International Journal of Engineering Research and General Science Volume 3, Issue 3, May-June, 2015
ISSN 2091-2730

A Trend Towards Optimistic Community Discovery Algorithms For Web Services

Mini Singh Ahuja (Assistant professor), Sonia Jasrotia (student)

M tech (CSE) Guru Nanak Dev University Regional Campus Gurdaspur,

jasrotiasonia@ymail.com and 7508410749

Abstract— Community structure is the most common topological structure in complex networks. An important feature of complex networks is that they are generally composed of highly inner connected sub networks called community. It is very useful and significant to understand the features of these communities. In this paper, a survey on complex networks has been done in web services. From the survey it has been concluded that none of the techniques performs effectively in all the fields. Therefore the paper ends with the future scope to overcome these issues.

Keywords— SOC, Web services, Complex Networks, Community Structure , SLA, Community discovery, community discovery algorithms.

INTRODUCTION

Service Oriented Computing (SOC) is the computing paradigm that utilizes services as fundamental elements for developing applications/solutions. Services are self describing, platform-agnostic computational elements that support little cost composition of distributed applications , fast. Services perform functions, which can be anything from simple process to complicated processes. Services allow organizations to expose their core competencies programmatically over the Internet using standard languages. The objective of Services Computing is to enable IT services and computing technology to perform business services more effectively and efficiently. The Service Oriented Computing (SOC) paradigm refers to the set of principles, methods and concepts that represent computing in Service Oriented Architecture (SOA) in which software applications are constructed based on independent component services with standard interfaces. SOC separates software development into three independent parties: Application builders (by software engineers), service providers (by programmers), and service brokers (joint effort from standard organizations, computer industry, and government).

Service providers: They use a traditional a programming language such as Java, C++, or C# to write programs and all work will be wrapped with open standard interfaces, call services, or Web services. Application builders can easily and simply use the services without further communication with the service providers if the services are available over the internet. The same services can be used by many applications.

Service brokers: Service brokers allow services to be published and registered for public access. It help application builders to find services they need.

Application builders: The application builders represent the end users to specify the application logic in a high level specification language, using standard services as components instead of constructing software from scratch using basic programming language constructs. Application builders are software engineers who have a good understanding of the application domain and software architecture.

WEB SERVICES

Web services are software components that communicate using constant, standards based Web technologies including HTTP and XML based messaging. Web services are planned to be accessed by other applications and differ in complexity from easy task to task. A Web service is a well defined concept of a set of computational or physical activities involving a number of resources, designed to fulfill the needs of customers or for business requirement. With the technology of Web services, companies or enterprises are able to represent their internal business processes as services and make them reachable via the Internet. Nowadays, companies such as Face book, Google, Twitter, and Amazon have offered Web services to provide simple access to some of their resources, enabling third parties to combine and reuse their services. According to a recent research, there are 28,606 Web services available on the Web, offered by 7739 different providers. With the fast acceptance of Cloud computing, social networks, and Web of Things further

accelerate the increase of available Web services on the Internet [9]. A Web Service uses standard web technologies to interact with other services. It is a modular and self described application. WS is limited to relatively simple functionalities, so it is necessary to combine a set of individual WS to obtain a more complex one, namely a composite WS. The WS composition problem aims is to selecting a set of existing WSs or composite WS which can satisfy the user's functional and non functional requirements [8]. Every service comes with a contract between customer and provider to make sure that both the parties understand the level of service quality that they should in that order expect and provide. In addition, contract also gives benefits to both customers and providers in such a way that customers get what they expect for their paid electronic solutions, while providers can prepare effective resources planning and avoid from over committing the resources to certain services. This contract is known as Service Level Agreement (SLA) [7].

COMPLEX NETWORK

A network is a set of items, which are vertices and nodes and the connections between them is called edges. Systems taking the form of networks (also called graphs) we abundant in the world like Internet, the World Wide Web, social networks, organizational networks and business networks and relationship between companies, metabolic networks, neural networks etc. A complex network is a graph (network) with non trivial topological features that do not occur in simple networks such as lattices or random graphs but often occur in graphs modeling real systems.

