CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

Grid computing is a model of distributed computing that uses geographically and administratively disparate resources. In Grid computing, individual users can access computers and data transparently, without considering location, operating system, account administration, and other details. In Grid computing, the details are abstracted, and the resources are virtualized.

Resources in Grid system are heterogeneous, geographically distributed, belong to different administrative domains and apply different management policies [45]. The main goal of Grid computing is to enable collaborative and secure resource sharing over multiple organizations which are geographically distributed.

Resource Discovery can be defined as a directory service directed to the spontaneous network’s environment [26]. In these networks, resources can enter or leave at any time and this mechanism’s proposal is to provide information about the resources available in a specific moment. In a Grid environment, resource discovery is a very complex task basically for two reasons. First, the resources that are potentially interconnected to a Grid are not only computers, but also software, instruments and data, among others. This adds a very high degree of heterogeneity that should be taken into account. Secondly, a Grid can have a very huge number of resources, spread over multiple administrative domains that are geographically distributed. In this scenario, scalability issues must be taken into account. An efficient resource discovery mechanism is the fundamental requirement for Grid computing systems, as it supports resource
management and scheduling of applications. Till recently, researchers have proposed many improved schemes to enhance resource discovery process by reducing the average number of hops, messages, memory consumed and average communication time.

This chapter focuses on the review of literature for Grid Resource Discovery mechanisms like hierarchical, centralized and hybrid systems. Further, DHT based resources discovery mechanisms and P2P resource discovery approaches have also been reviewed. Plenty of ideas implemented by researchers regarding the improvement of the lookup performance of DHT based Chord protocol in Grid environment are finally reviewed in this chapter.

### 2.2 RESOURCE DISCOVERY IN GRID COMPUTING

In recent years various resource discovery mechanisms have been implemented in Grid environment [48, 29]. Table 2.1 summarizes the different Resource discovery mechanisms including extensible resource discovery scheme, hierarchical and centralized methods, page rank techniques, super-peer based resource discovery method and hybrid resource mechanism.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Carlo Mastroianni et al. 2005 [18]</td>
<td>Examined how the super-peer model can handle membership management and resource discovery services in a multi-organizational Grid.</td>
</tr>
<tr>
<td>2</td>
<td>Juan Li and Son Vuong 2005 [69]</td>
<td>Exploited a comprehensive semantic based resource discovery framework, which performs an effective searching according to the semantic properties of resources being searched.</td>
</tr>
</tbody>
</table>

Table 2.1: Resource Discovery in Grid Computing Environment

Cont…
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Kashif Ali et al. 2005 [72]</td>
<td>Classified nodes as consumers and producers, depending on whether they consume or produce more jobs and connects all producer nodes using a overlay network that is a small-world graph. The consumer nodes hang off the small world graph and the producer nodes are forced to take part in resource cataloging and resource discovery.</td>
</tr>
<tr>
<td>4</td>
<td>Noorisyam Hamid et al. 2006 [95]</td>
<td>Concerned with efficient and quality-based resource discovery using Condor ClassAd and PageRank technique in order to achieve a quality resource matching and discussed how quality of users and resources are determined and considered in the discovery process prior to allocating jobs to resources.</td>
</tr>
<tr>
<td>5</td>
<td>Tania Gomes Ramos et al. 2006 [133]</td>
<td>Evaluated an extensible resource discovery mechanism for Grids, where the basic information retrieval can be extended to include user defined specific searches.</td>
</tr>
<tr>
<td>6</td>
<td>Vassilios V. Dimakopoulos and Evaggelia Pitoura 2006 [138]</td>
<td>Conducted flooding-based resource discovery in distributed systems and three different search strategies namely flooding, teeming and random paths strategies are implemented.</td>
</tr>
<tr>
<td>7</td>
<td>Xuan Wang and Ling-Fu Kong 2007 [152]</td>
<td>Introduced a decentralized resource discovery scheme based on the Peer-to-Peer protocol and the idea of resource clustering to overcome the deficiencies of the centralized resource discovery mechanism.</td>
</tr>
<tr>
<td>S.No</td>
<td>Author(s)</td>
<td>Description(s)</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>10</td>
<td>MahamatIssa Hassan and Azween Abdullah 2008 [83]</td>
<td>Facilitated a model for resource discovery method based on self-organizing the resources with the capability of handling the dynamic resource status on the Grid environment.</td>
</tr>
<tr>
<td>11</td>
<td>N.Malarvizhi V.R.Uthariaraj 2008 [84]</td>
<td>Proposed Grid Resource Discovery approach based on client agents which act on behalf of Grid users and search for resources in a network of resource brokers that are registries for various Grid resources.</td>
</tr>
<tr>
<td>12</td>
<td>Rafael Moreno and Vozmediano 2008 [110]</td>
<td>Presented new resource discovery mechanism based on a hybrid peer-to-peer approach based on the concept of discovery zone and solves discovery delays, bandwidth consumption, adaptation to changing conditions, and management complexity.</td>
</tr>
<tr>
<td>13</td>
<td>O. A. Rahmeh and P.Johnson 2008 [111]</td>
<td>Applied a distributed and scalable load balancing framework for Grid Networks and the generated network system is self-organized and depends only on local information for load distribution and resource discovery.</td>
</tr>
<tr>
<td>14</td>
<td>Youchan Zhu et al 2008 [156]</td>
<td>Introduced a resource discovery method which combines the centralized and distributed method.</td>
</tr>
<tr>
<td>15</td>
<td>ZHENG Xiu-Ying et al. 2008 [165]</td>
<td>Recommended a resource-centric P2P model for Grid Resource Discovery and to deal with heterogeneous resource allocation policy, a society of agents is proposed.</td>
</tr>
<tr>
<td>16</td>
<td>ZHENG Xiu-Ying et al. 2008 [166]</td>
<td>Introduced a model for resource discovery among Grids based on the community categorized by application domain.</td>
</tr>
<tr>
<td>17</td>
<td>Honggang Xia and Hongwei Zhao 2009 [59]</td>
<td>Designed a novel resource discovery mechanism in Grid in order to implement the balanced distribution of resource and to improve the utilization ratio of the resource.</td>
</tr>
</tbody>
</table>

Cont…

29
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Iraj Ataollahi and Morteza Analoui 2009 [63]</td>
<td>Dedicated a solution for the vagueness and uncertainty between advertised resources and requested resources. Problems based on rough set theory.</td>
</tr>
<tr>
<td>19</td>
<td>Wu-Chun Chung and Ruay-Shiung Chang 2009 [147]</td>
<td>Provided a novel information retrieving protocol, called the Grid Resource Information Retrieving (GRIR) protocol, to obtain the precise resource status.</td>
</tr>
<tr>
<td>21</td>
<td>Deniz Cokuslu et al. 2010 [31]</td>
<td>Synthesized and analyzed some recent Grid Resource Discovery methods which are based on centralized and hierarchical systems and evaluated them by defining some qualitative criteria, and compared different classes of methods with each other.</td>
</tr>
<tr>
<td>22</td>
<td>Juan Li 2010 [68]</td>
<td>Dedicated an efficient discovery framework which organizes a Grid network by a semantically linked overlay representing the semantic relationships between Grid participants and used semantics-aware topology construction method to group similar nodes to form a semantic small world.</td>
</tr>
<tr>
<td>23</td>
<td>Shaohui Ma et al. 2010 [122]</td>
<td>Produced a novel Grid Resource Discovery algorithm, which adopts Keywords Match algorithm based on Hash Table, employees TF_IDF to product resource keywords, locates and searches resource based on Consistent Hashing Chord, access lower latency and higher credibility peer by ant colony algorithm based on adaptively adjusting pheromone.</td>
</tr>
</tbody>
</table>

