Vision, applications and future challenges of Internet of Things: A bibliometric study of the recent literature

Mishra, D

http://hdl.handle.net/10026.1/5033

10.1108/IMDS-11-2015-0478
Industrial Management & Data Systems
Emerald

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
Vision, applications and future challenges of Internet of Things: A bibliometric study of the recent literature

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>Industrial Management &amp; Data Systems</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>IMDS-11-2015-0478.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Internet of Things, Vision, Applications, Bibliometrics, Network analytics</td>
</tr>
</tbody>
</table>
Ref. No.: IMDS-11-2015-0478

Title: Vision, applications and future challenges of Internet of Things: A bibliometric study of the recent literature

We would like to thank the reviewers for their comments that have contributed significantly in the improvement of our manuscript. We have rewritten and restructured the manuscript significantly and we hope that the reviewers like the updated version.

Reviewer: 1

Recommendation: Major Revision

Date: 17 December 2015

Manuscript number: IMDS-11-2015-0478

Manuscript title: Vision, applications and future challenges of Internet of Things: A bibliometric study of the recent literature

1. Originality:
The linkage between the manuscript’s objective and why the research was necessary was not well established e.g., Why do you want to review the literature using citation/co-citation analysis? Author also needs to justify the importance of understanding the vision of IoT. The motivation of study thus needs to be re-written to show stronger linkage. It would be nice to see a stronger connection between your findings and the theme of the journal. How does this understanding help organizations and the industry to make better decisions? The managerial and theoretical implication is missing at the moment. Thus the reviewer cannot conclude if the manuscript has contributed fully to the existing body of knowledge.

We would like to thank the reviewer for this comment. We have re-written the motivation of the study in a paragraph in section 1 (introduction), in page 3: “In recent years, scholars (Borgia, 2014; Whitmore et al., 2014; Madaanam et al., 2015; Russo et al., 2015)… for considering IoTs vision and applications in our study”. We have also added the managerial and theoretical implications as requested by the reviewer (please see section 4 now named discussion and subsections 4.1 and 4.2). In the start of section 2 (now research methodology) we have included the three sources that we base our study on (Tranfield et al., 2003; Saunders et al., 2009; Fahimnia et al., 2015).

2. Relationship to literature:
The literature section mainly is on the history and definition of IoT from different authors. In addition there was also some brief explanation on the applications of IoT. Apart from the definition and some brief explanation of where IoT is applicable, the reviewer is not able to see the contribution of your literature section. The reviewer is in the opinion that some of the points here lack of critical arguments. The lack of critical arguments, did not reflect well on the research objectives thus making the manuscript less convincing.

We would like to thank the reviewer for this comment. We have rewritten and restructured the literature review section (please see section 3).
3. Methodology:
The methodology is reasonable with the objectives of the manuscript. We would like to thank the reviewer for this comment.

4. Results:
Author at the moment only provides a brief explanation on some of the findings e.g "For instance, a thick arc is connecting Miorandi et al. (2012) and Gubbi et al. (2013), thereby reflecting the common thoughts the articles share. A similar pattern can be noticed between Ashton (2009) and Atzori et al. (2010). However, the thin line between Miorandi et al. (2012) and Weber and Weber (2010) reveal that their common ideas are limited compared to the ones by Miorandi et al. (2012) and Gubbi et al. (2013)". Without much elaboration, the write up does not add much value for readers. The discussion should include further discussion on the previous findings in relation to the existing ones. E.g, the differences and your contributions.

We would like to thank the reviewer for this comment. We have re-written the discussion section and added theoretical and managerial implications of our research (Please see sections 4.1 and 4.2 respectively).

5. Implications
In terms of contributions, the paper was presented in a limited manner. The reviewer cannot see how the research gaps were being addressed in the paper as the managerial and practical implication is missing at the moment. Without this information, it will be difficult to assess if the paper has bridge the gap between theory and practice.

We would like to thank the reviewer for this comment. We have added theoretical and managerial implications at the discussion section (Please see sections 4.1 and 4.2 respectively).

6. Quality of communication:
The paper is reasonably written in terms of language usage and structurally well presented. However, there are still some typo and grammatically errors. Please do thorough grammatical checking to the manuscript to improve the readability of this study.

We would like to thank the reviewer for this comment. We have proofread the paper to improve the readability.

Additional Questions:
1. Originality: Does the paper contain new and significant information adequate to justify publication?: The linkage between the manuscript’s objective and why the research was necessary was not well established e.g. Why do you want to review the literature using citation/co-citation analysis? Author also needs to justify the importance of understanding the vision of IoT. The motivation of study thus needs to be re-written to show stronger linkage. It would be nice to see a stronger connection between your findings and the theme of the journal. How does this understanding help organizations and the industry to make better decisions? The managerial and theoretical implication is missing at the moment. Thus the reviewer cannot conclude if the manuscript has contributed fully to the existing body of knowledge.

We would like to thank the reviewer for this comment. We have added theoretical and managerial implications in the discussion section (Please see sections 4.1 and 4.2 respectively).

2. Relationship to Literature: Does the paper demonstrate an adequate understanding of the relevant literature in the field and cite an appropriate range of literature sources? Is any significant work ignored?: The literature section mainly is on the history and definition of IoT from different authors. In addition
there was also some brief explanation on the applications of IoT. Apart from the definition and some brief explanation of where IoT is applicable, the reviewer is not able to see the contribution of your literature section. The reviewer is in the opinion that some of the points here lack of critical arguments. The lack of critical arguments, did not reflect well on the research objectives thus making the manuscript less convincing.

*We would like to thank the reviewer for this comment. We have rewritten parts of the paper to make it flow better and have increased the critical arguments/objectives.*

### 3. Methodology:
Is the paper’s argument built on an appropriate base of theory, concepts, or other ideas? Has the research or equivalent intellectual work on which the paper is based been well designed? Are the methods employed appropriate? The methodology is reasonable with the objectives of the manuscript.

*Thank you for this comment.*

### 4. Results:
Are results presented clearly and analysed appropriately? Do the conclusions adequately tie together the other elements of the paper?: Author at the moment only provides a brief explanation on some of the findings e.g. "For instance, a thick arc is connecting Miorandi et al. (2012) and Gubbi et al. (2013), thereby reflecting the common thoughts the articles share. A similar pattern can be noticed between Ashton (2009) and Atzori et al. (2010). However, the thin line between Miorandi et al. (2012) and Weber and Weber (2010) reveal that their common ideas are limited compared to the ones by Miorandi et al. (2012) and Gubbi et al. (2013)". Without much elaboration, the write up does not add much value for readers. The discussion should include further discussion on the previous findings in relation to the existing ones. E.g, the differences and your contributions.

*We would like to thank the reviewer for this comment. We have rewritten and restructured the paper significantly. Please see section 4 (now discussion section).*

### 5. Implications for research, practice and/or society
Does the paper identify clearly between any implications for research, practice and/or society? Does the paper bridge the gap between theory and practice? How can the research be used in practice (economic and commercial impact), in teaching, to influence public policy, in research (contributing to the body of knowledge)? What is the impact upon society (influencing public attitudes, affecting quality of life)? Are these implications consistent with the findings and conclusions of the paper?: In terms of contributions, the paper was presented in a limited manner. The reviewer cannot see how the research gaps were being addressed in the paper as the managerial and practical implication is missing at the moment. Without this information, it will be difficult to assess if the paper has bridge the gap between theory and practice.

