

# Study on Dissemination of Scholarly Papers with Open Source Options about Regional in-Service Courses

Masami Yoshida

**Abstract**—The study in this article aims to draw scholarly information communication of designers of in-service courses. In Japan, the system of periodical compulsory training for in-service teachers was introduced. These trainings were designed by local universities. Then, the author collected scholarly papers with open source options deployed in Google Scholar, and introduced into social graph analysis by using reference lists in papers. In addition, the author introduced extensive analyses. As result, there is limited relationships among extracted papers. 34.5% of papers reported about specific subject areas, and there was limited cases to report commonly happen events. It was proposed that method of academic communication should be developed to share common interests over the difference of personal academic field.

**Index Terms**—In-service training, open source, scholarly papers, social graph.

## I. INTRODUCTION

The revised teacher's license law was approved by the Central Education Council of Japan in 2007. The purpose of this revised teacher's license law was "to ensure teachers systematically acquire up-to-date knowledge and skills to maintain the professional competencies necessary for today's educators, teach with confidence and pride, and gain respect and trust from the public." Until this revision, teachers in Japan were granted a lifetime license after students completed pre-service training [1]. However, due to this revision, Ministry of Education, Culture, Sports, Science and Technology in Japan (hereinafter named as MEXT) changed teachers' license system from 2009. Licenses should be renewed every ten years. In addition, teachers are required to take license renewal courses totaling over 30 hours of instruction over a two year period prior to the date of license expiry. Contents of teaching license renewal course (hereinafter abbreviated as TLRC) was designated into two parts, where a) training on current educational issues (more than 12 hours) and b) training on subject teaching, student counseling and guidance, or issues related to enhance teaching (more than 18 hours) were included (see Fig. 1).

In Japan, public school education is operated by local government offices. And, MEXT encourages boards of education in each local administrative area to undertake initiatives. The accredited universities in each local area started to design TLRC independently in accordance with the

request of the board of education.



Fig. 1. In-service training for license renewal.

However, MEXT had not presented a concrete standard necessary contents for the TLRC. Even each accredited university plans TLRC in response to this policy development, but it is extremely limited chances for course designers to exchange concrete ideas of TLRC in nationwide scale.

## II. INFORMATION COMMUNICATION TECHNOLOGY AND TEACHER COMPETENCY

The globalization and internationalization are broadly deployed. At this moment, we can feel that the world is going to be flat [2] not only in economical transactions and cultural exchanges, but also in scholarly information. Also, a new term 'glocalization' popularized by Robertson [3] indicated importance that a new outcome of local conditions was toward global pressures. This means that local culture assigns meaning to global influences, and then it can blur the boundaries between the local and the global.

In addition, the influence of Web2.0 reaches a stage in post-industrial societies where ICT is sufficiently powerful and low cost to be used even by local academics. The integration of pervasive technologies with changes processed the ways how we communicate and learn from information. And, social relationships have been altered by dissemination of online information in terms of scale and size.

As initial investigation to view the scholarly communication among course designers of TLRC in Japan, the author made a survey study of scholarly papers circulated with open source option. Google Scholar was recognized as a conspicuous scholarly site, especially known as service of open access. Here, papers were collected by using keywords of ICT related areas where all of designers in the area can access without any troubles. Collected papers with open source option and including reference lists were key information of TLRC, because these scholarly papers can be led initiative prospective changes in course designing.

For analysis, the author introduced social graph theory and its method for data collection that include information of each document. This implied not only collect information of TLRC, but also analyze related studies with relation of academic

Manuscript received September 28, 2014; revised January 7, 2015. This work was supported in part by the research project. Grant-in-Aid for scientific research (B) of JSPS, project number 26301035.

M. Yoshida is with Graduate School of Social Sciences and Humanities, Chiba University, Chiba, Japan (e-mail: yoshida-m@faculty.chiba-u.jp).

papers that were easily accessed by course designers in online.

