THE UNIVERSITY OF TOKYO

DOCTORAL THESIS

博士論文

# Consensus-Building on Citations in Peer-to-Peer Systems

(Peer-to-Peerシステムにおける引用の合意形成について)

*Author:*

Kensuke ITO

伊東 謙介

*Supervisor:*

Dr. Noboru KOSHIZUKA

越塚 登

*A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy*

*in the*

Socio-information and communication studies course

Graduate School of Interdisciplinary Information Studies

August 27, 2021

# Declaration of Authorship

I, Kensuke ITO, declare that this thesis titled, "Consensus-Building on Citations in Peer-to-Peer Systems" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my work.
- I have acknowledged all main sources of help.
- The thesis is based on my work done jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---

*"Quis custodiet ipsos custodes?"* (Who watches the watchmen?)

Juvenal (A.D. 60–130), *Satires* vi, 347.

THE UNIVERSITY OF TOKYO

# *Abstract*

Socio-information and communication studies course

Graduate School of Interdisciplinary Information Studies

Doctor of Philosophy

**Consensus-Building on Citations in Peer-to-Peer Systems**

by Kensuke ITO

This thesis aims at consensus-building on citations in Peer-to-Peer (P2P) systems. Citations, a source of various quantitative measures for intellectual products (e.g., scientific publications, patents, web pages), are more robust and productive if autonomous peers in a P2P system can determine and construct their true structure. However, this consensus-building has remained unreliable due to three problems that preceding studies have not addressed simultaneously: *free-riding*, *strategic misreporting*, and *reviewer assignment*. Therefore, we combined *random walks on graphs* with *peer prediction methods* and proposed two incentive mechanisms (ex-ante and ex-post consensus) that reward reviewers who participated in consensus-building. Experimental studies support the usefulness of the two incentive mechanisms for all three problems, by showing that peers can (i) be reviewers more often as they get higher PageRank scores and (ii) maximize the expected rewards per review by always reporting true beliefs. Our proposal—rewards from the consensus-building on citation relationships—also contributes to open-access intellectual products as an alternative scheme to grants, royalties, and advertisements. On the other hand, potential applications require future studies to prevent spamming and Sybil attacks and make the reward a sufficient incentive.

*Keywords*: citation analysis, random walks on graphs, peer prediction methods

# *Acknowledgements*

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Noboru KOSHIZUKA and ex-supervisor Dr. Hideyuki TANAKA for their continuous support of my Ph.D. study. Their guidance helped me in all the time of research and writing of this thesis.

Besides, I would like to thank the rest of my thesis committee: Dr. Osamu SUDO, Dr. Jun REKIMOTO, Dr. Akihiro NAKAO, Dr. Junichi KISHIGAMI, Dr. Soichiro TAKAGI, and Dr. Shohei OHSAWA, for their insightful comments and encouragement, but also for the hard question to widen my research from various perspectives.

I would like to extend my sincere thanks to Dr. Toyotaka SAKAI and Dr. Kyohei SHIBANO, for a number of insightful comments to this thesis.

My sincere gratitude also goes to all my colleagues in KOSHIZUKA laboratory, TANAKA laboratory, and Startbahn, Inc., who have been together in the last six years and have had a variety of stimulating discussions about my research.

Finally, this thesis is supported by the Graduate Program for Social ICT Global Creative Leaders (GCL). I gratefully acknowledge the assistance of all associates, program staffs, and friends of the GCL.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **ACP** | **A**tmospheric **C**hemistry and **P**hysics |
| **AHCI** | **A**rts and **H**umanities **C**itation **I**ndex |
| **AMD** | **A**lgorithmic **M**echanism **D**esign |
| **BIP** | **B**itcoin **I**mprovement **P**roposal |
| **CC** | **C**reative **C**ommons |
| **CoI** | **C**onflicts-**o**f-**I**nterest |
| **DAG** | **D**irected **A**cyclic **G**raph |
| **DAMD** | **D**istributed **A**lgorithmic **M**echanism **D**esign |
| **DApps** | **D**ecentralized **App**lication**s** |
| **DG13** | **D**asgupata and **G**hosh in 20**13** |
| **DOAJ** | **D**irectory of **O**pen **A**ccess **J**ournals |
| **HITS** | **H**yperlink-**I**nduced **T**opic **S**earch |
| **NBER** | **N**ational **B**ureau of **E**conomic **R**esearch |
| **OA** | **O**pen **A**ccess |
| **OPR** | **O**pen **P**eer **R**eview |
| **OSS** | **O**pen **S**ource **S**oftware |
| **PPR** | **P**ersonalized **P**age**R**ank |
| **PR** | **P**age**R**ank |
| **P2P** | **P**eer-to-**P**eer |
| **RAP** | **R**eviewer **A**ssignment **P**roblem |
| **RQ** | **R**esearch **Q**uestion |
| **SCI** | **S**cience **C**itation **I**ndex |
| **SNAP** | **S**tanford **N**etwork **A**nalysis **P**roject |
| **SPoF** | **S**ingle **P**oint **o**f **F**ailure |
| **SSCI** | **S**ocial **S**ciences **C**itation **I**ndex |
| **TCR** | **T**oken **C**urated **R**egistry |

| | |
|---|---|
| **UNESCO** | **U**nited **N**ations **E**ducational, **S**cientific and **C**ultural **O**rganization |
| **USPTO** | The **US** **P**atent **T**rademark **O**ffice |
| **WWW** | **W**orld **W**ide **W**eb |

# List of Symbols

| | | |
|---|---|---|
| $\boldsymbol{A}$ | Adjacency matrix | Chapters 1, 4 |
| $a_{ij}$ | An element of $\boldsymbol{A}$ | Chapter 1 |
| $\boldsymbol{B}$ | $\lvert V\rvert \times \lvert V\rvert$ matrix for PPR algorithm | Chapter 2 |
| $b_{ij}$ | An element of $\boldsymbol{B}$ | Chapter 2 |
| $\dot{C}_t$ | Set of reviewers selected in period $t$ | Chapters 3, 4 |
| $E$ | Set of edges | Chapter 1 |
| $E_k$ | Set of out-edges directed from $k$ to $V_k$ | Chapters 3, 4 |
| $E_t$ | Set of edges in period $t$ | Chapters 1, 3, 4 |
| $G$ | Graph, or DAG unless otherwise specified | Chapters 1, 2 |
| $G_t$ | Citations as a growing DAG in period $t$ | Chapters 1, 2, 3, 4 |
| $\dot{G}_t$ | A proposal of new citations in period $t$ | Chapters 3, 4 |
| $\dot{G}_{t'}$ | A proposal of new citations in a period other than $t$ | Chapter 3 |
| $\dot{G}_{t''}$ | A proposal of new citations in a period other than $t$ and $t'$ | Chapter 3 |
| $I$ | Importance measure in the two-path mechanism | Chapter 2 |
| $i$ | An element of $V$ | Chapters 1, 2, 3, 4 |
| $j$ | Another element of $V$ | Chapters 1, 2, 3, 4 |
| $k$ | A new vertex coming to DAG | Chapters 3, 4 |
| $M_i$ | Set of tasks assigned to $i$ | Chapter 2 |
| $M_j$ | Set of tasks assigned to $j$ | Chapter 2 |
| $M^*$ | $M_i \cap M_j$ | Chapter 2 |
| $m^*$ | An element of $M^*$ | Chapters 2, 3, 4 |
| $n$ | An element of $M_i \setminus \{m^*\}$ | Chapters 2, 3, 4 |
| $n'$ | An element of $M_j \setminus \{m^*\}$ | Chapters 2, 3, 4 |
| $\boldsymbol{P}$ | Probability matrix | Chapters 1, 2, 3, 4 |
| $p_{ij}$ | An element of $\boldsymbol{P}$ | Chapters 1, 2 |

| | | |
|---|---|---|
| $\boldsymbol{P_{PPR}}$ | Probability matrix for PPR algorithm | Chapters 2, 3 |
| $\boldsymbol{P_{PR}}$ | Probability matrix for PR algorithm | Chapter 2 |
| $P_1$ | One random path | Chapter 4 |
| $P_2$ | The other random path | Chapter 4 |
| $q$ | The number of state transitions | Chapters 1, 3, 4 |
| $R_i$ | Random variable for $i$'s report | Chapter 2 |
| $r_i$ | Realization of $R_i$ | Chapter 2 |
| $r_i^{m^*}$ | Realization of $R_i$ for $m^*$ | Chapters 2, 3, 4 |
| $r_i^n$ | Realization of $R_i$ for $n$ | Chapters 2, 3, 4 |
| $r_i^{n'}$ | Realization of $R_i$ for $n'$ | Chapters 2, 3, 4 |
| $R_t$ | Set of the stock of all reports until period $t$ | Chapter 3 |
| $\dot{R}_t$ | Set of reports elicited from $\dot{C}_t$ | Chapters 3, 4 |
| $R_1$ | Set of reports on $P_1$'s out-edges elicited from $P_1$ | Chapter 4 |
| $R_2$ | Set of reports on $P_2$'s out-edges elicited from $P_2$ | Chapter 4 |
| $R_3$ | Set of reports on $P_1$ (or $P_2$)'s out-edges elicited from $P_2$ (or $P_1$) | Chapter 4 |
| $S$ | Random variable for sending-signals allocated to each vertex | Chapter 3 |
| $S_i$ | Random variable for $i$'s received-signals | Chapter 2 |
| $s_i$ | Realization of $S_i$ | Chapter 2 |
| $t$ | Period | Chapters 1, 2, 3, 4 |
| $U$ | Set of vertices as mark by the two-path mechanism | Chapter 4 |
| $V$ | Set of vertices | Chapters 1, 2 |
| $V_b$ | Set of base vertices for PPR algorithm | Chapters 2, 3 |
| $V_k$ | Set of base vertices for $k$ | Chapters 3, 4 |
| $V_t$ | Set of vertices in period $t$ | Chapters 1, 3, 4 |
| $X_i$ | Random variable for $i$'s reward | Chapter 2 |
| $x_i$ | Realization of $X_i$ | Chapter 2 |
| $x_i^{m^*}$ | Realization of $X_i$ for $m^*$ | Chapters 2, 3, 4 |
| $\dot{X}_t$ | Set of rewards for $\dot{R}_t$ | Chapters 2, 3, 4 |
| $\alpha$ | Damping factor for PR algorithm | Chapters 2, 3, 4 |
| $\beta$ | Exogenous parameter for the number of two-path mechanism | Chapter 4 |
| $\gamma$ | Exogenous parameter for uninformative strategies | Chapter 4 |

| | | |
|---|---|---|
| $\delta$ | Kronecker delta | Chapters 2, 3, 4 |
| $\epsilon$ | Exogenous parameter for the randomness of strategy allocation | Chapters 3, 4 |
| $\lambda$ | Exogenous parameter $(\geq 2)$ for the number of reviewers | Chapter 3 |
| $\mu$ | Exogenous parameter $(\leq \lambda)$ for the difficulty | Chapter 3 |
| $\sigma_i$ | Strategy by $i$ | Chapters 2, 3, 4 |
| $\sigma_i^*$ | Truthtelling strategy by $i$ | Chapters 2, 3, 4 |
| $\sigma_{-i}$ | Pair of Strategies by all vertices except $i$ | Chapter 2 |
| $\mathbf{1}$ | $\|V\| \times \|V\|$ matrix for PR algorithm | Chapter 2 |

*Dedicated to my parents*

# Chapter 1

# Introduction

The purpose of this thesis, as the title implies, is consensus-building on citations in *Peer-to-Peer* (P2P) systems[1]. In this Chapter 1, the author introduces some backgrounds and preliminaries of this purpose, while answering the following questions:

- *Why are citations important?* (1.1)
- *Why are citations in a P2P system important?* (1.2)
- *Why are citations in a P2P system difficult?* (1.3)
- *What is the Research Question (RQ) of this thesis?* (1.4)
- *What are academic contributions of the RQ?* (1.3–1.4)

Finally, this chapter introduces the thesis outline (1.5), which indicates that the RQ is examined through the proposal of *incentive mechanism with ex-ante consensus* and *incentive mechanism with ex-post consensus*.

## 1.1 The Importance of Citations

Citation is "the representation of a decision made by an author who wants to show the relation between the documents he is writing and the work of another" (Mahapatra, 2009, pp. 17–18), and its accumulation can engender "networks of relatedness of subject matter" (Newman, 2010, p. 68). Citations are now used in a wide range of intellectual products, including scientific papers that refer to previous papers as a

[1]See Definition 1.2.1, for the usage of the term P2P system in this thesis.

source of information, patents that specify prior arts to prove the novelty of invention, and web pages that connect to other web pages via hyperlinks[2]. This section describes the importance of citations from three perspectives: as a source of quantitative measures, a growing network, and a graph structure.

### 1.1.1 Citations as a Source of Quantitative Measures

First and foremost, citations are essential as they can be a source of various quantitative measures for the quality of intellectual products. We can confirm this claim through a brief overview of the studies on citation analysis[3].

Studies relevant to citation analysis have a long tradition, and the oldest seems to be Gross and Gross (1927) that counted the number of citations to evaluate the importance of scientific journals[4]. A representative work contributed to the early citation analysis was Garfield (1955) that proposed an idea of the database for the citation relationships among published scientific papers. Since the idea became available as *Science Citation Index* (SCI) in the early 1960s, the SCI has provided large-scale, the computable dataset to several influential studies[5]. For example, Price (1965) revealed that new citations tend to concentrate on the paper which already has a number of citations. This trend was later generalized as *cumulative advantage* (Price, 1976) or *preferential attachment* (Barabási & Albert, 1999). Garfield (1955, 1972) introduced *impact factor* as a tool to facilitate literature-searching in the SCI, which is still widely used to evaluate scientific journals. Small (1973) proposed *co-citation analysis* as a complementary concept of *bibliographic coupling* (Kessler, 1963), which computes the similarity between two papers by the number of subsequent papers citing them together. Citation analysis stems from these attempts to offer quantitative measures

[2]In these examples, web page citations differ from those of scientific papers and patents in some ways. This is the reason why this thesis proposes two types of mechanisms. The author explains the differences later in 1.1.3 and 2.1.1.

[3]According to Garfield et al. (1983), "Citation analysis is a bibliometric method that uses reference citations found in scientific papers as the primary analytical tool" (p.581). See Guidera (2009), Meho (2007) for other definitions of citation analysis.

[4]Note that the term citation analysis appears to have appeared in the 1950s. Gross and Gross (1927) was, at that time, regarded as a study in the field of *library science*.

[5]The SCI later expanded its scale through the development of the *Social Sciences Citation Index* (SSCI; in 1973) and the *Arts and Humanities Citation Index* (AHCI; in 1978). It is now managed by Clarivate Analytics as a part of *Web of Science* (https://clarivate.com/products/web-of-science/, accessed August 4, 2019).

for scientific publications[6].

