

Cloud Computing Research in the IS Discipline: A Citation/Co-Citation Analysis

Nianxin Wang

School of Economics and Management
Jiangsu University of Science and Technology

nianxin.wang@gmail.com

Huigang Liang*

College of Business
East Carolina University

huigang.liang@gmail.com

Yu Jia

School of Economics and Management
Jiangsu University of Science and Technology

rebitth@live.cn

Shilun Ge*

School of Economics and Management
Jiangsu University of Science and Technology

jzgsj@jzerp.com

Yajiong Xue

College of Business
East Carolina University

yajiong.xue@gmail.com

Zhining Wang

School of Management
China University of Mining and Technology

wzncumt@163.com

*Corresponding author

Huigang Liang, Tel: +1 252-737-1462, Email: huigang.liang@gmail.com

Shilun Ge, Tel: +86 511-84401164, Email: jzgsj@jzerp.com

Cloud Computing Research in the IS Discipline: A Citation/Co-Citation Analysis

ABSTRACT

Cloud computing is one of the hottest topics in the field of information systems. In this paper, we conduct a citation and co-citation analysis on cloud computing research published in the 11-year period from 2004 to 2014. A total of 214 papers were selected from 20 leading academic journals in IS and management and 2 prominent international IS conferences for our analysis. In the citation analysis, we rely on the degree centrality and betweenness centrality to identify 41 important papers. In addition, our main path analysis reveals three development stages of cloud computing research: the incubation stage, the exploration stage, and the burgeoning stage. In the co-citation analysis, we employ a principal component factor analysis of the co-citation matrix to identify six major research themes: foundations, SaaS model, security and risk, literature review, adoption and impacts, and modeling. This is among the first studies to examine the knowledge structure of cloud computing research in the IS discipline by using evidence-based analysis methods. Recommendations for future research directions in cloud computing are provided based on our analysis.

Keywords: Cloud computing; Literature review; Citation analysis; Co-citation analysis; Social network analysis; Information system research

1. INTRODUCTION

Cloud computing refers to the delivery of computing as a service rather than a product, i.e., shared resources, software, and information are provided to computers and other devices as a utility over the Internet [1-3]. It has five essential characteristics: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service[2]. Since cloud computing can help firms continuously improve their strategic agility while reducing the complexity of business and IT operations to stay competitive in today's fast-changing environments [4, 5], it has been touted as one of the most promising IT advancements that could fundamentally change how IT solutions are delivered [6]. Gartner predicted that the public cloud computing market will exceed USD 180 billion in 2015 and will occupy most of new IT investments in 2016 [7].

The rapid development of cloud computing markets has attracted much attention from information systems (IS) academics [6]. In recent years, some attempts have been made to summarize existing cloud computing research, map its intellectual structure, and predict its future directions. For example, Venters and Whitley [8] reviewed cloud computing research in technology and service dimensions. Yang and Tate [9] summarized existing cloud computing research themes on cloud computing including technology, commerce, concept, and application. Hoberg et al. [10] structured current cloud computing research into four dimensions from a business perspective: characteristics, adoption determinants, governance mechanisms, and business impact. Ermakova et al. [11] analyzed research progress of cloud computing in the medical field.

These reviews provide useful information of current research on cloud computing, facilitate the accumulation of cloud computing knowledge, and indicate that a phase of critical introspection has begun. This kind of introspection and self-reflection can be viewed as a sign of maturity of cloud computing research. However, we believe that an updated review is warranted based on two considerations. First, the rapid growth of cloud computing research requires periodic review to keep researchers up to date. The existing reviews are mostly published in 2012 or before [8-10]. The only review published in 2013 is limited to cloud computing in healthcare [11]. Considering that reviews

published in 2012 are unlikely to cover papers published in 2012, cloud computing research published since 2012 have not been reviewed in a systematic manner. In fact, our bibliometric analysis shows that cloud computing research entered a new stage in 2012, which provides retrospective evidence for the necessity of this current review.

Second, the existing reviews are mainly based on subjective analysis of experienced scholars in the research field and the modern bibliometric methodology has not been leveraged to compensate for human subjectivity. Some degree of human subjectivity is indispensable to carry out literature review. Yet, reviews purely based on subjective analysis might be constrained by their authors' limited time, energy, and cognitive capacity, and their interpretation of the literature is inevitably influenced by their personal perspectives [12]. It is possible that important papers are omitted or misinterpreted to fit with the authors' own research interests. Unlike these reviews, in this paper we use citation, co-citation, and main path analyses to examine the intellectual structure of cloud computing research. Citation, co-citation, and main path analyses are bibliometric methods that can validate and complement judgements made by human researchers. With the advantages of being objective and quantifiable, these bibliometric methods can provide an empirically duplicable review of the existing cloud computing research. While we still need to interpret the results of the bibliometric analysis and cannot completely eliminate subjectivity, the chances of making human errors can be greatly reduced and a more realistic depiction of cloud computing research can be produced. Thus, subjective review and objective review are complementary to each other and should be used together to improve the quality of literature reviews.

There are two goals for our research. The first is to identify the influential papers of cloud computing research in the IS field, and the second is to delineate the themes that constitute the intellectual structure of cloud computing research in the IS discipline and map the relationships among the themes. We provide recommendations for future research on a variety of issues related to cloud computing based on our citation/co-citation analysis.

2. METHODOLOGY

Figure 1 summarizes the steps of our study. Multiple research methodologies, including citation analysis, co-citation analysis, and social network analysis (SNA), are employed. These methodologies and the rationale for their use are described next.

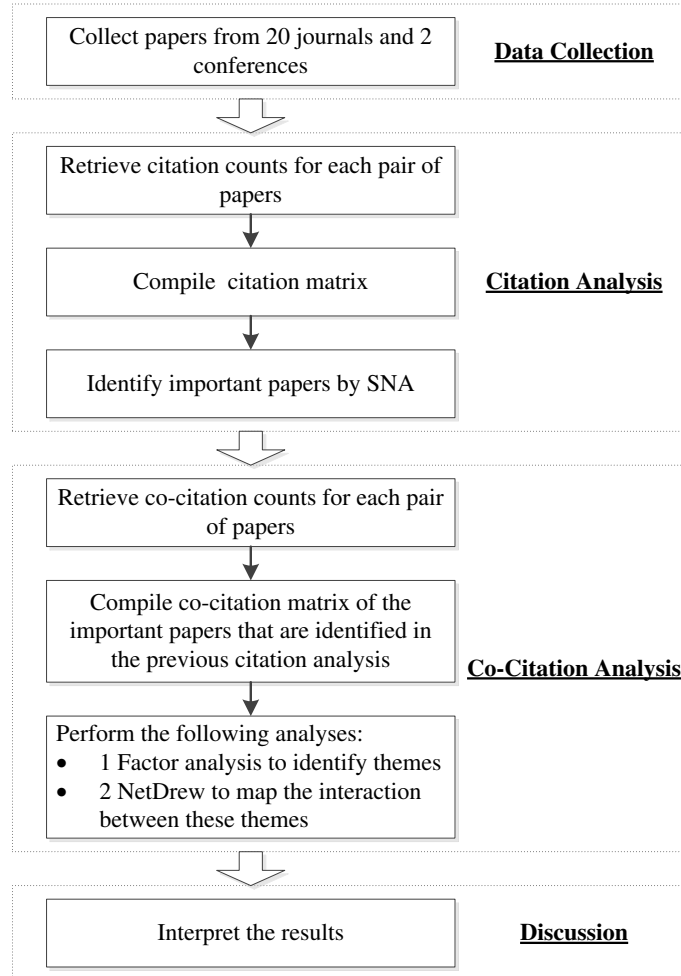


Figure 1. Steps used in this study

2.1 Citation Analysis

A citation occurs when one paper mentions or refers to another paper known as the source paper. As shown in Figure 2A, a citation relationship exists between Paper A and Papers C, D and E. The citation analysis can provide information on the identity of papers which make and receive citations as well as information on the total number citations those papers make or receive. The citation analysis can be used to identify source papers, influential papers, and inheritance relationships among related papers. It has been extensively used to investigate the intellectual structure of many disciplines of social sciences and

natural sciences [13].

2.2 Co-citation Analysis

The co-citation analysis, first introduced by Henry Small in 1973 [14], is an evaluation of semantic similarity of papers that share citations. Co-citation is defined as an occurrence in which two papers are cited together by another paper. The more co-citations two documents receive, the higher their co-citation strength, and the more likely they are semantically related [14]. As shown in Figure 2B, Paper A and Paper B are co-cited by Paper C, Paper D, and Paper E. Thus, Paper A and Paper B have a co-citation strength of 3.

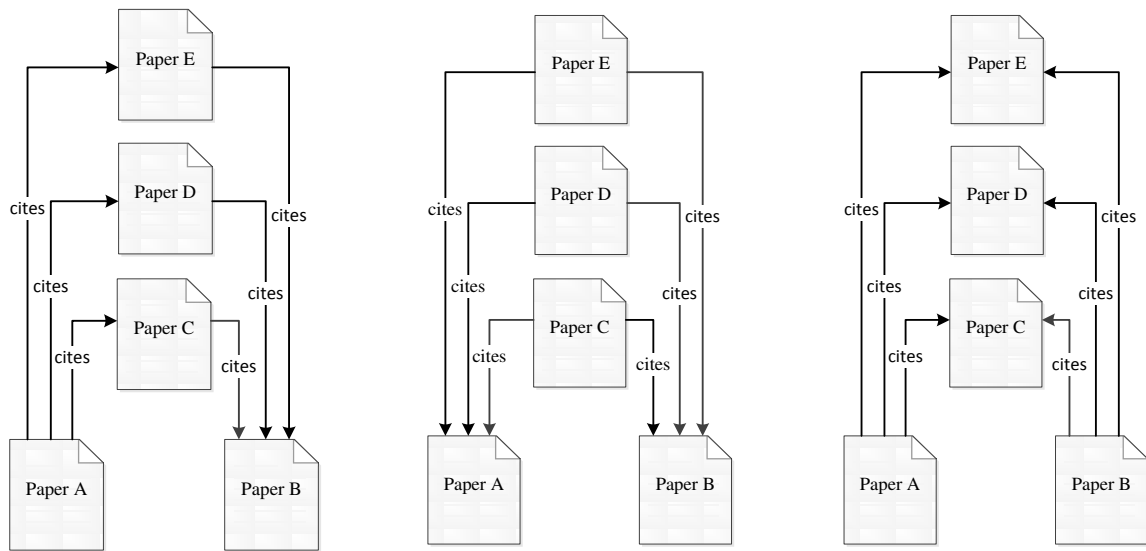


Figure 2A Citation relationship

Figure 2B Co-citation relationship

Figure 2C Bibliographic coupling

Figure 2. Citation, co-citation and bibliographic coupling

Bibliographic coupling that uses citation analysis to establish a similarity relationship between papers is a similar measure as co-citation. The concept of bibliographic coupling was introduced by Kessler in 1963 [15]. Two papers are bibliographically coupled if they both cite one or more papers in common. The coupling strength of two given papers is higher the more citations to other papers they share. As shown in Figure 2C, both Paper A and Paper B cite Paper C, Paper D, and Paper E. Thus, Paper A and Paper B have a bibliographic coupling strength of 3.