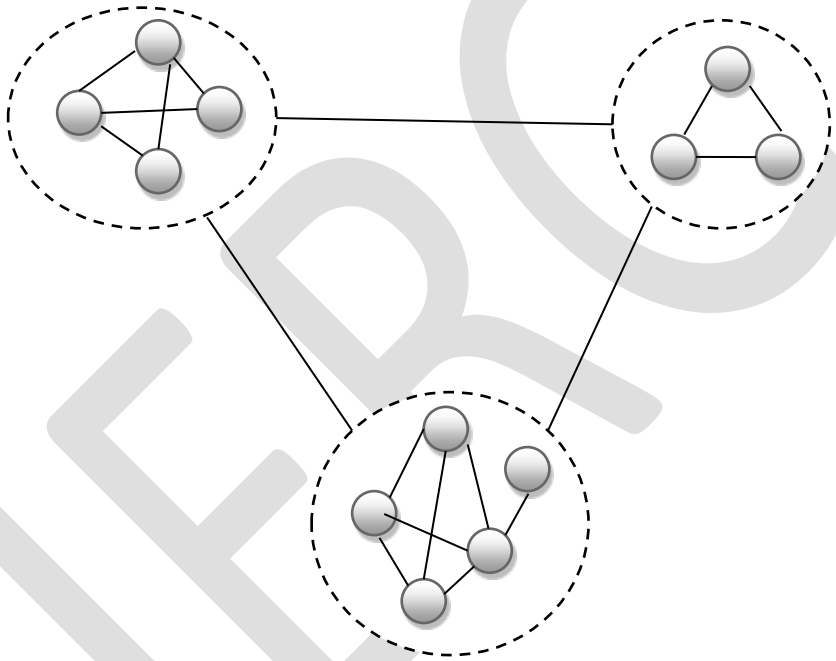

FIG.1: A simple graph with three communities, enclosed by the dashed circles [15].

The study of complex networks is a young and active area of scientific research motivated largely by the observed study of real world networks such as computer networks and social networks. Complex networks are currently being studied across many fields of science and engineering stimulated by the fact that many systems in nature can be described by models of complex networks. It is a large set of interconnected nodes, in which a node is a basic unit usually information content. Examples include the Internet, Internet is a complex network of computers and routers connected by various physical or wireless links like the World Wide Web, which is an vast virtual network of web sites connected by hyperlinks; and various communication networks, food webs, biological neural networks, social, metabolic networks and etc [1].

COMMUNITY STRUCTURE

Communities are defined as collections of nodes where connection with the community are dense, but connection between connections between communities are infrequent. There are so many networks or real world networks which show community structure. Most of these networks are generally sparse in global yet dense in local. Community structure is an important network property which can reveal many hidden features of the given networks. For example, the communities in World Wide Web correspond

with interested topics. Individuals belonging to the same community tend to have properties in common, in social networks, [3] communities also play an important role in information networks. There are several community detecting algorithms for detecting community structure in large scale complex networks. To study the effects of community structure on network properties, the modeling of real networks with community structure is very important. Large scale complex networks with thousands to millions of nodes are everywhere across many dissimilar domain. Community structures in these networks are often of specific interest. For example, communities represent customers with similar interests in customer preferences databases, or regions of uniform long term climate variability in climate networks. Identifying communities in a network is a complex problem due to the existence of several definitions of community and the intractability of many community detection algorithms. In a simple way, a community is a densely connected group of nodes that is sparsely connected to the rest of the network [2,3]

COMMUNITY DISCOVERY PROBLEM

The study of networks (a set of nodes interconnected by links) has become a global topic in many branches of science. This is because many systems of interest can be represented in this way, for example, Internet, the WWW, food webs, neural networks, communication networks, social networks etc. Community detection is also considered to be used for improving the search engines. Hence, community identification is essential for discovering and also for understanding the overall structural and functional properties of a large network [3]. Large networks have different properties like small world effect, high network transitivity, power law degree distributions etc. Community structure is defined as the possibility of recognizing within the networks, subsets of nodes which are which are densely connected among themselves than rest of the network. If we can detect these structures then we can get the information of practical importance. The detection and characterization of community structure in networks is the main issue. For example, groups within the worldwide web might correspond to sets of web pages on related topics groups. Merely the finding that a network contains tightly knit groups at all can convey useful information: if a metabolic network were divided into such groups, for example, it could provide proof for a modular view of the network's dynamics, with different groups of nodes performing different functions with some degree of independence [5]. Community structure methods normally assume that the network of interest divides naturally into subgroups and the experimenter's job for finding those groups. The number and size of the groups are determined by the network itself.