Cont...
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Adil Yousif et al. 2011 [1]</td>
<td>Presented a taxonomy that facilitates identifying and classifying the mechanisms used in the implementation of Grid resource selection process, as well as describing the most significant features of Grid resource selection mechanisms</td>
</tr>
<tr>
<td>27</td>
<td>Ali Sarhadi et al. 2012 [5]</td>
<td>Illustrated the dynamic structure of GIS based on peer-to-peer model using learning automata to increase efficiency of proposed model and generate dynamic structuring of peer-to-peer organization for GIS.</td>
</tr>
<tr>
<td>28</td>
<td>Regnard Viven and Shamila Ebenezer 2012 [114]</td>
<td>Examined favorably all the resource discovery methods in static and dynamic Grid applications.</td>
</tr>
</tbody>
</table>

Noorsiyam Hamid et al. proposed resource discovery using page rank technique in Grid environment with efficient and quality-based resource discovery using Condor, ClassAd and PageRank technique [95] in order to achieve a quality resource matching. Quality-based Grid Resource Discovery (Q-GreD) is developed which aims at providing a Grid Resource Discovery which takes into account the quality and reliability of both users and resources. Q-GReD provides a better resource discovery in a challenging large-scale Grid environment. Q-GReD is concerned with selecting quality and reliable resources for job allocation.
Tania Gomes Ramos et al. proposed and evaluated an extensible resource discovery mechanism for Grid environments [133]. The possibility of extending a resource discovery mechanism, by adding new resource searches automatically and “on-the-fly”, is very useful for the user/administrator, who need not manually modify the resource discovery Grid service. An extensible search allows the dynamic and transparent creation of a new resource search. This feature is particularly interesting in a Grid environment where new types of resources can be added at any time and new sets of information arise constantly.

Bin Lu and Juan Chen reviewed the mechanism and architecture of Grid resource management model from hyper topology space [11], and researched how to describe resources, abstract resources, provide unified system images and discovering resources. And then a formal model of Grid resources with three layers, Primitive Resource Layer, Aggregating Resource Layer and Service Resource Layer, is presented.

Mahamat Issa Hassan and Azween Abdullah proposed a model of resource discovery method that is based on classifying the nodes of resources according to their importance to a Grid application [83]. Each group has a leader that collects the information of the group members. This method may satisfy the five characteristics of resource discovery which are scalability, decentralization, tolerating the dynamics of the Grid, supporting attribute-based discovery and multi-attribute range queries.

The Grid provides mechanisms to share dynamic, heterogeneous, distributed resources spanned across multiple administrative domains. The dynamic nature of resources on computational Grids raises the question of an efficient scheduling process, which mainly consists of resource discovery as well as resource selection. So, appropriate resource discovery and
selection mechanism are important aspects of Grid Computing. N.Malarvizhi and V.R.Uthariaraj proposed [84] a novel approach for resource discovery and resource selection. This approach is based on client agents which act on behalf of Grid users and search for resources in a network of resource brokers that are registries for various Grid resources. After suitable resources are discovered, client agents determine the willing resources and use the current load information and memory availability of the resources to determine the optimum set of resources for the tasks in hand. Discovering and selecting resources in this way can improve the performance of Grids.

A new resource discovery model is proposed by Youchan Zhu et al. [156] which partitioned the resource cluster according to the resource type to construct the layered resource model and introduced a resource discovery method which combines the centralized and distributed method and take the classification of the resource service as the basic thought as well as unify the resources organization and discovery mechanism of the P2P correlation technique to design a resource discovery method based on the resources classification and designed the resource organization, transmit strategy based on the P2P DHT, in-cluster and adjacent clusters two level resource discover mechanism, achieves the auto disposition, independent management, dynamic discovery and the fault-tolerant request of the resources discovery mechanism.

For better use of the resources inside existing Grids, ZHENG Xiu-Ying et al. [166] considered the available resources within one Grid to be used by other Grids. To many Grids, pooling resources together to form a virtual resource house seems natural to enhance resource utilization as well as to reduce costs. The authors discussed a scalable model for resource discovery among Grids based on the application domain. On the top of the community concept, for the efficiency they used a central model for resource discovery inside community while using
the structured P2P network to achieve the scalability of the communities. In the work of resource discovery among Grids, the ontology-based semantic description of resource should be used, and multiple agents are also used for independence on different platforms.

A novel resource discovery mechanism is proposed by Honggang Xia and Hongwei Zhao [59], followed by the analysis of the specific model and technology related to Grid, together with a Hierarchical resource management model, and then a Discovery Strategy has been designed and realized. The main purpose of this Strategy is to improve the efficiency of the implementation of Grid system. This system expresses semantic information with the application of ontology theory and Semantic Web-related specifications, acquire a common understanding form on the concept based on these semantic information, filter out irrelevant information in order to narrow the scope of query, and begin intelligent reasoning with the use of concept correlation in order to achieve the generalization of query request, thus improving the keyword-based query approach of UDDI with significantly recall rate and precision rate. This scheduling system can improve the efficiency of the resource and the utilization ratio of the resources.

Iraj Ataollahi and Morteza Analoui [63] proposed a solution for the vagueness and uncertain problems based on rough set theory. Resource discovery entity uses two steps to find resources set which satisfy the user requirements. In the first step, dependent properties have been reduced, and in the next step resource matching has been done. After the reduction of irrelevant properties, the remained properties will be sent to the dependent properties reduction entity to reduce dependent properties. This algorithm dynamically reduces dependent properties in terms of a resource request. So this algorithm can deal with uncertain properties and fined a set of resource which may maximally satisfy the needs of requested resource.
A Grid Resource Information Monitoring (GRIM) prototype is introduced by Wu-Chun Chung and Ruay-Shiung Chang [147]. To support the constantly changing resource states in the GRIM prototype, the push-based data delivery protocol named Grid Resource Information Retrieving (GRIR) is provided. There is a trade-off between information fidelity and updating transmission cost. The more frequent the reporting is, the more precise the information will be. But there will be more overheads. The offset sensitive mechanism, the time-sensitive mechanism, and the hybrid mechanism in GRIR are used to achieve a high degree of data accuracy while decreasing the cost of updating messages. Experimental results show that the proposal alleviates both the update transmission cost and the loss of data accuracy compared to prior methods.

Anand Padmanabhan et al. presented a self-organized grouping (SOG) framework [7] that achieves efficient Grid Resource Discovery by forming and maintaining autonomous resource groups. Each group dynamically aggregates a set of resources together with respect to similarity metrics of resource characteristics. The SOG framework takes advantage of the strengths of both centralized and decentralized approaches that were previously developed for Grid/P2P resource discovery. The design of SOG minimizes the overhead incurred by the process of group formation and maximizes the performance of resource discovery. The way SOG approach handles resource discovery queries is metaphorically similar to searching for a word in an English dictionary, by identifying its alphabetical group at the first place, and then performing a lexical search within the group. Because multi attribute range queries represent an important aspect of resource discovery, the authors devised a generalized approach using a space-filling curve in conjunction with the SOG framework. They also exploit the Hilbert space-filling curve’s locality preserving and dimension reducing mapping. This mapping provides a 1-
dimensional grouping attribute to be used by the SOG framework. Experiments show that the SOG framework achieves superior lookup performance that is more scalable, stable and efficient than other existing approaches.

Deniz Cokuslu et al. carried out a survey of recent Grid Resource Discovery studies based on centralized and hierarchical systems and also provided synthesis, analysis and evaluation of these studies by classification [31]. It has given a comparative study of different classes proposed and concluded that the centralized methods are not suitable for the large scale environments. But they might be well suited to the systems in which the scale is small and indexing server is reliable. In such cases centralized systems can be used effectively. On the other hand, hierarchical methods are more suitable for environments in which scale is bigger since the load is distributed to many locations. Finally, it is concluded that even the load is hierarchically distributed and they still suffer from bottleneck problems in a large scale.