*We would like to thank the reviewer for this comment. We have added the managerial implications in the discussion section of the paper (Please see section 4.2).*

### 6. Quality of Communication:
Does the paper clearly express its case, measured against the technical language of the field and the expected knowledge of the journal’s readership? Has attention been paid to the clarity of expression and readability, such as sentence structure, jargon use, acronyms, etc.: The paper is reasonably written in terms of language usage and structurally well presented. However, there are still some typo and grammatically errors. Please do thorough grammatical checking to the manuscript to improve the readability of this study.

*We would to thank the reviewer for this comment. We have proofread the paper.*
Reviewer: 2

Recommendation: Reject

Comments:
A sample of bibliometric review paper could be referred to here in this paper: Big names in innovation research: a bibliometric overview by Cancino, Merigo and Coronado, 2015.

Additional Questions:

1. Originality: Does the paper contain new and significant information adequate to justify publication?:
The Internet of Thing is a relatively new area of research since the world only adopting into internet working and lifestyle actively in the last 15 years or so.

The idea of bibliometric study may or may not come at the appropriate time; as it is fairly new and not stable area of study, needs further refinement. Bibliometric study will inject further refinement to the topic of study, however the authors have missed this important argument completely.

We would like to thank the reviewer for this observation. However, we have provided evidence and conducted a bibliometric study for reasons, inter alia, that the reviewer refers to. We believe our study contributes to the further refinement of the topic of study, and we have revised our manuscript to fit with this contribution. We have added a paragraph in section 1 (introduction), page 3: "In recent years, scholars (Borgia, 2014; Whitmore et al., 2014; Madaham et al., 2015; Rasso et al., 2015)... for considering IoT's vision and applications in our study".

2. Relationship to Literature: Does the paper demonstrate an adequate understanding of the relevant literature in the field and cite an appropriate range of literature sources? Is any significant work ignored?:
The authors have search results of 1777 papers, and filtered with relevancy to 146 papers. However, the reviewer does not see the relationship between literature review and the 146 papers analysed. Instead the authors were more focus on the analysis of commonality within the selected papers and authors.

We would like to thank the reviewer for this comment. We have rewritten parts of the literature review to address the comment of the reviewer.

3. Methodology: Is the paper’s argument built on an appropriate base of theory, concepts, or other ideas? Has the research or equivalent intellectual work on which the paper is based been well designed? Are the methods employed appropriate?: The reviewer opined that paper needs improvement in flow and argument to support on the method of analysis preferred.

We would like to thank the reviewer for this comment. We have rewritten and restructured the paper so that it flows better.

In Sec 2.2, pg 6, the authors abruptly divided IoT applications into four category with no substantiations and without relevance to the literature review in earlier sections.

We would like to thank the reviewer for this comment. We have revised the section (now section 3).

The authors do take care of the bias tendency of self-citation when performing the analysis. The authors also took into consideration of co-author status.

We would like to thank the reviewer for this comment.

4. Results: Are results presented clearly and analysed appropriately? Do the conclusions adequately tie together the other elements of the paper?: The results presented were adequate in terms of showing the more popular authors, so what do all the nodes show in the relationship diagram? The authors could have elaborated further.

Let's just be clear, are the authors analysing popularity and influence journals or authors, or topics?
We would like to thank the reviewer for this comment. We are looking into analysing influence of researchers, but more importantly the popularity of particular articles and subsequently topics of research. We follow Fahimnia et al. (2015) and their review of green supply chain management literature. Our aim was rewritten in the abstract, introduction, and across the paper to ensure consistency.

5. Implications for research, practice and/or society Does the paper identify clearly between any implications for research, practice and/or society? Does the paper bridge the gap between theory and practice? How can the research be used in practice (economic and commercial impact), in teaching, to influence public policy, in research (contributing to the body of knowledge)? What is the impact upon society (influencing public attitudes, affecting quality of life)? Are these implications consistent with the findings and conclusions of the paper?: The Sec. 5 on the future trends has no bearing of whatsoever link to the result of the analysis. The reviewer can’t see the need this section or else, the authors would have to show this section of elaborating future trends are of the results from the bibliometric analysis.

Instead, it is suggested that the authors should delved more on the key papers identified, what are the contents and why they are popular and widely cited? What are the common topics instead of just common authors.

The conclusion could be shorten and more crisp in essence.

We would like to thank the reviewer for this comment. We have rewritten parts of the paper to address the comments by the reviewer. We incorporated the literature requested and found through our literature review.

6. Quality of Communication: Does the paper clearly express its case, measured against the technical language of the field and the expected knowledge of the journal's readership? Has attention been paid to the clarity of expression and readability, such as sentence structure, jargon use, acronyms, etc.: The paper needs to be thoroughly proof read again.

For examples:

1) Pg. 4, line 28, "IThe past decade has.... and published reports on various ...."

2) Pg. 4, line 37-39. "The phrase 'Internet of Thing' originated..........held at Proctor & Gamble (Ashton, 2009)" is too mouthful.

3) page 6, line 26 - 28, "For simplicity reasons"

And many more which the reviewer could not list them all.

We would like to thank the reviewer for this comment. We have proofread our manuscript.
Reviewer: 3

Recommendation: Major Revision

Overview:
The purpose of this study is to review the current literature on Internet of Thing (IoT), provide a thorough literature landscape on the domain, and offer insights into future research. Citation and Co-citation analyses on 146 identified IoT articles, published between 2000 and 2015, were conducted. The results indicate that although the number of journals exploring IoT is increasing, highly-cited papers are majority conceptual yet limited.

Overall Comments:
The paper has several strengths. First, the authors investigated an important topic- Internet of Thing (IoT). The topic is interesting and timely. Second, the paper is well-written and easy to follow. Third, the conceptualization of IoT is clear. Part 2 is nicely done. While the paper attempts to address a gap in academic and provide insight for researchers and practitioners, the quality of the paper could be improved. At present, the manuscript has the potential; however, in my humble view the paper still has several weaknesses that have to be overcome before a clear path to publication becomes evident. I listed my specific concerns as follows.

Research Motivations:
The research motivations of this study could be better articulated. It is not clear to me why we need a bibliometric study of the recent IoT research, despite the fact that scholars have already attempted to review the literature not long ago (Borgia, 2014; Madakam et al 2015; Whihtmore et al., 2014). The arguments derived from citing “no study has systemically reviewed the literature using citation/co-citation analysis” is not sufficient. Weber and Watson (2002) commented that a good literature review should identify critical knowledge gap. More importantly, a good piece of literature review also identifies systematic theoretical or methodological biases in a field and suggests fundamental reorientation for understanding the problem (Alvesson and Sandberg 2011). In that sense, the authors could illustrate in detail how the current investigation could help to add knowledge/value to the domain.

We would like to thank the reviewer for this comment. We have made our contributions both in terms of knowledge gap and methodological considerations in the subsection of theoretical contribution in the discussion section as well as in the conclusion section (please see throughout the paper— we have restructured and rewritten parts of the paper significantly). We have also made more explicit our impetus for this study in the introduction section (please see section 1).

Methodology:
One of my major concerns is the robustness of the paper search and identification processes. It is important for the authors to provide details on the process as it is difficult to evaluate the validity of the results without having known how the articles were identified in the first place. The authors are suggested to provide in greater details: (1) all indexed databases used and why they were chosen, (2) inclusion and exclusion criteria, (3) backward search used?

We would like to thank the reviewer for this comment. We have restructured significantly our manuscript and methodology, (now section 2) and included the necessary details as requested.

Analysis:
The argument on whether citation analysis should include self-citation has been going on since the early days of citation analysis. Early citation studies tended to exclude self-citation, but today’s Journal Impact Factor includes them. Thus, I would suggest that the authors could take into consideration the effect of self-citations and see if current pattern changes by excluding or discounting the self-citation effects.

