersect with varied schemas, whilst we also believe that it would act as a reasonable alternative for lightweight and real-time applications that extract partial information from records created in different but more complicated metadata schema. However, [20] indicates that it can not serve as a core mediator among several heterogeneous schemas because many tests [10] have forces us to think that Dublin Core would be *problematic* in this role. The main arguments are that, firstly, the limited and simple terms in Dublin Core lead to the loss of precision unavoidably. The improvement in the computation capability counteracts its expressive power more or less. Secondly, Godby and et. al. [20] also find that the goal of interoperability is even compromised by the ambiguities and overlapping in the fifteen elements.

[4] http://www.fedora.info

We propose to alleviate the aforementioned problems by adopting the concept of upper-level ontology. Basically, if exact semantic identity is lacking, terms can be unified in a higher-level, and information that is possibly related can be retrieved as well [15]. Concretely, Guarino indicate four different types of ontologies as described in Fig.6.

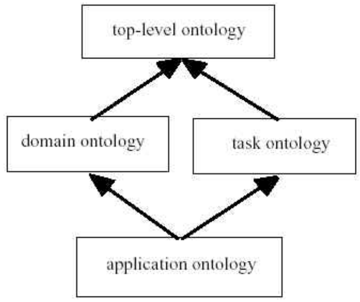

Fig. 6 Kinds of ontologies, according to their level of dependence on a particular task or point of view. Thick arrows represent specialization relationships. From [22].

The relationships are described briefly as follows:

– Top-level ontologies are independent of a particular problem or domain: it seems therefore reasonable to have unified top-level ontologies for large communities of users.

– Domain ontologies contain concepts which represent a specific domain defined in the top-level ontologies.

– Task ontologies describe concepts/controlled vocabularies for a special task or activity which is defined in the top-level ontologies.

– Application ontologies describe concepts depending both on a particular domain and task, which are often specializations of both the related ontologies.

Under such framework, and especially for such a scenario of P2P-based libraries, we adopt in our project a comparably complicated/complete metadata schema, say, standard MARC, which serves the role of *core mediator* in a specific domain whilst Dublin Core is adopted as the *upper-level mediator*. The arguments come as follows: It is generally resided in the super-peer because it is supposed to have more computational resources. The aforementioned problem of the loss of precision will be alleviated during the mapping procedure because the more self-contained schema will definitely have more capabilities to differentiate tiny distinctions. In another word, the items in various schemas can be linked according to certain rules or semantics, such as synonym, related terms, et al.. The advantage of adopting Dublin Core again comes from its partial mapping interoperability between heterogeneous schemas. In fact, if an average user searches in schema differed but content relevant collections, he may be just interested in several critical metadata information instead of the completed one from the affiliated collections. The simplicity and flexibility of Dublin Core elements for the partial mapping between heterogeneous schemas can fulfill such requirements. In the practical project, we rewrite Dublin Core in OWL to strengthen the upper-level ontology capabilities. Actually, we also cover some qualified ones. As well, the relationships of Dublin Core elements and other schema elements can be illustrated in OWL.

```
<owl:Ontology
rdf:about="http://purl.org/dc/elements/
1.1/">
...
     <dc:title xml:lang="en-US">
The Dublin Core Element Set v1.1 namespace
providing access to its content by means
of an OWL DL Ontology
     </dc:title>
     <dc:publisher xml:lang="en-US">
The Dublin Core Metadata Initiative
     </dc:publisher>
     <dc:description xml:lang="en-US">
The Dublin Core Element Set v1.1 namespace
provides URIs for the Dublin Core Elements
v1.1. Entries are declared using OWL
language to support OWL applications.
     </dc:description>
     <dc:language xml:lang="en-US">
     English
     </dc:language>
     <dcterms:issued>
     1999-07-02
     </dcterms:issued>
     <dcterms:modified>
     2003-09-2
     </dcterms:modified>
     <dc:source
rdf:resource="http://dublincore.org/doc
uments/dces/"/>
     <dc:source
rdf:resource="http://dublincore.org/usa
ge/decisions/"/>
     <dcterms:isReferencedBy
rdf:resource="http://www.dublincore.org
/documents/2001/10/26/dcmi-namespace/"/
>
     <dcterms:isRequiredBy
rdf:resource="http://purl.org/dc/terms/
"/>
     <dcterms:isReferencedBy
rdf:resource="http://purl.org/dc/dcmity
pe/"/>
</owl:Ontology>
```

Currently, we use the JENA, an inference engine under the Semantic Web framework, to achieve automatic deduction to a certain extent. The general framework is described in Fig.7.