Shaohui Ma et al. presented a Grid Resource Discovery model [122] with different layers, which adopts P2P (Peer-to-Peer) technology in Grid Resource Discovery mechanism. The model includes three layers based on three different peers: the Super peer_Agent, Super Peer and Ordinary peer. According to the model, this paper produces a novel Grid Resource Discovery algorithm, which adopts Keywords Mach algorithm based on Hash Table, employees TF_IDF to product resource keywords, locates and search s resource based on Consistent Hashing Chord, access lower latency and higher credibility peer by ant colony algorithm based on adaptively adjusting pheromone. It is testified by simulation implement that the algorithm can find the most suitable resources and improve resource discovery efficiency.
The taxonomy presented by Adil Yousif et al. [1] helps in identifying and classifying the mechanisms used in the implementation of Grid resource selection process, as well as describing the most significant features of Grid resource selection mechanisms. In this taxonomy, the authors aimed to highlight the main aspects of the selection mechanisms, which can benefit researchers and developers of Grid resource management systems. The taxonomy has mainly focused on selection decisions, orientation, models, organization and mapping methods of resource selection mechanisms. This Taxonomy describes, in a simple way, the basic features of resource selection mechanisms, which can help researchers and developers of Grid resource management systems to enhance the Grid resource allocation process.

Doreen Hephzibah Miriam and K.S.EaswaraKumar developed a P2P based Grid Resource Discovery model which uses parameterized HPGRID algorithm [35], to optimize Grid Resource Discovery and reaches all the Grid nodes during searching process even in the presence of non-alive nodes. The HPGRID nodes are partitioned isomorphically listing the available resources according to their zones which aid the user to select the needed resource to execute its job rather than traversing the whole Grid nodes. The scheme can adapt to a complex, heterogeneous and dynamic resources of the Grid environment, and has better scalability.

Saeed ebadi and Leyli Mohammad Khanli proposed a new distributed and hierarchical mechanism for improving fault tolerance and speed of service discovery in a Grid environment [19]. This approach has five layers and a root layer is used as coverage for the service index of under layers. For improvement fault tolerance in the proposed approach, after the requested service is found, the action of service discovery will not stop. It continues to find several instances of the requested service. These services will be sent to the cache of the institution layer.
that will be replaced immediately, if necessary. Also for speed-up service discovery, nodes in the same level, which have the same parent, send queries to each together directly. This approach achieves good efficiency, fault tolerance and consistency.

2.3 DHT BASED PEER-TO-PEER RESOURCE DISCOVERY

A peer-to-peer (P2P) [86, 126] computer network relies primarily on the computing power and bandwidth of the network participants rather than concentrating on these in a relatively few servers. P2P networks typically link nodes via largely ad hoc connections. An important fact of peer-to-peer networks is that all clients provide bandwidth, storage space, and computing resources. P2P means a network of peer nodes acting both as servers and clients. The P2P paradigm is based on the principle that every component of the system has the same responsibilities acting simultaneously as a client and as a server, as opposed to the traditional client-server model. P2P systems are divided into two main categories based on the connection protocol they employ and the way peers are organized-structured and unstructured. Structured P2P systems employ a rigid structure to interconnect the peers and to organize the file indices, while in unstructured systems each peer is randomly connected to a fixed number of other peers and there is no information about the location of files. For structured P2P systems, distributed hash tables (DHTs) are widely used. DHT-based systems [49,113,128] arrange <attribute; attribute-value>, that is, <key; value> pairs in multiple locations across the network. A query message is forwarded towards the node that is responsible for the key in a limited number of hops. The result is guaranteed, if such a key exists in the system. As compared to the flooding technique, however, DHT-based approaches need intensive maintenance on hash table updates. Table 2.2 summarizes various resource discovery approaches based on peer-to-peer techniques.
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M. Kelaskar et al. 2002 [73]</td>
<td>Evaluated four discovery mechanisms flooding and the forward routing algorithms CHORD, Pastry and CAN against the requirements of three prevalent classes of peer-to-peer applications, and investigate the suitability of these mechanisms for the applications.</td>
</tr>
<tr>
<td>2</td>
<td>Sujoy Basu et al. 2005 [130]</td>
<td>Presented the design of NodeWiz, a Grid Information Service that allows multi-attribute range queries to be performed efficiently in a distributed manner.</td>
</tr>
<tr>
<td>3</td>
<td>Domenico Talia and Paolo Trunfio 2006 [32]</td>
<td>Proposed a DHT-based framework leveraging different P2P resource discovery techniques to address the discovery of multiple resources and to support discovery of dynamic resources and arbitrary queries in Grids.</td>
</tr>
<tr>
<td>4</td>
<td>P. Trunfio et al. 2006 [135]</td>
<td>Conducted a review of the most promising Grid systems that use P2P techniques to facilitate resource discovery.</td>
</tr>
<tr>
<td>5</td>
<td>Domenico Talia et al. 2007 [33]</td>
<td>Introduced A DHT-based P2P framework to address the variety and dynamism of Grid resources.</td>
</tr>
<tr>
<td>6</td>
<td>Haiying (Helen) Shen et al. 2007 [55]</td>
<td>Presented a DHT-based Low-overhead Range-query Multi-resource manager, LORM, which is built on Cycloid DHT and LORM relies on a single DHT with constant maintenance overhead to achieve multi resource management with low overhead.</td>
</tr>
<tr>
<td>7</td>
<td>M. Marzolla et al. 2007 [86]</td>
<td>Exploited two P2P systems based on Routing Indexes, which are used to efficiently route queries and update messages in the presence of highly variable data.</td>
</tr>
<tr>
<td>S.No</td>
<td>Author(s)</td>
<td>Description(s)</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>8</td>
<td>Yuh-Jzer Joung et al. 2007 [158]</td>
<td>Designed a general keyword index and search scheme for structured P2P networks that avoid unbalanced load, hot spots, fault tolerance, storage redundancy and inability to facilitate ranking.</td>
</tr>
<tr>
<td>9</td>
<td>Chao-Tung Yang et al. 2008 [23]</td>
<td>Presented a peer-to-peer file sharing platform combined with Data Grid system for mobile devices to deal with the problems of file location discovery and client certification.</td>
</tr>
<tr>
<td>10</td>
<td>Rajiv Ranjan et al. 2008 [112]</td>
<td>Presented a detailed survey on Peer-to-Peer-based resource discovery in global Grids</td>
</tr>
<tr>
<td>11</td>
<td>Antonella Di Stefano et al. 2009 [8]</td>
<td>Dedicated an architecture that integrates the P2P interaction model in Grid environments, so as to build an open cooperative model wherein Grid entities are composed in a decentralized way.</td>
</tr>
<tr>
<td>12</td>
<td>Xu Ke et al. 2009 [151]</td>
<td>Enhanced a lookup model from human psychology process which can help to analyze and improve the P2P lookup protocols.</td>
</tr>
<tr>
<td>13</td>
<td>Mohammad Hassan Khoobkar and Mehregan Mahdabi 2009 [89]</td>
<td>Proposed a system based on variable-size Routing Indexes (RIs) that boosts the accuracy of discovery and also minimizes the average number of traversed network hops for each query.</td>
</tr>
<tr>
<td>14</td>
<td>Marc Sanchez-Artigas and Petro Garcia Lopez 2010 [85]</td>
<td>Presented Echo, a framework that benefit from the recursive structure of DHTs to embed clusters into their structures at without any cost which support a broader range of applications.</td>
</tr>
<tr>
<td>15</td>
<td>Mo Rai and Shuhang Guo 2010 [88]</td>
<td>Recommended a multiple-keyword search mechanism in structured P2P networks which groups peers with similar searches together. The formation and maintenance of groups are in a self-organizing fashion.</td>
</tr>
</tbody>
</table>

Cont…
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>H.M.N.Dilum Bandara and Anura P.Jayasumana 2011 [30]</td>
<td>Proved that super peer based architectures are the most promising resource discovery approaches.</td>
</tr>
<tr>
<td>17</td>
<td>Verdi March and Yong Meng Teo 2011 [139]</td>
<td>Proposed R-DHT (Read-only DHT), a DHT-based resource discovery scheme without distributing data items.</td>
</tr>
<tr>
<td>19</td>
<td>Srinivasan Arulanandam et al. 2012 [124]</td>
<td>Proposed a framework for building wireless digital library using Grid computing and P2P technology which stores large bytes of multimedia data across the network which is done efficiently and accessed easily from anywhere.</td>
</tr>
<tr>
<td>21</td>
<td>Carlos Pérez-Miguel et al. 2013 [20]</td>
<td>Proposed a high throughput computing system totally based upon the Peer-to-Peer (P2P) paradigm.</td>
</tr>
<tr>
<td>22</td>
<td>Kaigui Wu and Changze Wu 2013 [70]</td>
<td>Presented a novel efficient search algorithm, named State-Based Search (SBS), according to peer node’s state information which is essential for improving search performance in distributed systems.</td>
</tr>
</tbody>
</table>