We would like to thank the reviewer for this comment. We have inserted a comment on the citation and co-citation in our limitations (please see conclusions section of the manuscript).
Results and Implications:
The results part is rather descriptive. As mentioned earlier, a good piece of literature review identifies systematic theoretical or methodological biases in a field and suggests fundamental reorientation for understanding the problem (Alvesson and Sandberg 2011). Indeed, the authors have pointed out that there are limited number of IoT articles that attempt to discuss IoT using alternative theories and methodologies. Thus, I believe that the authors could go beyond the descriptive analysis and propose alternative theories or methods that could be used in the future.

We would like to thank the reviewer for this comment. Our paper provides particular research gaps with regards to the research that needs to be done, and research questions to help researchers acknowledge the diverse facets of IoT implementation and adoption. We highlight that more needs to be done towards alternative theories and case studies or mixed methods, but we are not suggesting any, because our findings do not suggest/propose such theories. We could have proposed alternative theories based on our experience, but we do not have evidence from our findings to suggest this. Furthermore, our aim is not to suggest alternative theories but to identify (i) top contributing authors (ii) key research topics related to the field (iii) the most influential works based on citations and PageRank and (iv) established and emerging research clusters. Our study is inspired by Fahimnia et al. (2015) and their review of the green supply chain management literature.

Contributions:
I have some reservation on the theoretical contribution of this study. The authors are advised to illustrate the novelty of the study and better justify the theoretical contributions of this paper.

We would like to thank the reviewer for this comment. We have rewritten parts of the paper in order to flesh out the contribution more explicitly. Please see throughout the paper and discussion section (now section 4), as well as subsections 4.1 and 4.2 that refer to theoretical contributions and managerial implications.

Minor Issues:
Some minor problem with spacing and typos (e.g. “published” or “Thag”)

We thank the reviewer for this comment. We have proofread the updated version and dealt with typos.
Vision, applications and future challenges of Internet of Things: A bibliometric study of the recent literature

Abstract

Purpose: The emergent field of Internet of Things (IoT) has been evolving rapidly with a geometric growth in the number of academic publications in this field. The purpose of this study is to review the literature of IoT in past 16 years using rigorous bibliometric and network analysis tools, offering at the same time future directions for the IoT research community and implications for managers and decision makers.

Design/methodology/approach: We adopted the techniques of bibliometric and network analysis. The paper reviewed the articles published on IoT from 2000 to 2015.

Findings: This study identifies (i) top contributing authors (ii) key research topics related to the field (iii) the most influential works based on citations and PageRank and (iv) established and emerging research clusters. Scholars are encouraged to further explore this topic.

Research limitations/implications: This study focuses only on vision and applications of IoT. Scholars may explore various other aspects of this area of research.

Originality/value: To the best of authors’ knowledge, this is the first study to review the literature on IoT by using bibliometric and network analysis techniques. The study is unique as it spans a long time period of 16 years (2000-2015). Our study proposes a five-cluster classification of research themes that may inform current and future research in IoT.

Keywords: Internet of Things, Vision, Applications, Bibliometrics, Network analysis.

Paper type: Literature review

1. Introduction

Recent years have witnessed the growing use of Internet as billions of people browse the web to access multimedia content and services, send and receive electronic mails, play games, and perform various tasks. This use creates a global platform for machines and smart objects to communicate, dialogue, compute, and coordinate (Miorandi et al., 2012), which in turn builds up a strong connection among the users of smart devices worldwide. Apart from connecting the users to the Internet, these devices play a crucial role in linking-up the physical world with the
cyber world (Conti et al., 2012). This has given birth to the next generation of embedded ICT systems, commonly known as Cyber-Physical Systems (CPS) (Poovendran, 2010; Park et al., 2012), which integrate computational devices with the physical environment. CPS is composed of four technologies: automation of knowledge work, Internet of Things, advanced robotics, and autonomous/near-autonomous vehicles. Looking at the economic value generated by these technologies, it can be clearly observed that IoT, with an estimated value of 36 trillion of dollars, creates the highest economic impact (McKinsey Global Institute, 2013).

The term “Internet of Things” (IoT) came into existence when Kevin Ashton used it for the first time in 1999 to represent the globally emerging Internet-based information service architecture (Ashton, 2009). Weber (2009) defined IoT as “an emerging global, Internet-based information service architecture facilitating the exchange of goods in global supply chain networks … on the technical basis of the present Domain Name System; drivers are private actors”. IoT facilitates a safe and trustworthy way of exchanging information related to goods and services in a global supply chain. It acts as a pillar for ubiquitous computing that opens the door for smart environments to spot and track items, and collect information from the Internet for their proper functioning. In doing so, members of MIT Auto-ID Center developed Electronic Product Code (EPC) that serves as a universal identifier for any specific item (Gama et al., 2012; EPC Global Inc., 2011). The main objective behind this development was to spread awareness about the use of Radio-Frequency Identification (RFID) globally. But, these days, the idea of “Thing” is not only restricted to RFID. It has expanded to include any real or physical object (e.g., RFID, sensor, actuator, smart item), “spime” data object as well as any virtual or digital system, which is capable of moving in time and space. These entities can be identified uniquely through the identification details (numbers, names and/or location addresses) assigned to them. Thus, the “Thing” can be read, recognized, located, addressed and controlled effortlessly by using Internet (Borgia, 2014).

Internet of Things has simplified our day-to-day lives by creating smart objects, applications and services, which ensure safety and security during the information exchange process. Indeed, IoT has the ability to influence economic activity across industries and affect their strategic decisions, investments and productivity (Borgia, 2014). Mandel (2014) visualized that US GDP will approximately increase by 2 to 5% by the end of 2025. At present, digital industries contribute about 20% of the GDP while the rest 80% comes mainly from physical industries, i.e., agriculture, construction, manufacturing, energy, transportation, and healthcare. Therefore, IoT aims to transform the way in which physical industries do business by connecting them to the computerized world.
In recent years, scholars (Borgia, 2014; Whitmore et al., 2014; Madaham et al., 2015; Russo et al., 2015) have attempted to review the literature on IoT by focusing on its vision, concepts, applications and features. Although these studies have provided insight into the field of IoT, they have not conducted additional analysis via rigorous bibliometric and network analytics tools. Such an analysis can help in refining the established and emerging areas of research, and in researchers acknowledging the different schools of thought and relevant applications of IoT. Moreover, the meaning of the term ‘IoT’ itself is continuously evolving since the technologies and ideas which drive it are also changing. These challenges signify the reason for considering IoTs vision and applications in our study.

To address this gap, this study reviews the literature from 2000-2015 on IoT using bibliometric and network analytics tools. We review, refine, and analyse a set of 1777 articles to obtain the most influential works, research themes, and researchers. We propose a five-cluster classification of research themes that provides additional insights on the current field and potential future research directions have been obtained.

In the next section, we present our methodological considerations and initial results of our review. Then we review the literature on vision and applications of IoT which is followed by a detailed analysis using the technique of bibliometric and network analysis. The paper ends with conclusion, limitations and future research directions.

2. Research methodology

Literature review maps and assesses the relevant literature in order to find out the possible research gaps which would be beneficial in further strengthening of the knowledge (Tranfield et al., 2003). In this paper we followed (i) Saunders et al. (2009) and their conceptualisation of literature review as an adaptive cycle which involves the process of defining relevant keywords, conducting literature search and finally, performing analysis, (ii) the approach proposed by Rowley and Slack (2004): scanning documents, making notes, structuring the literature review, writing the literature review, and building the bibliography, and (iii) Fahimnia et al. (2015) and their review of green supply chain literature using bibliometric and network (citation and co-citation) analysis.