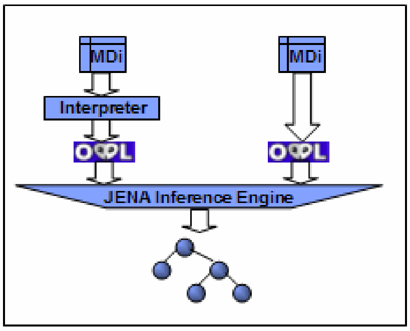

Fig. 7 Metadata Records Integrating Mechanism

The interpreter in Fig. 7 is used to transform metadata format into OWL (Web Ontology Language) [11]. Consequently, the explicit relationships among related items can be automatically generated by importing the candidate schemas.

Additionally, the explicit relationships among related terms can be generated according to the predefined rules. And, OWL also defines a special kind of resources called containers for representing collections of things. Combined with the aforementioned method, related records can also be clustered together for further information retrieval.

A simple example for the combination is shown in Fig. 8. Briefly, the pseudo-code is as follows:

```
 //read the RDF file
model1.read(new
InputStreamReader(in1),"");
```

```
model2.read(new
InputStreamReader(in2),"");


//merge the models
Model model= model1. union(model2);

//output the Model as RDF
model.write(system.out,
"RDF/XML-ABBREV")
```

Likewise, the *interaction* and *difference* of the model can also be computed in a similar manner.

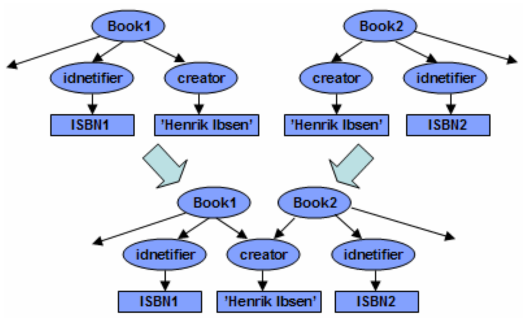

Fig. 8 Merging two metadata records into one

Furthermore, RDF/OWL [11,12] also defines a special kind of resources called *containers* for representing collections of things. Combined with the aforementioned method, related records can also be clustered together for further information retrieval.

Some other approaches appearing recently in line with the upper-level ontology-based application are: the ABC (A Boring Core) ontology [16] which is implemented for a domain-specific application on 'hydrogen economy' project [17]; the SUMO (Suggested Upper Merged Ontology) with an ambitious goal of developing a standard upper ontology that will promote data interoperability, information search and retrieval, automated deduction, and natural language processing. [18], and as well as DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) [19].

6 Integrating Heterogeneous Schema in Different Peer Communities

Coming back to the scenario III in Figure 3, where one peer in a peer community tries to discover the relevant resources in the other communities. The first and critical issue is how it can discover the heterogeneous but content related peers without losing the large-scale, decentralized and real-time search capabilities. Current approaches in integrating heterogeneous metadata records can be crudely classified as centralized and decentralized ones. The centralized approach includes the NSDL[5] library, which generate a global view for 9 different metadata formats. The Norwegian National Library, BIBSYS[6], is built upon NO-MARC which is regarded as main metadata standard each affiliated libraries have to conform to, but they are also allowed to have their own extensions. The distributed approaches are based on ad-hoc network, such as P2P system. Among the ebullient research activities, Edutella receives many attentions by adopting the concept of *Semantic Overlay Clusters* (SOC) for clustering the super-peer networks to enable the creation of context-specific, logical views over the physical P2P network topology [1].