A natural extension of the citation analysis is the quantification of patented inventions whose specifications cite earlier inventions (prior art) as proof of novelty[7]. The use of citations had already been proposed to *the US Patent Trademark Office* (USPTO) in 1949 (Hart & Goldsmith, 1949; Seidel, 1949), and Garfield (1957) conducted the first experiment leading to the SCI on 5,000 chemical patents[8]. Patent citation analysis has so far revealed that the number (or frequency) of citations correlates with various proxy measures for the quality of inventions, including awards from academic journals (Carpenter et al., 1981), opinions from in-house senior staff (Albert et al., 1991), honors from public institutions (Breitzman & Narin, 1996), and renewal decisions by patent owners (Thomas, 1999). Moreover, studies focusing on the patents as a result of R&D activities have found correlations between citation-based indicators and several economic indicators, such as the social value of innovation[9] (Trajtenberg, 1989, 1990), stock performance (Breitzman & Narin, 2001; Deng et al., 1999), and Tobin's q (Hall et al., 2001, 2005). These preceding studies indicate that citations are informative for the quantification of patented inventions as well.

Another important subject for current citation analysis is the quantification of web pages[10]: the documents which can be interconnected via hyperlink on *World Wide Web* (WWW; Berners-Lee & Cailliau, 1990). For example, Larson (1996) did co-citation analysis for over 30 gigabytes of web pages collected by web-crawler; Abraham (1997) addressed citation analysis for visualizing the tree structure of WWW[11]; Ingwersen (1998) proposed *web impact factors* as an application of the impact factor to

[6]These researches are often classified as *bibliometrics* (Pritchard et al., 1969), or *scientometrics* (Nalimov & Mulchenko, 1969) when focusing on scientific publications. See review articles such as Osareh (1996) and Mingers and Leydesdorff (2015), for further details on the studies that follow these classifications.

[7]Strictly speaking, there are two types of patent citations: applicant citation and examiner citation. The former cites both patents and scientific papers ex-ante to show the novelty to the examiner, and the latter cites only patents ex-post to supplement incomplete applicant citations. See empirical studies (e.g., Alcácer & Gittelman, 2006; Alcácer et al., 2009; Criscuolo & Verspagen, 2008) for the actual distribution of the two types of citations.

[8]According to Garfield and Merton (1979), SCI started to cover patent citations in 1964, but was forced to remove them in 1966 due to the economic difficulties of extracting (applicant) citation data from patents held by private industry.

[9]It is estimated from price and product attributes.

[10]As an analogy of bibliometrics or scientmetrics, studies related to this subject are often classified as *webometrics* (Almind & Ingwersen, 1997) or *cybermetrics* (Aguillo, 1997). For example, see Björneborn and Ingwersen (2004) for the relations among these classifications.

[11]Björneborn and Ingwersen (2004), Björneborn et al. (2004), Thelwall (2002) classified the methods of citation analysis for the structure of WWW.

web pages. In addition, the quantification of web pages has been extensively studied in computer science, reflecting the need for efficient search. *PageRank* (PR; Brin & Page, 1998; Page et al., 1999) and *Hyperlink-Induced Topic Search* (HITS; Kleinberg, 1999) are the two representative algorithms for scoring relative importance among web pages, where the former uses the concepts of eigenvalues and *random walks on graphs* (e.g., Lovász et al., 1993)[12], and the latter uses the number of both out- and in-links (hubs and authorities) as criteria. The PR algorithm is of particular significance as the core technology of the *Google* search engine. Its score has recently been applied to citation analyses on scientific publications[13] (e.g., Bergstrom, 2007; Bollen et al., 2006; González-Pereira et al., 2010; Ma et al., 2008) and patents (e.g., Bruck et al., 2016; Dechezleprêtre et al., 2013; Lukach & Lukach, 2007; Shaffer, 2011) as an alternative to raw citation counts. Note that this thesis also uses the PR score as the quantitative measure (see Assumption 1.4.3).

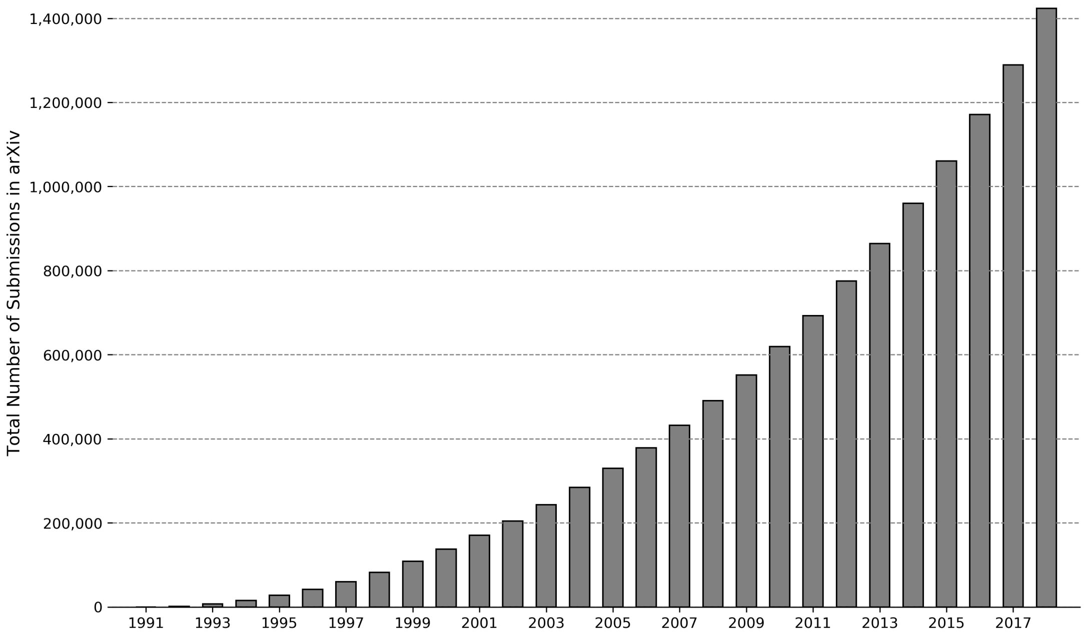
Thus, citations have long been analyzed because they are important sources of various quantitative measures for the quality of intellectual products, such as scientific papers, patents, and web pages.

### 1.1.2 Citations as a Growing Network

One of the features of such quantitative measures is that they become more important as the network structure of citations grows, i.e., as "new nodes are born over time with forming attachments to existing nodes when they are born" (Jackson, 2010, p. 124). Citations as a growing network reflect the cumulative nature of intellectual products whose total number is ever-increasing (as long as they are correctly archived), and this nature strengthens the utility of quantitative measures that can efficiently search and evaluate a large number of accumulated intellectual products. This is by no means a new opinion—editorial statements of the journal *Scientometrics* vol.1 (Price, 1978), for example, mentioned the tremendous increase of

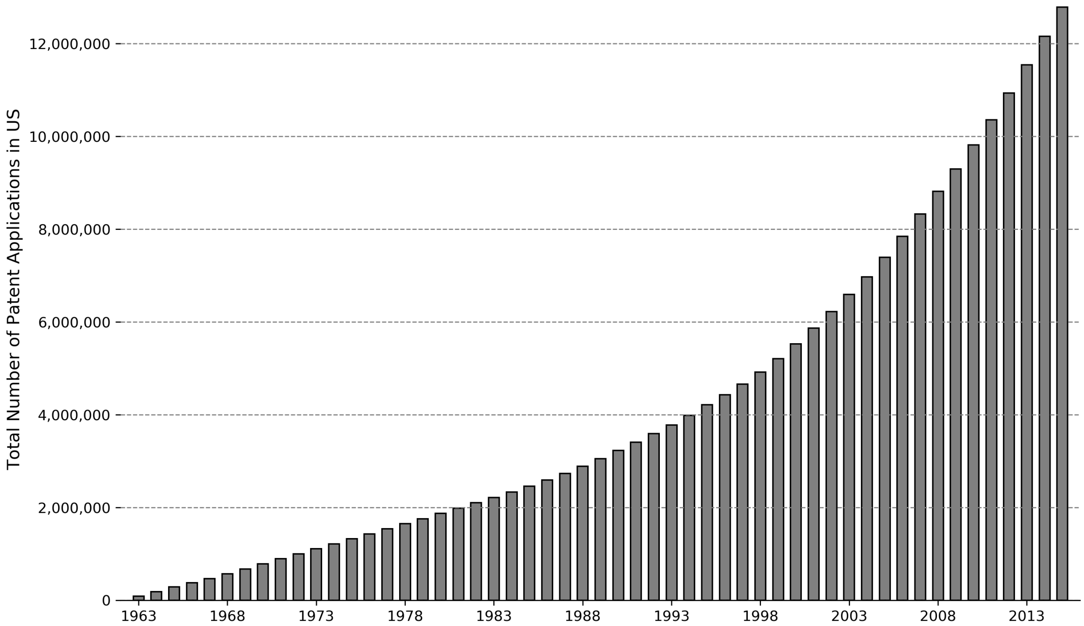
[12] We will confirm the details of random walks on graphs in Section 2.2.

[13] It should be emphasized here that Pinski and Narin (1976) has already proposed a method close to the PR algorithm as an influence measure for scientific journals. Still, it was confronted with a computational restriction at that time.




(a) Total Number of Pre-print Papers Submitted to arXiv (1991 – 2018)

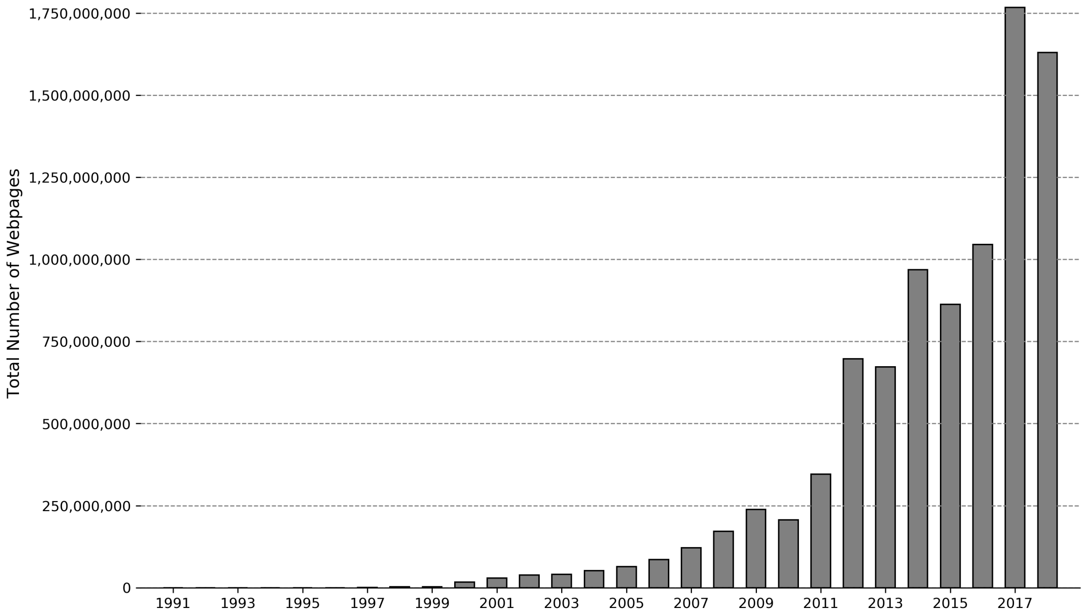
*Source*: arXiv Monthly Submission Rates (https://arxiv.org/stats/monthly_submissions, accessed July 23, 2019), created by the author.




(b) Total Number of Patent Applications in the USPTO (1963 – 2015)

*Source*: U.S. Patent Statistics Chart Calendar Years 1963 – 2015 (https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm, accessed July 23, 2019), created by the author.

FIGURE 1.1: **Exponential Growth in Total Numbers of Scientific Papers and Patents.** Figure 1.1a shows the total (cumulative) number of the pre-print papers submitted to arXiv from 1991 to 2018, and Figure 1.1b shows that of the patent applications in the USPTO from 1963 to 2015, respectively. These figures both depict a trend of exponential growth, implying that total numbers of scientific papers and patents are exponentially increasing.




*Source*: NetCraft and Internet Live Stats (https://www.internetlivestats.com, accessed July 23, 2019), created by the author.

**FIGURE 1.2: Exponential Growth in the Total Number of Web Pages.** This figure shows the total number of web pages from 1991 to 2018, where *NetCraft and Internet Live Stats* provides original dataset by counting the number of unique hostnames associated with IP address. The total number apparently has the trend of exponential growth, although it occasionally drops due to some reasons including the update of aggregation methods.

the scientific production is the necessity for quantitative evaluation that can facilitate research activity (by providing more economical and balanced utilization of the available funds). Needless to say, this feature is the same for patents and web pages.

Citations, at least those for the aforementioned three intellectual products, are not just growing but may be exponentially increasing. Figure 1.1a and 1.1b each depict time-series data related to scientific papers and patents[14]; the former is the total number of pre-print papers submitted to *arXiv* (Ginsparg, 1994, 2011)[15] from 1991 to 2018, and the latter is that of patent applications in the USPTO[16] from 1963 to 2015. We can see that both numbers have a trend of exponential growth. This trend is more pronounced for web pages. Figure 1.2 depicts time-series data on the total number of web pages from 1991 to 2018[17], which has the trend of exponential

[14]Note that Figure 1.1a and 1.1b depict the cumulative number, whereas original datasets both display non-cumulative number.

[15]https://arxiv.org/stats/monthly_submissions, accessed July 23, 2019.

[16]https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm, accessed July 23, 2019.

[17]The author retrieved original dataset from *NetCraft and Internet Live Stats* (https://www.internetlivestats.com, accessed July 23, 2019). According to the source, the dataset is computed by regarding each web page as a unique hostname.

growth even though it has occasionally dropped due to some reasons including the update of aggregation methods[18]. These statistics, indicating exponential growth in intellectual products, imply the further importance of citation-based quantitative measures.

Thus, citations are important because the quantitative measures they generate are beneficial for intellectual products whose total number is (exponentially) increasing.

### 1.1.3 Citations as a Graph Structure

It should also be emphasized that we can develop various quantitative measures because the network structure of citations itself is easy to quantify as *graph* which is "a collection of vertices joined by edges" (Newman, 2010, p. 109). Here, *vertices* and *edges* are the terms in graph theory[19], corresponding to nodes and links in the context of computer science. To my knowledge, Garner et al. (1967) provides the first graph-theoretic citation analysis, with scientific papers as vertices and their citation relationships as edges.

Specifically, this thesis focuses on *directed acyclic graph* (DAG) as a graphical representation of citations. For the DAG, we use the following set of definitions (and notations) which are based on Diestel (2012), Thulasiraman and Swamy (1992), and Uhlir (2017):

**Definition 1.1.1** (Graph)**.** A *graph* is a pair $G = (V, E)$, where $V$ is a non-empty and finite set of *vertices* and $E \subseteq V \times V$ is a finite set of *edges*,

**Definition 1.1.2** (Directed graph)**.** A *directed graph* is the graph $G = (V, E)$ where all elements in $E$ are ordered pairs of vertices,

**Definition 1.1.3** (Walk)**.** A *walk* is a sequence of vertices $v_0, v_1, v_2, \cdots, v_k$ in a graph $G = (V, E)$, where $\{(v_{i-1}, v_i) \mid 1 \leq i \leq k\} \in E$ holds,

**Definition 1.1.4** (Closed walk)**.** A *closed walk* is the walk whose starting and ending vertices are the same,

[18] https://www.internetlivestats.com/total-number-of-websites/#sources, accessed July 23, 2019.
[19] See Biggs et al. (1986), West et al. (1996) for example, regarding the history and outline of graph theory.