2.1 Keyword search and data collection
The articles were collected using Scopus database only. The reason is that Scopus is the largest abstract and citation database covering more than 20,000 peer-reviewed journals in the fields of science, technology, medicine, social sciences, and arts and humanities, which belong to, inter alia, the publishing houses of Elsevier, Emerald, Informs, Taylor and Francis, Springer and Inderscience (Fahimnia et al., 2015). According to Yong-Hak (2013), Scopus database is more comprehensive than Web-of-Science (WoS) database because WoS includes only ISI indexed journals which is further limited to only 12,000 titles. In fact, Chicksand et al. (2012) noted that Scopus is a good source of supply chain peer reviewed articles.

Keeping in mind the objective of this paper, we chose the keywords which fully cover IoT vision and applications. Hence, we used the following keywords for the process of data collection: ‘Internet of Things’, ‘Vision of Internet of Things’, and ‘Applications of Internet of Things’. Through these keywords, three different combinations were made: (1) Internet of Things, (2) Vision AND IoT and (3) Applications AND IoT. We searched for the aforementioned keywords in “title, abstract, keywords” of articles belonging to Scopus database. The initial search resulted in 1777 articles. The number of articles obtained for each combination of keywords is shown in Table 1. The results containing the necessary information such as title of the paper, authors’ names and affiliations, abstract, keywords and references, were then saved in RIS format.

Table 1: Initial results

<table>
<thead>
<tr>
<th>Search keywords</th>
<th>Search results (no. of papers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet of Things</td>
<td>986</td>
</tr>
<tr>
<td>Vision AND IoT</td>
<td>426</td>
</tr>
<tr>
<td>Applications AND IoT</td>
<td>365</td>
</tr>
<tr>
<td>Total</td>
<td>1777</td>
</tr>
</tbody>
</table>

While refining the search results, we removed the duplicates as there is a possibility that few articles may belong to more than one combination of keywords. On eliminating such duplications, we were left with 1556 papers. Since Rodriguez et al. (2004) categorised articles and reviews as “certified knowledge”, we restricted ourselves to only scientific publications (articles and reviews) that appeared in peer reviewed journals. Unpublished articles, working papers and magazine articles were excluded during data purification process. This search resulted in 923 relevant documents, published during 16-year period i.e., 2000-2015. Table 2 shows the breakdown of refined search results for each of the three combination of keywords. For carrying
out these refinements in the RIS file, Endnote bibliography software was used. Then, the final
RIS data file was stored for future analysis.

Table 2: Refined search results

<table>
<thead>
<tr>
<th>Search keywords</th>
<th>Search results (no. of papers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet of Things</td>
<td>458</td>
</tr>
<tr>
<td>Vision AND IoT</td>
<td>270</td>
</tr>
<tr>
<td>Applications AND IoT</td>
<td>195</td>
</tr>
<tr>
<td>Total</td>
<td>923</td>
</tr>
</tbody>
</table>

In the next step, we excluded those articles that were not included in the well-known journals i.
It was found that these journals have published 146 articles. The number of articles published
per journal is shown in Table 3.

Table 3: Journal wise publication breakdown table

<table>
<thead>
<tr>
<th>Journals</th>
<th>Number of articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal and Ubiquitous Computing</td>
<td>25</td>
</tr>
<tr>
<td>International Journal of Production Economics</td>
<td>11</td>
</tr>
<tr>
<td>International Journal of Production Research</td>
<td>5</td>
</tr>
<tr>
<td>Computers and Security</td>
<td>6</td>
</tr>
<tr>
<td>IEEE Security and Privacy</td>
<td>5</td>
</tr>
<tr>
<td>IEEE Internet of Things Journal</td>
<td>47</td>
</tr>
<tr>
<td>Network Security</td>
<td>7</td>
</tr>
<tr>
<td>Computer Law and Security Review</td>
<td>9</td>
</tr>
<tr>
<td>Electronic Design</td>
<td>13</td>
</tr>
<tr>
<td>Journal of Information and Computational Science</td>
<td>9</td>
</tr>
<tr>
<td>Logistics Journal</td>
<td>6</td>
</tr>
<tr>
<td>Computer Networks</td>
<td>6</td>
</tr>
<tr>
<td>Future Generation Computer Systems</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>146</td>
</tr>
</tbody>
</table>

Figure 1 demonstrates the changing pattern of publications in each year, starting from 2009 to
2015. It can be clearly seen from the figure that the number of publications on IoT before 2013
increased slowly, but in the last 3 years, it has been increasing dramatically.
2.4 Data analysis

The output of our data analysis was used to conduct our analysis with bibliometric and network analytics tools. To conduct bibliometric analysis, different software packages are available with own capabilities and limitations. The most commonly used software for this purpose are ‘Publish or Perish’, ‘HistCite’, and ‘BibExcel’. There were two main reasons for selecting BibExcel software in this study. First, it is highly flexible in altering the data imported from databases such as Scopus and WoS, and second, it is able to offer an extensive data analysis which can be further used by network analysis tools. Other tools, such as HistCite can only work with data imported from WoS while Publish or Perish works with Google Scholar and Microsoft Academic Search. It is worth mentioning here that apart from BibExcel, the other tools do not generate data for future network analysis.

The analysis of data for the bibliometric analysis was conducted in two stages. In the first stage, bibliometric analysis was performed using BibExcel software which provides data statistics containing author, affiliation and keyword statistics. We opted for BibExcel software because it is flexible enough to handle huge amount of data and is also compatible with other applications such as, Excel, Pajek and Gephi (Persson et al., 2009). The data entered in BibExcel is in RIS format and contains all the necessary bibliographic information related to the papers. In our analysis, we mainly concentrated on information regarding authors, title, journal, publication year, keywords, affiliations, and references. During these analyses, the RIS file is converted into
different formats and, as a result, various file types are produced. To get a thorough knowledge about the processes and applications of BibExcel, readers may refer to Paloviita (2009) and Persson et al. (2009).

In the second stage, network analysis was done which makes use of the data prepared in BibExcel software. To conduct this analysis, Gephi was chosen. Besides Gephi, the most widely used tools for conducting any network analysis are Pajek (Batagelj and Mrvar, 2011), VOSviewer (van Eck and Waltman, 2013), and HistCite Graph Maker. For carrying out this study, we preferred Gephi because it provides flexible visual aids, powerful filtering techniques, inherent toolkit for network analysis and capability to handle different data formats. In addition, the other network analysis software lack one or the other property of Gephi. For instance, HistCite graph maker accepts WoS data files, Pajek can only handle .Net files and VOSviewer has limited tools for performing network analysis. Gephi is an important open source software package which makes use of a 3D render engine to make large networks in real time (Gephi, 2013). Owing to its flexible and multi-task architecture, it can easily handle complicated datasets and generate insightful visualization. Bastian et al. (2009) noted that Gephi provides “easy and broad access to network data and assist in specializing, filtering, navigating, manipulating and clustering of data”. Before going for visualization and mapping in Gephi, a dataset containing published papers, which is denoted by nodes, and their citations, represented by the arcs or edges between the nodes, must be prepared. Thus, the bibliographic data which is downloaded from Scopus and saved in RIS format cannot be used directly, and in this situation, BibExcel software acts as a mediator which reformats the original data file to graph dataset or .NET file. This file is saved for future network analysis in Gephi.

3. Review of the literature on Internet of Things

This section is broadly divided into two major areas of literature, that is, vision of IoT, and applications of IoT.