### III. GRAPH

In this study, social graph theory is widely used to analyze data. A social graph illustrates the personal relations of online users (see Fig. 2).

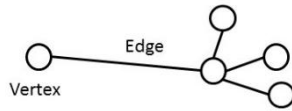


Fig. 2. Graph.

The graph is made up of “vertices” and “edges” that connect them. A vertex usually corresponds to a person, and an edge usually means a message sent to a clear target person. The natural graph involves hubs of vertices with many edges, “bridges” of edges that connect two separate groups, and “clusters” that are groups of vertices. These characteristics are identical in the following “complex networks.”

#### A. Complex Networks

The field of complex networks involves studying graphs with non-trivial topological features. A complex network is a graph with features that do not occur in simple networks such as regular graphs or random graphs but often occur in real graphs. Most social, biological, and technological networks display patterns of connection between their elements that are neither purely regular nor purely random. There are three known characteristics in complex network research.

##### 1) Scale free network

The graph of a Gaussian distribution shows a bell-shaped curve, and most scientific studies on education use this tool. However, Barabási *et al.* [4] found that human social activities followed *power law*. A power law distribution has more data with extreme values than a Gaussian distribution, indicating the existence of hubs. Such a network is called a Scale Free Network. In addition, the power law has increased magnitude of skewness of the distribution in the ranking graph.

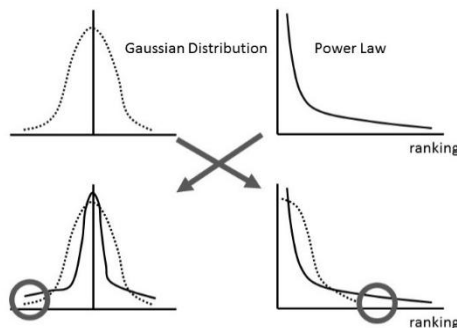


Fig. 3. Comparison of gaussian distribution and power law.

When the conventional Gaussian distribution and power law are drawn together on the same graph, remarkable differences can be found in the portion of the distribution with a large number of occurrences far from the central part (see Fig. 3). This characteristic was popularized by Anderson [5] of Wired.com in 2004 as the long tail distribution, which is

recognized as a particular type of heavy-tailed distribution.

##### 2) Small world phenomenon

Milgram [6] reported that people who may be very far apart physically and socially are still connected with relatively small paths. He showed evidence of the phenomenon by investigation and called the phenomenon “six degrees of separation.” Thus, many vertices in a social graph are connected by a small distance to one another under the small world phenomenon.

Alternatively, Granovetter [7] created theories to explain the meaning of possibly reaching an edge with an unknown vertex. Particularly, he explained the power of spreading information in social networks. He surveyed people in a suburban community about their channels of mobility information, and reported that over 80% of people selected their jobs based on information from unknown people. This was concluded as “the strength of weak ties,” which plays an important role in affecting social cohesion.

##### 3) Cluster

In complex networks, a process of “preferential attachment” appears in which new network vertices prefer to make a connection to the more popular existing vertices. This is called as the Watts-Strogatz model and explains high clustering [8].

### IV. METHOD

In this study, scholarly papers in Google Scholar were retrieved, and especially downloadable full text of papers with open source options were collected. Then, edges were coded based on scholarly papers in the reference list of each extracted paper. Thus, data of social graph were developed based on papers and reference data of these papers. Here, collected data and processing analysis have the following characteristics.

- Social Graph is directed graph, because of a paper is eventual product. Also, there is no reciprocal edge pair.
- There is no duplicated edge.
- If a paper has no reference list, and this data are not appeared in the social graph.
- A paper and including papers in a reference list construct a cluster.

#### A. Google Scholar

In this study, free of charge service of Google Scholar is used to collect data. Google Scholar is the service to retrieve relevant work of keywords across the world of scholarly research. In fact, distribution of records for source items in Scopus or Web of Science are larger than Google Scholar [9], but Google Scholar involves records of scholarly papers written by Japanese-language.