In this paper, we prefer co-citation analysis to bibliographic coupling because the usefulness of bibliographic coupling has been questioned. Bibliographic coupling, as a retrospective similarity

measure [16, 17], can only use the past information to establish the similarity relationship between papers, and the coupling strength cannot change over time. Co-citation is able to overcome this problem by considering two papers' incoming citations to assess their similarity, a measure that can change over time. Additionally, the co-citation measure reflects the opinion of many other authors and thus is a more reliable indicator of subject similarity [14]. Co-citation analysis has been used in investigating the foundations of specific fields in the management discipline. For example, Pilkington and Meredith [18] employed citation analysis combined with a network analysis of co-citation data from three major operations management (OM) journals to reveal the intellectual structure of the OM field.

2.3 Social Network Analysis

The flow of communication and exchange of ideas through citations and co-citations can be further understood by combining the above mentioned bibliometric techniques with the SNA technique. Combining mathematics, graph theory, and computer science, SNA can enable the measurement, evaluation, and visualization of relationships and relationship patterns. A social network consists of a finite set of social actors and the relations among them [19]. SNA is a method that investigates the relationships between the social actors through analysis of the network structure by using the relational data. These are the contacts, ties, or information exchanged between actors, which relate one actor to another. The measures of the relationships could include influence, affinity, communication patterns, or cohesion between the actors [19].

In bibliometric analysis, we define a network of papers, authors, or the academic journals. In the case of the network of papers, SNA could be utilized to better understand the relationships among actors (papers) by studying the information exchanged among the members (citation and co-citation) and provide insights into how knowledge is spread throughout the academic community. Therefore, the citation and co-citation analyses can be combined with the SNA to understand the intellectual structure of cloud computing research in the IS discipline by revealing how it is shaped by communication patterns among papers. Scholars have applied SNA to understand the intellectual structure of many fields such as operation management [18], strategic management [20], supply chain management [21], and information systems [22].

2.4 Data Collection

Cloud computing research resides in an interdisciplinary area that includes technological, behavioral, managerial, and social dimensions. Existing reviews mainly focus on papers that study technological issues [23], whereas papers that examine other issues are rarely reviewed. To analyze the current state of cloud computing research that concerns IS researchers, we select 13 top IS journals (*Management Information Systems Quarterly*, *Information Systems Research*, *Journal of Management Information Systems*, *Information Systems Journal*, *Journal of the Association for Information Systems*, *Journal of Strategic Information Systems*, *Journal of Information Technology*, *European Journal of Information Systems*, *Information & Management*, *Decision Support Systems*, *Communications of the Association for Information Systems*, *Communications of the ACM*, and *ACM Transactions on Management Information Systems*), 2 top electronic commerce journals (*International Journal of Electronic Commerce* and *Electronic Commerce Research and Applications*), 5 top management journals (*Service Science*, *Management Science*, *Decision Sciences*, *Harvard Business Review*, and *Sloan Management Review*), and 2 premier IS conference proceedings (*International Conference on Information Systems* and *Americas Conference on Information Systems*) as the publication sources. These sources were searched by using the keywords including cloud computing, cloud service, on-demand service, IaaS, PaaS, and SaaS. The time span is from 2004 to 2014. A total of 208 papers about cloud computing are found.

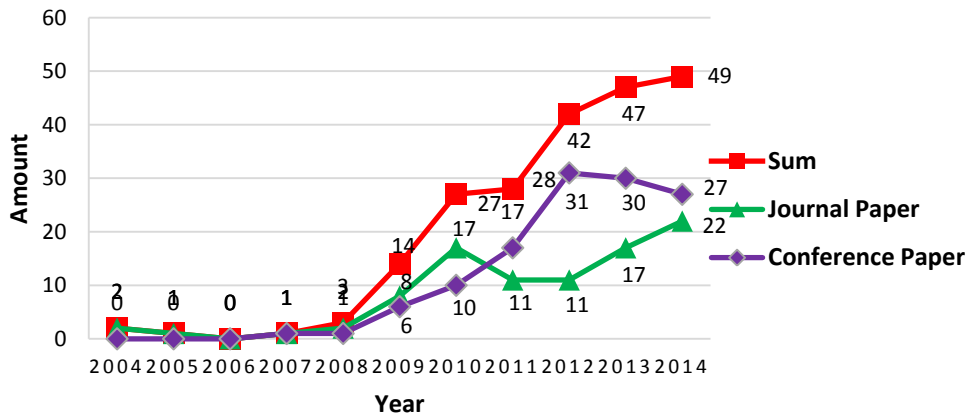


Figure 3. Annual distribution of cloud computing paper

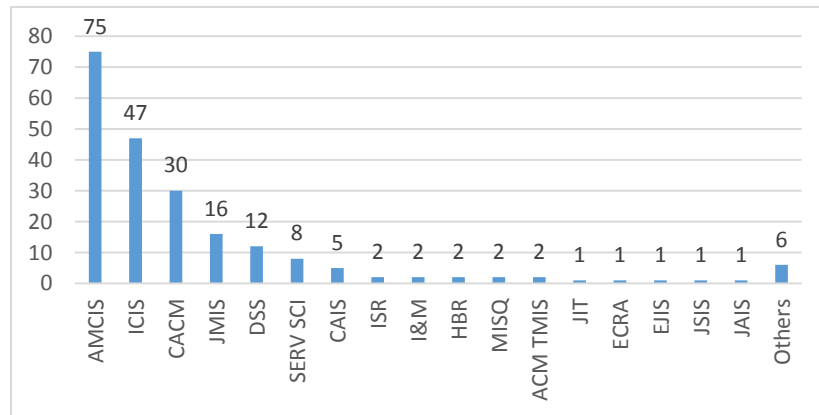


Figure 4. Source distribution of cloud computing papers

After reviewing the 208 papers, 6 additional papers outside the 208 papers were identified. These 6 papers were cited at least 5 times by the original 208 papers. Due to their high citations, we added them into our analysis. Thus, our sample had a total of 214 papers. Among the 6 papers, one is the NIST definition of cloud computing, two are research reports from UC Berkeley and the European Network and Information Security Agency, one is from the 2008 Grid Computing Environments Workshop, and the other 2 are journal papers from Future Generation Computer Systems and Business & Information Systems Engineering. The annual distribution of the number of papers is shown in Figure 3. After a slight decline from 2004 to 2006, the number of publications shows a steady increase from 2006 to 2014. After 2011, papers on cloud computing in ICIS and AMCIS mushroomed, exceeding the number of journal publications. This confirms that it is essential to include conference proceedings in the review. These conference papers represent the latest research trend and they cannot be published on journals quick enough because of the lengthy review process for top IS journals. Figure 4 shows the distribution of cloud computing papers across publication sources. The two conferences have produced 132 papers, which greatly outnumber the 82 journal papers. That is, conference papers account for the majority of the cloud computing literature, which further supports the necessity of including conference papers in our review. Regarding journals sources, CACM, JMIS, and DSS are the top three journals that have published the most cloud computing papers, publishing 30, 16, and 12 papers, respectively. To achieve an in-depth understanding of the current research on cloud computing, we use the citation analysis to identify the important papers in the 214 papers and the co-citation analysis to investigate the research themes of the important papers.

3. RESULTS

3.1 Citation Analysis

Based on the references of the 214 papers, the citation relationship among them is obtained and a 214×214 matrix is generated. The matrix is then imported into two SNA tools, namely, UCINET and PAJEK to identify important papers and conduct main path analyses.

3.1.1 Citation Network

The citation network of the 214 papers is plotted in Figure 5, and the citation relationships between the papers are sparse. One important metric for the citation networks is density, or the number of connections between nodes in the network. If there are no connections between any of the nodes in a network then density is zero. If each node is connected to every other node (the graph is fully connected) then the density is one. Thus, density ranges between 0 and 1. Abrahamson and Rosenkopf suggest that a score above 0.5 indicates high density and below 0.5 indicates low density [24]. For the citation network of cloud computing, the network density is 0.0104. After 44 isolated nodes are removed, the density is 0.0166. The low network density indicates that cloud computing, as an emerging research field, is still in the early stage and that the connections between papers are scant.