3.1 Internet of Things: Vision

The phrase “Internet of Things” originated at MIT Auto-ID Center and Kevin Ashton was the first to introduce it in 1999 during a presentation held at Procter & Gamble (Ashton, 2009). Ashton visualized that the physical world can be connected to the Internet via sensors and actuators which are capable of providing real time information and hence benefit our lives in
several ways. This concept came into public eye when International Telecommunications Union (ITU) published its first report on this subject in 2005. By adopting an integrated and comprehensive approach, ITU suggested that “Internet of Things will connect the world’s objects in both a sensory and intelligent manner through combining technological developments in item identification (“tagging things”), sensors and wireless sensor networks (“feeling things”), embedded systems (“thinking things”) and nanotechnology (“shrinking things”). In 2009, the Cluster of European Research projects (CERP) gave its vision on IoT by combining different ideas and technical components of Pervasive Computing, Ubiquitous Computing and Ambient Intelligence. They defined IoT as “a dynamic global network infrastructure with self-capabilities based on standard and interoperable communication protocols where physical and virtual “things” have identities, physical attributes, virtual personalities and use intelligent interfaces, and are seamlessly integrated into the information network” (Jain et al., 2009).

The RFID group views IoT as “the worldwide network of interconnected objects uniquely addressable based on standard communication protocols”. In their work on IoT, the CERP expected “Things” to “become active participants in business, information and social processes where they are enabled to interact and communicate among themselves and with the environment by exchanging data and information “sensed” about the environment, while reacting autonomously to the “real/physical world” events and influencing it by running processes that trigger actions and create services with or without direct human intervention”. CERP’s vision of IoT has been extended by Uckelmann et al. (2011) to form a blend of two different concepts: Web 2.0 and self-sustainability. Specifically, Web 2.0 technology uses simple and instinctive interfaces that enable users to make their web contributions, irrespective of their technical capabilities. This interaction between Things and users is of central importance because it will be one of the key issues in the future Web of Things. Uckelmann et al. (2011) combined these concepts and gave their own vision of IoT: “the future Internet of Things links uniquely identifiable things to their virtual representations in the Internet containing or linking to additional information on their identity, status, location or any other business, social or privately relevant information at a financial or non-financial pay-off that exceeds the efforts of information provisioning and offers information access to non-predefined participants. The provided accurate and appropriate information may be accessed in the right quantity and condition, at the right time and place at the right price. The Internet of Things is not synonymous with ubiquitous/pervasive computing, the Internet Protocol (IP), communication technology, embedded devices, its applications, the Internet of People or the Intranet/Extranet of Things, yet it combines aspects and technologies of all of these approaches”.

Atzori et al. (2010) pin-pointed three viewpoints for defining IoT: Things-oriented (sensors), Internet-oriented (middleware) and Semantic-oriented (knowledge). From a Things-oriented perspective, IoT is not merely the identification of objects but provides a much broader vision. The internet-oriented perspective emphasizes that efficient links should be established between
devices by taking advantage of the IP protocol and focusing on the networking paradigm. The semantic-oriented perspective aims at using semantic technologies for handling the large amount of data which is being generated from various IoT objects (Borgia, 2014). With reference to Atzori’s vision of IoT, Gubbi et al. (2013) noted that benefits of IoT can be realised only when these three paradigms coincide. In the context of smart environments, Gubbi et al. (2013) defined IoT as the “Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless ubiquitous sensing, data analytics and information representation with Cloud computing as the unifying framework”. Recently, Borgia (2014) suggested that a complete vision of IoT can be observed via 6As, that is, “Anytime-Anywhere”, “Anyone-Anything” and “Any path/network-Any service”.

3.2 Internet of Things: Applications

Following our review, we have categorized IoT applications into four major domains, that is, ‘Industry domain’, ‘Healthcare domain’, ‘Smart environments domain’, and ‘Personal and Social domain’.

3.2.1 Industry domain

The real time information provided by RFID and Near Field Communication (NFC) technology helps in keeping track of every activity in a supply chain, starting from product design to distribution and then final delivery of products to the end users. In doing so, organizations can obtain accurate and timely information related to the products that can help organizations respond to the market changes in shortest possible time. As an outcome, smart/advanced organizations (e.g. Wal-Mart and Metro) can meet changing customer requests promptly and with zero safety, stock whereas traditional organizations take approximately 120 days to meet this demand (Yuan et al., 2007). According to Karpischek (2009), shop assistants can provide up-to-date product information to the customers by having real time access to the ERP system. The real time information provided by RFID-based objects and smart shelves helps smart systems in reducing the level of material wastage, thereby saving cost and increasing profit margin. IoT applications can also be seen in the automobile industry. For instance, sensors installed in the vehicles can monitor its each and every detail (such as, tire pressure, motor data, fuel
consumption, location, speed, distance from other vehicles) and then transfer the gathered data to the central system (Hank et al., 2013).

### 3.2.2 Healthcare domain

Internet of Things has several potential benefits in medical and healthcare sectors. Smart tracking devices help in detecting a moving person or item. IoT involves real time location tracking as well as movement tracking at choke points, where the former may be used to identify and track the location of a patient in a hospital, and the latter may help in monitoring the movement of patients through entry and exit points of a ward. In addition, these devices help in continuously managing the inventory status and monitoring the movement of materials within a hospital (Atzori et al., 2010). Other relevant applications aim at identifying patients and infants and at avoiding incidents such as infant mismatching, wrong dosage of medicines, and incorrect procedures. These incidents can be minimized by maintaining an electronic medical record system that contains information of all in- and out-patients. In fact, patients’ conditions can be analysed by using sensor devices that help in obtaining real time information related to patients’ health. The data generated through these devices can be then transferred to medical staff for further diagnosis by using communication technologies (such as, Bluetooth, Zig Bee, Wireless HART, and ISA100).

### 3.2.3 Smart environments domain

Internet of Things may enhance the quality of people’s life in several ways. Nowadays, vehicles with mobile sensors get detailed information related to traffic density or surface conditions of the road as compared to the fixed sensors which were used earlier (Ganti et al., 2011). Moreover, the data gathered from these sensors can be then transmitted to control centers via vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems. Additionally, Polycarpou et al. (2013) identified the application of IoT in parking systems. Its application may help drivers in finding a parking lot as per their convenience and preference, thereby saving time and fuel, while bringing down the level of carbon footprint. Sensors located at parking lots ease the work of municipalities by detecting the illegally parked vehicles which can be then towed away. The payment systems at toll booths and parking lots can be made easy and smooth. The drivers may adopt NFC technology in their mobile phones for payments at parking and use RFID-based electronic system for toll collection (Qadeer et al., 2009). In addition, IoT may find
its applications in transforming the traditional gym to smart gym. The gym trainer can feed the exercise description in the training machine for each trainee. The RFID tag in the machine can then automatically identify the trainee and monitor the health parameters throughout the training session (Atzori et al., 2010). Further applications can be observed in entertainment and tourism sectors. In this regard, Amato et al. (2012) mentioned that smart phone users can obtain information related to monuments and tourist places.

3.3.4 Personal and social domain

Many benefits are provided by IoT to the personal and social domain. A broad range of applications can be generated by combining sensors and smart devices (e.g., broadband gateways, mobile phones, laptops, PCs, TV, speakers, appliances, plugs, surveillance cameras and lights). Computerized home systems enable residents to control every activity remotely via web applications. Chen et al. (2013) suggested that users can live a comfortable life if their smart phones act as a remote control for managing all the household appliances and their habits are continuously monitored by tracking their mobile phones. As an example, by analysing the information flow, a system can learn a person’s schedule, and thus perform automatic functions such as unlocking the door and switching on the lights. In the context of loss and theft, a web-based RFID application acts as a search engine for Things. It assists users in finding the lost objects by following their previous location records. Additionally, it alerts the user if any specific object is displaced from its original position. For instance, the user may receive an SMS via application if the stolen objects are taken out of any restricted area. Moreover, social domain applications allow easy communication among people so that they can build and maintain strong social relationships. Social networking enables an automatic update of our social activities on social networking sites such as Twitter and Facebook.