The Algorithms

Algorithm of Girvan and Newman: It is the first algorithm of the modern age of community detection in graphs. It is a hierarchical divisive algorithm, in which links are iteratively removed based on the value of their betweenness, which expresses the number of shortest paths between pairs of nodes that pass through the link The Girvan–Newman algorithm detects communities by progressively removing edges from the original network. The connected components of the remaining network are the communities. Instead of trying to construct a measure that tells us which edges are the most central to communities, the Girvan–Newman algorithm focuses on edges that are most likely "between" communities. The algorithm's steps for community detection are summarized below:

1. The betweenness of all existing edges in the network is calculated first.
2. The edge with the highest betweenness is removed.
3. The betweenness of all edges affected by the removal is recalculated.
4. Steps 2 and 3 are repeated until no edges remain.

The Louvain Method: The Louvain Method for community detection is a method to extract communities from large networks created by Vincent Blondel. The method is a greedy optimization method that appears to run in time O(n log n). The Louvain method is a simple, efficient and easy-to-implement method for identifying communities in large networks. The method is a greedy optimization method that attempts to optimize the "modularity" of a partition of the network. The optimization is performed in two steps. First, the method looks for "small" communities by optimizing modularity locally. Second, it aggregates nodes belonging to the same community and builds a new network whose nodes are the communities. These steps are repeated iteratively until a maximum of modularity is attained and a hierarchy of communities is produced [21].

Expectation maximization algorithm: An expectation maximization (EM) algorithm is an iterative method for finding maximum likelihood or maximum a posteriori (MAP) estimates of parameters in statistical models, where the model depends on unobserved latent variables. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which

computes parameters maximizing the expected log likelihood found on the *E* step. These parameter estimates are then used to determine the distribution of the latent variables in the next E step.

Label Propagation Algorithm: This algorithm uses the concept of node neighborhood and the diffusion of information in the network to identify communities. Initially, each node is labeled with a unique value. Then an iterative process takes place, where each node takes the label which is the most spread in its neighborhood. This process goes on until one of several conditions is met, for in-stance no label change. The resulting communities are defined by the last label values [22].

LITERATURE REVIEW

Chunguang Li and Philip K Maini (2005), have theoretically analyzed the scaling properties of the network by using a mean-field approach and proposed an evolving network model with community structure. The analytical and numerical results indicate that the network can produce community structure

Filippo Radicchi et al. (2003), has discussed about the communities, community detection algorithm. He studied the Divisive and agglomerative algorithms for detecting the communities in the complex network.

Gunce Keziban Orman and Vincent Labatut (2011), have discussed the properties of complex network and compare the five community detection algorithms by using a set of artificial networks.

M. E. J. Newman (2006), has discussed about the community detection methods, the methods of optimal modularity, further techniques of modularity maximization and the methods for dividing the network into more than two communalities.

Michelle Girvan and M. E. J. Newman (2001), have discussed the traditional community detecting algorithms and also discussed the new ”edge betweenness” algorithm for detecting the community structure on computer-generated graphs and on real-world networks.

Nan Du, Bai Wang and Bin Wu (2008), have proposed a new method Com Tector for the community detection in complex networks. This algorithm is based on the overlapping nature of cliques in real world, this algorithm can be applied to many large sparse graph.

Quan Z. Sheng et al. (2014), have studied about the state of the art of Web services composition. They abstract a generic model for the life cycle of Web services composition, which is used to compare different research prototypes based a set of assessment criteria. They compared number of Web services composition standards and services composition platforms.

Steve Harenberg et al. (2014), have discussed the community detection algorithms for overlapping and disjoint community detection on large scale real world networks. The algorithms evaluated by measuring the structural properties of the communities and their performance. Their results show that two types of measures are not equivalent.

V. Gabrel , M. Manouvrier, C. Murat (2014), have discussed about the workflow based WS composition problem and present the complexity analysis of QoS (quality of service)aware workflow based WS composition.

Xizhe Zhang et al. (2013), have studied about the web services, complex structure, community structure and the community structure of structural service networks formed by public web services available on the Internet.

LIMITATIONS OF EARLIER TECHNIQUES

1. The use of communicate discovery during classification and retrieval of cloud web services has been ignored.
2. The effect of community discovery on the different natured users i.e. those who demands multiple things at a time has also ignored.