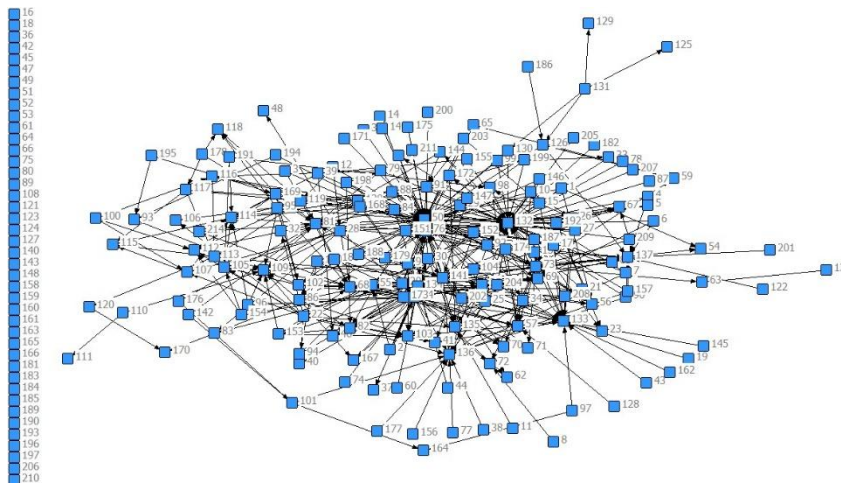


Figure 5. Citation network of the 214 papers

3.1.2 Important Papers

The importance of a paper can be determined by its influence in the citation network, which can be

measured by two indexes, namely, degree centrality and betweenness centrality. Degree centrality is measured as the number of direct ties that a node in the network has [25]. The more ties the node has, the more active, or more central, the node is. Those on the periphery will be disadvantaged compared with the central nodes because they have fewer ties. For directed networks like citation networks, there are two different types of degree centrality. One is in-degree centrality, measuring the number of papers that cites this focal paper. The other is out-degree centrality, measuring the number of papers that this focal paper cites. We use in-degree centrality to show a paper's importance because it reflects how the paper is recognized by other authors.

Betweenness centrality is concerned with the extent to which one node exists on the shortest path (the geodesic distance) between other nodes [25]. It is the number of times that a given node needs a (different) given node to reach any other node by the shortest path. In a communication network, nodes with high betweenness scores can control the flow of information and thus may be able to take on the role of gatekeeper or broker. Betweenness values can also indicate which nodes are viewed most often as leaders [26]. Since the citation network is a type of directed graph, we followed White and Borgatti [27] to compute the betweenness centrality scores of all 214 papers.

Previous studies only employed citation count or in-degree centrality to identify important papers [28, 29]. Compared with these studies, combining in-degree centrality with betweenness centrality can better identify the important papers. In this study, the paper with in-degree centrality of 5 and above or betweenness centrality of 10 and above is considered as an important paper. A total of 41 papers meet this criteria (see Table 1). These 41 papers occupy almost 20% ($41/214=0.1916$) of the 214 papers. According to the "80/20 rule" [30], we contend that 80% of the information about cloud computing studies comes from the 20% of the important papers. Focusing on the 41 important papers can keep the most valuable information and reduce the complexity of data analysis. By reviewing the sources of these 41 papers, we found that ICIS, JMIS, CACM, AMCIS, and DSS contain 11, 8, 6, 3, and 3 highly-cited papers respectively, suggesting that CACM, JMIS, DSS, and the two conferences are important outlets for cloud computing studies to date. In addition, we categorized these papers based on research type. As Figure 6 shows, 28, or 68% of the important papers are non-empirical (18 conceptual + 5 literature

review + 5 modelling), suggesting that scholars are still trying to make sense of cloud computing by elaborating on the basic concepts surrounding this new technology. Only 13, or 32%, important papers report empirical studies, and most of them are survey-based. This suggests that researchers are prudent in making efforts to quantitatively investigate cloud computing questions before they can sufficiently understand the concepts related to cloud computing and are able to theoretically define and empirically measure these concepts.

Table 1. Important papers of citation network analysis

No	Author	Title	Research Type	In Degree	Betweenness Centrality
50	Armbrust et al.[1]	A View of Cloud Computing	Conceptual	57	84.12
132	Mell and Grance[2]	The NIST Definition of Cloud Computing	Conceptual	54	0.00
76	Marston et al.[31]	Cloud Computing - The Business Perspective	Conceptual, Literature Review	32	155.65
134	Armbrust et al.[32]	Above the Clouds: A Berkeley View of Cloud Computing	Conceptual	29	0.00
133	Buyya et al.[33]	Cloud Computing and Emerging IT platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility	Conceptual	20	0.00
136	Weinhardt et al.[34]	Cloud Computing - A Classification, Business Models, and Research Directions	Conceptual	17	0.00
109	Xin and Levina[35]	Software-as-a Service Model: Elaborating Client-side Adoption Factors	Conceptual	14	0.00
68	Cusuman[36]	Cloud Computing and SaaS as New Computing Platforms	Conceptual	13	62.00
137	Foster et al.[37]	Cloud Computing and Grid Computing 360-Degree Compared	Conceptual	13	0.00
81	Benlian and Hess[38]	Opportunities and Risks of Software-as-a-Service: Findings from a Survey of IT Executives	Quantitative, Survey	12	215.67
114	Choudhary[39]	Comparison of Software Quality Under Perpetual Licensing and Software as a Service	Theoretical, Modeling	11	58.00
69	Durkee[40]	Why Cloud Computing will Never Be free	Conceptual	10	1.00

55	Brynjolfsson et al.[41]	Cloud Computing and Electricity: Beyond the Utility Model	Conceptual	9	6.67
135	Catteddu[42]	Cloud Computing Benefits, Risks and Recommendations for Information Security	Conceptual	9	0.00
57	Hayes[43]	Cloud Computing	Conceptual	8	0.00
82	McAfee[44]	What Every CEO Needs to Know about the Cloud	Conceptual	8	0.00
56	Ryan[45]	Cloud Computing Privacy Concerns on our Doorstep	Conceptual	7	0.00
103	Koehler et al.[46]	Customer Heterogeneity and Tariff Biases in Cloud Computing	Quantitative, Survey	7	17.82
112	Susarla et al.[47]	A Transaction Cost Perspective of The “Software as a Service” Business Model	Quantitative, Survey	6	0.00
102	Winkler et al.[48]	The Impact of Software as a Service on IS Authority – A Contingency Perspective	Qualitative, Case Study	5	39.12
105	Benlian et al.[49]	The Role of SaaS Service Quality for Continued SaaS Use: Empirical Insights from SaaS Using Firms	Quantitative, Survey	5	22.00
141	Venters and Whitley[8]	A Critical Review of Cloud Computing: Researching Desires and Realities	Literature Review	5	156.17
73	Yang and Tate[9]	A Descriptive Literature Review and Classification of Cloud Computing Research	Literature Review	4	106.67
91	Ackermann et al.[50]	Perceived IT Security Risks of Cloud Computing: Conceptualization and Scale Development	Quantitative, Survey	4	107.00
113	Bardhan et al.[51]	An Interdisciplinary Perspective on IT Services Management and Service Science	Conceptual	4	103.33
116	Demirkan[52]	Coordination Strategies in an SaaS Supply Chain	Theoretical, Modeling	4	26.00
24	Hoberg et al.[10]	The Business Perspective on Cloud Computing - A Literature Review of Research on Cloud Computing	Literature Review	3	139.07
86	Malladi and Krishnan[53]	Cloud Computing Adoption and its Implications for CIO Strategic Focus – An Empirical Analysis	Quantitative, Survey	3	24.80
107	Huang and	Firm-Level Productivity Analysis for	Quantitative,	3	10.67

	Wang[54]	Software as a Service Companies	Secondary Data		
117	Sen et al.[55]	Demand Information Sharing in Heterogeneous IT Services Environments	Theoretical, Modeling and Simulation	3	16.00
35	Martens and Teuteberg[56]	Risk and Compliance Management for Cloud Computing Services: Designing A Reference Model	Theoretical, Modeling	2	12.80
98	Su[57]	Emergence of Cloud Computing: An Institutional Innovation Perspective	Conceptual	2	14.00
119	Benlian[58]	Service Quality in Software-as-a-Service: Developing the SaaS-Qual Measure and Examining its Role in Usage Continuance	Quantitative, Survey	2	84.67
126	Iyoob et al.[59]	Cloud Computing Operations Research	Conceptual	2	14.00
28	Walther et al.[60]	Success Factors and Value Propositions of Software as a Service Providers - A Literature Review and Classification	Literature Review, Meta-Analysis	1	16.67
79	Demirkan and Delen[61]	Leveraging the Capabilities of Service-oriented Decision Support Systems: Putting Analytics and Big Data in Cloud	Conceptual	1	23.00
92	Giessmann and Stanoevska[62]	Platform as a Service – A Conjoint Study on Consumers’ Preferences	Quantitative, Survey	1	12.00
96	Zainuddin and Gonzalez[63]	Configurability, Maturity, and Value Co-Creation in SaaS: An Exploratory Case Study	Qualitative, Case Study	1	14.00
120	Benaroch et al.[64]	Should We Go Our Own Way? Backsourcing Flexibility in IT Services Contracts	Theoretical, Modeling	1	42.00
144	Loske et al.[65]	Cloud Computing Providers’ Unrealistic Optimism Regarding IT Security Risks: A Threat to Users?	Quantitative, Survey	1	22.00
169	Winkler and Brown[66]	Horizontal Allocation of Decision Rights for On-premise Applications and Software-as-a-Service	Quantitative, Survey	1	28.00

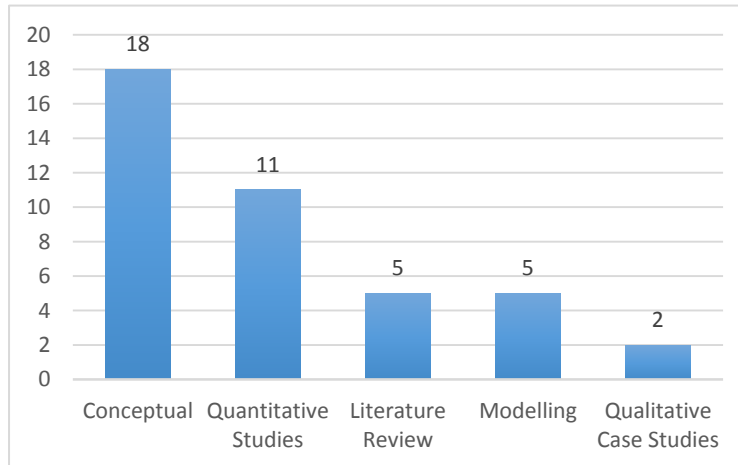


Figure 6. Research types of the 41 important papers

Table 1 shows that the in-degree centrality of Papers 50, 132, 135, 76, 134, 133, 136, and 109 is relatively high (≥ 14), indicating that these papers are major knowledge sources for cloud computing studies. Although the in-degree centrality of Papers 73, 91, 113, 24, and 119 is low (≤ 4), their betweenness centralities are relatively high (> 84). Thus, these papers are vital to the dissemination of knowledge in cloud computing research by bridging between other papers.