3.2 Internet of Things: influence of researchers

To analyse the influence of particular researchers, the author field was first extracted from the RIS data file and then the frequency of occurrence of each of these authors was noted. Table 4 shows the top ten contributing authors along-with their number of publications. It can be clearly observed that Weber and Wang with 6 publications dominate the list, and is followed by Jara with 4 publications.
3.3 Keyword statistics

A similar analysis was performed in order to identify the most commonly used words in the paper titles and the list of keywords. Table 5 and Table 6, show the top 20 keywords used in the paper titles and most popular keywords from the list of keywords, respectively. By comparing these two tables, it can be observed that there is a uniformity in the use of keywords in the title and the list of keywords. For instance, in both tables the top keywords include a combination of Internet of Things, vision and its applications. It is to be noted here that the most popular keywords which occur in Table 5 are actually the search keywords which we chose for this study.

Table 4: Top 10 contributing authors

<table>
<thead>
<tr>
<th>Author</th>
<th>Number of published articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weber, R.H.</td>
<td>6</td>
</tr>
<tr>
<td>Wang, H.</td>
<td>6</td>
</tr>
<tr>
<td>Jara, A.J.</td>
<td>4</td>
</tr>
<tr>
<td>Guo, L.</td>
<td>3</td>
</tr>
<tr>
<td>Yang, J.</td>
<td>3</td>
</tr>
<tr>
<td>Ten Hompel, M.</td>
<td>3</td>
</tr>
<tr>
<td>Tao, F.</td>
<td>2</td>
</tr>
<tr>
<td>Li, J.</td>
<td>2</td>
</tr>
<tr>
<td>Ferrari, G.</td>
<td>2</td>
</tr>
<tr>
<td>Huang, G.Q.</td>
<td>2</td>
</tr>
</tbody>
</table>
3.4 Network Analysis

3.4.1 Citation analysis

The aim of citation analysis is to examine the citation frequency of a particular document. Garfield (1972) mentioned that the total number of citations on a scientific journal is an indication of its significance in that area of research. It has also been emphasized that the impact of heavily cited articles on scientific research is greater than that of less cited articles (Sharplin and Marby, 1985; Culnan, 1986). Through citation analysis, researchers can determine the time period during which the major articles in a field were published and how their popularity has evolved over time, and hence if an article is still useful for current research (Pilkington and Meredith, 2009). Although citation analysis has received a lot of criticism, it is regarded as one of the most commonly used techniques for analysing literature and identifying the most influential author, journal, or work in that particular area of research (Mac Roberts and Mac Roberts 1989, 2010; Vokurka 1996).

Table 6: Top 20 commonly used words in titles

<table>
<thead>
<tr>
<th>Word</th>
<th>Frequency</th>
<th>Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Things</td>
<td>60</td>
<td>Systems</td>
<td>8</td>
</tr>
<tr>
<td>Internet</td>
<td>58</td>
<td>New</td>
<td>8</td>
</tr>
<tr>
<td>IoT</td>
<td>10</td>
<td>Information</td>
<td>8</td>
</tr>
<tr>
<td>RFID</td>
<td>10</td>
<td>Data</td>
<td>8</td>
</tr>
<tr>
<td>Supply</td>
<td>10</td>
<td>Management</td>
<td>5</td>
</tr>
<tr>
<td>Approach</td>
<td>10</td>
<td>Networks</td>
<td>5</td>
</tr>
<tr>
<td>Applications</td>
<td>10</td>
<td>Chain</td>
<td>7</td>
</tr>
<tr>
<td>Service</td>
<td>9</td>
<td>Security</td>
<td>7</td>
</tr>
<tr>
<td>Vision</td>
<td>8</td>
<td>Architecture</td>
<td>7</td>
</tr>
<tr>
<td>Privacy</td>
<td>8</td>
<td>Design</td>
<td>7</td>
</tr>
</tbody>
</table>
Figure 2: Frequency distribution of top 10 cited articles

Figure 2 demonstrates the top ten influential works published between 2000 and 2015. The most influential article during this period, having received 764 citations, is the work published by Atzori et al., (2010). These authors introduced and compared different visions of the IoT paradigm, that is, ‘Internet oriented’, ‘Things oriented’, and ‘Semantic oriented’. In addition, they discussed the enabling technologies and the potential applications of IoT. Another important contribution was by Weber (2010) who devoted to the study of privacy and security challenges related to IoT. This work received 153 citations which reflects the significance of the article in this field. Gubbi et al. (2013) provided a Cloud centric vision for IoT. They presented a case study of data analytics based on the ‘Aneka’ cloud platform which is based on interaction of private and public Clouds. The article received 136 citations and became the third most influential work of this period. Furthermore, the article by Xu (2011) which has been cited 75 times, increased the awareness of readers in terms of product quality and gave suggestions to explore the roles of service-oriented architecture, RFID, agents, work flow management, and IoT as enablers so that customer value can be improved in new product development. The impact of the work can be identified from the fact that until now 75 scientific articles have been published based on their work. A piece of work by Gunasekaran et al. (2009) that received 50 citations becomes the fifth most important article of this time period. The main motive behind their work was to understand the concept of e-procurement in SMEs. They developed a framework for the successful adoption of e-procurement. Table 7 shows the numbers of citations received by the influential articles.

Table 7: Top ten articles based on citations
3.4.2 PageRank analysis

To measure the importance of any article, several methods are available; one is Citation analysis which has been discussed above (Cronin and Ding, 2011). Ding et al. (2009) emphasized that significance of an article cannot be determined only by measuring the number of citations. Besides popularity, prestige which reflects that how many times an article has been cited by highly cited papers, is an important criteria. Although these measures may be positively correlated in some cases, it is not mandatory that a highly cited paper is also a prestigious paper. Brin and Page (1998) introduced PageRank as a measure for both popularity and prestige, and as an excellent way to prioritize the results of web keyword searches.

Suppose that article $A$ has been cited by papers $T_1, \ldots, T_n$. Define a parameter $d$ as the damping factor, which represents the fraction of random walks that continue to propagate along the citations, and whose value is fixed between 0 and 1. Now, define $C(T_i)$ as the number of times paper $T_i$ has cited other papers. The PageRank of paper $A$, denoted by $PR(A)$, in a network of $N$ papers is calculated as follows:

$$PR(A) = \frac{(1 - d)}{N} + d \left( \frac{PR(T_1)}{C(T_1)} + \cdots + \frac{PR(T_n)}{C(T_n)} \right)$$

It is important to note that if $C(T_i) = 0$, then $PR(T_i)$ will be divided to the number of papers instead of $C(T_i)$. The value of parameter $d$ has always been a point of debate. Brin and Page (1998) argued that in the original Google PageRank algorithm, the value of parameter $d$ was
fixed at 0.85, while Chen et al. (2007) claimed that \( d = 0.5 \) is a more appropriate choice for carrying out PageRank analysis in citation networks.

Table 8 extrapolates the top 10 papers using PageRank analysis. When comparing Table 7 and Table 8, it is observed that the topmost paper based on citations is Atzori et al. (2010). It has still remained on the first position in the list of top ten high-PageRank papers. The second highly cited paper (that is, Weber (2010)) is not present in the list whereas, the third highly cited paper (Gubbi et al. (2013)) came down to the fourth position in Table 8. Atzori et al. (2012), ranked ninth in Table 7, is second in Table 8.