**Definition 1.1.5** (Cycle)**.** A *cycle* is the closed walk whose vertices are all distinct except the starting and ending ones,

**Definition 1.1.6** (Directed acyclic graph)**.** A *directed acyclic graph* (DAG) is the directed graph which does not contain any cycles.

Citations as the DAG are accordingly $G = (V, E)$[20], where all vertices (intellectual products) in $V$ have no cycles, and all edges (citation relationships) in $E$ have the citing-to-cited directions[21][22]. This structure reflects the property that scientific publications and patents can only cite existing ones (i.e., all edges have new-to-old directions). Note that web pages can cite subsequent (newer) web pages, which means that their citations are not the DAG but a directed graph. Nevertheless, the DAG is still a good approximation because we can convert citations on web pages into the DAG by removing all old-to-new edges from cycles[23].

To express the growth of citations, we further assume that DAGs have a given $q$ number of state transitions $(G_t)_{t=0}^{q} = (G_0, G_1, \cdots, G_q)$ that successively attaches one new vertex (with directed edges to existing vertices) in each period[24]. Figure 1.3 depicts an example of the one-step state transition in a simple DAG, which is from $G_t$ (Figure 1.3a) to $G_{t+1}$ (Figure 1.3b). Since $G_{t+1}$ receives a new vertex $\{6\}$ with the citation to existing vertices $\{3, 4\}$, the state transition can be described as $V_{t+1} = V_t \cup \{6\}$ and $E_{t+1} = E_t \cup \{(6, 3), (6, 4)\}$. This thesis will consistently deal with citations formalized as such the growing DAG.

Graphs are helpful in the citation analysis as we can quantify their entire structure in the form of $|V| \times |V|$ matrix with each vertex arranged in rows and columns[25]. Among a variety of matrices, this thesis only uses *adjacency matrix* $\boldsymbol{A} = (a_{ij})$ and

[20] In the case of Figure 1.3a, since $G_t = (V_t, E_t)$ has 5 nodes and 7 edges, $V_t = \{1, 2, 3, 4, 5\}$ and $E_t = \{(2, 1), (3, 1), (4, 1), (4, 2), (4, 3), (5, 2), (5, 4)\}$ hold.

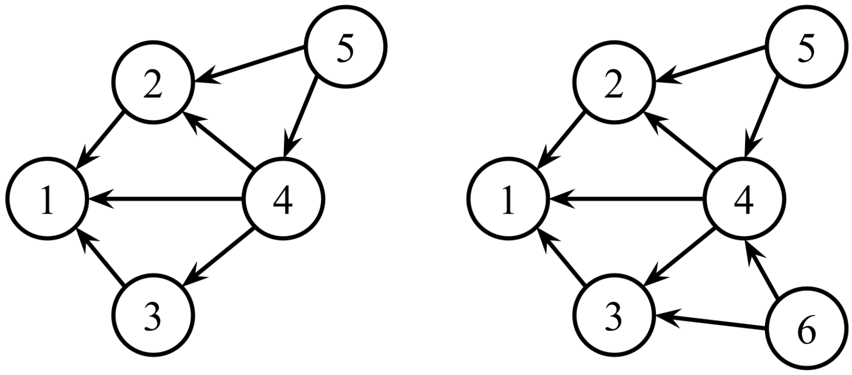
[21] Edges with directions are called as *directed edges* or *arcs* or *arrows*.

[22] Here, $(v_a, v_b) \in E$ designates that an intellectual product $v_a$ cites another product $v_b$.

[23] This will not fundamentally change the graph structure, given that the main part of citations is made at the same time with the creation of web pages, and about 85% of web pages are not acting according to the survey by *Netcraft* (https://news.netcraft.com/archives/category/web-server-survey/, accessed August 29, 2019).

[24] For simplicity, we assume that the initial state $G_0 = (V_0, E_0)$ has a sufficient number of vertices and edges.

[25] In most cases (including this thesis), when the graph is growing, its row and column arrange each vertex in the order of attachment.




(a) $G_t = (V_t, E_t)$ (b) $G_{t+1} = (V_{t+1}, E_{t+1})$

**FIGURE 1.3: Growing DAG for the Process of Citations.** We can approximate the network structure of citations as the DAG: $G = (V, E)$, where $V = \{1, 2, \cdots\}$ denotes intellectual products as the set of vertices and $E \subseteq V \times V$ denotes their citation relationships as the set of edges. Furthermore, to express the growth of citations, this thesis introduces the state transition that successively attaches one new vertex to the DAG in each period. For example, $V_{t+1} = V_t \cup \{6\}$ and $E_{t+1} = E_t \cup \{(6,3),(6,4)\}$ hold in the state transition from $G_t$ in Figure 1.3a to $G_{t+1}$ in Figure 1.3b.

*probability matrix*[26] $\boldsymbol{P} = (p_{ij})$—each represents $G_t$ in Figure 1.3a as follows:

$$\boldsymbol{A}(G_t) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \quad \boldsymbol{P}(G_t) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} \frac{1}{5} & 1 & 1 & \frac{1}{3} & 0 \\ \frac{1}{5} & 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{5} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{5} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{5} & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{1.1}$$

where $a_{ij}$ designates the existence of an edge directed from vertex $j$ to vertex $i$ with binary values $\{0, 1\}$; $p_{ij}$ designates the probability of Markov-chain transition from vertex $j$ to vertex $i$[27]. Such matrix representations have contributed to the quantitative measures (e.g., the PR algorithm), which can evaluate intellectual products considering the entire structure of citations.

Thus, citations are important because their structure as the growing DAG has the matrix representations that can facilitate various quantitative measures for intellectual products.

[26] Probability matrix is also called as *stochastic matrix* or *transition matrix*.

[27] Note that, since $\boldsymbol{P}$ requires $\sum_i p_{ij} = 1$ by definition, $\boldsymbol{P}(G_t)$ assumes that the Markov-chain transition jumps to one of the existing vertices uniformly at random when it reaches a vertex with no out-edges (i.e., vertex 1). This is often referred to as *stochasticity adjustment*.

In this section, we confirmed the importance of citations from three perspectives: as a source of quantitative measures, a growing network, and a graph structure. To summarize, citations are important because their network structure as the growing DAG can be a source of quantitative measures to evaluate ever-increasing intellectual products.

## 1.2 The Importance of Citations in a P2P System

For these important citations, which determines their true network structure in the first place, and how? This thesis considers such *consensus-building on citations*. To prevent misbehavior (see Section 1.3.1 for specific examples), consensus-building on citations has so far relied on some centralized authority, such as an editorial board to review (or assign peer-reviewers to) scientific publications, a patent examiner to evaluate submitted inventions through prior art searches, and a search engine to rank a large number of web pages[28]. However, we cannot dismiss the possibility that the centralized authority itself will misbehave (e.g., manipulate citation relationships or reject worth-registering intellectual products) for its benefit; in other words, consensus-building on citations has the question of *who watches the watchmen*. This is the reason why we need citations in a P2P system[29]:

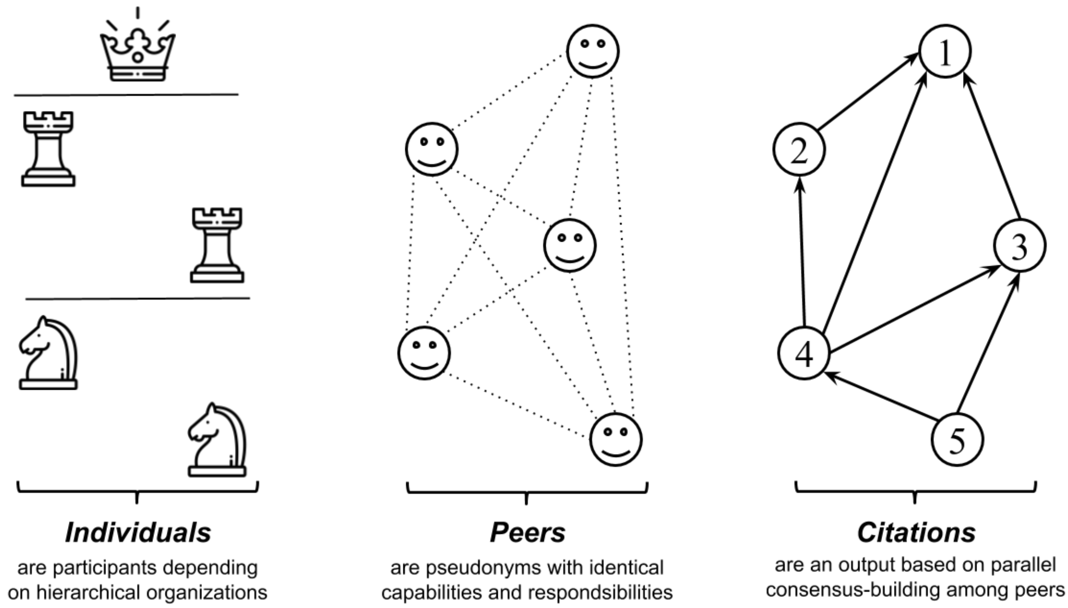
**Definition 1.2.1** (P2P system)**.** A *P2P system* is a self-organizing system consisting of autonomous peers with identical capabilities and responsibilities, which depends on neither centralized control nor hierarchical organization[30],

where peers are a kind of pseudonym for individuals participating in the P2P system. We can extend Definition 1.2.1 to our main subject as follows:

[28] Unlike scientific publications and patents, WWW allows anyone to cite other web pages without authoritative review; however, it still relies on centralized search engines for consensus-building on the validity of citation relationships. This thesis proposes two incentive mechanisms because consensus-building can occur either ex-ante or ex-post citations (the former corresponds to scientific publications and patents, and the latter corresponds to web pages). See Section 2.1.1 for its details.

[29] The term P2P system has a range of definitions because of interdisciplinary discussions from computer science to socio-economics. Here, Definition 1.2.1 is based on definitions by Haase et al. (2008), Rowstron and Druschel (2001), Steinmetz and Wehrle (2005).

[30] According to the traditional classification by Schollmeier (2001), Definition 1.2.1 denotes *Pure Peer-to-Peer system* in that it does not allow any central entities. See Androutsellis-Theotokis and Spinellis (2004), Steinmetz and Wehrle (2005), Wang and Sun (2008) and Koskela et al. (2013) for other (more detailed) classifications of P2P systems.




*Note*: chess pieces are designed by Freepik from Flaticon (https://www.flaticon.com/home, accessed January 14, 2020).

**FIGURE 1.4: Three Layers in a P2P Citation System.** P2P citation systems (Definition 1.2.2) assume three layers: individuals under hierarchical organizations, peers with identical capabilities and responsibilities, and citations $G_t = (V_t, E_t)$ as an output. Citations are constructed through parallel consensus-building among peers; e.g., one group of peers reviews $\{(4,1),(4,2),(4,3)\} \subset E_t$ while another reviews $\{(5,3),(5,4)\} \subset E_t$.

**Definition 1.2.2** (P2P citation system)**.** A *P2P citation system* is the P2P system that can construct citations through parallel consensus-building among peers.

Figure 1.4 illustrates three layers assumed in Definition 1.2.2: individuals under hierarchical organizations, peers with identical capabilities and responsibilities, and citations $G_t = (V_t, E_t)$ as an output[31]. Namely, this thesis considers citations where individuals as peers autonomously build consensus on the validity of each network structure (as a growing DAG) in parallel; e.g., in Figure 1.4, one group of peers reviews $\{(4,1),(4,2),(4,3)\} \subset E_t$ while another reviews $\{(5,3),(5,4)\} \subset E_t$[32]. In addition to the (aforementioned) independence of possibly misbehaved centralized authorities, this section introduces two considerable importance of such citations in a P2P system: robustness and productivity.

[31]For simplicity, this thesis consistently assumes that individuals-to-peers, peers-to-products ($V_t$), and thus individuals-to-products are all one-to-one correspondence. This will be introduced in Section 1.4.1, especially as Assumption 1.4.1 and Figure 1.6.

[32]Note that the scope of the review (and consensus-building) differs between the two incentive mechanisms this thesis proposes. The incentive mechanism with ex-ante consensus, mainly for scientific publications and patents, reviews each subgraph, e.g., $\{\{3,4,5\},\{(5,3),(5,4)\}\}$. On the other hand, the incentive mechanism with ex-post consensus, mainly for web pages, reviews each out-edge, e.g., $(5,3)$. See Chapters 3 and 4 for details.

### 1.2.1 P2P Systems for Robustness

From the perspective of computer science, P2P systems can provide citations with robustness[33] against random failures and attacks. This is because P2P systems, consisting of peers with identical capabilities and responsibilities, have no *single point of failure* (SPoF) where a malfunction leads to a failure of the entire system.

Historically, research on systems with no SPoF has been active since the court-ordered shutdown of *Napster* (Napster, 1999, 2000) in 2001[34]. Despite its reputation as the first popular P2P file-sharing system, Napster used a centralized server for file searching (even though not used for file transfer), which allowed a U.S. court to shut down the entire system just by interfering with the server. This SPoF, exposed by the shutdown, encouraged subsequent systems, e.g., *Freenet* (Clarke et al., 2002; Clarke et al., 2001) and *Gnutella 0.4* (Specification, 2003), to develop an alternative *flooding search* (e.g., Lv et al., 2002) that relies not on centralized servers but on query propagation among peers. File searching with no SPoF continues to develop new methods for efficiency, such as propagating queries only to those neighboring peers with high potential (Crespo & Garcia-Molina, 2002; Tsoumakos & Roussopoulos, 2003), and having some peers maintain a portion of index on file locations as *distributed hash tables* (Ratnasamy et al., 2001; Rowstron & Druschel, 2001; Stoica et al., 2001; Zhao et al., 2004)[35]. In this way, P2P systems have enhanced their robustness by delegating critical data processing (e.g., file searching) from SPoF to distributed peers. These precedents imply that citations would be robust against the failure of centralized authorities (e.g., editorial boards, patent examiners, and search engines) if P2P systems let distributed peers do consensus-building (as a critical data processing) on the validity of citation relationships[36].

[33]In this thesis, robustness follows the definition by *IEEE standard glossary of software engineering terminology*: "The degree to which a system or component can function correctly in the presence of invalid inputs or stressful environmental conditions" (IEEE et al., 1990, p. 64).

[34]See A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001), e.g., https://www.copyright.gov/fair-use/summaries/a&mrecords-napster-9thcir2001.pdf, accessed October 24, 2019.

[35]P2P systems based on the distributed hash table are often distinguished from other (unstructured) P2P systems as *structured P2P systems* because they impose constraints on the topology of overlay network for selecting neighboring peers.