3.1.3 Main Path Analysis

Main path analysis is a powerful bibliometric tool that has just begun to be applied by IS researchers to conduct literature review recently [67]. By constructing the position of each paper in terms of its citing and cited papers, main path analysis enable us to make visible the structural backbone of a body of literature [68]. In the course of development of cloud computing research, new articles acquire information from previous articles and add new ideas of their own. The dyadic knowledge flow between the citing and cited articles can be revealed by the citation relationships. The citation network thus can be viewed as a directional network showing scientific knowledge flows. A citation network includes many nodes (articles) and links (citation relationships). A main path refers to a single path starting with an early article that has no cited articles in the field and ending with an article that has not been cited. A citation network can have a large number of main paths. The global main path refers to the main path that has the greatest weight in a citation network, which represents the backbone of

knowledge dissemination in the field [69]. The weight of each link on a global main path reflects the importance of the citation relationship. A link shared by many paths is deemed more crucial than a link rarely shared [70]. The weight of each link is indicated by traversal count which measures the frequency a citation link has been traversed if one exhausts all main paths in a citation network [69]. There are three algorithms for identifying the traversal count: the Node Pair Projection Count (NPPC), which accounts for the number of times each link is involved in connecting all node pairs; the Search Path Link Count (SPLC), which accounts for the number of all possible search paths through the network emanating from an origin; and the Search Path Node Pair (SPNP), which accounts for all connected vertex pairs along the paths. Following the guideline recommended by Diana and Leydesdorff [71], we use the Search Path Link Count (SPLC) algorithm, which accounts for the number of all possible search paths through the network emanating from an origin, to identify the global main path of the citation network of the 214 cloud computing articles. Therefore, this global main path can approximately reveal how cloud computing in the IS discipline has evolved over time.

The global main path contains 12 papers published between 2009 and 2014. As shown in Figure 7, each paper on the global main path is marked by the first author and publication year, and the arrow points to the cited paper. Among the 12 papers, nine are among the 41 important papers, indicating that the results of our citation analysis and main path analysis are highly consistent. After analyzing the 12 papers on the global main path, we contend that cloud computing research evolved through three stages: incubation stage (2004-2008), the exploration stage (2009-2011), and the burgeoning stage (2012-2014). These stages can be confirmed by inspecting Figure 3: there are clearly two jumps in total number of publications from 2008 to 2009 and from 2011 to 2012, suggesting that cloud computing research entered a new stage in 2009 and 2012, respectively.

Incubation stage (2004-2008): This stage includes 7 papers. During this stage, except in a workshop paper [37], the term of cloud computing has not been explicitly defined. But some similar constructs were employed, such as computing as utility [72], on-demand web service [73], service orientation model of software [74], Software as a Service [35], and the computed cloud (cloud computing) [43]. The commonality of these similar constructs suggests the paradigm shift from locally

installed programs to internet enabled services.

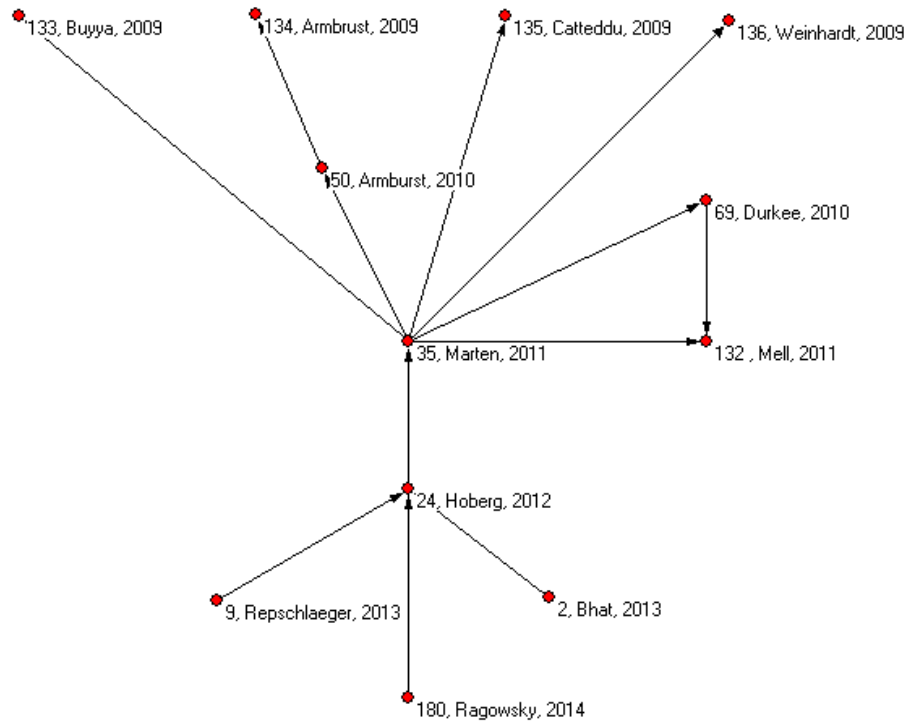


Figure 7. Main path analysis of cloud computing research

Exploration stage (2009-2011): This stage includes 69 papers, most of which are focused on ontological issues of cloud computing, including definition, technical feature, opportunity for further development, and cloud computing challenges. During this period, cloud computing was still a novel notion and scholars were trying to figure out what it is, what it means, what it contains, and what can be done with it. Hence, most papers were conceptual pieces intended to understand cloud computing. For example, in Paper 133, Buyya et al. [33] propose the concept of market-oriented cloud computing which includes service architecture, cloud resource management strategy, and resources pricing and allocation. Moreover, they boldly claim that cloud computing should be considered the fifth public utility after water, electricity, gas, and telephone. Paper 136 [34] conducted a full comparison between cloud computing and grid computing, and proposed a business model and the future direction of cloud computing. A widely accepted definition of cloud computing was given by Paper 132 [2], and the technical features were also elaborated. Papers 30 and 134 [32, 75] analyzed the opportunities and

challenges faced by the development of cloud computing. Paper 135 [42] analyzed issues such as earnings, risks, and information security in cloud computing. In Paper 35 [56], a reference model was designed with a systematic literature review and demand analysis on cloud computing services to help the management of risks and compliance for enterprises.

Burgeoning stage (2012-2014): During this stage, 138 papers on cloud computing were published. Researchers' attention has switched from pure conceptualization to investigations of specific research questions related to cloud computing. Paper 24 [10], a literature review, identified four dimensions of cloud computing research, including the features, determinants of adoption, management mechanisms, and commercial impacts, from a business perspective. More studies began to examine business and technology related issues of cloud computing, e.g., cloud computing adoption, capacity planning of cloud computing, and business impact of cloud computing. For example, Paper 174 [76] investigated the determinants of cloud computing adoption in the manufacturing and services sectors, the capacity planning problem of a service vendor providing a business process characterized by volatile demand to its customers were analyzed in Paper 77 [77], and Paper 110 [78] considered cloud computing services and its impact on market structure, firm profitability, and consumer welfare. The research focus on cloud computing has switched from conceptual development and exploration in the second stage to business and technology related issues in the third stage.

3.2 Co-Citation Analysis

The co-citation analysis was used to address our second goal: to identify and illustrate the knowledge groups of cloud computing research in the IS discipline and the relationships among them. By analyzing the references of the 41 important articles on cloud computing in IS discipline, we can determine if any two articles are commonly co-cited. If a set of articles are frequently co-cited, then they constitute a structural knowledge group. These groups and the relationships among them constitute the intellectual structure of a field [79].

We followed the method of Nerur et al. [20] to determine the major knowledge groups of cloud computing research in the IS discipline. A principal component factor analysis of the co-citation matrix was employed. Using SPSS16.0, we analyzed the correlation matrix of the 41 important papers

identified by the citation analysis based on a varimax rotation. The KMO (Kaiser-Meyer-Olkin) measure of sampling adequacy for the co-citation matrix was acceptable at 0.825 and Bartlett’s test was significant at the 0.001 level, which indicated that principal components analysis was applicable. Eleven factors with a minimum eigenvalue of 1 were extracted, which together explained 81.96% of the variance in the correlation matrix. A paper was retained only if its loadings’ absolute value is above 0.4 [18, 20]. As a result, 6 factors were kept for further analysis. Table 2 lists the 6 factors which explain 69.01% of the variance.

Table 2. Factors extracted for cloud computing research

	Factor 1 Foundations	Factor 2 SaaS Model	Factor 3 Security and Risk	Factor 4 Literature Review	Factor 5 Adoption and Impacts	Factor 6 Modeling
	<i>No 50 [1]</i>	No 96[63]	<i>No 35[56]</i>	<i>No 24[10]</i>	<i>No 28[60]</i>	<i>No 114[39]</i>
	No 55[41]	<i>No 105[49]</i>	No 81[38]	No 73[9]	<i>No 86[53]</i>	<i>No 116[52]</i>
	<i>No 56[45]</i>	<i>No 107[54]</i>	No 91[50]	No 79[61]	No 102[48]	<i>No 117[55]</i>
	No 57[43]	No 109[35]	No 135[42]	<i>No 141[8]</i>	<i>No 119[58]</i>	
	No 68[36]	<i>No 112[47]</i>	<i>No 144[65]</i>			
	<i>No 69[40]</i>					
	No 76[31]					
	No 82[44]					
	<i>No 103[46]</i>					
	No 113[51]					
	<i>No 132[2]</i>					
	<i>No 133[33]</i>					
	<i>No 134[32]</i>					
	<i>No 136[34]</i>					
	<i>No 137[37]</i>					
Variance explained	17.04	3.34	3.03	2.05	1.47	1.37
Percent of variance explained	41.56	8.14	7.40	4.99	3.58	3.34

Total variance explained: 69.01%.