### Table 8: Top 10 articles based on PageRank

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Page Rank</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atzori et al., 2010</td>
<td>0.02149</td>
<td>764</td>
</tr>
<tr>
<td>Atzori et al., 2012</td>
<td>0.01093</td>
<td>38</td>
</tr>
<tr>
<td>Chui et al., 2010</td>
<td>0.00891</td>
<td>67</td>
</tr>
<tr>
<td>Gubbi et al., 2013</td>
<td>0.00885</td>
<td>136</td>
</tr>
<tr>
<td>Yu et al., 2010</td>
<td>0.00843</td>
<td>45</td>
</tr>
<tr>
<td>Gomez and Paradells, 2010</td>
<td>0.00809</td>
<td>119</td>
</tr>
<tr>
<td>Kranz et al., 2010</td>
<td>0.00806</td>
<td>89</td>
</tr>
<tr>
<td>Katasonov et al., 2008</td>
<td>0.00783</td>
<td>45</td>
</tr>
<tr>
<td>Miorandi et al., 2012</td>
<td>0.00768</td>
<td>343</td>
</tr>
<tr>
<td>Guinard et al., 2010</td>
<td>0.00763</td>
<td>256</td>
</tr>
</tbody>
</table>

#### 3.4.3 Co-citation analysis

Co-citation analysis is a way to investigate the relationships between authors, topics, journals or keywords, thus explaining how these groups are related with each other (Small, 1973; Pilkington and Liston Heyes, 1999). It can be conducted either on the basis of authors, which helps in manifesting the social structure, or on the basis of publications, which reveals the intellectual structure of research field (Chen et al., 2010). Through co-citation analysis, the major research
clusters within a particular field and how they evolve and vary across different journals over time can be determined. According to Leydesdorff and Vaughan (2006: in Pilkington and Meredith, 2009), the data received through co-citation “can be considered as such linkage data among texts, while cited references are variables attributed to texts…one should realize that network data are different from attributes as data. From a network perspective, for example, one may wish to focus on how the network develops structurally over time.”

Figure 3: Force Atlas layout of 172 nodes (articles)

To perform co-citation analysis, a .NET file obtained for 146 articles in BibExcel is opened in Gephi. When the .NET file is opened for the first time, Gephi generates a random map which has no visible pattern. However, different layouts can be created by using various algorithms of Gephi. In this study, we used Force Atlas layout which is highly recommended by developers as it is easy to understand. In these networks, edges attract and nodes repulse each other. It is worth mentioning here that the values of repulsion strength, gravity, speed, node size and other characteristics can be altered manually (Bastian et al., 2009).

On performing co-citation mapping for the first time in Gephi, it was found that 186 articles out of a total of 492 have been co-cited by other papers within this sample. When the .NET file is initially opened, Gephi generates a random map which has no visible pattern. However, different layouts can be created by using various algorithms of Gephi. In this study, we used Force Atlas, a force driven algorithm which is highly recommended by developers as it is easy to understand (Fahimnia et al., 2015). In these networks, edges attract and nodes repulse each other. It is worth mentioning here that the values of repulsion strength, gravity, speed, node size and other
characteristics can be altered manually (Bastian et al., 2009). By using this algorithm, the nodes which are strongly connected move to the center of the network whereas, the less connected nodes move out to the boundaries.

The Force Atlas layout of 172 node co-citation map is shown in Fig. 3. The co-cited articles are connected with each other while, the poorly connected nodes shift away from the center. Moreover, the nodes which are isolated from rest of the network, also termed as ‘outliers’, are excluded for the purpose of data clustering, done in the next section. On excluding these outliers we are left with a network having 172 nodes and 862 edges.

3.4.4 Data clustering: identifying research themes in the literature of Internet of Things

Data clustering is a technique to group a similar set of articles together (Radicchi et al., 2004). This can be done by grouping the nodes into clusters such that the edges between the nodes of the same cluster are denser as compared to those of different clusters (Clauset et al., 2004; Leydesdorff, 2011; Radicchi et al., 2004). Blondel et al. (2008) argued that the density of links inside communities versus the links between communities can be measured by Modularity. The default modularity tool in Gephi is based on Louvain algorithm, and the value of modularity index varies between -1 and +1. Blondel et al. (2008) gave the formula for calculating modularity index as follows:

\[
Q = \frac{1}{2m} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j),
\]

where \( A_{ij} \) represents the weight of the edge between nodes i and j, \( k_i \) is the sum of the weights of the edges attached to node i (\( k_i = \sum_j A_{ij} \)), \( c_i \) is the community to which vertex i is assigned, \( \delta(u, v) \) is equal to 1 if \( u = v \) and 0 otherwise, and finally \( m = \left( \frac{1}{2} \right) \sum_{ij} A_{ij} \).
On applying this algorithm to 172-node network, five major clusters were created. The positioning of and interaction among these clusters is depicted in Figure 4. It can be observed that the thickness of the arcs between the nodes vary from each other, where the thickness reflects the degree frequency for co-occurrence of any two papers in the reference list of other papers. The modularity index for this network was found to be 0.19. This indicates strong inter-relationship between the nodes of each cluster, as well as, between the nodes of different clusters. In view of Hjørland (2013), when two or more papers are often cited together, they probably share similar area of interest. Hence, a detailed analysis of papers belonging to one cluster can help in identifying the research area of that cluster. Since the number of papers in each cluster is high, we considered only the top publications of each cluster, on the basis of their co-citation PageRank. Table 9 shows the top publications of each cluster.

Table 9: Top 10 papers of each cluster: co-citation PageRank measure
In order to find out the area of research focus of each cluster, we carefully examined the contents and research areas of the leading papers. We found that research belonging to 1st cluster is mostly theoretical and conceptual. Researchers in this cluster review the literature and outline current and future challenges (e.g., Atzori et al., 2010; 2012). The aim of the 2nd cluster is to move ahead with well-established concepts and theories and implement them in different fields, including, for instance, IoT in smart cities and hospitals (e.g., Yu et al., 2010). Authors in 3rd cluster are mainly interested in studying the applications of IoT in logistics and supply chain (Luo et al., 2007; Krantz et al., 2010). Researchers belonging to 4th cluster concentrate at designing and planning of IoT whereas, the 5th cluster is devoted to study the security and privacy aspects of IoT (e.g., Katsanov et al., 2008; Zorzi et al., 2010). It can be observed that the first cluster is the most popular one, while there is a scope of future work in cluster 4th and 5th. Therefore, so far literature has mainly focused on reviewing the literature on IoT and suggesting potential applications in different contexts. Scholars are yet to conduct and report findings on case studies focusing on the adoption of IoT in these contexts, as well as the challenges that may come to the forefront during IoT adoption. Such studies would be important, since it is of crucial importance that information systems (IS) research and practice associates technology innovation with the context within which it is embedded (Avgerou, 2001). Furthermore, there are so far no studies focusing on providing particular frameworks or models on how IoT could be adopted, as well as whether/how IoT is different than other adoption processes of Information Systems, given that the number and type of IoT technology (and devices) is increasing exponentially every year (Guinard and Vlad, 2009). Finally, since IoT adoption would need to consider the wider socio-organizational context in which it will be embedded, there are yet studies to examine IoT.
considering the local organizational, but also the national and international context, as well as both the technical/rational decisions and actions involved in the adoption process and the cultural, social and cognitive forces of such a process.