[36]In addition to robustness, P2P citation systems, where the burden of reviewing proposed citations is not concentrated on a centralized entity, can also contribute to *scalability* that means "the ability of a system to accommodate an increasing number of elements or objects, to process growing volumes of work gracefully, and/or to be susceptible to enlargement" (Bondi, 2000, p. 1).

Moreover, for P2P systems involving record-keeping, SPoF is in data processing and data storage. To avoid failures caused by centralized data storage, such systems have distributed peers who share the same record of state transitions. A representative example of this approach is *the Bitcoin protocol* (Nakamoto et al., 2008) that aims for a P2P electronic cash system. Specifically, it eliminates the SPoF in existing cash systems by letting peers share all transaction records of bitcoin with a format later called *blockchain*[37], which also plays an important role in consensus-building on the validity of transaction records (as will be mentioned in Section 1.3.2). The Bitcoin protocol has spawned a variety of alternative protocols, such as *Ethereum* (Buterin et al., 2014; Wood et al., 2014) that aims for application development on P2P systems, by generalizing the handling information from transaction records (of bitcoin) to triggers for executing specific programs[38]. These blockchain-based precedents imply that citations would be further robust if P2P systems let distributed peers share the record of state transitions $(G_t)_{t=0}^{q}$ after the consensus-building on their validity.

Thus, citations in a P2P system are important because they become more robust against random failures and attacks by eliminating SPoFs in data processing or data storage.

### 1.2.2 P2P Systems for Productivity

From the perspective of socio-economics, P2P systems can provide citations with productivity (i.e., an efficient process to construct a growing DAG). This is because P2P citation systems, assuming parallel consensus-building, can be interpreted as extended *crowd-sourcing*—one of the preceding concepts for productivity.

With the widespread adoption of Napster, *Wikipedia*[39], and *Open Source Software* (OSS) *Development* (Feller, Fitzgerald, et al., 2002), P2P systems have become a subject of socio-economic studies as well (e.g., Bauwens, 2005a, 2005b). These studies tend to focus on individuals (as peers) with no hierarchical organization rather than

[37] Strictly speaking, in the current Bitcoin protocol, only peers classified as full nodes have all transaction records. Despite this heterogeneity of peers, this thesis regards the Bitcoin protocol as a P2P system in Definition 1.2.1, as this protocol permits anyone to build full nodes.

[38] Applications and programs developed on Ethereum are often referred to as *Decentralized Applications* (DApps; e.g., Raval, 2016) and *smart contracts* (e.g., Hileman & Rauchs, 2017; Szabo, 1997), respectively.

[39] https://en.wikipedia.org/wiki/Wikipedia, accessed December 29, 2019.

on system design with no SPoF[40] and particularly have attempted to characterize a new type of productivity stemming from their parallel, cooperative behavior. Starting with *peer-production* (Benkler, 2002a, 2006), defined as "a process by which many individuals, whose actions are coordinated neither by managers nor by price signals in the market, contribute to a joint effort that effectively produces a unit of information or culture" (Benkler, 2002b, p. 1256)[41], this attempt produced a variety of similar (or interchangeable) concepts, such as *parallel development* (Feller, Fitzgerald, et al., 2002), *open innovation* (Chesbrough, 2003), *mass collaboration* (e.g., Elliott, 2007; Tapscott & Williams, 2008), and *hyper-productivity* (Bauwens, 2009).

Citations for the above three intellectual products have, in this context, increased the productivity of their construction primarily through one of these concepts, *crowdsourcing* (Howe, 2006a)[42]. In scientific publications, for example, *crowd-sourced review*—"a public review process in which any community member may contribute to the article review" (Ford, 2013, p. 315)[43]—has been adopted in various online systems including journals (Pöschl, 2004)[44], conference managements (Soergel et al., 2013)[45], and pre-print servers (Berthaud et al., 2014)[46]. In addition to the main purpose of shortening the time between submission and publication, this may also have a positive effect on the quality of reviews and articles (Bornmann et al., 2011; Fitzpatrick, 2010; Prug, 2010). In patents, several organizations, such as *Peer To Patent*[47]

[40]Accordingly, P2P systems discussed in socio-economic studies do not always eliminate SPoFs (as a web application). Wikipedia, for instance, depends on a centralized server. However, its open, editable property is often introduced as an example of social P2P processes.

[41]See also another following definition: "production systems that depend on individual action that is self-selected and decentralized, rather than hierarchically assigned" (Benkler, 2006, p. 62).

[42]Specifically, the original definition is as follows:

> Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the wide network of potential laborers. (Howe, 2006a, para. 3)

See Estellés-Arolas and González-Ladrón-De-Guevara (2012), Howe (2006b), Schenk and Guittard (2011) for other definitions and discussions on crowd-sourcing.

[43]Ford (2013) introduces crowd-sourced review as one of the categories in *Open Peer Review (OPR)*, a broader concept that includes even a simple non-blind style in which the editorial board assigns peer reviewers.

[44]*Atmospheric Chemistry and Physics* (ACP; https://www.atmospheric-chemistry-and-physics.net/index.html, accessed November 15, 2019). Note that ACP was founded in September 2001, while Pöschl (2004) described its review process in the form of academic article.

[45]*OpenReview.net* (https://openreview.net/, accessed November 15, 2019).

[46]*Episciences* (https://www.episciences.org/, accessed November 15, 2019).

[47]https://www.peertopatent.org/, accessed November 15, 2019.

(Bestor & Hamp, 2010; Noveck, 2006) and *Article One Partners* (Malone, 2011), have adopted crowdsourcing to efficiently find prior arts for submitted inventions. These attempts to lessen the burden of patent examiners are referred to as *crowd-sourced prior art search* (Ghafele et al., 2011). Finally, in web pages, crowdsourcing has contributed to a number of related systems (Doan et al., 2011; Yuen et al., 2011), especially as *crowd-sourced human-based computing* (Wightman, 2010). Even when we focus on search engines, crowdsourcing has enhanced their information retrieval as a part of many features, such as relevance assessment (Alonso et al., 2008; Grady & Lease, 2010), spam detection (McCreadie et al., 2012), and personalization based on social networks (Bozzon et al., 2012)[48].

One of the limitations of these precedents is that centralized authority remains in their consensus-building. Crowd-sourcing only facilitates centralized authorities (editorial boards, patent examiners, search engines) to evaluate intellectual products. Even other concepts often require some authority (e.g., administrator, project manager) to build consensus within a community[49] (Kreiss et al., 2011; O'Neil, 2014). Centralized authorities compromise not only the system's robustness but also its productivity since they prevent parallel consensus-building[50]. This implies citations would be more productive if P2P systems, as extended crowd-sourcing, realized parallel consensus-building among distributed peers.

Thus, citations in a P2P system are important because their construction becomes more productive by delegating consensus-building on the validity of citation relationships from centralized authorities to distributed peers.

In this section, we confirmed the importance of citations in a P2P system. To summarize, citations in a P2P system are important because they are independent of possibly misbehaved centralized authorities and can obtain robustness by no SPoF and productivity by parallel consensus-building.

[48] There are also *open-source search engines* (e.g., Middleton & Baeza-Yates, 2007; Trotman et al., 2012) in which the development itself follows crowdsourcing (or other concepts such as peer-production).

[49] As an example, it is often mentioned that Wikipedia is managed by a hierarchical organization consisting of *stewards*, *bureaucrats*, *administrators*, etc. (Butler et al., 2008; O'Neil, 2011).

[50] Again, this claim relates to scalability issues in that the burden of consensus-building is concentrated on a small number of centralized entities.

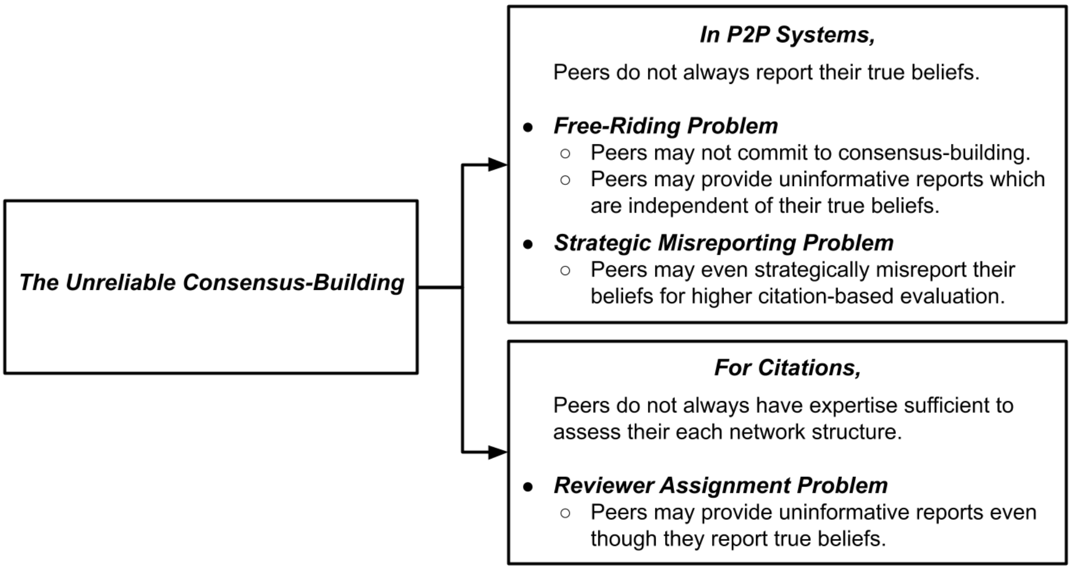



FIGURE 1.5: **Three Problems behind the Unreliable Consensus-Building.** This thesis focuses on the unreliable consensus-building that stems from three problems: free-riding and strategic misreporting in P2P systems and reviewer assignment in citations. To our knowledge, preceding studies have not addressed the three problems simultaneously.

## 1.3 Problem Statement

Despite this importance, however, citations in a P2P system are difficult because their consensus-building has remained unreliable. In this section, the author first introduces the reasons for such *unreliable consensus-building* in P2P citation systems. Then he shows addressing this problem has academic contributions to consensus-building in P2P systems, reviewer assignment problems, and network formation.

### 1.3.1 The Unreliable Consensus-Building

As Figure 1.5 depicts, the unreliable consensus-building stems from three problems: *free-riding*, *strategic misreporting*, and *reviewer assignment*, with the first two problems attributed to P2P systems, and the last to citations.

First and foremost, in P2P systems, peers do not always report their true beliefs. Given the time and effort to assess the validity of each citation relationship, peers may not commit to consensus-building in the first place, or they may provide uninformative reports which are independent of true beliefs (e.g., automatically sending the same report). Such behavior of "an individual user who uses the system resources without contributing anything to the system" (Ramaswamy & Liu, 2003,

p. 1) is common in P2P systems as a free-riding problem[51]. To make matters worse, peers may even strategically misreport their beliefs. An intuitive example would be *link spamming* (e.g., Henzinger et al., 2002), where web pages strategically manipulate their links to improve the result of search engines with link-based ranking algorithms (e.g., PR and HITS)[52]. Namely, peers, who each desire a higher citation-based evaluation, may strategically misreport citation relationships whenever they commit to consensus-building or register new intellectual products[53]. This problem is severe in P2P systems because, unlike search engines for web pages, they cannot rely on a centralized authority to keep the details of their ranking (or consensus) algorithm secret. In this way, peers do not always report their true beliefs under P2P systems due to the possibility of free-riding and strategic misreporting.

Furthermore, for citations, peers do not always have sufficient expertise to assess each network structure. Citations, especially those for scientific publications and patents, represent the relationship of technical knowledge based on expertise in specific subjects, which is another reason why their assessment has relied on some centralized authority (e.g., editorial boards, patent examiners) as a group of experts. This means that even if all peers report their true beliefs, consensus-building may be untrustworthy unless they have sufficient expertise corresponding to the assigned part of citations (i.e., truthful reports from randomly selected peers are not always informative reports). To solve the unreliable consensus-building, we, therefore, need to assign appropriate peers (as reviewers) by taking their expertise into account before eliciting their true beliefs.

Thus, consensus-building is unreliable because peers do not always report their true beliefs (i.e., free-riding and strategic misreporting in P2P systems) or have sufficient expertise (i.e., reviewer assignment in citations).

---

[51] System resources are intellectual products and their citations. Note that, for P2P citation systems, free-riding exists not only in consensus-building but also in registration, i.e., peers may not register new intellectual products (and their citations) without some incentive. Thus, this thesis considers the reward mechanism which can encourage both reliable consensus-building and new registration.

[52] Link spamming is performed mainly through *link farms* (Wu & Davison, 2005) that artificially produces several new referring pages; however, as will be discussed in Section 1.4.1, this thesis assumes that individuals-to-peers, peers-to-products ($V_t$), and thus individuals-to-products are all one-to-one correspondence.

[53] Such strategic behavior also exists in other two intellectual products. For scientific publications, researchers may submit a number of incomplete papers for a higher citation-based index (Hernández-Alvarez & Gomez, 2016), and peer-reviewers may skew their opinion due to the *conflicts-of-interest* (CoI; Resnik & Elmore, 2018). Also, for patents, applicants may withhold their citations to broaden the scope of patent rights (Lampe, 2012).

### 1.3.2 State of the Art: Consensus-Building in P2P systems

To our best knowledge, preceding studies have not addressed free-riding, strategic misreporting, and reviewer assignment simultaneously. On the other hand, several research topics cover a part of the three problems, which allows us to present academic contributions of addressing unreliable consensus-building from existing research topics.

First of all, the attempt to incorporate expertise has an academic contribution to consensus-building in P2P systems.

Discussions on this research topic originate from P2P file-sharing systems. Ever since Adar and Huberman (2000), Saroiu et al. (2001) pointed out the rampant free-riding in Napster and Gnutella[54], P2P file-sharing systems have induced peers to upload their files through some incentive mechanism based on *rewards* and *game-theoretic concepts*. For example, *Mojo Nation* (Wilcox-O'Hearn, 2002) rewards peers who have uploaded files with an internal currency[55], while *KaZaA* (Leibowitz et al., 2003) rewards them with a higher reputation which facilitates file downloads[56]; moreover, to prevent the strategic misreporting, *BitTorrent* (Cohen, 2003) adopts the game-theoretic tit-for-tat strategy to its file-sharing process[57].

The Bitcoin protocol is the first practical one that applied such the incentive mechanism (based on rewards and game-theoretic concepts) to consensus-building in P2P systems[58]. Specifically, it enables consensus-building on transaction records among strategically distributed peers—a long-standing challenge for P2P electronic cash systems—with the following main rules: (i) transaction records of bitcoin are sequentially stored in blocks, and peers share the identical chain of blocks as a result of consensus-building (blockchain), (ii) peers can create a new block and connect it to

[54]Free-riding in P2P file-sharing systems means that peers do not upload their files but only download files uploaded by other peers.

[55]Strictly speaking, original article says "Mojo Nation is not a file-sharing system (like Gnutella or Napster), but a file store, in which the storage, transfer, and naming of files is performed in a distributed manner, independent of any individual node" (Wilcox-O'Hearn, 2002, p. 1).

[56]See also some survey papers, such as Gupta et al. (2003), Karakaya et al. (2009), Rahman (2009).