In tackling the problem in scenario III, how to build a SOC for dynamically locating appropriate super-peers is highly appreciated. Here the point is how to generate, index and store peers' capabilities information in such a distributed environment. In order to meet such requirements, a DHT approach can be adopted for locating peers in the overall system architecture. Furthermore, in stead of simply mapping peers' identifiers to indices, a Hilbert Space Filling Curve (HSFC) [8] -based index space are used to indicate relevant super-peers. A HSFC is a one dimensional curve which visits every point within a two dimensional space. It may be thought of as the limit of a sequence of curves which are traced through the space. Curve H1 has four vertices at the center of each quarter of the unit square. Curve H2 has 16 vertices each at the centre of a sixteenth of the unit square (Fig. 9).

Simply, one unit square can be regarded as a subject category in an average library system, such as, music, history, social science, etc. There is no constraint in content of the

---

5 http://www.nsdl.org

6 http://www.bibsys.no

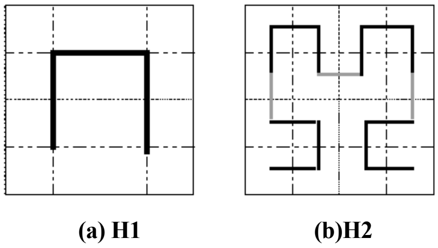

Fig. 9 HSFC approximation for 2 dimensions.

subject from the semantics level. The key issue here is how to map the functionalities of one super peer into specific subjects. Each super peer is responsible for collecting all of the content information of the sub-peers in its community, and then each community may contain one or more subjects which may match certain query. Consequently, the relevant peers are clustered in the HSFC index space. The query processing procedure can be carried out in two steps: (1) finding the relevant clusters of the HSFC based index space according to specific query; (2) querying the selected peers in the overlay network for metadata information.

In order to test the scalability of the HSFC method, a simulation experiment is carried out to test the system performance. The results are shown in Fig. 10.

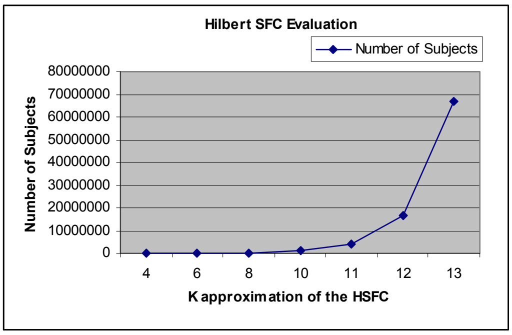

Fig. 10 Relationship between the number of subject and K approximation

In the simulation, the K approximation is increased from 4 to 13, indicating the number of subject increases from 2^{4*2} =256 to 2^{13*2}= 67108864 ('2' in the exponent means 2 dimension). The time required to process the HSFC index space increases sharply after K=12 in term of 2^{12*2}= 16777216. That is, if the number of subjects in the P2P network is less than 16777216 (2^{12*2}), this method is applicable. However, it is hard to find any real-world large-scale application which has more than 16777216 (2^{12*2}). Furthermore, as figure 10 (a) plots, the response time is promising when the subject number is less than 2^{11*2}=4194304. In short, the HSFC-based index space method is as expected when adopting into the DHT-based system architecture.

7 Conclusion and Future Work

In this paper, the current approaches have been described in the field of integrating/harvesting metadata records in distributed digital libraries. Additionally, a general framework in harvesting metadata in a super-peer based P2P network has been introduced. Three major scenarios have also been presented. Accordingly, the critical issues in these scenarios have been indicated and proposed solutions, such as extending OAI protocol for P2P network, mapping heterogeneous metadata schemas by adopting novel inference engine, have also been presented. Finally, a feasible solution of creating Hilbert Space Filling Curves index space for clustering relevant peers has been introduced. A simulation experiment also proved it is feasible. However, there are also many untouched issues, such as security and reliability in P2P network. Furthermore, there is no control of the qualities of the data providers and many trivial but hard to solve problems [10] in the conventional OAI-PMH applications will not avoidable in current system. The expectations, however, from the data providers to re-check and re-annotate their won documents are very unlikely. So, how to automatically annotate the collection with ontology-based semantics will be the future work. Additionally, many efforts will be put into improving the query mechanism in order to allow users to choose their favorite data providers.

References

1. Alexander Löser, Martin Wolpers, Wolf Siberski, Wolfgang Nejdl: Efficient data store discovery in a scientific P2P network. Semantic Web Technologies for Searching and Retrieving Scientific Data workshop in conjunction with the 2nd International Semantic Web Conference.