M.Kelaskar et al. studied four popular decentralized discovery mechanisms based on flooding and request forwarding techniques [73] – Flooding, Chord, Pastry, and CAN. They also studied three classes of peer-to-peer applications namely distributed file sharing, person-to-
person messaging and distributed computing. It is concluded that accurate discovery of peers is not guaranteed in flooding mechanisms and Chord needs $O(\log N)$ hops to locate a key where $N$ is the total number of nodes in the system. Pastry attempts to minimize the distance traveled by the message by taking into account network locality and CAN design is fault-tolerant, robust, and self-organizing. The hash based request-forwarding algorithms outperform the flooding algorithm by orders of magnitude with an additional cost of routing table maintenance.

A DHT based P2P resource discovery framework [32] is proposed by Domenico Talia and Paolo Trunfio to manage various Grid resources and complex queries. The approach proposed for dynamic resource discovery used a DHT for broadcasting queries to all nodes without redundant messages, and adopts a similar “incremental” approach of dynamic query. It reduces the number of exchanged messages and response time. To be implemented over different platforms and spanned through wide area networks, this framework needs to be seamlessly integrated with the Open Grid Services Architecture (OGSA) and the Web Service Resource Framework (WSRF), which provides standard mechanisms for the implementation of Grid systems and applications. Moreover, the framework needs to explicitly address XML-based queries to support the discovery of WSRF-enabled Web services.

P. Trunfio et al. presented a report to serve as a review of the most promising Grid systems that incorporate P2P resource location methods [135], in order to perform a qualitative comparison of the existing approaches and to draw conclusions about their advantages and their weaknesses. The systems discussed in this report make use of different P2P approaches, ranging from unstructured to structured ones. Along with the presentation of the systems, their capabilities in terms of scalability, reliability, efficiency, ease of implementation, ease of use,
self-organization, fault-tolerance, security, robustness are also discussed. Furthermore, P2P-based approaches for building scalable Grid Resource Discovery services based on semantic information are discussed to improve the precision of resource discovery.

Domenico Talia et al. discussed the characteristics of Grid resources and identified critical problems of resource discovery in Grids [33]. A DHT-based P2P framework has been introduced to address the variety and dynamism of Grid resources. It exploits multiple DHT and existing P2P techniques for multiple static resources and implements an incremental resource discovery approach for dynamic resources. As compared to the original strategy, the incremental approach generates reduced number of messages and experiences lower response time in large-scale Grids.

Most current P2P-based resource management approaches support multi-attribute range queries at a high cost. They either depend on multiple P2P networks with each P2P network responsible for a single attribute, or they keep the resource information of all attributes in a single node. Haiying (Helen) Shen et al. [55] presented a low-overhead range query multi-attribute P2P-based resource management approach, LORM. Unlike other P2P-based approaches, it relies on a single P2P network and allocates resource information to different nodes based on resource attributes and values. Moreover, it has high capability to handle the large-scale and dynamic characteristics of resources in Grids.

The convergence of the Grid and Peer-to-Peer (P2P) worlds has led to many solutions that try to efficiently solve the problem of resource discovery on Grids. Some of these solutions are extensions of P2P DHT-based networks. Based on the concept that the DHT based P2P solutions are poorly flexible and efficient in handling them, M. Marzolla et al. presented two P2P
Both are based on Routing Indexes, which are used to efficiently route queries and update messages in the presence of highly variable data. The first system uses a tree-shaped overlay network and the second one is an evolution of the first, and is based on a two-level hierarchical network topology, where tree topologies must only be maintained at the lower level of the hierarchy, i.e., within the various node groups making up the network. It is shown that both the solutions, the bitwise RIs are very effective in limiting the message span and number of hops.

Chao-Tung Yang et al. presented a peer-to-peer file sharing platform combined with Data Grid system for mobile devices [23]. Facing the limit of network bandwidth and unstable signal strength in wireless network, wireless peer-to-peer architecture is now widely accepted as a new computing paradigm for wireless devices. In wireless P2P architecture, one can speed up parallel file transfer and enhance the communication stability by establishing multiple concurrent P2P connections. This paper mainly applies Data Grid system to deal with the problems of file location discovery and client certification. An architecture that uses GridFTP protocol to establish a video sharing platform for mobile devices is proposed. Combined with data Grid, the system can perform video analysis and classification upon its computing resources, keeping the files in the newest state at all time.

Rajiv Ranjan et al. [112] presented a summary of the current state of the art in Grid Resource Discovery, resource taxonomy with focus on the computational Grid paradigm, P2P taxonomy with a focus on extending the current structured systems (e.g., distributed hash tables) for indexing d-dimensional Grid resource queries, a detailed survey of existing work that can
support d-dimensional Grid resource queries, and classification of the surveyed approaches based on the proposed P2P taxonomy.

Xu ke et al. [151] attempted to analyze and classify the methods of P2P resource discovery from the perspective of thinking process and summary up a general framework of P2P lookup, based on which more innovative strategies can be created and evaluated easier. It is summarized that on lookup psychological process, there are two key points, the first is that the heurist strategy relies on a global framework which decides the lookup efficiency and optimization and the other is that before the tentative strategy can be practiced, its candidate collection for search must be optimized and reduced by history experience or comparison. This improved P2P lookup model from Chord is based on the binary tree framework.

Fundamental design choices for distributed resource advertising and querying are evaluated in the context of existing practical systems. A generic model for cost of resource discovery is presented and multi-attribute resource and query characteristics from Planet-Lab and SETI@home are presented [30]. It is observed that attributes of both resources and queries are highly skewed, correlated, queries are less specific, and Generalized Pareto distribution is suitable for capturing the distribution of most dynamic attributes and their rate of change. Based on these observations, different design choices are evaluated for resource discovery in terms of their cost of advertising/querying, latency, load balancing, and routing table size. It is concluded that super peer-based architectures have the potential to support large-scale resource aggregation as they simultaneously balance the cost and load.

Carlos Perez-Miguel et al. proposed a novel HTC system based on P2P networks [20]. The goal is that this method is to build a totally decentralized HTC system based on highly
scalable and reliable P2P storage maintaining the target characteristics, lack of central management points, disconnected operation and FCFS execution order of jobs. For this purpose, a distributed queue system that manages the execution of jobs over a group of nodes in a totally decentralized way is implemented. Also explained the main characteristics of the design, and have implemented a prototype based on the Cassandra P2P storage system.

2.4 CHORD PROTOCOL IN RESOURCE DISCOVERY

Structured P2P systems (such as Chord) employ a rigid structure to interconnect the peers and to administrate the file indices. Grids have significant similarities and differences with P2P systems. The main similarity is that the network changes dynamically, and this makes maintaining centralized global views of the network unrealistic in both of them, while the main difference is resources shared change dynamically for Grids but not for P2P systems. In fact, P2P networks can work well when nodes are joining and leaving the network. In a Chord-based resource discovery model, the Grid agencies can be used as the network nodes to form a Chord ring topology. Chord protocol plays a vital role in Grid Resource Discovery. Table 2.3 summarizes the overview of Chord protocol in resource discovery process.