[57]Studies on the incentive mechanism for P2P file-sharing systems are categorized as *algorithmic mechanism design* (AMD; Nisan & Ronen, 2001; Nisan et al., 2007), or as *distributed algorithmic mechanism design* (DAMD; Feigenbaum et al., 2000; Feigenbaum & Shenker, 2004) when emphasizing their distributed aspects. Both of these categories use game-theoretic concepts to address several problems in communication systems, such as routing and load balancing.

[58]Consensus-building in P2P systems is often studied in the framework of *Byzantine Generals Problem* (Lamport et al., 1982; Pease et al., 1980) which considers the possibility of reliable consensus in communication systems where some components may send conflicting information due to malfunction.

any block in the existing chain, but this task requires computational resources (*proof-of-work*; Dwork & Naor, 1992; Jakobsson & Juels, 1999), (iii) if the chain branches to more than one path, the longest chain is considered to be the consensus (*Nakamoto consensus*), (iv) peers who create a block in the longest chain will be rewarded with newly issued bitcoins (rewards for contributors). See Liu et al. (2019) for several game-theoretic analyses on this Bitcoin protocol. Furthermore, subsequent applications, mostly developed on Ethereum, try to extend the scope of consensus-building to data outside the blockchain[59]. For example, *Augur* (Peterson et al., 2015), *Gnosis* (Gnosis, 2017), and *Stox* (Stox, 2017) are platforms for *decentralized prediction markets*, in which peers do consensus-building even on the actual outcome of predicted subjects (e.g., weather in a given location, election results); *AdChain registry* (Goldin et al., 2017) and *Ocean Protocol* (Ocean Protocol, 2019) are platforms for the *Token-Curated Registry* (TCR; Goldin, 2017a, 2017b), in which peers curate a high-quality, reliable list of any content (e.g., restaurants, universities, web pages) as a decentralized recommender system. Although these applications have different design patterns (Ito, 2018; Lockyer, 2018), their consensus-building generally uses the following *token-staking scheme*[60]: (i) peers can stake their reward tokens on a binary choice {accept, reject} before putting new data in the blockchain, (ii) consensus is the selection that obtains more tokens compared to another selection after a certain period, (iii) all staked tokens are redistributed among peers who stake their tokens on the consensus side. See Asgaonkar and Krishnamachari (2018), Falk and Tsoukalas (2018), Wang and Krishnamachari (2018), for game-theoretic analyses on this token-staking scheme[61].

For this research topic, the problem of unreliable consensus-building is a new attempt to add citations to the incentive mechanism, thereby incorporating the expertise of distributed peers into consensus-building in P2P systems. This is worth addressing because it can further extend the scope of consensus-building to technical content (e.g., scientific publications, patents). Our proposal, the two incentive

[59]The system resulting from this extension is often referred to as *decentralized oracle* or *consensus-based oracle*. See Voshmgir (2019) for the term *oracle* in the context of blockchain.

[60]We will confirm limitations of the token-staking scheme in Section 2.1.2.

[61]The series of studies introduced in this paragraph, focusing on the incentive mechanism for consensus-building in P2P systems, is nowadays often referred to as *cryptoeconomics* (Buterin, 2015; Davidson et al., 2016; Zamfir, 2015) or *token economy* (Voshmgir, 2019).

mechanisms detailed in Chapters 3 and 4, selects appropriate peers according to their citation-based expertise, rather than using the proof-of-work or token-staking scheme that accepts any peer with a certain amount of computational resources or reward tokens. Thus, addressing this problem has an academic contribution in that it incorporates expertise into the consensus-building in P2P systems.

### 1.3.3 State of the Art: Reviewer Assignment Problem

On the contrary, if we take the unreliable consensus-building not from P2P systems but from citations, the attempt to cover strategic peers has an academic contribution to the research topic called the *Reviewer Assignment Problem* (RAP).

RAP originates from Dumais and Nielsen (1992), which proposed "automated means of assigning the submitted manuscripts to appropriate members of the review committee" (Dumais & Nielsen, 1992, p.1). To achieve automated paper-to-reviewer assignment, the RAP considers a system that can quantitatively measure the expertise of reviewer candidates (mostly authors of other papers) with a variety of data from submitted documents, such as authors, bibliographies, abstracts, and keywords[62]. (Wang et al., 2008; Wang et al., 2010). This is particularly effective for academic conferences where the organizing chair needs to assign a large number of submitted papers to reviewers; indeed, several systems, e.g., *GRAPE* (Di Mauro et al., 2005) and *Toronto paper matching system* (Charlin & Zemel, 2013), have been used to lessen the burden of conference management.

In this context, the assignment of peers for consensus-building in P2P citation systems is a RAP that focuses on citation analysis under one-to-one correspondence between papers and reviewers (see Assumption 1.4.1)[63]. Although RAPs tend to use multi-disciplinary methods (e.g., text mining, artificial intelligence[64]), citation analysis remains one of the most important methods for finding experts on a given paper. For example, Yarowsky and Florian (1999) counts how much one author

[62] In addition to the expertise, recent studies (e.g., Ahmed et al., 2017; Liu et al., 2014; Long et al., 2013; Tang et al., 2012) take into account the diversity of reviewers to prevent the biased-review due to CoI. We will mention this issue again in Section 5.2.3.

[63] Because of the assumption of one-to-one correspondence, such simplified RAP is also relevant to the research topic of *research-paper recommender systems* (Beel et al., 2016) that considers user-to-paper assignment rather than paper-to-reviewer assignment. In particular, Gori and Pucci (2006) is related to the mechanism this thesis proposes in Chapter 3, in that it uses the PPR algorithm for the recommendation.

[64] For example, see Kolasa and Krol (2011).

cites the other author's paper to measure the relevance of two authors; Küçüktunç et al. (2012) uses *Personalized PageRank* (PPR; Haveliwala, 2002)—an algorithm this thesis also employs—for citation networks, to find papers that are both relevant and important for a submitted paper[65]; Li and Watanabe (2013) incorporates the method of co-citation analysis (Small, 1973) into its measurement[66].

For this research topic, the problem of unreliable consensus-building is a new attempt to extend the RAP to a group of strategic peers. This is worth addressing because it can provide the RAP with robustness and productivity as a P2P system. To our knowledge, Xu et al. (2019) is the closest to this subject, but their game-theoretic model does not take into account free-riding, as it is not intended for operation on P2P systems. Thus, addressing this problem has an academic contribution in that it extends the RAP to a group of strategic peers who may do free-riding and misreporting.

### 1.3.4 State of the Art: Network Formation

Moreover, if we consider citations as one of the social networks[67], the attempt to construct a growing network among strategic peers has an academic contribution to researches on network formation.

Network formation is a research topic that models the emergence of social networks, such as the friendship between individuals, political alliances between nations, and (of course) citations between documents (e.g., De Paula, 2020; Jackson, 2005a, 2010). Jackson (2005b) classified this research topic into two approaches. One approach is to model *How* a certain type of social network appears, assuming that vertices randomly form edges according to a predetermined probability distribution. Despite the original model with a fixed number of vertices (Erdős & Rényi, 1959, 1960, 1961), the How approach focuses on the growing networks (Section

[65]Note that, in Küçüktunç et al. (2012), the PPR algorithm is referred to as *random walk with re-start* (RWR) algorithm.

[66]Furthermore, when considering not only citations but also graph in general, we can find a number of graph-based studies on the RAP. For example, Liu et al. (2014), Rodriguez and Bollen (2006, 2008) use co-authorship graph, in which vertices and edges each represent authors (i.e., reviewer candidates) and their co-authorship; Watanabe et al. (2005) uses the graph with a preferential attachment (Barabási & Albert, 1999; Price, 1976) nature, in which vertices represent keywords of papers and reviewers (whose number is ever-increasing), and edges represent their co-occurrence.

[67]One of the definitions of the term social network is "a set of social relationships for which there is no common boundary" (Bott & Spillius, 2014, p. 59), which was originally proposed in Bott (1957). For a history of social network analysis, see Freeman (2004).

1.1.2)[68]. For example, Barabási and Albert (1999) proposed the preferential attachment model where each newborn vertex forms edges according to the degree distribution of existing vertices[69]; Kleinberg et al. (1999), Kumar et al. (2000) proposed another *copying* model where each newborn vertex forms edges to neighbours of a randomly-selected vertex. These growing but non-strategic (random) models have since improved their fitness with large empirical data (e.g., Leskovec et al., 2008; Leskovec et al., 2005), or addressed more specific subjects such as link prediction (Liben-Nowell & Kleinberg, 2007) and community detection (Parthasarathy et al., 2011)[70].

The other approach is to model *Why* a certain type of social network that appears, assuming that vertices strategically form edges to maximize their utility. Following the original game-theoretic model (Aumann & Myerson, 1988)[71], the Why approach focuses on the network formation as a game between a fixed number of vertices. For example, Jackson and Wolinsky (1996) proposed a non-cooperative game where each vertex can obtain utility from both direct and indirect connections with other vertices, but direct connection (edge) needs maintenance cost[72]; Bala and Goyal (2000), Watts (2001) each extended this game to a dynamic model, where the former allows vertices to form edges unilaterally (e.g., citation), and the latter does not allow edges to form without the agreement between two vertices (e.g., matching). These strategic (game-theoretic) but non-growing models have since normatively described network stability and efficiency or addressed more specific topics such as PageRank games (Hopcroft & Sheldon, 2008) among web pages.

For this research topic, the problem of unreliable consensus-building is a new attempt to model growing and strategic networks, thereby bridge How and Why approaches. This is worth addressing because the two approaches are fragmented

[68] Even today, the How approach covers models with a fixed number of vertices, especially in the framework of *exponential random graph models* (e.g., Robins et al., 2007).

[69] Note that Price (1976) and Barabási and Albert (1999) are different in that the former assumes directed graph while the latter assumes undirected graph. See survey papers (e.g., Albert & Barabási, 2002; Boccaletti et al., 2006) for further studies based on the preferential attachment.

[70] See also Aggarwal and Subbian (2014).

[71] This study follows Myerson (1977) which is the first study to incorporate graph-theory into the discussion of resource allocation under the cooperative game.

[72] Specifically, Jackson and Wolinsky (1996) presented two models, which they called *the connections model* and *the co-author model*, respectively.

(Jackson, 2005b, 2010)[73], even though many social networks (including citations) have both growing and strategic aspects. To our knowledge, Avin et al. (2018) is the closest to this subject, but their game-theoretic model focuses only on the preferential attachment without incentive mechanisms. Thus, addressing this problem has an academic contribution in that it bridges How and Why approaches on network formation through incentive mechanisms for the growing and strategic network.

In this section, we confirmed the difficulty of citations in a P2P system. To summarize, citations in a P2P system are difficult because their consensus-building has remained unreliable due to three problems that preceding studies have not addressed simultaneously: free-riding, strategic misreporting, and reviewer assignment. Addressing this problem has academic contributions in that it (i) incorporates expertise into the consensus-building in P2P systems, (ii) extends the RAP to a group of strategic peers, (iii) bridges How and Why approaches on network formation.

## 1.4 Research Question

Then, can we solve this problem by some incentive mechanism, as preceding studies on consensus-building in P2P systems have done? This is what this thesis specifically examines as the RQ.

**Definition 1.4.1** (Research question)**.** The *Research Question (RQ)* of this thesis is: *Can we design some incentive mechanism to solve the unreliable consensus-building in P2P citation systems?*[74]

Sections 1.3.1 and 1.3.2 imply that Definition 1.4.1 is to address free-riding, strategic misreporting, and reviewer assignment simultaneously by using rewards and game-theoretic concepts. We propose two incentive mechanisms that provide game-theoretically computed rewards to peers who contribute to consensus-building as

[73] See pp. 9-10, p. 29 in Jackson (2005b); pp. 153-154, p. 459 in Jackson (2010). Note that, although it is a non-strategic model, Jackson and Rogers (2007) proposed a growing network consisting of agents (vertices) who can form edges to maximize their utility function.

[74] See Figure 1.6 for the term unreliable consensus-building and Definition 1.2.2 for the term P2P citation systems, respectively.

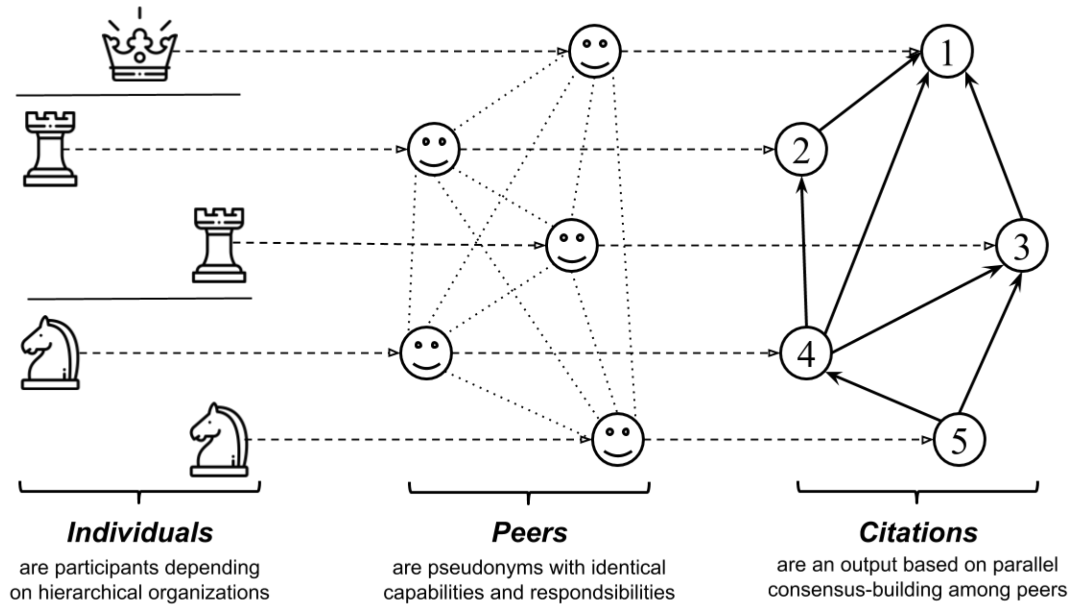



**FIGURE 1.6: One-to-One Correspondences Assumed in Three Layers.** This thesis assumes that individuals-to-peers, peers-to-products, and thus individuals-to-products are all one-to-one correspondence. This assumption frees incentive mechanisms from spamming and the Sybil attack, thereby allowing our discussion to focus on citations $G_t = (V_t, E_t)$.

reviewers (see Chapters 3 and 4 for details). This section first introduces three assumptions to examine the RQ and then confirms the RQ's academic and social contributions.

### 1.4.1 Assumptions of the RQ

As a first step to examine the RQ, this thesis sets three assumptions, where the first and second ones are to simplify the problem of unreliable consensus-building, and the third one is to define the goal of this thesis.

**Assumption 1.4.1** (One-to-one correspondence)**.** In P2P citation systems (Definition 1.2.2), individuals-to-peers, peers-to-products, and thus individuals-to-products are all one-to-one correspondence.