Papers with loadings $\geq \pm 0.7$ are shown in italics.

The first factor in Table 2 appears to define the foundation of cloud computing. It is the most important among the six factors, explaining 41.56% of the total variance. This factor has 15 papers, and most of these papers are conceptual studies. A careful review of these papers reveals six topics:

definition, core technologies, opportunities and obstacles, computing utility, pricing, and research directions.

There are two influential definitions for cloud computing. From the NIST perspective, “Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [2]. From the Berkeley view, “Cloud computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services” [1]. Both definitions imply or explicitly acknowledge that cloud computing refers to an IT service model over the Internet and has five essential characteristics (on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service), three service models (Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS)), and four deployment models (private cloud, community cloud, public cloud, and hybrid cloud).

Marston et al. [31] state that three core technologies, virtualization, multitenancy and Web services support the growth of cloud computing. Foster et al. [37] argue that cloud computing is not a completely new concept; it has intricate relationship with Grid Computing and other relevant technologies such as cluster computing, utility computing, and distributed systems. Many studies compared cloud computing with other computing paradigms. For example, Weinhardt et al. [34] and Foster et al. [37] compare cloud computing with grid computing from various perspectives such as virtualization, types of application, development of applications, access, business model, SLAs/Liability, control, and switching cost. Most of these studies agree that cloud computing is derived from previous computing paradigms and has its own novelties. Armbrust et al. [32] demonstrate that three aspects are new in cloud computing from a hardware point of view: (1) infinite computing resources available on demand; (2) the elimination of an up-front commitment by cloud users; and (3) the ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed.

As Marston et al. [31] illustrate, cloud computing offers several advantages such as low cost of entry

for small firms, facilitating collaboration, immediate access to computing resources without upfront capital investments, making possible new classes of applications and delivering services that were not possible before, and mining insights from data. However, the shift to cloud computing has been slow. McAfee [44] explain that security and privacy, business continuity and service availability, data lock-in, data confidentiality/auditability, data transfer bottlenecks, performance unpredictability, scalable storage, and bugs in large-scale distributed systems, are the top obstacles of cloud computing development and diffusion [1, 45, 80].

Buyya et al. [33] posit that cloud computing can be considered as the 5th Utility to provide the basic level of computing service, since cloud computing refers to a model consisting of services that are commoditized and delivered in a manner similar to traditional utilities such as water, electricity, gas, and telephone. The utility model has received some support [33, 72]. Yet Brynjolfsson et al. [41] assert that the utility model cannot be used to describe cloud computing, because “cloud computing cannot achieve the plug-and-play simplicity of electricity, at least, not as long as the pace of innovation, both within cloud computing itself, and in the myriad applications and business models it enables, continues at such a rapid pace”. They think that cloud computing plays a catalyst role in innovation, and the opportunities for combinational innovation will grow with cheaper and more ubiquitous cloud computing services.

Pricing is another fundamental problem in cloud computing research, since it has direct impact on revenues for cloud computing service providers [59]. Weinhardt et al. [34] classifies the pricing model into free, pay-per-use, subscription, and dynamic pricing. Among the four pricing models, the most frequently used is pay-per-use, in which the user pays a static price for a used unit, often per hour, GB, CPU-hour etc. Subscription is a similar but different pricing to pay-per-use, where the user subscribes (signs a contract) for using a pre-selected combination of service units for a fixed price and longer time frame, usually monthly or yearly. In the dynamic pricing model, the target service price is established as a result of dynamic supply and demand. Dynamic pricing model can achieve more economically efficient allocations and prices of differentiated high-value services. Although cloud computing providers in the perfect competition market may drive prices downward, Durkee [40] contends that cloud computing will never be free, since cloud computing service providers will adopt differentiation

strategies to develop and price their services. Cusuman [36] suggests that although the pricing and delivery model of cloud computing have many advantages, it will not eliminate the traditional software products anytime soon, because users have many customized applications and data stored in proprietary databases and they would have difficulty in switching to a cloud computing platform quickly. Koehler et al. [46] found that customer preferences for cloud services are heterogeneous and suggested the application of second degree price discrimination.

As a new IT service delivery model, cloud computing transforms the IT artifact from IT resources into IT services, which exerts a profound impact on IS research. Scholars have tried to provide directions to guide future research in this largely uncharted territory. Bardhan et al. [51] suggest that with the popularity of cloud computing, service-oriented thinking is a fast-growing main paradigm in IS research, and a new science titled as service science, management, and engineering (SSME), should be adopted as a fundamental area for IS research. Marston et al. [31] divide the IS research agenda in cloud computing into six categories from the business perspective, including cloud computing economics, cloud computing and IT strategy/policy issues (including security), technology adoption and implementation issues, cloud computing and green IT, and regulatory issues.

The second factor is labeled as SaaS Model which focuses on the issues of software as a service model, e.g., contract design, adoption factor, continued use, value co-creation, and productivity of SaaS vendor. This factor explains 8.14% of total variance and includes 5 papers. Susarla et al. [47] employed the perspective of transaction cost economics to design SaaS contract and select price mechanism (time and materials V.S. fixed price. Xin and Levina [35] investigated client side determinants of adopting the SaaS model. Benlian et al. [49] examined the importance of SaaS service quality factors for shaping customer satisfaction and SaaS continuance intentions. Zainuddin and Gonzalez [63] conducted a case study to investigate how value co-creation components change over time as SaaS configurability moves toward maturity. Huang and Wang [54] explored scale economies of pure-SaaS firms, non-SaaS firms, and mixed-SaaS firms by examining 179 publicly listed software companies in the United States, and found that the presence of significant diseconomies of scale in pure-SaaS firms and SaaS firms are more productive only in utilizing capital assets.

The third factor is named as Security and Risk, explaining 7.40% of the total variance and having 5 papers. Security plays a critical role in cloud computing adoption and diffusion [65]. Benlian and Hess [38] conducted a survey of 349 IT executives at German companies, and found that security threats are the dominant factor influencing IT executives' overall risk perceptions. Ackermann et al. [50] developed a scale for perceived cloud computing security risks including six dimensions (confidentiality, integrity, availability, performance, accountability, and maintainability) and empirically validated the scale as a second-order construct based on data collected from 356 organizations. Martens and Teuteberg [56] developed a reference model that serves to support firms in managing and reducing cloud computing risk and compliance efforts. Based on the psychological theory of “unrealistic optimism” and a longitudinal mixed-methods study, Loske et al. [65] revealed that cloud service providers suffer from “unrealistic optimism” and therefore significantly underestimate their services’ exposure to IT security risks, which in turn reduces the propensity to implement necessary IT security measures in the cloud computing.

The fourth factor is titled as Literature Review. This factor explains 4.99% of total variance and includes 4 papers. All of these 4 papers review the cloud computing literature. Venters and Whitley [8] reviewed cloud computing research in technology and service desires. Hoberg et al. [10] structured current cloud computing research into four dimensions from a business perspective: characteristics, adoption determinants, governance mechanisms, and business impact. Yang and Tate [9] summarized existing cloud computing research themes on cloud computing including technology, commerce, concept, and application. Demirkan and Delen [61] proposed a conceptual framework for decision support systems in cloud and suggested future research directions.

The fifth factor is termed as Adoption and Impacts which focuses on questions about cloud computing adoption and business impacts. This factor explains 3.58% of total variance and includes 4 papers. Benlian et al. [58] developed a service quality measurement specifically for SaaS solutions, which can be used as a diagnostic tool by SaaS providers and users to spot strengths and weaknesses in the delivery of SaaS solutions. They also applied the measurement in a study to examine the effect of service quality on SaaS usage continuance [58]. Several studies paid attention to the business impacts of

cloud computing [48, 53]. For example, Malladi and Krishnan [53] found that cloud computing can enable CIO strategic focus, and complementarities in process and systems capabilities and organizational learning can maximize the value of cloud computing. Winkler et al.[48] took a contingency approach to examine how SaaS adoption affects the arrangements between business and IT departments, and found that in most cases there exist dominant and reinforcing contingencies determining a definite mode of SaaS governance.

The sixth factor explains 3.34% of total variance and is named as Modeling because all of its three papers use the analytic modeling method to examine cloud computing problems. Choudhary [39] modelled differences in how new software features are disseminated in SaaS and perpetual licensing, and showed that these differences affect the provider's incentive to invest in product development. He found that the SaaS licensing model leads to greater investment in product development under most conditions, which leads to higher software quality in equilibrium under SaaS as compared to perpetual licensing, and the software providers earn greater profits and social welfare is higher under SaaS under these conditions. Demirkan [52] examined the performance of an SaaS setup under different coordination strategies between the application service providers (ASPs) and the application infrastructure providers (AIPs), and the analysis indicates that coordination between the monopoly ASP and the AIP can result in an outcome with the same overall surplus. Sen et al. [55] developed a pricing heuristic and tested it under different levels of information accuracy and granularity, and the heuristic was shown to provide better economic welfare for both participants. They also thought that information is effective in providing stable service levels, and encouraged collaborations between customers and service providers.