Table 2.3: An Overview of Chord Protocol in Resource Discovery

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ion Stoica et al. 2001 [128]</td>
<td>Developed a scalable protocol for lookup in a dynamic peer-to-peer system with frequent node arrivals and departures.</td>
</tr>
<tr>
<td>2</td>
<td>David R. Karger and M. Ruhl 2004 [71]</td>
<td>Proposed a version of the Chord peer-to-peer protocol that allows any subset of nodes in the network to jointly offer a service without forming their own Chord ring.</td>
</tr>
</tbody>
</table>

Cont…
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>J.Salter and N.Antonopoulos 2004 [121]</td>
<td>Presented a resource discovery system capable of resolving keyword-value pair queries, based on a two-layered Chord ring architecture.</td>
</tr>
<tr>
<td>4</td>
<td>Gerald Kunzmann et al. 2005 [51]</td>
<td>Analyzed the efficiency of Chord's stabilization algorithm by measuring the number of nodes with erroneous neighbor lists and the resulting costs.</td>
</tr>
<tr>
<td>5</td>
<td>Jing Wang et al. 2006 [67]</td>
<td>Presented a bidirectional query Chord system based on latency-sensitivity to decrease the latency and cost of locating resources.</td>
</tr>
<tr>
<td>6</td>
<td>LIU et al. 2008 [80]</td>
<td>Introduced the super peer concept as the topological structure based on P2P network of semantic information.</td>
</tr>
<tr>
<td>7</td>
<td>Yi-Chun Wu et al. 2008 [155]</td>
<td>Proposed two protocols based on Chord namely Enhanced Bidirectional Chord (EB-Chord) and Enhanced Bidirectional Chord with Lookup-Parasitic Random Sampling (EB-Chord-LPRS) to greatly reduce the average latency.</td>
</tr>
<tr>
<td>8</td>
<td>Xiao-Jin Ren et al. 2008 [148]</td>
<td>Introduced an iterative join algorithm for Chord that is suitable for highly dynamic environment.</td>
</tr>
<tr>
<td>9</td>
<td>Wang Biqing 2009 [142]</td>
<td>Enhanced the Chord routing algorithm by not increasing the length of routing table, repetitious entries are deleted and anticlockwise routing in the same amount are added.</td>
</tr>
<tr>
<td>11</td>
<td>Zhigang Wang et al. 2009 [167]</td>
<td>Carried out a research to reduce the hops of a lookup message while searching resource in Chord, the pre-node list and the successor list of a node would be used as routing information and the routing information in the clockwise second half ring will be added to the finger table of the node.</td>
</tr>
<tr>
<td>S.No</td>
<td>Author(s)</td>
<td>Description(s)</td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>12</td>
<td>Agostino Forestiero et al. 2010 [3]</td>
<td>Implemented Self-Chord, a set of ant-inspired mobile agents which move and reorder the resource keys in a ring of peers in a self-organizing fashion without any predetermined association between keys and peers.</td>
</tr>
<tr>
<td>13</td>
<td>Ashok Kumar Ojha and Krishna Kant 2010 [9]</td>
<td>Dedicated Sub-ring Based Chord model and a lookup algorithm in which the lookup process keeps only $O(\log k)$ routing information where $k$ is the number of sub-rings.</td>
</tr>
<tr>
<td>15</td>
<td>Lei Xu and Fu Yang 2010 [78]</td>
<td>Presented virtual two-layered Chord in which node's identifier is assigned by hashing the node's Gateway address and IP address and it makes the same network segment nodes which are logically adjacent.</td>
</tr>
<tr>
<td>16</td>
<td>Qin Dong and Jingling Zhao 2010 [109]</td>
<td>Proposed an improved Chord model to deal with ignorance of the semantic property of the content and physical topology information in P2P and the system facilitates the landmark +RTT method to generate topology information.</td>
</tr>
<tr>
<td>17</td>
<td>Wang Biqing 2010 [143]</td>
<td>Recommended a two-layer routing model for Chord(T-Chord) by logically constructing inter-domain transmitting network on the Chord, the target node which is not in the same half ring with the originating lookup node can possibly be found in only one hop.</td>
</tr>
<tr>
<td>18</td>
<td>Wang Ziyang and Yyin Siqing 2010 [144]</td>
<td>Implemented the model by using positive and negative direction routing tables to search together in which hop times are reduced overall to improve the search efficiency.</td>
</tr>
</tbody>
</table>

Cont…
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Panagiotis Gouvas et al. 2011 [98]</td>
<td>Described a generic framework, Ubi-Chord that facilitates the design and development of autonomic and decentralized services in ad-hoc networks.</td>
</tr>
<tr>
<td>20</td>
<td>Thiago Q. de Oliveira et al. 2011 [134]</td>
<td>Constructed a hybrid bio-inspired architecture for P2P networks, called WChord which adds concepts of Chord and the dominance hierarchy of wasps P. dominulus.</td>
</tr>
<tr>
<td>21</td>
<td>Wenming Jiang et al. 2011 [146]</td>
<td>Developed a novel optimized Chord (opChord) algorithm by adding second level successor node's information to finger table structure and designing a two-hops query method, opChord can decrease the number of forwarding messages between nodes and satisfy the efficient search requirements of mobile P2P network.</td>
</tr>
<tr>
<td>22</td>
<td>Yuta Shimano and Fumiaki Sato 2011 [159]</td>
<td>Proposed a method to reduce the query response time by the rearrangement and replacement of the node ID of the Chord.</td>
</tr>
</tbody>
</table>

Ion Stocia et al. presented Chord, a distributed lookup protocol that efficiently locates the node that stores a particular data item [128]. Chord provides support for just one operation, given a key; it maps the key onto a node. Data location can be easily implemented on top of Chord by associating a key with each data item, and storing the key/data item pair at the node to which the key maps. Chord adapts efficiently as nodes join and leave the system, and can answer queries even if the system is continuously changing. Results from theoretical analysis, simulations, and
experiments show that Chord is scalable, with communication cost and the state maintained by each node scaling logarithmically with the number of Chord nodes.

David R.Karger and M.Ruhl stated and analyzed version of the Chord P2P protocol [71] that allows any subset of nodes in the network to jointly offer a service without forming their own Chord ring. These subgroups are useful for efficiently carrying out computations or functions that do not require the involvement of all nodes. This protocol utilizes the routing functionality of the existing Chord ring, so that subgroups can be implemented more efficiently than by creating a separate routing infrastructure for each subgroup. This variant supports the same efficient join/leave/insert/delete operations that subgroup would get if they did form their own separate peer-to-peer network, but significantly less resources than the separate network would.

J.Salter and N.Antonopoulos [121] developed P2P network based on a set of Chord rings and designed to provide a list of machines hosting matching resources in response to a query composed of keyword-value pairs. Each ring is responsible for a single keyword. Within each keyword ring, nodes are responsible for indexing resources matching one or more values associated with the ring’s keyword. A further ring named the super ring organises a group of super nodes that host pointers to each keyword ring. Each super node can be responsible for pointing to multiple keyword rings. A node can become a member of multiple keyword rings, for up to as many as it has capacity for. Nodes which are members of popular keyword rings must be able to handle more queries than a node in a less popular keyword ring.

The clockwise searching scheme of original Chord doesn’t make good use of the anti-clockwise neighbors’ information. All these increase the latency and cost of locating resources.
A bidirectional Chord system based on latency-sensitivity is proposed by Jing Wang et al. [67] to decrease the latency and cost of locating resources. This system partitions the nodes according to the latencies among these nodes and the landmarks to decrease the average response time, and by the means of bidirectional query algorithm on Chord, the search path length is reduced.

Routing is essential in P2P applications. Chord routing table can only cover half of the Chord ring. As long as target node is in the half ring which is not covered by the routing table, the target node must be found via not less than one medium node, namely, two hops. So, Chord is not very efficient. Therefore, a two-layer routing model for Chord (T-Chord) is proposed by Wang Biqing et al. [142]. By logically constructing inter-domain transmitting network on the Chord, the target node which is not in the same half ring with the originating lookup node can possibly be found in only one hop. Thus, the covering problem of Chord routing table can be solved well.