The first assumption concerns the structure of P2P citation systems. As Figure 1.6 depicts, Assumption 1.4.1 leads to a simplified environment in which all individuals—now synonymous with peers—can neither post more than one intellectual product nor share one intellectual product as a co-author. In other words, each individual (= peer) independently does one-shot registration. This setting frees incentive mechanisms from several problems, such as *spamming* (e.g., Hayati et al., 2010; Sahami et al., 1998) and a *Sybil attack* (Douceur, 2002)[75], thereby allowing our discussion to focus on citations $G_t = (V_t, E_t)$.

**Assumption 1.4.2** (Expected rewards as objective)**.** Peers aim to maximize the total amount of their expected rewards.

The second assumption concerns the objective of peers, which can free incentive mechanisms from the two issues required in practice. One is to make rewards sufficient incentives for peers. For example, the Bitcoin protocol sustains the value of bitcoin with several mechanisms, such as the proof-of-work to impose computational resources on the issue of new bitcoin, *difficulty adjustment* to stabilize in-flow of the computational resources, and *block-reward halving* to fix the total supply of bitcoin (Nakamoto et al., 2008). The other is to consider the incentive outside P2P systems, which remains to be discussed so much in the context of consensus-building in P2P systems. For example, the Bitcoin protocol is still at the risk of *Goldfinger attack* (Kroll et al., 2013) in which peers, even though they receive bitcoin as a reward, attempt to damage its value for incentives outside the protocol, such as short-selling or holding alternative assets. Assumption 1.4.2 allows our discussion to leave out these two issues.

Thus, Assumptions 1.4.1 and 1.4.2 simplify the unreliable consensus-building into a reward maximization problem that is closed within a given growing DAG. How to relax Assumptions 1.4.1 and 1.4.2 will be discussed in Chapter 5.

**Assumption 1.4.3** (Reliable consensus-building)**.** Consensus-building is reliable if peers can (i) be reviewers more often as they get higher PR scores and (ii) maximize the amount of expected rewards per review by always reporting true beliefs.

Finally, the third assumption defines the goal of this thesis by setting the conditions for reliable consensus-building. Note that Assumption 1.4.3 covers reviewer assignment in that condition (i) measures expertise of peers with PR scores, free-riding in that condition (ii) rewards reviewers, and strategic misreporting in that

[75] Spamming is "the act of spreading unsolicited and unrelated content" (Hayati et al., 2010, p. 1), and the Sybil attack is "the forging of multiple identities" (Douceur, 2002, p. 251). In the framework of P2P citation systems, spamming and the Sybil attack can be interpreted as (extreme) one-to-many correspondences in individual-to-products and individual-to-peers, respectively.

condition (ii) provides maximized expected rewards when reviewers always report their true beliefs, respectively[76]. Intuitively, peers (= individuals = intellectual products) should register intellectual products that will attract more citations in the future and always report their own and others' citation relationships truthfully. This thesis intends for the P2P citation system where peers autonomously build a reliable consensus through such an environment. Assumption 1.4.3 is directly related to the conclusion of this thesis: if our proposal—two incentive mechanisms in Chapters 3 and 4—satisfies conditions (i) and (ii), then the answer to the RQ will be *Yes*; otherwise, it will be *No*. We will confirm this through a combination of theory (Chapter 2) and experiments (Chapters 3 and 4).

These are the assumptions to examine the RQ. On the other hand, this thesis needs further assumptions regarding *peer prediction method* (Miller et al., 2005)—a game-theoretic method our incentive mechanisms employ for reward computation. The author will describe them in Chapter 2 and summarize all assumptions when answering the RQ in Chapter 6.

### 1.4.2 Contributions of the RQ

For academic contributions, we have already confirmed them in Sections 1.3.2–1.3.4 on the unreliable consensus-building; namely,

- incorporating expertise into the consensus-building in P2P systems,
- extending the RAP to a group of strategic peers,
- bridging How and Why approaches on network formation.

In addition, the RQ leads to the methodology that has the following other academic contributions:

- providing strong truthfulness for random walks on graphs,
- leveraging graphs to make peer prediction practical.

Chapter 2 will detail these academic contributions of methodology.

Furthermore, given potential applications of the two incentive mechanisms, the RQ has the following social contribution:

[76]We will define what a true belief is in Section 2.3.

- developing a new reward source for open-access intellectual products.

Chapter 5 will detail this social contribution of the RQ.

Thus, the RQ has academic and social contributions, with part of the former stemming from the methodology. These contributions are the background and motivation for the writing of this thesis.

In this section, we confirmed what the specific RQ of this thesis is. To summarize, the RQ is: *Can we design some incentive mechanism to solve the unreliable consensus-building in P2P citation systems?*, which includes both academic and social contributions (Section 1.4.2) under Assumptions 1.4.1–1.4.3.

## 1.5 Thesis Outline

To consider the RQ, this thesis consists of six chapters, including this Chapter 1. Figure 1.7 illustrates an outline of the chapters and their relationships, where Chapters 3 and 4 discuss two different incentive mechanisms.

Chapter 2 covers the methodology, which first highlights that this thesis proposes two incentive mechanisms—an *incentive mechanism with ex-ante consensus* (Chapter 3) for scientific publications and patents, and an *incentive mechanism with ex-post consensus* (Chapter 4) for web pages[77]. Despite algorithmic differences, both mechanisms combine random walks on graphs and peer prediction methods, where the former is specifically the PPR algorithm (Haveliwala, 2002) or *two-path mechanism* (Babichenko et al., 2018) and the latter is *multi-task peer prediction* (DG13; Dasgupta & Ghosh, 2013)[78]. We here review these three components, including the academic contribution of their combination. Furthermore, this chapter mentions experimental datasets retrieved from three real-world citations, corresponding to scientific publications (from arXiv), patents (from USPTO), and web pages (from Google).

Chapter 3 introduces the incentive mechanism with ex-ante consensus, which is mainly for citations on scientific publications and patents. This chapter is based on

[77]Note that since the only essential difference between the two incentive mechanisms is whether or not there is a case for rejection (see Chapters 3 and 4 for details), we can use these mechanisms for other intellectual products (in the extreme, we can use the former for web pages and the latter for scientific publications and patents).

[78]See Table 2.1.

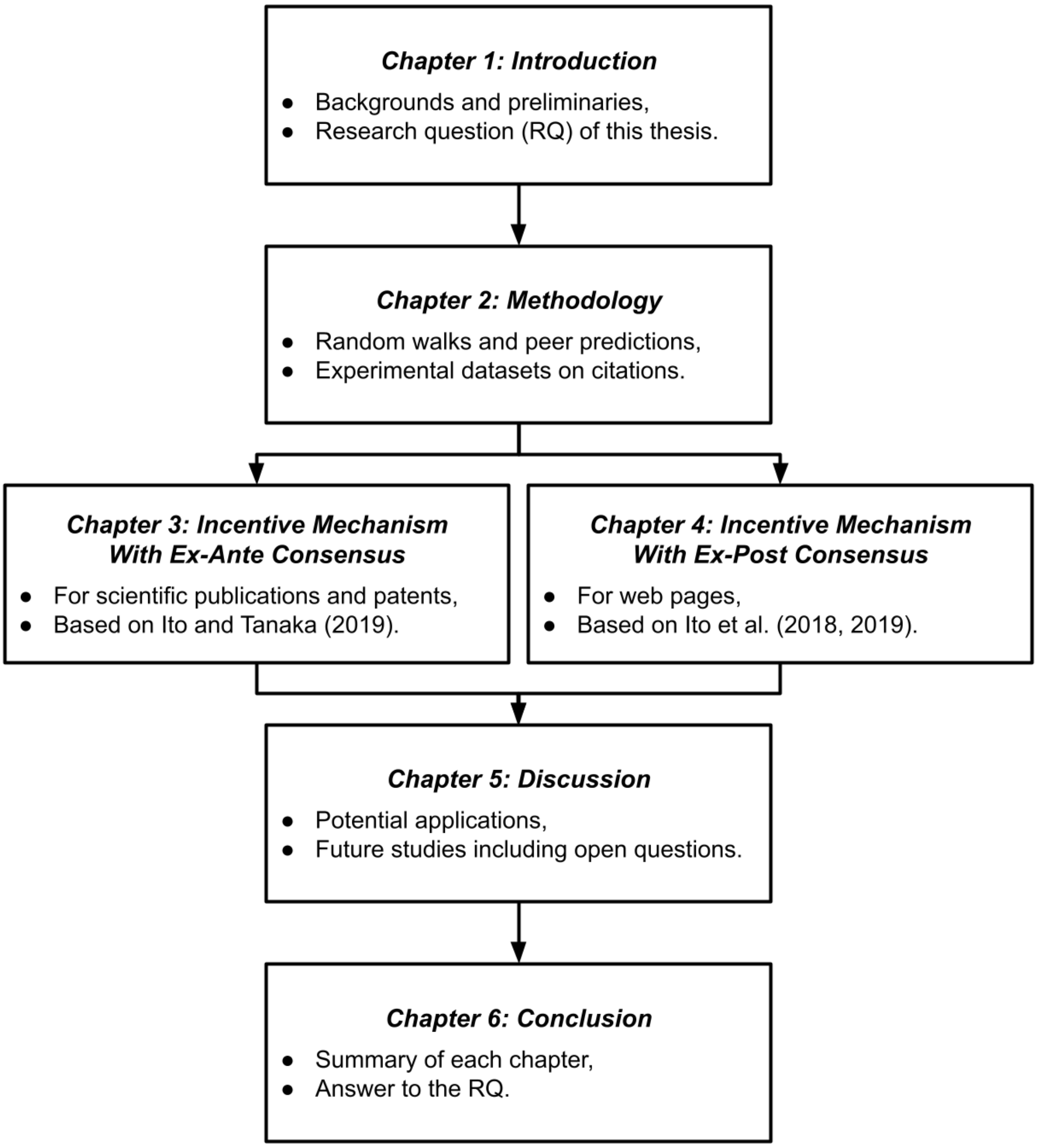



FIGURE 1.7: **Thesis Outline.** This thesis consists of six chapters including this Chapter 1. After describing methodology in Chapter 2, Chapters 3 and 4 each propose two different mechanisms: incentive mechanism with ex-ante consensus and incentive mechanism with ex-post consensus. Discussions on potential applications and future studies of these two mechanisms are together in Chapter 5. Finally, Chapter 6 concludes this thesis with the summary of each chapter and the answer to the RQ.

Ito and Tanaka (2019) that combines Haveliwala (2002) and Dasgupta and Ghosh (2013). Since scientific publications and patents are generally published after peer-review, the mechanism first builds consensus on the validity of intellectual products (and their citations) *ex-ante* then accepts only the peer-reviewed products into the system. Here, the PPR algorithm (Haveliwala, 2002) assigns appropriate reviewers (vertices) to a newly arrived intellectual product (vertex), and the multi-task peer prediction (Dasgupta & Ghosh, 2013) provides maximum expected rewards for reviewers who always report their true beliefs. Experiments confirm Assumption 1.4.3 using real-world datasets retrieved from arXiv and USPTO.

Chapter 4 introduces the incentive mechanism with ex-post consensus, which is mainly for citations on web pages. This chapter is based on Ito et al. (2018, 2019) that combines Babichenko et al. (2018) and Dasgupta and Ghosh (2013). Since web pages are generally published without peer-review, the mechanism first accepts all products coming into the system then builds consensus on the validity of intellectual products (and their citations) *ex-post*. Here, the two-path mechanism (Babichenko et al., 2018) randomly draws two paths on reviewers (vertices) to search an important intellectual product (vertex), and the multi-task peer prediction (Dasgupta & Ghosh, 2013) provides maximum expected rewards for reviewers who always report their true beliefs. Experiments confirm Assumption 1.4.3 using real-world datasets retrieved from Google.

Chapter 5 discusses potential applications and future studies of the two incentive mechanisms. The first half of this chapter shows that potential applications are (as Section 1.2 implies) to make crowdsourcing robust and productive, while introducing related systems and synergies with *Creative Commons* (CC; Lessig, 2004), *Semantic Web* (Berners-Lee et al., 2001), and *Linked Open Data* (Berners-Lee, 2006). We also confirm that potential applications lead to the social contribution of developing a new reward source for open-access intellectual products[79]. On the other hand, the second half of this chapter considers how to relax the strong Assumptions 1.4.1 and 1.4.2 as future studies, while introducing useful preceding studies (e.g., Goel et al., 2020a; Iwamura et al., 2019; Nakamoto et al., 2008). Future studies other than relaxing these assumptions are summarized as open questions.

Chapter 6 finally concludes this thesis with a summary of each chapter and the answer to the RQ. This thesis clarified that subject to several assumptions (including Assumptions 1.4.1–1.4.3), the answer to the RQ is *Yes*.

[79]See Chapter 5 for a definition of the term open-access.

## 1.6 Summary of This Chapter

This chapter introduced some backgrounds and preliminaries of this thesis which aims at consensus-building in P2P systems. We can summarize all arguments in Chapter 1 as answers to the questions presented at the beginning:

- *Why are citations important?* (1.1) — Citations are important because their structure as a growing DAG can be a source of quantitative measures to evaluate the ever-increasing intellectual products efficiently,
- *Why are citations in a P2P system important?* (1.2) — Citations in a P2P system are important because they are independent of possibly misbehaved centralized authorities and can obtain both robustness and productivity,
- *Why are citations in a P2P system difficult?* (1.3) — Citations in a P2P system are difficult because their consensus-building has remained unreliable due to three problems: free-riding, strategic misreporting, and reviewer assignment.
- *What is the Research Question (RQ) of this thesis?* (1.4) — The RQ of this thesis is: *Can we design some incentive mechanism to solve the unreliable consensus-building in P2P citation systems?*,
- *What are academic contributions of the RQ?* (1.3–1.4) — Academic contributions of the RQ are (i) incorporating expertise into the consensus-building in P2P systems, (ii) extending the RAP to a group of strategic peers, (iii) bridging How and Why approaches on network formation.

Based on the above, the next Chapter 2 covers the methodology of this thesis, which examines the RQ through two incentive mechanisms (with ex-ante or ex-post consensus) consisting of the same research fields—random walks on graphs and peer prediction methods.

# Chapter 2

# Methodology

To examine the RQ, as mentioned in Section 1.5, this thesis proposes two incentive mechanisms (with ex-ante or ex-post consensus) consisting of the same research fields—random walks on graphs and peer prediction methods. Chapter 2 covers the details of such methodology, while answering the following questions:

- *Why are the two incentive mechanisms important?* (2.1.1),
- *Why are random walks on graphs important?* (2.1.2),
- *Why are peer prediction methods important?* (2.1.2),
- *What are academic contributions of the methodology?* (2.2–2.3).

Furthermore, this chapter mentions experimental datasets (2.4) which were retrieved from three real-world citations, corresponding to scientific publications (from arXiv), patents (from USPTO), and web pages (from Google).

## 2.1 The Two Incentive Mechanisms

The two incentive mechanisms this thesis proposes are the *incentive mechanism with ex-ante consensus* and the *incentive mechanism with ex-post consensus*, where the former is based on Ito and Tanaka (2019), and the latter is based on Ito et al. (2018, 2019). This first section briefly introduces the importance of these two incentive mechanisms and their components, according to Table 2.1 below.