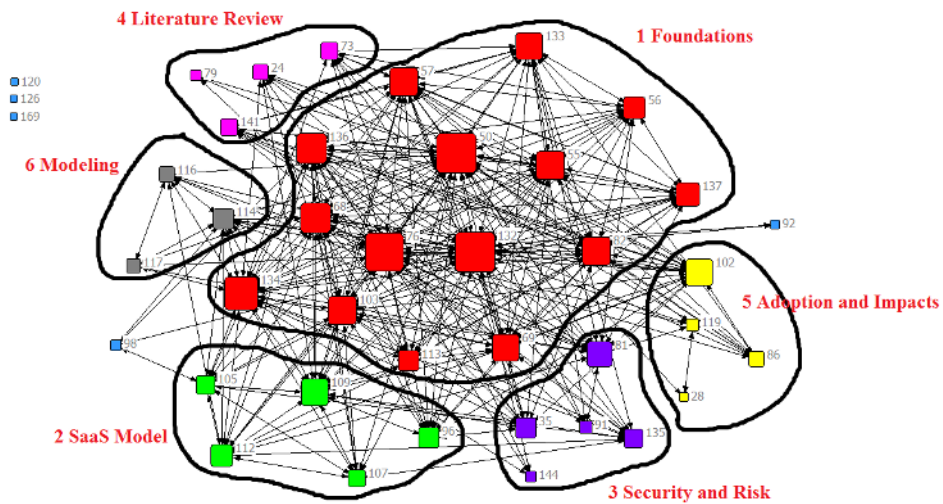


Figure 8. Relationship among the 5 knowledge groups of cloud computing research in the IS discipline

We applied NetDraw, a Windows program for visualizing social network data, to depict the relationships among the 6 factors and identify their network structure. In Figure 8, the size of a node representing a paper is proportionate to the paper's co-citation count. Papers that are co-cited with many other papers will tend to be centrally located while those that are co-cited with fewer other papers will lie toward the outskirts of the diagram. As shown in Figure 8, Foundations is the central and largest group, and papers in this group interact more strongly among themselves than with papers in the other three groups, indicating that most efforts have been devoted to the basic propositions of cloud computing research and the papers in Foundations lay the foundation for the studies in other knowledge groups. Group 2 (SaaS Model), Group 3 (Security and Risk), Group 4 (Literature Review), Group 5 (Adoption and Impacts), and Group 6 (Modeling) are peripherally located. They interact primarily with Factor 1 and rarely with each other, suggesting that studies in these five groups have been independently conducted and few papers have integrated multiple perspectives that need to cite papers from these groups simultaneously. An exception is Paper 81 in Group 3 that is co-cited three times with the papers in Group 5, indicating that security is an important antecedent of cloud computing adoption and impacts. Papers in Group 3 have few internal links, showing that the research topics in Security and Risk are diverse and independent.

4. FUTURE RESEARCH DIRECTIONS

In this section, we will discuss future directions in cloud computing research based on the co-citation analysis findings. Given the increased knowledge of cloud computing and rapid growth of cloud computing research, we believe that conceptual papers that only focus on conceptualizations of cloud computing will greatly decrease and more attention will be paid to substantive and empirical issues, including SaaS Model, Security and Risk, Literature Review, Adoption and Impacts, and Modeling.

4.1 SaaS Model

Since cloud computing services are provided by professional IT vendors, quality of service (QoS) and Service Level Agreement (SLA) become two important concepts in the context of SaaS model. Our analysis shows that previous studies in SaaS Model focus on service quality [49], contract design [47], and determinants of adoption [35]. These studies make no explicit distinctions between SaaS and other cloud computing service models (IaaS and PaaS). However, the nature of service quality and the antecedents and consequences of SaaS adoption could be different from other service models and thus require separate investigation. In addition, factors that influence SaaS adoption may differ in various cultural contexts. Further research on service quality and adoption of SaaS should consider culture as a contextual factor.

As one kind of outsourcing relationship, two prevailing mechanisms for managing inter-organizational relationships, formal controls and relational governance [81], could also be suitable for SaaS model. The management of the relationship with SaaS service providers should be an important issue for cloud users. Many important topics remain understudied. Some important topics include: How to design the formal contract to guarantee service providers' SLA? What can be used for organization users to improve the service quality of SaaS service providers? Which governance approach is more effective, formal control or relational governance? What's the relationship between formal controls and relational governance in the context of SaaS model, substitute or complement?

4.2 Security and Risk

When an organization uses cloud services, its business data that used to be stored locally are moved to the cloud. This change presents new challenges for the security and reliability of the data and

information systems. Moreover, the risk of security attacks and the resulting loss will be amplified because the data and computing resources are highly concentrated on the cloud. This puts a lot of pressures on security management. Prior studies have investigated the features of security management, security risk, security management framework, and security strategy. Some design issues are worthy of further research. The first issue is related to the features and content of the cloud security management. A multi-level model of cloud security management, including data security, application security, and service security, should be designed to address the unique challenges of cloud computing. The second issue is concerned with the security risks and loopholes brought by virtualization, openness and service-orientation of cloud computing and dynamic identification approaches on security management and technical risks in the mode of multi-party participation. The mechanism underlying the occurrence and transmission of risks should be studied, and a risk evaluation model should be established under the circumstance of dynamic evolution using an objective model. The third issue pertains to the motivations of multi-party participation, user behavior, and measurement and estimation methods for the risks of multi-party participation, considering the features of randomness, uncertainty, and concurrency of multi-party participation. Moreover, a coordinative monitoring system needs to be constructed by combining a technical framework with hazard identification as the core and a management framework with incentive mechanisms. Meanwhile, a secure audit method based on rule configuration should be studied from the perspective of the life cycle of security management.

4.3 Literature Review

Literature review on a regular basis plays a critical role in evaluating the progress of a research field, disseminating accumulated knowledge among researchers, and facilitating further knowledge creation and theory building in the field [82]. It has been a consensus across most scientific disciplines that reviewing the relevant literature must be done to understand the context, boundary, constructs, relationships, and competing theories of the focal phenomenon before a research project can be carried out to extend or build theory [83]. However, existing reviews of cloud computing research are limited in this regard. We recommend that future literature reviews should address two important issues: theory building and methodology rigor.

First, we contend that the uniqueness of cloud computing warrants indigenous theories and such theories should emerge from comprehensive literature reviews. According to our review, there is no single indigenous theory that has been developed for cloud computing. We propose that cloud computing theories can be built from several perspectives. For example, the changing relationships among members in a cloud computing enabled value chain challenge applicability of value chain theory. Traditional value chain theory is “product centered”, and firms are engaged in exchange relations. In contrast, the feature of cloud-based IS has caused the firm’s operation and management logic into a “service-dominant logic”, and firms now share coordination relations [84, 85]. The openness of cloud computing extends the boundary of the firm and its focus of IS management. In traditional client-based IS, firms pay most attention to their own strategies and the enabling and supporting roles of IS [86]. Yet in the planning of cloud computing deployment, firms should consider business co-creation mechanisms from the perspective of cloud computing enabled value chain. Another area where cloud computing theory can be built is IT strategy. The service features of cloud computing question the applicability of resource-based view (RBV). While RBV can be used to analyze the strategic and business value of traditional IS which are usually integrated into an organization’s business processes and become its private resources, it is difficult to apply RBV on cloud computing because cloud computing is essentially a service provided by vendors and not the organization’s private resources. A new theory is needed to explain and predict how cloud computing influences firms’ IT strategies and creates strategic benefits.

Second, as mentioned in the Introduction section, the existing literature reviews are mainly based on subjective analysis, and rigorous methodology is lacking. Quantitative methods, e.g., bibliometric analysis and meta-analysis, should be more widely applied to conduct literature reviews on cloud computing studies in a more objective manner. This paper is one of the first attempts to quantitatively review cloud computing research in the IS discipline by using citation and co-citation analysis. We call for more literature reviews based on bibliometric methods in the future and believe that findings of such rigorous reviews will provide accurate pictures of the field and facilitate building indigenous theories of cloud computing. Although the present paper is not intended for theory building per se, it provides an

accurate understanding of the cloud computing field that lays the foundation for building cloud computing theories later. Our findings reveal definitions, features, core technologies, and opportunities and obstacles of cloud computing, as well as what have been done in specific research themes. These findings can help other researchers identify what is lacking and where an indigenous theory is needed the most. Future literature reviews can take a step further, e.g., by using meta-analysis, to select an interesting perspective of cloud computing and identify important constructs and their relationships examined in previous studies to synthesize theories unique to cloud computing.

4.4 Adoption and Impacts

By employing various theories, such as diffusion of innovation theory, TOE framework, institutional theory, and information process view, researchers find that many technological, organizational, and environmental factors influence firms' cloud computing adoption or their intention to adopt cloud computing, and the determinants of cloud computing adoption vary across different countries and different industries [38, 76, 87-89]. Just adopting cloud computing is not enough, and cloud computing must be appropriately assimilated into the adopting firm's business processes to realize its potential benefits after its adoption [90]. To better guide firms to realize business value of cloud computing, it is imperative to advance the research focus from adoption to assimilation, and understand the conditions that enable or inhibit the assimilation of cloud computing.

Although a few studies have examined the strategic impacts of cloud computing [53], the understanding in this area is lacking. Given that cloud computing represents a paradigm shift in IT strategy management, many traditional IT strategy management issues in the context of cloud computing will appear as new research questions and deserve rigorous academic investigations. These questions include: Does cloud computing has the potential to be a source of competitive advantage? What are the underlying mechanisms of IT business value co-creation in the cloud computing enabled value chain? What is the dynamics of cloud computing enabled value chain? Whether and how dose cloud computing enable or support business model innovation? How can service-dominant logic be used to examine the strategic and business value of cloud computing? Does cloud computing enhance business-IT alignment and organizational agility, and how?

Cloud computing has great impacts on IT functions. For example, operation and maintenance of cloud computing is provided by external service providers, and the job for IT staff will switch to the governing of relationships and contracts with service providers. Unfortunately, the new role of the IT function in the cloud context, have not been investigated. We propose that the following questions deserve to be answered. First, how to redesign IT functions in terms of organization mode, function reform, basic tasks, service process, and performance indicator? Second, which IT capabilities are needed in realizing the business value of cloud computing? Third, what mechanism should be used to govern the relationships with cloud computing service providers: relationship, contract, or both? Fourth, how do the subscribing organization and cloud computing vendors maintain effective cooperative relationship in the post-adoption stage?