CONCLUSION AND FUTURE SCOPE

In this paper, a survey on various complex networks has been done. From the survey, it has been found that the useage of communicate discovery during classification and retrieval of cloud web services has been ignored. Moreover the effect of community discovery on the different natured users i.e. those who demands multiple things at a time has also ignored.

Therefore in near future, the enhancement of the cloud web services can be done further by using the nodes i.e. high end server and respective servers based upon the community structure to improve the results.

REFERENCES:

[1] Steve Harenberg, Gonzalo Bello, L. Gjeltema, Stephen Ranshous, Jitendra Harlalka, Ramona Seay, Kanchana Padmanabhan and Nagiza Samatova (2014), "Community detection in large-scale networks: a survey and empirical evaluation", doi: 10.1002/wics.1319.

[2] Chunguang Li and Philip K Maini (2005), "An evolving network model with community structure**",** journal of physics a: mathematical and general, doi:10.1088/0305-4470/38/45/002.

[3] Nan Du, Bai Wang and Bin Wu (2008), "Community detection in complex networks". Journal of computer science and technology, *Vol.23, No.4*.

[4] Filippo Radicchi, Claudio Castellano†, Federico Cecconi, Vittorio Loreto†, and Domenico Parisi (2004). "Defining and identifying communities in networks",PNAS , vol.101, no. 9, 2658–2663, March 2, 2004.

[5] M. E. J. Newman (2006), "Modularity and community structure in networks", PNAS, vol. 103, no. 23, 8577–8582.

[6] M. E. J. Newman (2008), "Detecting community structure in networks", University of Michigan, Ann Arbor, MI 48109-1120.

[7] Mohd Hilmi Hasan, Jafreezal Jaafar, Mohd Fadzil Hassan (2012), "Monitoring web services' quality of service: a literature review", Springer, DOI 10.1007/s10462-012-9358-7.

[8] V. Gabrel, M. Manouvrier, C. Murat (2014), "Web services composition: Complexity and models", ScienceDirect, http://dx.doi.org/10.1016/j.dam.2014.10.020.

[9] Quan Z. Sheng, Xiaoqiang Qiao, Athanasios V. Vasilako, Claudia Szabo, Scott Bourne, Xiaofei Xu (2014), "Web services composition: A decade's overview", ScienceDirect, http://dx.doi.org/10.1016/j.ins.2014.04.054.

[10] Xizhe Zhang, Shuai Feng, Ying Yin, Bin Zhang (2013), "Community Discovery of Public Cloud Web Services based on Structural Networks" IEEE, 978-1-4673-4714-3/13.

[11] Andrea Lancichinetti and Santo Fortunato (2009), "Community detection algorithms: A comparative analysis", DOI: 10.1103/PhysRevE.80.056117.

[12] Jierui Xie, Stephen Kelley, Boleslaw K. Szymanski (2013), "Overlapping Community Detection in Networks: The State of the Art and Comparative Study", ACM Computing Surveys, Vol. 45, No. 4, Article 43.

[13] M. Girvan and M. E. J. Newman (2002), "Community structure in social and biological networks", PNAS, vol. 99, no. 12, 7821–7826.

[14] M. E. J. Newman, "The structure and function of complex networks", Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A.

[15] Santo Fortunato, "Community detection in graphs", Complex Networks and Systems Lagrange Laboratory, ISI Foundation, Viale S. Severo 65, 10133, Torino, I-ITALY.

[16] Natali gulbahce and sune lehmann, "The art of community detection", Center for Complex Networks Research, Northeastern University, Boston, MA 02115, USA.

[17] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, Ploutarchos Spyridonos (2011), "Community detection in Social Media Performance and application considerations", Springer, DOI 10.1007/s10618-011-0224-z.

[18] "Web Services Conceptual Architecture (WSCA 1.0)" by Heather Kreger IBM Software Group, (2001).

[19] Erin Cavanaugh, "Web services: Benefits, challenges, and a unique, visual development solution".

[20] Esther Galbrun, Aristides Gionis, Nikolaj Tatti (2014), "Overlapping community detection in labeled graphs", Springer, DOI 10.1007/s10618-014-0373-y.

[21] http://perso.uclouvain.be/vincent.blondel/research/louvain.html.

[22] Günce Keziban Orman and Vincent Labatut (2009), "A Comparison of Community Detection Algorithms on Artificial Networks" Springer LNAI 5808, pp. 242–256, 2009