Florida, USA, 2003.
2. Andreas Paepcke, Kevin Chen-Chuan Chang, Hector Garcia-Molina, Terry Winograd: Interoperability for Digital Libraries Worldwide. Communications of the ACM 41(4): 33-43 (1998)
3. Andy Oram (Ed.). Peer-to-Peer: Harness the Power of Disruptive Technologies, O'Reilly, 2001.
4. Benjamin Ahlborn, Wolfgang Nejdl and Wolf Siberski. OAI-P2P: A Peer-to-Peer Network for Open Archives. Workshop on Distributed Computing Architectures for Digital Libraries, 31st Intl. Conference on Parallel Processing (ICPP2002), Vancouver, Canada, August 2002.
5. Beverly Yang, Hector Garcia-Monlina. Designing a super-peer network. ICDE 2003.
6. Carl Lagoze, Herbert Van de Sompel. The OAI and OAI-PMH: where to go from here? Presentation at the CERN Workshop Series on Innovations in Scholarly Communication: Implementing the benefits of OAI (OAI3). Available at: http://agenda.cern.ch/fullAgenda.php?ida=a035925
7. Carl Lagoze, William Y. Arms, Stoney Gan, Diane Hillmann, Christopher Ingram, Dean B. Krafft, Richard J. Marisa, Jon Phipps, John Saylor, Carol Terrizzi, Walter Hoehn, David Millman, James Allan, Sergio Guzman-Lara, Tom Kalt: Core services in the architecture of the national science digital library (NSDL). *JCDL 2002*: 201-209.
8. H.Sagan: Space-Filling Curves, Springer-Verlag, 1994.
9. L.Xu and D. Embley: Combining the best of global-as-view and local-as-view for data integration. Available at: http://www.deg.byu.edu/papers/PODS.integration.pdf
10. Martin Halbert, Joanne Kaczmerk and Kat Hagedorn: Findings from the Melon Metadata Harvesting Initiative. Pp. 58-69, ECDL2003, Trondheim, Norway.
11. OWL Web Ontology Language, http://www.w3.org/TR/owl-guide/
12. Resource Description Framework (RDF), Available at: http://www.w3.org/RDF/
13. S.H. Kwok. Decentralized knowledge reuse with peer-to-peer technology. Proceeding of the First Workshop on e-Business. Web2002.
14. Papazoglou, M.P., Service -oriented computing: Concepts, characteristics and directions. In Fourth International Conference on Web Information Systems Engineering (WISE'03), 2003.
15. S.Staab and R.Studer, Handbook on Ontologies, Springer-Verlag Berlin Heidelberg 2004, ISBN:3-540-40834-7.
16. Carl Lagoze and Jane Hunter, The ABC Ontology and Model (v3.0), J.Digital Information, vol.2, no.2, 6 Nov 2001, Article no. 77; http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze.
17. Jane Hunter, J. Drennan, and Suzanne Little, Realizing the Hydrogen Economy through Semantic Web Technology, IEEE Intelligent System, January/February. 2004.
18. I.Niles and A.Pease Towards a Standard Upper Ontology, Proc. 2nd Int'l Conf. Formal Ontology in Information Systems (FOIS), ACM Press, 2001, pp2-9.
19. A.Gangemi et al., Sweetening Ontologies with DOLCE, Proc. 13th Int'l Conf. Knowledge Eng. And Knowledge Management (EKAW 2002), LNCS 2473, Springer-Verlag, 2002, pp.166-181.
20. Carol Jean Godby, Devon Smith and Eric Childress Two Paths to Interoperable metadata,sa0 Dublin Core Conference 2003, Seattle, Washington, USA.
21. Harren, M., Hellerstein, M., Huebsch, R., Loo, B., Shenker,and S., Stoica, I.: Complex Queries in DHT-based Peer-to-Peer Networks. Proc. IPTPS (2002)
22. Guarino, N.: Formal Ontology and Information Systems. In proceeding of FOIS'98, Trento, June 1998, pp3-15.
23. Gysenns M., Paredaens J., Van den Bussche J. and Van Gucht D. A graph-oriented object database model. In IEEE Transaction on KDE, Vol.6, No.4, pp.572-586, 1994.