“Self-Chord,” a peer-to-peer (P2P) system is proposed that inherits the ability of Chord-like structured systems for the construction and maintenance of an overlay of peers, but features enhanced functionalities deriving from ant-inspired algorithms, such as autonomous behavior, self-organization, and capacity to adapt to a changing environment. Here, resource indexing and placement is uncorrelated with network structure and topology, and resource keys are organized and managed by self-organizing mobile agents through simple local operations driven by probabilistic choices. Agostino Forestiero et al. [3] also aim to be a step towards the mathematical analysis of bio-inspired systems and swarm intelligence phenomena. The modeling of these systems through fluid-like differential equations has proved to be a valid approach.

The routing table in Chord suffers from serious information redundancy and it is not very efficient. Wang Biqing [143] proposed an improved Chord routing algorithm. On the premise of
not increasing the length of routing table, repetitious entries are deleted and anticlockwise routing in the same amount is added. According to interval between key and node that originated lookup, a lookup direction will be chosen and lookup efficiency must be improved greatly. It is shown that the algorithm eliminates information redundancy, reduces the average lookup path length and gets higher efficiency so as to coordinate lookup efficiency and routing table length correctly.

Wang Ziyang and Yyin Siqing [144] presented an improved method on the structured P2P algorithm of Chord. By increasing the reverse route table for each node in each of the search process, because both positive and negative direction routing tables search together, the find hop times are reduced overall to improve the search efficiency. The simultaneous search of positive and negative routing tables increases the burden on the node, but because of significant increase in computer performance now, the burden on the node brought by the algorithm can be ignored basically. The analysis of examples shows that the logical hop times of improved algorithm is less than the original Chord algorithm hop times, and can effectively improve the routing efficiency of Chord.

Future networks are becoming larger in scale, more dynamic and heterogeneous, imposing new requirements to the design and provision of advanced services. In order to cope with these requirements, an approach for efficiently designing services is proposed [98], based on an existing generic framework that facilitates the design and development of autonomic and decentralized services in ad-hoc networks. The design of the services is realized independently from the underlying physical network while specific functions are provided to application developers for deploying useful applications. Through a set of emulations over a reference
implementation, it is demonstrated that services may be efficiently and reliably delivered over a dynamic network.

Thiago Q. de Oliveira et al. described a new architecture for P2P networks based on a bio-inspired model [134], proving its viability in relation to the Chord model. This infrastructure explores the advantages offered by hybrid P2P networks and P2P networks based on DHT. It also explains an existing bioinspired model of self-organization in the hierarchy of wasps P. Dominulus.

With the rapid increase in the number of devices, mobile P2P network (MP2P) has emerged as a state-of-the-art technology for large-scale resources sharing in wireless networks. Chord is a typical P2P network structure and widely used in the Internet. However, with the characteristics of the limited bandwidth and dynamic nature of mobile environment, the wired-based Chord algorithm cannot adapt itself to MP2P network. Wenming Jiang et al. [146] proposed a novel optimized Chord (opChord) algorithm. By adding second level successor node's information to rmger table structure and designing a two-hops query method, opChord can decrease the number of forwarding messages between nodes and satisfy the efficient search requirements of mobile P2P networks, and then can improve the overall performance of Chord network.

Distances between entries in routing table are too long and only half of the identifier space can be covered in base Chord. To solve this problem, a higher order Chord routing algorithm is proposed [12]. By introducing a new routing table structure formula, the improved algorithm augments neighbors number which need to be maintained, increases distribution density of routing table entries, extends routing searching coverage areas to the entire identifier
space, and gets higher efficiency under the premise of keeping low cost. This method proved that higher order Chord reduces the average lookup path length and gets higher efficiency.

2.5 RESOURCE DISCOVERY BASED ON RECENTLY VISITED NODE HISTORY

Now-a-days numerous successful implementations of resource discovery approaches using P2P technique are available. Table 2.4 summarizes the resource discovery in Grid computing using recently visited node history.

Table 2.4: An Overview of Resource Discovery in Grid Using Recent History of Nodes

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tsutomu Inaba et al. 2008 [136]</td>
<td>Proposed a new logical link reconnection algorithm that records the access orders of peers by a user.</td>
</tr>
<tr>
<td>2</td>
<td>Chunling Cheng et al. 2010 [24]</td>
<td>Proposed a strategy that uses the node's counterclockwise node information to replace the redundant information and also proposes a strategy to solve the problem in repeated search in P2P network by establishing an objective resource table.</td>
</tr>
<tr>
<td>3</td>
<td>Pethalakshmi and Jeyabharathi 2013 [108]</td>
<td>Proposed RVN-Chord (Recently Visited Node) which modifies the original Chord’s finger table by adding a new column which stores the ID of Recently visited node.</td>
</tr>
</tbody>
</table>

Chunling Cheng et al. propose a strategy that using the node's counterclockwise node information to replace the redundant information [24], the strategy uses the space effectively which occupied by the redundant information. They also proposed a strategy to solve the problem that repeated search in P2P network by establishing an objective resource table. Then designed an advanced chord routing algorithm based on redundant information replaced and objective resource table (AChord). This method uses LRU algorithm to manage the objective...
resource node table. The so-called objective resource table is that which stores the objective resource (resource which is frequently accessed by a node) in the source node. This table is useful for positioning the requested resource quickly. AChord has a very good compromise both in space complexity and time complexity.

Pethalakshmi and Jeyabharathi proposed to use the modified chord protocol by adding an additional entry into the Base Chord finger table to store the location of Recently Visited Node (RVN) [108]. The next lookup will use this id to locate its key, if the match is found. The primary aim of any protocol is the efficient and fast lookup of nodes containing keys. Generally the performance of Chord like algorithm can be analyzed in terms of three metrics: the size of the finger table of every node, the number of hops a request needs to travel in the worst case, and the average number of hops. In the original Chord, when a node requests a key, it has to search its own finger table first and it may find the successor of the key in that table if the match is found. Otherwise, the node sends messages to other nodes whose node id is less than or equal to the searched key and it may need a few hops to locate it. After locating the successor of the key, the result is returned to the node who started the search and the lookup process is successfully done. This protocol modifies the finger table of the original Chord to add a new entry in the finger table which stores the recently visited node’s id.

2.6 RESOURCE DISCOVERY BASED ON FEATURE SELECTION

Feature Selection is the process of selecting a subset of relevant features for use in model construction. The central assumption while using a feature selection technique is that the data contains many redundant or irrelevant features. Redundant features are those which provide no new information than the currently selected features, and irrelevant features provide no useful
information in any context. Feature selection techniques are a subset of the more general field of feature extraction. Feature extraction creates new features from functions of the original features, whereas feature selection returns a subset of the features. Application of feature selection technique in Chord protocol for the improvement of resource discovery process is discussed in this section. The methods for resource discovery in Grid computing using feature selection are summarized in Table 2.5.

Table 2.5: An Overview of Feature Selection in Grid Resource Discovery

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ertao Lv et al. 2007 [38]</td>
<td>Proposed a two-layer hybrid P2P network (HP2P) to reduce the performance degradation caused by the reorganizations.</td>
</tr>
<tr>
<td>2</td>
<td>Eric Jui-LinLu et al. 2008 [37]</td>
<td>Designed a multi-layered P2P resource sharing model, called ML-Chord that assigns nodes into Chord-like layers based on the categories of shared resources.</td>
</tr>
<tr>
<td>3</td>
<td>Pethalakshmi and Jeyabharathi 2014 [100]</td>
<td>Presented new method based on feature selection which isolates the dominant nodes and non-dominant nodes in the Chord ring.</td>
</tr>
<tr>
<td>4</td>
<td>Pethalashmi and Jeyabharathi 2014 [102]</td>
<td>Applied feature selection technique in DHT based P2P Chord protocol (FS-CHORD) to lookup resources in large scale Grid.</td>
</tr>
</tbody>
</table>

Ertao Lv et al. proposed [38] a two-layer hybrid P2P network (HP2P). The upper layer is Chord network and the lower layer is cluster in HP2P. Peers are organized into clusters, and the network reorganizations are restricted within the clusters where a peer can join or leave. Besides, each cluster is also enhanced through some specific mechanisms such as Supernode, metadata redundancy and gossip flooding. Consequently, the two layers of P2P become more robust.
Further, the high efficiency of Distributed Hash Tables (DHTs) can be retained, and the churn caused by excessive node joining and leaving can be reduced.