**TABLE 2.1: Differences between the Two Incentive Mechanisms.**

| | Incentive Mechanisms | |
|---|---|---|
| | with *ex-ante* consensus | with *ex-post* consensus |
| Discussed in | Chapter 3 | Chapter 4 |
| Based on | Ito and Tanaka (2019) | Ito et al. (2018, 2019) |
| **Main Scopes** | | |
| Citations with | peer-review | no peer-review |
| Citations on | scientific publications, patents | web pages |
| **Components** | | |
| Random walk by | Haveliwala (2002)<br>*Personalized PageRank* | Babichenko et al. (2018)<br>*Two path mechanism* |
| Peer prediction by | Dasgupta and Ghosh (2013)<br>*Multi-task peer prediction* | Dasgupta and Ghosh (2013)<br>*Multi-task peer prediction* |

### 2.1.1 Main Scopes

First of all, we consider the two incentive mechanisms to cover different types of citations: citations with peer-review and citations without peer-review.

Precisely, the former corresponds to citations on scientific papers and patents, which, in general, cannot be published unless they have passed peer-review by centralized authorities (i.e., editorial boards, patent examiners). To construct such citations in a P2P system, the incentive mechanism has to first build consensus on the validity of intellectual products (and their citations) *ex-ante* and then accept only peer-reviewed products into the system. See Chapter 3 for the detail of this ex-ante consensus.

On the other hand, the latter corresponds to citations on web pages which, in general, can be published without peer-review but will be subject to a ranking by a centralized authority (i.e., search engine). To construct such citations in a P2P system, the incentive mechanism has to first accepts all products coming into the system and then build consensus on the validity of intellectual products (and their citations) *ex-post*. See Chapter 4 for the detail of this ex-post consensus.

Thus, the two incentive mechanisms are important because they allow us to cover citations with peer-review (e.g., those on scientific publications and patents) and citations without peer-review (e.g., those on web pages).

### 2.1.2 Components

Despite their different main scopes, both incentive mechanisms consist of the same research fields: random walks on graphs and peer prediction methods.

Random walks on graphs are a graphical representation of Markov-chain transitions, which repeatedly move from a vertex to one of its neighbors at random (Lovász et al., 1993). For the two incentive mechanisms, this component is useful to address reviewer assignment (Figure 1.5) as a network-based importance measure. Random walks on graphs, although originally intended to represent state transitions[1], have been applied to measure the importance of each vertex, especially since Brin and Page (1998) and Page et al. (1999) developed the PR algorithm[2]. As Table 2.1 shows, the incentive mechanism with ex-ante consensus uses the PPR algorithm (Haveliwala, 2002), while the incentive mechanism with ex-post consensus uses the two-path mechanism (Babichenko et al., 2018); both of which are extensions of the PR algorithm. We will review their details in Section 2.2.

Peer prediction is a reward-based game-theoretic method that aims to elicit true beliefs from peers who report on tasks with no ground truth[3] (e.g., peer-review of scientific publications, customer review in online shopping) (Miller et al., 2005). For the two incentive mechanisms, this component is useful to address free-riding and strategic misreporting (Figure 1.5), as an alternative to the existing token-staking scheme (Ito, 2018)[4]. Peer prediction methods, although quite new to consensus-building in P2P systems[5], have more potential than the token-staking scheme where peers may not participate in consensus-building to avoid losing tokens (i.e., free-riding)[6] or may stake tokens based on the prediction of other peers' beliefs rather than their own (i.e., strategic misreporting)[7]. As Table 2.1 shows, both of the two

[1]One of the common examples is shuffling a deck of cards. By assuming a graph whose vertices are all permutations of the deck, and edges are directed to vertices that each vertex can reach in a single shuffle, we can represent repeated shuffle moves by the random walk on the graph.

[2]See Section 1.1.1 and Section 1.3.3 for the application of the PR algorithm in the context of citation analysis and RAP, respectively.

[3]This concept is referred to as *information elicitation without verification* (Waggoner & Chen, 2014).

[4]See Section 1.3.2 for the details of the token-staking scheme.

[5]To our knowledge, only Goel et al. (2020a), Goel et al. (2020b) use peer prediction methods into consensus-building in P2P systems.

[6]See Appendix A for reward computation in the token-staking scheme.

[7]This is referred to as *Keynesian beauty contest*. Recently, this concept has been generalized as the *p-Beauty contest game* (e.g., Moulin, 1986; Nagel, 1995): a number-guessing game in which players predict mean value (of submitted numbers) multiplied by $p \in (0, 1]$. If the game guesses the mean value (i.e., $p = 1$), there exist multiple Nash equilibria as the number of choices.

incentive mechanisms use a *multi-task peer prediction* by Dasgupta and Ghosh (2013, *DG13*) because of its simplicity and strong solution concept. We will review its detail in Section 2.3.

Thus, random walks on graphs and peer prediction methods are important for the two incentive mechanisms because the former is useful to address reviewer assignments in citations. The latter is useful to address free-riding and strategic misreporting in P2P systems.

In this section, the author briefly introduced the importance of the two incentive mechanisms and their components. To summarize, the two incentive mechanisms are important because they allow us to cover citations with peer-review (e.g., those on scientific publications and patents) and citations without peer-review (e.g., those on web pages); moreover, for both incentive mechanisms, random walks on graphs and peer prediction methods are important to address reviewer assignment and free-riding and strategic misreporting.

## 2.2 Random Walks on Graphs

Random walks on graphs—one component of the two incentive mechanisms—can be quantified with the probability matrix $\boldsymbol{P}$ whose element $p_{ij}$ designates the probability of Markov-chain transition from vertex $j$ to vertex $i$ (e.g., equation 1.1).

**Definition 2.2.1** (Random walk)**.** A *random walk* is the walk[8] whose sequence of vertices are stochastically determined according to $\boldsymbol{P}(G)$[9].

One advantage of the matrix representation is that we can consider the importance of each vertex in $G$ as $\boldsymbol{P}(G)$'s dominant (right) eigenvector (corresponding to eigenvalue 1), which is known to indicate the stationary distribution of iterative random walks (e.g., Axelsson, 1996; Pillai et al., 2005).

Based on this property, the PR algorithm (Brin & Page, 1998; Page et al., 1999) computes the dominant eigenvector of the following probability matrix $\boldsymbol{P}_{PR}$:

[8]See Definition 1.1.3.

[9]Note again that $\boldsymbol{P}(G)$ makes the stochasticity adjustment to a DAG. In other words, even though the random walk ends up reaching the vertex with no out-edges, it is iterated by jumping to one of the existing vertices uniformly at random.

$$\boldsymbol{P}_{PR} = (1-\alpha)\boldsymbol{P} + \alpha\frac{1}{|V|}\mathbf{1}, \tag{2.1}$$

where $\mathbf{1}$ is $|V| \times |V|$ matrix whose elements are all 1, and $\alpha \in [0, 1]$ is an exogenous parameter called *damping factor*[10] (i.e., $\boldsymbol{P}_{PR}$ is the linear combination of two probability matrices: $\boldsymbol{P}$ and $\mathbf{1}/|V|$[11]). Namely, $\boldsymbol{P}_{PR}$ quantifies the iterative random walks which, with probability $\alpha$, jump to one of all existing vertices uniformly at random[12]. This modification is to make the PR algorithm work even in the directed graph, including dead-end loops.

### 2.2.1 Personalized PageRank Algorithm (Haveliwala, 2002)

The PPR algorithm (Haveliwala, 2002)[13] is an extension of the PR algorithm, which aims to apply the random walk to recommender systems such as user-to-paper assignment (Gori & Pucci, 2006) and paper-to-reviewer assignment (Küçüktunç et al., 2012; Liu et al., 2014). While the PR algorithm computes the importance of each vertex in $G$ from the viewpoint of the entire graph structure, the PPR algorithm computes it from the viewpoint of given *base vertices* $V_b \subset V$. Specifically, the PPR algorithm computes the dominant eigenvector of the following probability matrix $\boldsymbol{P}_{PPR}$, which is slightly different from $\boldsymbol{P}_{PR}$:

$$\boldsymbol{P}_{PPR} = (1-\alpha)\boldsymbol{P} + \alpha\frac{1}{|V_b|}\boldsymbol{B}, \tag{2.2}$$

where $\boldsymbol{B}$ is $|V| \times |V|$ matrix whose element $b_{ij}$ becomes 1 if $i$ is included in base vertices $V_b$; otherwise, it becomes 0 (i.e., $\boldsymbol{P}_{PPR}$ is the linear combination of the two probability matrices: $\boldsymbol{P}$ and $\boldsymbol{B}/|V_b|$). Namely, $\boldsymbol{P}_{PPR}$ quantifies the iterative random walks, which, with probability $\alpha$, jump to one of the base vertices uniformly at random[14]. This modification allows the PPR algorithm to incorporate the relevance of $V_b$ into its importance measure.

[10] In most cases, $\alpha = 0.15$.
[11] The probability matrix $\mathbf{1}/|V|$ is often referred to as *teleportation matrix*.
[12] For this property, the PR algorithm is often categorized as *random surfer models*.
[13] This algorithm is originally named *Topic-sensitive PageRank*. The term Personalized PageRank became common after Jeh and Widom (2003) allowed Haveliwala (2002) to set up a larger number of base vertices $V_b$.
[14] For this property, the PPR algorithm is often categorized as *random walk with restart* (RWR) models.

As we will confirm in Chapter 3, the incentive mechanism with ex-ante consensus uses the PPR algorithm to assign reviewers (vertices) to a newly submitted intellectual product (vertex). Furthermore, experimental studies will show that this PPR-based reviewer assignment is positively correlated with the PR score, i.e., the incentive mechanism with ex-ante consensus satisfies condition (i) of Assumption 1.4.3.

### 2.2.2 Two Path Mechanism (Babichenko et al., 2018)

The two-path mechanism (Babichenko et al., 2018) is another extension of the PR algorithm, which aims to protect the random walk from strategic misreporting by vertices (peers). As the name implies, this mechanism leverages *path*—a particular type of walk:

**Definition 2.2.2** (Path)**.** A *path* is the walk whose vertices are all distinct[15].

**Definition 2.2.3** (Random path)**.** A *random path* is the path whose sequence of vertices are stochastically determined according to $\boldsymbol{P}(G)$.

Instead of using iterative random walks, the two-path mechanism regards an important vertex as the first intersection of two independent random paths drawn by letting each vertex sequentially report out-edges (the detailed algorithm will be described in Chapter 4).

The gist of this simple mechanism is that any vertex can no longer manipulate its probability to be the first intersection at reporting out-edges. We can formulate this as *weak truthfulness*—a solution concept to show "agents are indifferent between lying and truthtelling" (Dasgupta & Ghosh, 2013, p. 321)—by assuming that any vertex $i \in V(G)$ reports its out-edges according to *strategy* $\sigma_i$ (see Assumption 2.3.4) and obtains rewards as a random variable $X_i$ that takes some fixed value when $i$ becomes the first intersection and zero otherwise:

**Definition 2.2.4** (Weak truthfulness)**.** A mechanism satisfies *weak truthfulness* in expectation, if $\mathbb{E}\left[X_i \mid \sigma_i^*, \sigma_{-i}\right] = \mathbb{E}\left[X_i \mid \sigma_i, \sigma_{-i}\right]$ holds for every $i, \sigma_i, \sigma_{-i}$,

[15] Thus, path differs from walk in that the drawing process stops when it reaches a previously visited vertex, which implies that path and walk are identical concepts on DAGs.

where $\sigma_{-i} = (\sigma_j)_{j \in V(G) \setminus \{i\}}$ denotes the pair of strategies by all vertices except $i$, and $\sigma_i^*$ denotes the truthtelling strategy by $i$. Thus, $\sigma_i$ has no effect on $\mathbb{E}[X_i]$, irrespective of $\sigma_{-i}$[16]. Babichenko et al. (2018) proved that the two-path mechanism satisfies such weak truthfulness on DAGs and also has a good approximation to the importance measure associated with the PR algorithm[17].

**Proposition 2.2.1** (Weak truthfulness in the two-path mechanism)**.** For DAGs, the two-path mechanism satisfies weak truthfulness.

See proposition 3.1 in Babichenko et al. (2018), for the proof of this proposition.

As we will confirm in Chapter 4, the incentive mechanism with ex-post consensus uses the two-path mechanism to assign reviewers (vertices) to out-edges of existing intellectual products (vertices). Furthermore, experimental studies will show that this two-path reviewer assignment is positively correlated with the PR score, i.e., the incentive mechanism with ex-post consensus satisfies condition (i) of Assumption 1.4.3.

### 2.2.3 State of the Art: Random Walks on Graphs

Random walks on graphs were, to the best of our knowledge, first proposed by Göbel and Jagers (1974) as a hybrid of Markov-chain transitions and graph theory. According to surveys (e.g., Aldous & Fill, 1995; Lovász et al., 1993), most of the early study was theoretical (perhaps due to technical constraints), such as estimating *hitting times* (the number of steps to reach a vertex) or *mixing times* (the number of steps to converge to the stationary distribution) for a given graph[18].

Studies with implementation have flourished, especially since Brin and Page (1998) and Page et al. (1999) developed the PR algorithm in computer science[19]; their

[16]Note that *truthfulness*, also known as *strategy-proofness* or *incentive-compatibility* (Nisan et al., 2007), would be more popular solution concept. Specifically, a mechanism satisfies truthfulness in expectation, if $\mathbb{E}\left[X_i \mid \sigma_i^*, \sigma_{-i}\right] \geq \mathbb{E}\left[X_i \mid \sigma_i, \sigma_{-i}\right]$ holds for every $i, \sigma_i, \sigma_{-i}$, meaning that no agent can obtain a higher expected utility by any possible strategy deviating from his/her true beliefs. Needless to say, weak truthfulness is a necessary condition of truthfulness.

[17]Babichenko et al. (2018) uses an importance measure $I(i)$ which designates the probability of visiting a vertex $i$ in one random path multiplied by $|V|$. With the adjacency matrix $\boldsymbol{A} = (a_{ij})$, this can be defined as $I(i) = \sum_j a_{ij} p_{ij} I(j) + 1$, while the PR score for $i$ can be described as $PR(i) = \sum_j a_{ij} p_{ij} PR(j)$ in the $\alpha = 1$ case.

[18]Before the advent of WWW, random walks on graphs assumed electrical networks as their main application (e.g., Chandra et al., 1996; Nash-Williams, 1959).

[19]Chung and Zhao (2010) discusses in detail the relevance of the PR algorithm and the early studies on random walks on graphs.

contribution was to introduce the aforementioned damping factor $\alpha$, thereby making the random walk practical as a measure of importance for web pages[20]. Random walks on graphs, thanks to the proliferation of the PR algorithm, were the basis for a variety of subsequent algorithms with different purposes, such as spam detection (e.g., Becchetti et al., 2008; Gyongyi et al., 2004), link prediction (e.g., Backstrom & Leskovec, 2011; Liben-Nowell & Kleinberg, 2007), and recommendation (e.g., Gori & Pucci, 2006; Haveliwala, 2002; Küçüktunç et al., 2012). See Gleich (2015) for a more comprehensive review of the PR algorithm and its applications.