In addition, Google Scholar is known as its strong retrieval function. And, it is possible to search matching keywords in the body text of papers where other academic database retrieve keywords from abstract of papers. Importantly, Google Scholar has links to the free full text versions for almost half of its find [9]. And, academics at a well-endowed library with good link resolver software and blessed with competent and motivated systems librarians could achieve a

better link-through rate. This means that Google Scholar can include links to the online collections of users' academic libraries. But, the author did not use domain specific 'library links' function of Google Scholar in this study, because it dominated only inside campus and was not commonly available service across universities.

The author used 'TLRC' and 'informatics education' as keywords of retrieval (used Japanese-language keywords of above terms). Here, a Japanese term of 'informatics education' is common in educational field than 'ICT education' or 'computer education', because 'informatics education' is the term of a school subject name.

The author did not use function to track citation counts for papers because results of this function are reported as not to be stable [10].

Date of retrieval: September 2nd, 2014

### B. Analysis of Social Graph

The NodeXL, an extendible toolkit for community exploration implement, an add-in to the Microsoft Excel was used for visualizing social graphs and calculating graph metrics. The NodeXL supports multiple social network data providers that import graph data (vertex and edge lists) into the Excel spreadsheet. However, the graph data were collected manually in this study because different file formats were used in scholarly papers to open to the public.

## V. RESULTS

The author could collect 29 full text of scholarly papers with open source option (see Table I) from 426 items in retrieval results of Google Scholar. Calculated rate of downloadable paper (6.8%) was lower than the average rate of the preceding study in English documents (46.96%, [11]).

TABLE I: COLLECTED DATA

Year	Retrieved	Downloadable	Rate (%)	No Reference List
2009	115	7	6.1	-
2010	71	5	7.0	-
2011	64	7	10.9	1
2012	75	3	4.0	-
2013	58	2	3.4	1
2014	43	5	11.6	2
Total	426	29	6.8	4

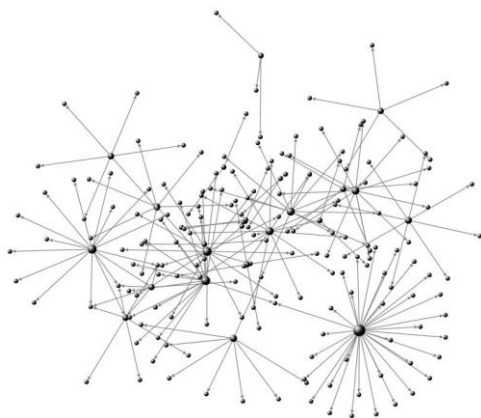


Fig. 4. Social graph of paper relationship, vertex size proportional to vertex out-degree statistics (Harel-Koren fast multiplex algorithm [12], [13]).

Beyond our belief, four papers inside 29 papers had no

reference list. Then, finally 25 unique papers are used for continuous analysis in this study.

Fig. 4 shows a calculated social graph. Papers in a reference are shown as vertices, and are connected with a vertex of a retrieved paper.

## VI. DISCUSSION

Table II shows basic metrics of calculated social graph. It turned out that only one fourth of vertices have been connected. But, retrieved 25 unique vertices involve a considerable number of edges. Then, it can be said that observed social graph is not random graph and included clusters, but it is still needed to develop *bridges* for clusters to become a *complex network*.

TABLE II: GRAPH METRICS

Metrics	Retrieved
Vertices	200
Unique Edges	186
Connected Components	22
Maximum Vertices in a Connected Component	50
Maximum Edges in a Connected Component	49
Maximum Out-degree	33
Maximum Diameter	4
Average Diameter	2.39
Graph Density	0.0046

As the observed characteristics of the Genesis of area for TLRC, relations among retrieved 25 unique vertices are limited even where each vertex have high out-degree. It should be clear that interdisciplinary research communication is inevitable in order to enhance this academic area. Therefore, the average diameter smaller than Milgram's six degree does not show the close relationship of vertices, but limited ties among clusters.