4.5 Modeling

Most modeling papers used economic models to investigate cloud computing related issues, e.g., software quality under SaaS and perpetual licensing [39], SaaS coordination strategy between ASP and AIP [52], pricing heuristic under different levels of information accuracy and granularity [55]. The existing studies did not examine behaviors of cloud computing users by using empirical data, and the models which were brought forward in these papers have power of explanation rather than prediction. In the era of big data, future studies can combine big data with these economic model, and big data can be used in the model generation and model testing. For example, many studies assume users' request follow a Poisson process, but more and more empirical studies found that the interval distribution of human behavior in Internet-based service is a power-law distribution [91], which primarily challenge the validity of the existing economic model.

5. LIMITATIONS

This study has a few limitations. First, our review covers a selected set of top IS journals and conferences as well as some top management journals. We believe that these publication sources can reasonably represent cloud computing research that is relevant to the IS community. However, the downside of this choice is that we might have missed some papers published on other outlets. Therefore, we limit the scope of this review to cloud computing research in the IS discipline. Second, the factor

analysis of co-citations is based on the 41 important papers rather than all of the 214 papers we can find. Generating co-citations for 214 papers is a daunting task, and we assume that the result based on the 214 papers is unlikely to be significantly different from our current findings. However, this needs to be verified in future research. Third, the factor analysis of co-citations provide statistics that have no substantive meanings. We interpret the results based on the content of the papers in each factor to identify the overarching theme. Therefore, the bibliometric approach is not completely objective and a little bit of human opinion is needed to make the results meaningful. We also notice that it is impossible and sometimes undesirable to completely eliminate human subjectivity from scientific research. The question is how to effectively integrate the bibliometric approach and subjective approach. Future research is needed to develop meaningful procedures to keep a suitable balance between objectivity and subjectivity in conducting literature reviews.

6. CONCLUSION

Cloud computing has greatly influenced the IT industry and attracted much attention from IS scholars. To understand the current literature on cloud computing, we conduct citation and co-citation analyses based on 214 papers published on 20 prestigious journals in IS and management and 2 top IS conferences from 2004 to 2014. In the citation analysis, in-degree centrality and between centrality are used to identify the important papers on cloud computing in IS discipline, and main path analysis is employed to make the structural backbone of the network consisted by these cloud computing papers. In the co-citation analysis, a matrix is compiled by retrieving co-citation counts for each pair of the important papers that identified in the citation analysis, and a principal component factor analysis is conduct to reveal the knowledge groups of cloud computing research.

Our review makes several contributions to the IS literature. First, we find that cloud computing research is still in its infancy and that the relationships between researches are loose. The research topics are scattered without a core paper group. This finding is consistent with the results of Yang and Tate [9] and Hoberg et al. [10]. Second, we identify 41 important papers based on citation analysis. Most of these important papers are about the conceptualization and review of cloud computing, and cloud computing research in information system heavily relies on publications in reference disciplines. Third, our main

path analysis shows that cloud computing research has evolved through three stages, namely, incubation, exploration, and burgeoning. Paper in the incubation stage doesn't explicitly define the term of cloud computing and used some similar constructs. Papers in the exploration stage mainly focus on the definition, technical features, opportunities, and challenges of the development of cloud computing. Papers in the burgeoning stage start to address specific research topics. This finding extend past literature reviews on cloud computing [9-12] from a cross-sectional perspective to an evolutionary perspective, which can provide insights for future studies. Fourth, different from previous reviews on cloud computing that employ a preset subjective framework to structure the related research themes, such as the works of Yang and Tate [9] and Hoberg et al.[10], we adopt a factor analysis of the co-citation matrix of 41 important papers and identify five major knowledge groups of cloud computing research. Foundations is the largest knowledge group, and the other five (SaaS Model, Security and Risk, Literature Review, Adoption and Use, and Modeling) interact primarily with Foundations and rarely with each other. Finally, we provide future research directions based on our review of the cloud computing literature. In summary, this is probably the first study that integrates citation analysis and co-citation analysis to systematically review the cloud computing literature. Our findings will help IS researchers gain an in-depth understanding of the current status of the cloud computing research field, and our recommendations for future research directions will assist researchers to decide what topics are important when they delve into this promising but largely uncharted territory.

Acknowledgements

This research is partly supported by the Key Program of National Science Foundation of China (grant number: 71331003) and the General Program of National Science Foundation of China (grant number: 71471079, 71471080, 71302140, and 71271104).

References

- [1] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, A view of cloud computing, *Communications of the ACM*, 53 (2010) 50-58.
- [2] P. Mell, T. Grance, The NIST definition of cloud computing, *National Institute of Standards and Technology*, 53 (2009) 50.

- [3] D.C. Plummer, T.J. Bittman, T. Austin, D.W. Cearley, D.M. Smith, Cloud computing: Defining and describing an emerging phenomenon, *Gartner*, June, 17 (2008).
- [4] S. Petter, W. DeLone, E.R. McLean, The past, present, and future of "IS Success", *Journal of the Association for Information Systems*, 13 (2012) 341-362.
- [5] D. Zisis, D. Lekkas, Addressing cloud computing security issues, *Future Generation computer systems*, 28 (2012) 583-592.
- [6] L. Kappelman, E. McLean, J. Luftman, V. Johnson, Key Issues of IT Organizations and Their Leadership: The 2013 SIM IT Trends Study, *MIS Quarterly Executive*, 12 (2013) 227-240.
- [7] Gartner, Gartner Says Worldwide IT Spending On Pace to Surpass \$3.4 Trillion in 2008, 2008.
- [8] W. Venters, E.A. Whitley, A critical review of cloud computing: researching desires and realities, *Journal of Information Technology*, 27 (2012) 179-197.
- [9] H. Yang, M. Tate, A descriptive literature review and classification of cloud computing research, *Communications of the Association for Information Systems*, 31 (2012) 35-60.
- [10] P. Hoberg, J. Wollersheim, H. Krcmar, The Business Perspective on Cloud Computing - A Literature Review of Research on Cloud Computing, 18th Americas Conference on Information Systems, Seattle, Washington, 2012.
- [11] T. Ermakova, K. Ereik, J. Huenges, Cloud Computing in Healthcare – a Literature Review on Current State of Research, 19th Americas Conference on Information Systems, Chicago, Illinois, 2013.
- [12] S. Raghuram, P. Tuertscher, R. Garud, Research Note—Mapping the Field of Virtual Work: A Cocitation Analysis, *Information Systems Research*, 21 (2010) 983-999.
- [13] E. Garfield, Is citation analysis a legitimate evaluation tool?, *Scientometrics*, 1 (1979) 359-375.
- [14] H. Small, Co-citation in the scientific literature: A new measure of the relationship between two documents, *Journal of the American Society for Information Science*, 24 (1973) 265-269.
- [15] M.M. Kessler, Bibliographic coupling between scientific papers, *American Documentation*, 14 (1963) 10-25.
- [16] E. Garfield, From bibliographic coupling to co-citation analysis via algorithmic historio-bibliography, speech delivered at Drexel University, Philadelphia, PA, 2001.
- [17] K.W. Boyack, R. Klavans, Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?, *Journal of the American Society for Information Science and Technology*, 61 (2010) 2389-2404.
- [18] A. Pilkington, J. Meredith, The evolution of the intellectual structure of operations management—1980–2006: A citation/co-citation analysis, *Journal of Operations Management*, 27 (2009) 185-202.
- [19] S. Wasserman, J. Galaskiewicz, *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*, SAGE1994.
- [20] S.P. Nerur, A.A. Rasheed, V. Natarajan, The intellectual structure of the strategic management field: An author co - citation analysis, *Strategic Management Journal*, 29 (2008) 319-336.
- [21] M. Giannakis, The intellectual structure of the supply chain management discipline: a citation and social network analysis, *Journal of Enterprise Information Management*, 25 (2012) 136-169.
- [22] G.L. Polites, R.T. Watson, Using social network analysis to analyze relationships among IS journals, *Journal of the Association for Information Systems*, 10 (2009) 595-636.
- [23] K. Birman, G. Chockler, R. van Renesse, Toward a cloud computing research agenda, *ACM SIGACT News*, 40 (2009) 68-80.
- [24] E. Abrahamson, L. Rosenkopf, Social network effects on the extent of innovation diffusion: A computer

simulation, *Organization science*, 8 (1997) 289-309.

[25] L.C. Freeman, Centrality in social networks conceptual clarification, *Social networks*, 1 (1979) 215-239.

[26] L.C. Freeman, D. Roeder, R.R. Mulholland, Centrality in social networks: II. Experimental results, *Social networks*, 2 (1980) 119-141.

[27] D.R. White, S.P. Borgatti, Betweenness centrality measures for directed graphs, *Social Networks*, 16 (1994) 335-346.

[28] C.W. Holsapple, L.E. Johnson, H. Manakyan, J. Tanner, A citation analysis of business computing research journals, *Information & Management*, 25 (1993) 231-244.

[29] S.B. Eom, Mapping the intellectual structure of research in decision support systems through author cocitation analysis (1971–1993), *Decision Support Systems*, 16 (1996) 315-338.

[30] T.E. Nisonger, The “80/20 Rule” and Core Journals, *The Serials Librarian*, 55 (2008) 62-84.

[31] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, Cloud computing — The business perspective, *Decision Support Systems*, 51 (2011) 176-189.

[32] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, M. Zaharia, *Above the Clouds: A Berkeley View of Cloud Computing*, University of California at Berkeley, 2009.

[33] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, Cloud computing and emerging IT platforms : Vision , hype , and reality for delivering computing as the 5th utility, *Future Generation Computer Systems*, 25 (2009) 599-616.

[34] C. Weinhardt, A. Anandasivam, B. Blau, N. Borissov, T. Meinl, W. Michalk, J. Stößer, Cloud computing—a classification, business models, and research directions, *Business & Information Systems Engineering*, 1 (2009) 391-399.

[35] M. Xin, N. Levina, *Software-as-a Service Model : Elaborating Client- Side Adoption Factors*, 29th International Conference on Information Systems, Paris, 2008.

[36] M. Cusumano, Cloud computing and SaaS as new computing platforms, *Communications of the ACM*, 53 (2010) 27-29.