ML-Chord [37] is a multi-layered P2P resource sharing model; based on the specific number of categories for domain or ontology of the network, it divides the entire network into a number of overlay layers. Each category layer is a Chord-like overlay network. A peer with good capabilities can be selected as a Bridge-peer (BP) and it is linked to all categories and the remaining nodes are normal nodes. Nodes in each layer are sorted by its ID number and connected to a successor and a predecessor. BP layer is formed by including only bridge peers in the Chord-like overlay network. Every new query is started from category layer to find the successor of the key and if the key is not found, BP layer is enquired. There are two tables maintained in this model, namely finger table and BP finger table. BP layer is maintained by normal peers as well as bridge peers in a different manner.

Pethalakshmi and Jeyabharathi [100] discussed RVN-Chord [108] which is based on Recently-Visited-Node history for efficient resource discovery and proposed feature selection technique to extract the valuable features of resources. The selected resources and the nodes which store these resources are taken into account.

Pethalakshmi and Jeyabharathi [102] proposed a new method based on the application of feature selection technique in Chord protocol for the improvement of resource discovery process. The aim of this method is to consider the valuable features of resources of each node based on the individual resource’s strength. To extract meaningful features of nodes, frequency of queries posted on a particular resource is calculated. Based on the strength and resource count of each node, nodes are taken for comparison. APriori algorithm is applied and dominant resource sets
are selected among the nodes. APriori is a classic algorithm for frequent itemset mining and
association rule learning over transactional databases. It proceeds by identifying the frequent
individual items in the database and extending them to larger and larger itemsets as long as those
itemsets appear sufficiently in the database. The frequent itemsets determined by APriori can be
used to create association rules which highlight general trends in the database. This has
applications in domains such as market basket analysis.

2.7 RESOURCE DISCOVERY BASED ON FUZZY LOGIC

A fuzzy set is any set that allows its members to have different grades of membership in
the interval [0, 1]. Fuzzy logic methodology can be utilized either in solving the problems that
are complex to be analyzed quantitatively or in the case of natural phenomena that are not easy
to be modeled mathematically. IF-THEN rules are used in the FIS (Fuzzy Inference System)
where the output is concluded from the fired rules based on the given inputs [160,161]. The
system parameters are modeled as linguistic variables based on expert knowledge; also the
corresponding membership functions are designed for each parameter. Therefore, fuzzy logic
theory can be employed even in the nonlinear systems suffering from uncertainty and
complexity. That complex uncertain system can be modeled effectively based on fuzzy logic
without any need for complicated mathematical models [94]. In Grid computing, fuzzy logic is
used in resource discovery, resource scheduling and load balancing. The various resource
discovery algorithms using fuzzy classification are summarized in Table 2.6.
Table 2.6: An Overview of Resource Discovery algorithms Using Fuzzy Logic

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Miao Yuting et al. 2010 [87]</td>
<td>Developed Hot-Chord algorithm which constructs an inner chord dynamically by using the nodes which hot resources are locating, and to decrease the inquiring space.</td>
</tr>
<tr>
<td>3</td>
<td>Codruț Grosu et al. 2011 [25]</td>
<td>Designed a Chord based architecture to achieve optimal implementation results for range queries, as well as provide a fault-tolerant substrate.</td>
</tr>
<tr>
<td>4</td>
<td>Duo Liu and Chung-Horng Lung 2011 [36]</td>
<td>Modified the objective function and the distance function that greatly reduces the computational complexity of FCM while keeping the clustering accurate.</td>
</tr>
<tr>
<td>5</td>
<td>Saeed Javanmardi 2013 [120]</td>
<td>Proposed a novel approach for Grid Resource Discovery which is using fuzzy theory for creating semantic Overlay Network.</td>
</tr>
<tr>
<td>6</td>
<td>Pethalakshmi and Jeyabharathi 2013 [106]</td>
<td>Applied fuzzy technique for resource discovery in Chord protocol to easily locate the Grid resources by reducing message and time complexity.</td>
</tr>
</tbody>
</table>

To support multi-attribute, multi-keyword like complex queries, MF-Chord [162] is proposed which increases the recall ratio of search process in limited hops. The two main issues considered in this model are indexing and mapping between keywords and nodes, and improving query performance. Along with single dimensional fingerprint which includes main keywords
information about attributes of resources, MF-Chord introduces an opposite fingerprint for every resource that represents partial reversal of the fingerprint to balance the query load effectively. MF-Chord dynamically generates the query-request-forwarding-tree based on the newly formatted finger table and query-request-message format. The performance of MF-Chord is verified by comparing it with various algorithms and it is proved that MF-Chord is suitable for large-scale networks.

Miano Yuting et al. presented Hot–Chord [87] in which an inner Chord is constructed dynamically by using the nodes, where the hot resources are located. Each inner node maintains an additional routing table, which records $O(\log Hm)$ other inner node’s information and the mapping relationship between node’s new Hash value and original Hash value. This algorithm decreases the inquiring space and average query hops and increases the query efficiency of the hot resources.

Fuzzy clustering is more flexible than hard clustering and is practical for P2P traffic identification because of the natural treatment of data using fuzzy clustering. Fuzzy C-Means clustering (FCM) is an iteratively optimal algorithm normally based on the least square method to partition data sets, which has high computational overhead. Duo Liu and Chung-Horng Lung proposed [36] modifications to the objective function and the distance function that greatly reduces the computational complexity of FCM while keeping the clustering accurate. The proposed FCM clustering technology can be incorporated into a Fuzzy Inference System (FIS) to implement real-time network traffic classification by updating the training data set continuously and efficiently.
Saeed Javanmardi proposed a novel approach for Grid Resource Discovery which uses fuzzy theory for creating semantic Overlay Network [120]. Fuzzy theory belongs to the intelligent approach which shows uncertainty in Phenomena. They also used fuzzy theory in resource discovery phase to discover the most adequate resources. Here, Peers are grouped together in the network space based on fuzzy theory with three parameters namely Delay, Bandwidth and characterization in the semantic space. Then the resource discovery is done by using some fuzzy input parameters like the computational capabilities of nodes and semantic similarity. The results of the experiments show the efficiency of the proposed approach in terms of scalability, precision, search expressiveness and response time.

Pethalakshmi and Jeyabharathi applied Fuzzy classification [106] which classifies elements into a fuzzy set and its membership function is defined by the truth value of a fuzzy propositional function. The Chord is constructed initially with $2^m$ nodes. The ring is divided into three rings according to the basic Fuzzy-rule. This algorithm restructures the Chord in a different manner that nodes with more number of resources and unique resources are selected and considered for lookup process. The search process is done simultaneously in all the rings and this method effectively reduces the required number of hops, messages and communication time.

Pethalakshmi and Jeyabharathi discussed multi-ring chord with parallel search for quick resource discovery in Grid Computing [107]. N nodes are grouped into three rings according to the classification procedure of decision tree algorithm with the help of basic fuzzy logic properties. Whenever a new node enters, the algorithm correctly predicts the corresponding ring and the node will join into the desired place. In this place Divide-and-Conquer learning concept
is applied to split the nodes into subsets and this process is recursively executed for the subset of nodes.

### 2.8 RESOURCE DISCOVERY BASED ON PHYSICAL LOCATION OF NODES

In base Chord protocol, the Ring is constructed by considering logical address of the nodes which do not use the physical location of nodes in the network. Various resource discovery mechanisms are proposed to use the geographical location of nodes to quickly locate the resources in Grid and to reduce the communication time. Table 2.7 presented an overview of various physical location based resource discovery mechanisms.