4. Discussion

In this paper we conducted a bibliometric and network analytics study of the literature related to IoT. This study was triggered because of two facts: firstly, the IoT literature is growing exponentially, but however the literature surrounding IoT is still underdeveloped; and secondly, IoT has attracted significant attentions from both academics and industry. However the majority of the literature stems from technology perspective. Research related to the adoption and applications of IoT in business –for instance in particular smart cities, hospitals and supply chains– are still underdeveloped. In the subsections that follow, we outline our theoretical contribution and the managerial implications of our work.

4.1 Theoretical contributions

The current study provides a bibliometric and network analytics review of the literature on IoT, inspired by Fahimnia et al. (2015) and their review of the green supply chain management literature. No matter if we conducted our study in a time span of 16 years, the majority of articles have been published over the last 5 years. It is also worth mentioning that the top influential studies (as our findings suggest) come from few researchers. Our contribution lies in (i) identifying top contributing authors in the field as well as the key research topics and influential works based on citation analysis and PageRank; and (ii) proposing a five-cluster classification of the IoT research themes based on data clustering. Such a clustering is important, we believe, since it enables researchers not only to acknowledge the diversity of research in the field, but also because it provides those areas where more research would need to be conducted. Our study, hence, differs from reviews such as Atzori et al. (2010) or Atzori et al. (2012) in that we are not only reporting different visions of IoT and enabling technologies or appropriate policies for the establishment and the management of social relationships through IoT. Research should not only focus on identifying the current and emerging technology solutions for IoT (Katasonov et al., 2008; Gomez and Paradells, 2010), but scholars should attend to the diverse socio-organizational, both local and international, context in which IoT is to be embedded (Avgerou, 2001). The study of Krantz et al. (2010) that investigates human-computer interaction as enabled by IoT and related technologies we believe is to the right direction since it shifts the interest
from the technology per se to how IoT is embedded within human interactions. Yet, there are studies to be conducted on the adoption process and the enablers, drivers, barriers, and models of IoT adoption by organizations and supply chains from both a technological and socio-organizational point of view. Furthermore, there are limited, if any, studies that look into the relationship between IoT adoption and increase of organizational and supply chain performance. Such studies are a necessity, given the recent focus on efficiency and sustainability within the supply chain, and the aim to use technological solutions that enable transparency and visibility at the lowest cost, energy consumption, and environmental footprint (e.g. Malhota et al., 2013; Dubey et al., 2016). Finally, paraphrasing the endorsement of scholars (e.g. Holmstrom et al., 2009; Taylor and Taylor, 2009; Ketokivi and Choi, 2014) who suggest the use of alternative lenses to the study of operations management related phenomena, we would like to stress the importance of using alternative theories and mechanisms that look into the wider implications of IoT implementation and adoption. Therefore, based on the findings of this research we identify and propose the following questions:

1. What are the drivers and barriers of IoT implementation and adoption?
2. How can we explain IoT implementation and adoption using alternative organizational theories?
3. How can we measure the impacts of IoT on organizational and supply chain performance?
4. Can we propose a holistic model that explains the acceptance of IoT applications?

4.2 Managerial implications

Our study has the following managerial implications: firstly, it enables practitioners to acknowledge the vision and different applications of IoT, as well as the different focus of research clusters; secondly, suggests that managerial attention should be not only on the selection of technologies, but also on the wider socio-organizational implications of the IoT adoption for organizations and supply chains; and thirdly, it enables managers and decision makers to gain a holistic understanding of the implications of IoT so that they make better decisions with regards to its adoption and the necessary resources that need to be in place to facilitate the transition to the IoT era and the implications of IoT for achieving superior performance.

5. Conclusions, limitations, and future research
The study has reviewed and examined articles published over a period of 16 years (2000-2015) by using bibliometric and network analyses. The main objective of this study was to identify highly cited and co-cited works related to IoT offering future research directions to the IoT research community and implications for managers and decision makers. We have also proposed a five-cluster overview of research themes across IoT. Our results and five-cluster classification of IoT research illustrate the increasing importance of IoT, but on the other hand the studies that acknowledge the applications of IoT for organizations and supply chains and the wider socio-organizational context that needs to be considered; such studies are missing from the literature. Hence, the majority of the highly cited and co-cited works in the field are dominated by conceptualisations and there are few applications of IoT that include case studies, which would provide a more in-depth understanding of how IoT emerges, how it is adopted, and what the advantages and challenges from its use are. Furthermore, our findings highlight the need for alternative theories and lenses to be used in order to study IoT related phenomena. The findings of this study may help scholars in understanding (i) the concept of IoT; (ii) the changing research trends in the field of IoT, and those articles that have been influential in shaping research in these years; and (iii) intellectual structure of the field.

The paper has the following limitations:

1. The findings of the review are based on 15 peer reviewed journals with a focus on last 17 years (2000-2016) publications. Certainly, this study may have missed articles published in other peer reviewed journals.
2. The study adopted bibliometric technique of citation and co-citation analysis for reviewing the literature (Pilkington and Meredith, 2009). There may be other methods to be used for citation and co-citation analysis.
3. We have used the method of citation analysis but did not focus on the impacts of self-citation. We acknowledge that citation studies tend to exclude self-citation, but today’s Journal Impact Factor includes them. It may be that other studies could consider the effect of self-citations and see if current pattern changes by excluding or discounting the self-citation effects.
4. We have used particular keywords (‘Internet of Things’) in our searches for abstract, title, and full text. However, the use of other keywords may generate different search results.
5. We have classified IoT application into four categories. This classification is by no means exhaustive, and other scholars could use in their studies different classifications.
6. The study has not taken into consideration the technologies and architectural elements of IoT (Gubbi et al, 2013; Borgia, 2014).
Despite the aforementioned limitations, we believe that our current attempt will offer motivation to undertake research to advance IoT and theories to explain IoT related phenomena. We believe that the existing IoT literature can be further enriched through research that examines IoT using behavioural and organizational theories. Furthermore, research on IoT is in nascent stage, and hence we believe that use of research methods such as case studies, ethnography, grounded theory and action research can provide alternative angles to explain the complexity surrounding IoT. Furthermore, studies that shift the focus from purely technological to the socio-organizational implications of IoT adoption, and that suggest holistic models of IoT adoption and implementation would benefit both researchers and managers who would like to further explore IoT. Finally, studies (both case studies and surveys) that explore the drivers and barriers of IoT implementation and adoption as well as its impact on the environment and performance would be strongly needed.

References


ITU Internet Reports (Nov., 2005), “The Internet of Things”.


life, business, and the global economy”, Executive Summary.

through RFID data”, in Proceedings of the 1st IEEE International Workshop on RFID
Data Management (RFDM). Cancun, Mexico, pp. 102–108.


Mohar, B. (1988), “The Laplacian spectrum of graphs”, In: Graph Theory, Combinatorics, and

Molloy, M., Reed, B. (1995), “A critical point for random graphs with a given degree sequence”,

solution using a novel material handling system”, Innovative Food Science and Emerging
Technologies”, Vol. 6, No. 1, pp. 91–105.

Vol. 8, No. 3, pp. 26-41.

256. doi: 10.1137/S003614450342480.

and their applications”, Physical Review E, Vol. 64, 6118.

Analysis & Defenses”, In Proceedings of the 3rd international symposium on Information

Vol. 46, No. 4, pp.46-53.


Paloviita, A. (2009),“Stakeholder perceptions of alternative food entrepreneurs”, World Review
of Entrepreneurship, Management and Sustainable Development, Vol. 5, No. 4, pp. 395–
406.

health monitoring and prognosis”, IEEE Transactions on Systems, Man and Cybernetics,
Part C (Applications and Reviews), Vol. 40, No. 1, pp. 1–12.