[37] I. Foster, Y. Zhao, I. Raicu, S. Lu, Cloud computing and grid computing 360-degree compared, *Grid Computing Environments Workshop*, 2008. GCE'08, IEEE, 2008, pp. 1-10.

[38] A. Benlian, T. Hess, Opportunities and risks of software-as-a-service: Findings from a survey of IT executives, *Decision Support Systems*, 52 (2011) 232-246.

[39] V. Choudhary, Comparison of software quality under perpetual licensing and software as a service, *Journal of Management Information Systems*, 24 (2007) 141-165.

[40] D. Durkee, Why cloud computing will never be free, *Communications of the ACM*, 53 (2010) 62-69.

[41] E. Brynjolfsson, P. Hofmann, J. Jordan, Cloud computing and electricity: beyond the utility model, *Communications of the ACM*, 53 (2010) 32-34.

[42] D. Catteddu, Cloud Computing: benefits, risks and recommendations for information security, in: C. Serrão, V.A. Díaz, F. Cerullo (Eds.) *Web Application Security*, Springer Berlin Heidelberg 2010.

[43] B. Hayes, Cloud computing, *Communications of ACM*, 51 (2008) 9-11.

[44] A. McAfee, What every CEO needs to know about the cloud, *Harvard Business Review*, 89 (2011) 124-132.

[45] M.D. Ryan, Cloud Computing Privacy Concerns on our Doorstep, *Communications of the ACM*, 54 (2011) 2-5.

[46] P. Koehler, A. Anandasivam, M.A. Dan, C. Weinhardt, Customer heterogeneity and tariff biases in cloud

- computing, 31st International Conference on Information Systems, Saint Louis, USA, 2010.
- [47] A. Susarla, A. Barua, A.B. Whinston, A transaction cost perspective of the " Software as a Service" business model, *Journal of Management Information Systems*, 26 (2009) 205-240.
- [48] T.J. Winkler, C. Goebel, A. Benlian, F. Bidault, O. Günther, The impact of software as a service on IS authority—a contingency perspective, 32nd International Conference on Information Systems, Shanghai, 2011.
- [49] A. Benlian, M. Koufaris, T. Hess, The Role of SaaS Service Quality for Continued SaaS Use: Empirical Insights from SaaS Using Firms, 31st International Conference on Information Systems, St. Louis, 2010.
- [50] T. Ackermann, T. Widjaja, A. Benlian, P. Buxmann, Perceived IT security risks of cloud computing: Conceptualization and scale development, 33rd International Conference on Information Systems, Orlando, 2012.
- [51] I.R. Bardhan, H. Demirkan, P. Kannan, R.J. Kauffman, R. Sougstad, An interdisciplinary perspective on IT services management and service science, *Journal of Management Information Systems*, 26 (2010) 13-64.
- [52] H. Demirkan, H.K. Cheng, S. Bandyopadhyay, Coordination strategies in an SaaS supply chain, *Journal of Management Information Systems*, 26 (2010) 119-143.
- [53] S. Malladi, M.S. Krishnan, Cloud Computing Adoption and its Implications for CIO Strategic Focus—An Empirical Analysis, 33rd International Conference on Information Systems, Orlando, 2012, pp. 1-19.
- [54] K.-W. Huang, M. Wang, Firm-Level Productivity Analysis for Software as a Service Companies, 30th International Conference on Information Systems, Phoenix, 2009.
- [55] S. Sen, T.S. Raghu, A. Vinze, Demand Information Sharing in Heterogeneous IT Services Environments, *Journal of Management Information Systems*, 26 (2010) 287-316.
- [56] B. Martens, F. Teuteberg, Risk and compliance management for cloud computing services: Designing a reference model, Seventeenth Americas Conference on Information Systems, Detroit, Michigan, 2011.
- [57] N. Su, Emergence of cloud computing: an institutional innovation perspective, Thirty Second International Conference on Information Systems,, Shanghai, 2011.
- [58] A. Benlian, M. Koufaris, T. Hess, Service quality in software-as-a-service: developing the SaaS-Qual measure and examining its role in usage continuance, *Journal of management information systems*, 28 (2011) 85-126.
- [59] I. Iyob, E. Zarifoglu, A.B. Dieker, Cloud Computing Operations Research, *Service Science*, 5 (2013) 88-101.
- [60] S. Walther, A. Plank, T. Eymann, N. Singh, G. Phadke, Success Factors and Value Propositions of Software as a Service Providers – A Literature Review and Classification, 18th Americas Conference on Information Systems, Seattle, Washington, 2012.
- [61] H. Demirkan, D. Delen, Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud, *Decision Support Systems*, 55 (2013) 412-421.
- [62] A. Giessmann, K. Stanoevska, Platform as a Service—A Conjoint Study on Consumers’ Preferences, Thirty Third International Conference on Information Systems, Orlando, 2012.
- [63] E. Zainuddin, P. Gonzalez, Configurability, maturity, and value co-creation in SaaS: an exploratory case study, Thirty Second International Conference on Information Systems, Shanghai, 2011.
- [64] M. Benaroch, Q. Dai, R.J. Kauffman, Should we go our own way? Backsourcing flexibility in IT services contracts, *Journal of Management Information Systems*, 26 (2010) 317-358.
- [65] A. Loske, T. Widjaja, P. Buxmann, Cloud Computing Providers’ Unrealistic Optimism regarding IT Security Risks: A Threat to Users?, DOI (2013).
- [66] T.J. Winkler, C.V. Brown, Horizontal allocation of decision rights for on-premise applications and

- software-as-a-service, *Journal of Management Information Systems*, 30 (2013) 13-48.
- [67] H. Liang, J.-J. Wang, Y. Xue, X. Cui, IT outsourcing research from 1992 to 2013: A literature review based on main path analysis, *Information & Management*, DOI (2015).
- [68] L.-C. Yin, H. Kretschmer, R.A. Hanneman, Z.-Y. Liu, The evolution of a citation network topology : The development of the journal scientometrics, *International Workshop on Webometrics, Informetrics and Scientometrics*, Nancy , France, 2006.
- [69] J.S. Liu, L.Y.Y. Lu, An integrated approach for main path analysis: Development of the Hirsch index as an example, *Journal of the American Society for Information Science and Technology*, 63 (2012) 528-542.
- [70] D. Lucio-Arias, L. Leydesdorff, Main-path analysis and path-dependent transitions in HistCite™-based historiograms, *Journal of the American Society for Information Science and Technology*, 59 (2008) 1948-1962.
- [71] D. Lucio - Arias, L. Leydesdorff, Main - path analysis and path - dependent transitions in HistCite™ - based historiograms, *Journal of the American Society for Information Science and Technology*, 59 (2008) 1948-1962.
- [72] H.K. Bhargava, S. Sundaresan, Computing as Utility: Managing Availability, Commitment, and Pricing Through Contingent Bid Auctions, *Journal of Management Information Systems*, 21 (2004) 201-227.
- [73] J.Y. Sayah, L.-J. Zhang, On-demand business collaboration enablement with web services, *Decision Support Systems*, 40 (2005) 107-127.
- [74] A. Elfatatry, P. Layzell, Negotiating in service-oriented environments, *Communications of the ACM*, 47 (2004) 103-108.
- [75] H. Nuseibeh, Adoption of Cloud Computing in Organizations, *17th Americas Conference on Information Systems*, Detroit, Michigan, 2011.
- [76] T. Oliveira, M. Thomas, M. Espadanal, Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors, *Information & Management*, 51 (2014) 497-510.
- [77] C. Dorsch, B. Häckel, Combining models of capacity supply to handle volatile demand: The economic impact of surplus capacity in cloud service environments, *Decision Support Systems*, 58 (2014) 3-14.
- [78] P.-y. Chen, S.-y. Wu, The Impact and Implications of On-Demand Services on Market Structure, *Information Systems Research*, 24 (2013) 750-767.
- [79] L. Leydesdorff, L. Vaughan, Co - occurrence matrices and their applications in information science: extending ACA to the web environment, *Journal of the American Society for Information Science and Technology*, 57 (2006) 1616-1628.
- [80] N. Leavitt, Is cloud computing really ready for prime time?, *Computer*, 42 (2009) 15-20.
- [81] J. Goo, R. Kishore, H.R. Rao, K. Nam, The role of service level agreements in relational management of information technology outsourcing: an empirical study, *MIS Quarterly*, 33 (2009) 119-145.
- [82] J. Webster, R.T. Watson, Analyzing the past to prepare for the future: Writing a literature review, *MIS Quarterly*, 26 (2002) Xiii-Xxiii.
- [83] A.H. Van de Ven, *Engaged scholarship: a guide for organizational and social research: a guide for organizational and social research*, OUP Oxford 2007.
- [84] S.L. Vargo, R.F. Lusch, Service-dominant logic: continuing the evolution, *Journal of the Academy of marketing Science*, 36 (2008) 1-10.
- [85] R.F.L.a.S. Nambisan, Service Innovation: A Service-Dominant Logic Perspective, *MIS Quarterly*, 39 (2015) 155-175.
- [86] G. Piccoli, B. Ives, Review: IT-dependent strategic initiatives and sustained competitive advantage: A review

and synthesis of the literature, *MIS Quarterly*, 29 (2005) 747-776.

[87] A. Bhattacharjee, S.C. Park, Why end-users move to the cloud: a migration-theoretic analysis, *European Journal of Information Systems*, 23 (2014) 357-372.

[88] C. Low, Y. Chen, M. Wu, Understanding the determinants of cloud computing adoption, *Industrial management & data systems*, 111 (2011) 1006-1023.

[89] A. Rai, D.S. Bajwa, An Empirical Investigation into Factors Relating to the Adoption of Executive Information Systems: An Analysis of EIS for Collaboration and Decision Support, *Decision Sciences*, 28 (1997) 939-974.

[90] H. Liang, N. Saraf, Q. Hu, Y. Xue, Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management, *MIS Quarterly*, 31 (2007) 59-87.

[91] A.-L. Barabasi, The origin of bursts and heavy tails in human dynamics, *Nature*, 435 (2005) 207-211.