**Table 2.7: An Overview of Physical Location Based Resource Discovery Methods**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alberto Colorni et al. 1991[4]</td>
<td>Explored the implications that the study of ants behaviour can have on problem solving and optimization and introduced a distributed problem solving environment and proposed its use to search for a solution to the travelling salesman problem.</td>
</tr>
<tr>
<td>2</td>
<td>Fazli Erbas 2004 [40]</td>
<td>Presented a novel network layer approach for quality of service (QoS) routing in mobile ad hoc networks.</td>
</tr>
<tr>
<td>4</td>
<td>Javad Taheri and Mohammad Hazem Akbari 2007 [65]</td>
<td>Designed A topology-aware Chord (TAC) that introduced the local ring concept by dividing the geographical space into smaller areas.</td>
</tr>
<tr>
<td>5</td>
<td>SUN Mingsong and Zhang Zhongqiu 2008 [131]</td>
<td>Proposed Quasi-Chord model which utilizes Global Network Position (GNP) system to coordinate the host on the physical layer and uses the Cantor space filling Curve to map the 2-dimensional geometrical space into 1-dimensional.</td>
</tr>
</tbody>
</table>

Cont…
<table>
<thead>
<tr>
<th>S.No</th>
<th>Author(s)</th>
<th>Description(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Jianwei Zhang et al. 2011 [66]</td>
<td>Proposed a new Chord model to utilize the global TSP problem to achieve the physical topology matching in Chord.</td>
</tr>
<tr>
<td>7</td>
<td>Jaber Karimpour et al. 2012 [64]</td>
<td>Proposed CLT-Chord in which the cache location tables are used in addition to the local finger tables for routing requests, in which each node stores nodes which are at its close geographical range.</td>
</tr>
<tr>
<td>8</td>
<td>Haiying Shen and Kai Hwang 2012 [57]</td>
<td>Presented a resource management scheme including hierarchical cycloid overlay architecture, resource clustering and discovery algorithms for wide-area distributed Grid systems.</td>
</tr>
<tr>
<td>9</td>
<td>Zhen He et al. 2012 [164]</td>
<td>Designed the physical topology aware Chord model (TSP-Chord) based on ant colony algorithm.</td>
</tr>
<tr>
<td>10</td>
<td>Pethalakshmi and Jeyabharathi 2013 [103]</td>
<td>Introduced the modified version of Chord which considers physical network topology through dividing the geographical space in which the nodes are distributed into smaller Regions based on their physical location</td>
</tr>
</tbody>
</table>

Liu Ye et al analyzed [79] the Hashing function’s properties and presented novel logical connections among the destination node, the Chord semantic routing relay node sequence, and the ID of clustering neighbor nodes. They also described TCS-Chord (an improved routing algorithm to Chord based on the Topology-aware Clustering in Self-organizing mode) to improve the efficiency of Chord routing. Since the clustering nodes only have local views in the self-organizing mode, some rules are applied for a node to learn other nodes’ physical topology-aware locations. The TCS-Chord’s routing algorithm is described completely, and the experiments indicate that TCS-Chord can improve the Chord semantic routing efficiently.
Javad Taheri and Mohammad Hazem Akbari introduced the TAC [65], a novel topology-aware protocol which is based on Chord. TAC introduces the local ring concept by dividing the geographical space into smaller areas. Through binding each new node to a proper local ring concerning its physical location, TAC considers the physical network topology of the overlay network to demonstrate more efficient key lookup.

Quasi-Chord [131] is in the light of the ideology of Chord. This model utilizes global network (GNP) system to coordinate the host on the physical layer and uses the Cantor space filling Curve to map the 2-dimensional geometrical space into 1-dimensional. Then Quasi-Chord model is built according to the Cantor Value. Quasi-Chord constructs the logical topology on the basis of the underlying topology information. So the nodes close in the underlying layer will be also close in the overlay network, which can greatly reduce the network traffic. To build up Quasi-Chord model, three steps are focused by the author. In the first step, GNP is used to get each host’s coordinates in P2P geometric space; In the second step, Cantor space filling curve is used to convert the 2-dimensional space into 1-dimensional; Quasi-Chord circle is constructed according to the host’s Cantor value. The two main advantages of the method are the correspondence relationship of physical layer and logic layer which are more precise and combine network positioning and space filling curve to construct the P2P network, which can greatly reduce the network traffic brought by the TTL.

On the basis of in-depth analysis of the mapping storage methods of the Identifier-Locator separation network, Jianwei Zhang et al. proposed a physical topology aware Chord model (Ant-Chord) which is based on ant colony algorithm [66]. The ideas of Ant-Chord is to regard the storage nodes in the whole Chord as a TSP problem and solve the TSP problem
quickly by using the ant colony algorithm, then to build the Chord with the obtained TSP solution, and "Luoyang shovel" method is applied to optimize the Ant-Chord’s routing hops. Adding proximity routing table and dividing proximity physical Region are the two main ways to achieve physical topology match. Adding proximity routing table has the minimum changes with the Chord structure, but increases the routing table overhead. Dividing proximity physical Region on the one hand, needs to solve the balances between the number of Regions and the number of nodes within a Region, on the other hand it is opaque between each Region, so the efficiency of inter-domain communication is not ideal. As a widespread concerned academic issue, TSP problem has been researched for a long time and is also a standard test platform for a number of new algorithms. Metaheuristic algorithm is very successful in solving TSP problem and ant colony optimization (ACO) is one of the classic meta-heuristic algorithms. Simply matching the physical topology has little effect on the logical routing hops of Chord, “Luoyang Shovel Method (LSM)” optimization method is designed which records the storage space information of successor nodes when generating the routing table of Chord nodes to increase the Chord routing flexibility and achieve the optimal routing hops.

In CLT Chord [64], the cache location tables are maintained to store the details of nodes which are at its close geographical range. These details are collected while the node is joining the network. For this reason, geographical range of a new node is attached to the join request and is sent to the server. After receiving the request, the server sends a response message to the new node in which IP address of its successor node and one of the nodes in the same geographical range are contained. Then the new node joins the network and communicates with the node that is in its geographical range to put entries in the cache location table. In this model, cache location table is stored in cache memory. Due to the continuous change of network, the nodes in
geographical range must communicate with each other in specified intervals, i.e. the nodes ping each other. Since node joins and leaves the network frequently, cache location tables should be updated to route the requests properly, cache location table arranges the nodes based on lifetime of a node, desired number of responses sent to other nodes, and time delay between the nodes. The node begins lookup requests from the first better node and if it does not receive its desired response, forwards the request to the second better node and so on. The algorithm continues to work like the basic Chord protocol.

Haiying Shen and Kai Hwang presented [57] a resource management scheme including a hierarchical cycloid overlay architecture, resource clustering and discovery algorithms for wide-area distributed Grid systems; established program/data locality by clustering resources based on their physical proximity and functional matching with user applications. Further, they developed dynamism-resilient resource management algorithm, cluster-token forwarding algorithm, and deadline-driven resource management algorithms. The advantage of the proposed scheme lies in low overhead, fast and dynamism-resilient multi resource discovery. In particular, it supports efficient resource clustering, reduces communication cost, and enhances resource discovery success rate in promoting large-scale distributed supercomputing applications.

Pethalakshmi and Jeyabharathi presented a Geo-Chord [103], which is based on geographical location of nodes. Based on the Euclidean distance among nodes, neighbors are identified and many Region-Rings are formed. Ring heads of Regions form the main Chord ring. Geo-Chord makes use of the physical network topology of the overlay network to demonstrate more efficient key lookup. Region finger table is maintained to store the information of proximate nodes (the nodes within same Region) and therefore lets the lookup process to be done
more efficiently. Geo-Chord performs better in terms of average number of messages, hops and average communication time.

2.9 SUMMARY

Resource discovery is an important and complex task in Grid computing. Various mechanisms have been implemented to quickly locate the Grid resources. This chapter reviewed the various resource discovery mechanisms such as hierarchical, centralized and Hybrid Systems. Algorithms related to P2P based Grid Resource Discovery and various improvements made on structured P2P Chord protocol are also reviewed in this chapter. Review of feature selection based resource discovery mechanisms, Fuzzy based mechanisms and geographical location based mechanisms are presented at the end of this chapter.