### A. Group Analysis

Fig. 5 shows the result of group analysis by using Girvan-Newman cluster algorithm [14]. It is obvious that connections between clusters are extremely limited and many of 25 unique papers do not interrelated. A box in the left bottom shows the sub-graph of three papers of same author, and it shows clear *preferential attachment*. And, this leads conclusion of independent academic activities of this social graph.

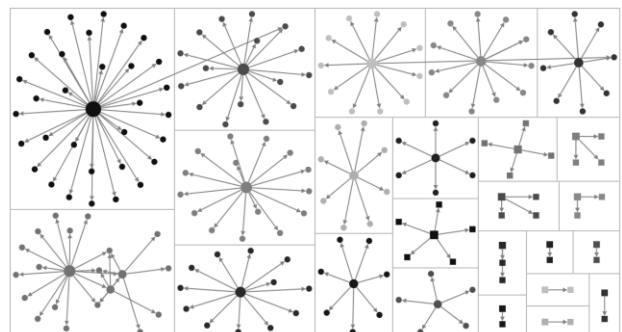


Fig. 5. Result of group analysis, Harel-Koren fast multiscale algorithm is used for layout, each of the cluster in its own box.

### B. Area Analysis

Table III shows the deployment of areas of extracted papers, and it is found out that 34.5% of papers report about subject area studies. Also, some of papers report administrative system of newly introduced TLRC. Space left for discussing commonly happen cases or issues for improving course designing that are intended to exchange ideas and information through interdisciplinary academic area is limited.

Area	Number
Lesson Designing	4
Review	1
Administration	2
Evaluation	4
+ Teaching Method	8
+ Materials Development	9
+ Content Development	1
Total	29

(+: Subject area study)

### C. PageRank Analysis

PageRank algorithm is the variant of eigenvector centrality. This is used to rank web pages in Google [14]. PageRank is a way of measuring the importance of website pages. According to Google:

“PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites. (Cited from Optimization Theory [15]).”

Decision in this indicator involves an idea that edges from important vertices are worth more than edges from unimportant vertices. All vertices starts off equal, but as the computation progresses, vertices with more edges start gaining importance. Their importance propagates out to the vertices to which they are connected. After re-computing many times, the values stabilize, other mathematical factors are used to modify data, such as the dampening factor, resulting in the final values for PageRank.

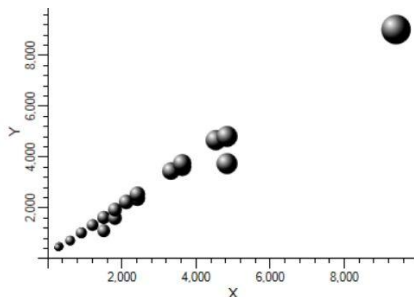


Fig. 6. Paper network mapping degree to the X axis and PageRank to the Y axis. Out-degree to vertex size. Axes are shown, and edges are hidden.

Fig. 6 shows close relationship among Degree, PageRank and Out-degree. This result also implies the situation of the social graph with independent clusters and limited citation rate but each paper includes different number of reference papers. At this moment, importance of each vertex is same level as the number of papers in a reference list. As long as

TLRC covers wide area of contents of training, this characteristic is appeared for a while. However, once emerging commonly discussing topics and connection across clusters, it is inferred that importance of papers changes drastically.

## VII. CONCLUSION

In fact, action of MEXT includes a contradiction; on one hand, advocating the decentralization of authority and respect for local governments; while, on the other hand, reinforcing its supervisory powers [16]. MEXT should show the national standards of professional competence for teachers. Standardizing teacher education programs will help to guarantee higher standards in quality of teachers and enhance communication among course designers.

As conclusion of findings in this study, there is limited relationship among papers of TLRC with open source option. In fact, it is easy for academics to report papers in existing academic societies, but there are limited chance for course designers to share common interests over the difference of major areas.

Then, in order to put scholarly communication of TLRC designers into effect, further enhancement should be introduced to make Granovetter's weak ties. It needs for course designers to be developed a new academic association as soon as possible.

## REFERENCES

- [1] MEXT. (Sep. 3, 2014). Improvement of the Quality of Teachers. [Online]. Available: <http://goo.gl/HTI5z2>
- [2] T. Freidman, *The World is Flat*, New York: Farrar, Straus and Giroux, 2005.
- [3] R. Robertson, “Glocalization: time-space and homogeneity-heterogeneity,” in *Global Modernities*, M. Featherstone, S. Lash, and R. Robertson, Eds., London: SAGE, 1995, pp. 25-44.
- [4] A.-L. Barabási, R. Albert, and H. Jeong, “Mean-field theory for scale-free random networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 272, pp. 173-187, 1999.
- [5] C. Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More*, Hachette Digital, Inc., 2006.
- [6] S. Milgram, “The small world problem,” *Psychology Today*, vol. 2, pp. 60-67, 1967.
- [7] M. S. Granovetter, “The strength of weak ties,” *American Journal of Sociology*, pp. 1360-1380, 1973.
- [8] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, pp. 440-442, 1998.
- [9] P. Jacsó, “As we may search-comparison of major features of the web of science, scopus, and google scholar citation-based and citation-enhanced databases,” *Current Science -Bangalore-*, vol. 89, p. 1537, 2005.
- [10] A. V. Kulkarni, B. Aziz, I. Shams, and J. W. Busse, “Comparisons of citations in web of science, scopus, and google scholar for articles published in general medical journals,” *JAMA*, vol. 302, pp. 1092-1096, 2009.
- [11] P. Jacsó, “Google scholar: The pros and the cons,” *Online Information Review*, vol. 29, pp. 208-214, 2005.
- [12] D. Harel and Y. Koren, “A fast multi-scale method for drawing large graphs,” in *Proc. 8th International Symposium on Graph Drawing*, 2001, pp. 183-196.
- [13] D. Harel and Y. Koren, “A fast multi-scale method for drawing large graphs,” *Journal of Graph Algorithms and Applications*, vol. 6, pp. 179-202, 2002.
- [14] J. Golbeck, *Analyzing the Social Web*, Newnes, 2013.
- [15] Optimization Theory. (Sep. 13, 2014). *How Page Rank Works*. [Online]. Available: <http://goo.gl/p09gi3>
- [16] N. Nakayama, A. Takagi, and H. Imamura, “Teacher certification renewal system: An analysis based on a nation-wide survey of Japanese

teachers of English,” *Educational Perspectives*, vol. 43, pp. 28-37, 2010.



**Masami Yoshida** was born in Osaka, Japan in 1957. His major is in educational technology and B.Sc. and M.Edu.

He used to work in the Faculty of Education, Toyama University, and did studies about pre-service trainings. Then, he moved to National Institute of Multimedia Education and engaged a research project for faculty development. He also had experiences to contribute to Japan International

Cooperation Agencies (JICA) as an expert and was dispatched to Thailand, Malaysia, and Papua New Guinea for educational development there.

Recently, he provided training for directors of educational TV from African, South American and Asian countries in a JICA training center.

Now, he works in the fields of public education, studies of public affairs, Graduate School of Social Sciences and Humanities, Chiba University, Japan. His recent publication is “Countermeasures and culture preparation of cyberbullying”, the center for research and development in teacher education, Faculty of Education, Chiba University (ed.), Tokyo: Hukumura Pub.

Prof. Yoshida is a member of the Japan Society for Educational Technology, Japan Association for Educational Media Society, Japan Society for Information and Systems in Education, and Japan Association for College and University Education.