For these studies, the methodology of this thesis has an academic contribution in that it provides *strong truthfulness* (Shnayder et al., 2016a, Definition 2.3.1) for random walks on graphs. One of the recent research topics is to make random walks on graphs satisfy truthfulness (or strategy-proofness or incentive-compatibility) that intuitively represents the situation where no agent can obtain a higher utility by any possible strategy deviating from their true beliefs (e.g., Nisan et al., 2007). This is especially important for P2P systems because peers, who know the details of ranking (or consensus) algorithms, may strategically misreport the graph structure (Section 1.3.1)[21]. Strategic misreporting in random walks has been studied since *EigenTrust* (Kamvar et al., 2003)[22]. The two-path mechanism (Babichenko et al., 2018) is the first to achieve the weak truthfulness (Definition 2.2.4)[23]. In this context, the methodology of this thesis—combining random walks on graphs with DG13—is an enhancement of their solution concept from weak to strong truthfulness by leveraging multi-task peer prediction.

In this section, we confirmed the details of the PPR algorithm (Haveliwala, 2002)

[20]Note again that, although it was neither practical at that time nor directly related to the early studies on random walks on graphs, Pinski and Narin (1976) already proposed the use of stationary distribution as a measure of importance for scientific publications.

[21]In other words, unless random walks on graphs satisfy truthfulness or relevant solution concepts, the aforementioned spam detection, link prediction, recommendation, etc. inevitably rely on some centralized authority to keep their details secret (e.g., the ranking algorithm in Google).

[22]EigenTrust is an algorithm for reputation management in P2P file-sharing systems, which leverages the random walk on the graph of reciprocal evaluation among peers. While this algorithm assumes P2P systems, it requires pre-trusted peers if there exists a group of malicious peers doing strategic misreporting.

[23]More recently, several studies (e.g., Wąs et al., 2019) have taken an axiomatic approach to the truthfulness in random walks on graphs.

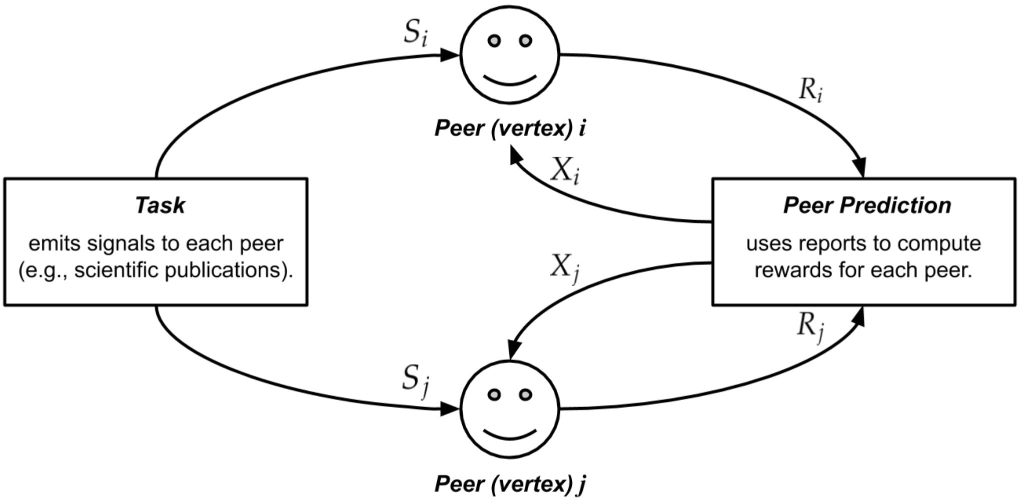



FIGURE 2.1: **Peer Prediction Methods.** In order to model true beliefs for tasks with no ground truth, peer prediction methods assume that peers (vertices) $i, j$ receive *signals* $S_i, S_j$ from task and report them as $R_i, R_j$ that will be used to compute rewards $X_i, X_j$. Here, expected rewards should be maximized when peers provide true beliefs—truthtelling of realized signals (i.e., $r_i = s_i, r_j = s_j$).

and the two-path mechanism (Babichenko et al., 2018). To summarize, the PPR algorithm and the two-path mechanism are both extensions of the PR algorithm (Page et al., 1999), which adds the concepts of relevance and weak-truthfulness, respectively. The methodology of this thesis—combining random walks on graphs with DG13—has an academic contribution in that it provides the strong truthfulness for random walks on graphs.

## 2.3 Peer Prediction Methods

Peer prediction methods—the other component of the two incentive mechanisms—require several assumptions to model true beliefs for tasks with no ground truth. First of all, the methods assume that there exists at least two non-cooperative peers (vertices) $i$ and $j$ who report *signals* emitted from tasks[24].

**Assumption 2.3.1** (Signal reporting)**.** Peers $i$ and $j$ each report what *signals* $S_i$ and $S_j$ were, which are discrete random variables emitted from the task.

Figure 2.1 illustrates the signal reporting and the role of peer prediction, in which peers $i, j$ receive signals $S_i, S_j$ and report them as $R_i, R_j$ that will be used by peer prediction to compute rewards $X_i, X_j$. Here, $S_i, S_j, R_i, R_j, X_i, X_j$ are all random variables,

[24]Note that we can apply peer prediction methods to more than two peers, just by picking $j$ randomly from peers assigned to the same task when computing $i$'s rewards.

and let $s_i, s_j, r_i, r_j, x_i, x_j$ denote their realizations. Peer prediction methods regard true beliefs as truthtelling of signals (i.e., $r_i = s_i, r_j = s_j$) and aim to provide maximum expected rewards $\mathbb{E}[X_i], \mathbb{E}[X_j]$ when peers keep truthtelling.

Furthermore, although several peer-prediction methods (e.g., Miller et al., 2005; Shnayder et al., 2016a) consider multiple signals, this thesis focuses on *binary signals* with *positive correlation*:

**Assumption 2.3.2** (Binary signals)**.** $s_i, s_j \in \{0, 1\}$.

In other words, the task is subject to a binary choice, such as {accept, reject} in academic peer-review and {agree, disagree} in a vote of confidence. Assumption 2.3.2 implies $r_i, r_j \in \{0, 1\}$ as well.

**Assumption 2.3.3** (Positive correlation)**.** Binary signals $\{0, 1\}$ to peers $i$ and $j$ are positively correlated; namely, $Pr(S_i = 0 \mid S_j = 0) > Pr(S_i = 0)$ and $Pr(S_i = 1 \mid S_j = 1) > Pr(S_i = 1)$[25].

Positive correlation intuitively means that *i* and *j* have, to some extent, similar beliefs on assigned task[26]. We will discuss Assumption 2.3.3 further in Section 2.3.2.

Finally, based on Assumptions 2.3.1–2.3.2, we can define strategy as follows:

**Assumption 2.3.4** (Strategies as probability matrices)**.** Peers $i$ and $j$ follow mixed strategies $\sigma_i$ and $\sigma_j$ that have probability matrices $\boldsymbol{P}(\sigma_i)$ and $\boldsymbol{P}(\sigma_j)$, respectively.

Figure 2.2 introduces several examples of $\boldsymbol{P}(\sigma_i)$—a probability matrix for $i$'s mixed strategy $\sigma_i$—whose element $p_{rs}$ designates the probability of reporting $r_i$ from $s_i$; accordingly, in $s_i \in \{0, 1\}$ case, $\boldsymbol{P}(\sigma_i)$ is $2 \times 2$ matrix. Figure 2.2a represents *truthtelling strategy* $\sigma_i^*$ that lets $i$ always report true signals, which corresponds to $\boldsymbol{P}(\sigma_i^*)$ as the identity matrix; Figure 2.2b represents *perverse strategy* that lets $i$ always report wrong signals, which corresponds to $\boldsymbol{P}(\sigma_i)$ as the flipped identity matrix; Figure 2.2c represents an example of *uninformative strategy* that lets $i$ stochastically report 0 or 1 independent of realized signals, which corresponds to $\boldsymbol{P}(\sigma_i)$ as matrices whose column vectors are all identical; then, Figure 2.2d represents an example of other possible strategies.

[25] Simultaneously, $Pr(S_i = 1 \mid S_j = 0) < Pr(S_i = 1)$ and $Pr(S_i = 0 \mid S_j = 1) < Pr(S_i = 0)$.

[26] This positive correlation is one of the simplest examples for *stochastic relevance* (Johnson et al., 2002) that means, for all $s' \neq s''$, there exists at least one signal $s$ such that $Pr(S_i = s \mid S_j = s') \neq Pr(S_i = s \mid S_j = s'')$ (Shnayder et al., 2016a). Stochastic relevance is a necessary condition for most peer prediction methods, including the model for multiple signals.

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}$$

(a) Truthtelling

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

(b) Perverse

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.3 & 0.3 \\ 1 & 0.7 & 0.7 \end{array}$$

(c) Uninformative (example)

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.1 & 0.6 \\ 1 & 0.9 & 0.4 \end{array}$$

(d) Others (example)

**FIGURE 2.2: Strategies as Probability Matrices.** Peer prediction methods often represent a peer $i$'s mixed strategy $\sigma_i$ as probability matrix $\boldsymbol{P}(\sigma_i)$ whose element $p_{rs}$ designates the probability of reporting $r_i$ from $s_i$; accordingly, in $s_i \in \{0, 1\}$ case, $\boldsymbol{P}(\sigma_i)$ is $2 \times 2$ matrix. Here, $\boldsymbol{P}(\sigma_i)$ becomes (a) the identity matrix if $i$ always report true signals (*truthtelling strategy*; $\sigma_i^*$), (b) the flipped identity matrix if $i$ always reports wrong signals (*perverse strategy*), (c) matrices whose column vectors are all identical if $i$ stochastically reports 0 or 1 independent of realized signals (*uninformative strategy*).

### 2.3.1 Multi-Task Peer Prediction (Dasgupta & Ghosh, 2013)

Multi-task peer prediction is an extension of peer prediction methods, which, as the name implies, assigns multiple tasks to a single peer for reward computation with stronger solution concept[27]. Specifically, we use DG13—a multi-task peer prediction that satisfies the following strong truthfulness (Shnayder et al., 2016a):

**Definition 2.3.1** (Strong truthfulness)**.** A mechanism satisfies *strong truthfulness* if $\mathbb{E}\left[X_i \mid \sigma_i^*, \sigma_j^*\right] \geq \mathbb{E}\left[X_i \mid \sigma_i, \sigma_j\right]$ holds for every $\sigma_i, \sigma_j$, where equality occurs only when both $i$ and $j$ adopt the perverse strategy[28].

In other words, strong truthfulness can provide strictly higher expected rewards for the pair of truthtelling strategies than for other realistic strategy pairs, which is important to satisfy condition (ii) of Assumption 1.4.3[29].

Now, we need additional notations to distinguish between multiple tasks. Let $M_i$ and $M_j$ each denote the sets of tasks assigned to peers $i$ and $j$, and let $M^* = M_i \cap M_j$

[27]See Section 2.3.2 for the limitation of preceding peer-prediction methods.

[28]The original definition by Shnayder et al. (2016a) generalizes both truthful and perverse strategies as *permutation strategy* to encompass the case of multiple signals.

[29]Note that, contrary to the name, strongly truthfulness is a weaker concept than truthfulness (and weak truthfulness) in that $i$'s best response depends on $j$'s strategy.

denote the set of *overlapped tasks* which are assigned to both $i$ and $j$; moreover, let $r_i^m, r_j^m, x_i^m, x_j^m$ denote realizations of $R_i, R_j, X_i, X_j$ for a task $m$, respectively[30].

DG13 computes rewards for any overlapped task $m^* \in M^*$ (thus, $\mathbb{E}[X_i]$ means the amount of $i$'s expected rewards per report for an overlapped task), with the following rule:

$$x_i^{m^*} = \delta(r_i^{m^*}, r_j^{m^*}) - \delta(r_i^n, r_j^{n'})^{31}, \tag{2.3}$$

where $n \in M_i \setminus \{m^*\}$ and $n' \in M_j \setminus \{m^*\}$ denote two tasks $(n \neq n')$ randomly selected from those to which each peer was assigned other than $m^*$[32], and $\delta$ is Kronecker delta denoting the following function:

$$\delta(a,b) = \begin{cases} 0 & \text{if } a \neq b, \\ 1 & \text{if } a = b. \end{cases}$$

$\delta(r_i^{m^*}, r_j^{m^*})$ is reward term, which becomes 1 if $i$ and $j$ have the same report for $m^*$, and 0 otherwise. On the other hand, $\delta(r_i^n, r_j^{n'})$ is penalty term, which randomly picks two different tasks $n \in M_i \setminus \{m^*\}$ and $n' \in M_j \setminus \{m^*\}$ then compares their reports in the same manner. Thus, $x_i^{m^*} \in \{-1, 0, 1\}$ holds for every $m^* \in M^*$.

For example, if peers $i$ and $j$ always report 1, i.e., $\boldsymbol{P}(\sigma_i) = \boldsymbol{P}(\sigma_j) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, then $\mathbb{E}[X_i] = 0$ holds because the values of reward and penalty terms in equation 2.3 both become 1. We can derive the same result from another uninformative strategy $\boldsymbol{P}(\sigma_i) = \boldsymbol{P}(\sigma_j) = \left(\begin{smallmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{smallmatrix}\right)$ as well, where the expected values of reward and penalty terms both become 0.5.

**Theorem 2.3.1** (Strong truthfulness in DG13)**.** DG13 satisfies strong truthfulness.

See Appendix B for the proof of this theorem.

As we will confirm in Chapters 3 and 4, the two incentive mechanisms both use DG13 for their reward computation. Furthermore, experimental studies will show that the strong truthfulness in DG13 works even combined with the PPR algorithm and the two-path mechanism, i.e., the two incentive mechanisms satisfy condition (ii) of Assumption 1.4.3.

[30] Note that Assumption 2.3.3 holds even across multiple tasks. Thus, multi-task peer predictions require similarity (to some extent) not only in the belief of peers but also in the type of assigned tasks.

[31] Here, because of symmetry, $x_i^{m^*} = x_j^{m^*}$ holds.

[32] Therefore, DG13 needs $|M^*| \geq 1$ and at least three tasks: $m^*, n, n'$.

### 2.3.2 State of the Art: Peer Prediction Methods

Peer prediction method was first introduced by Miller et al. (2005) as a hybrid of the *proper scoring rule* (e.g., Gneiting & Raftery, 2007) and game theory; specifically, it originally assumed a single task emitting stochastic but correlated signals and computed a peer $i$'s reward $x_i$ based on how much $r_i$ would affect $r_j$. One of the limitations for early peer-prediction methods, as Jurca and Faltings (2005) pointed out, was that they had *multiple Nash equilibria*, including those by uninformative strategies (Figure 2.2c)[33].

To overcome the limitation, Dasgupta and Ghosh (2013) proposed multi-task peer prediction (DG13) that achieved the aforementioned strong truthfulness (Definition 2.3.1) under binary signals $\{0, 1\}$. DG13's approach—assigning multiple tasks to a single peer—is simpler and easier to implement than other (single-task) approaches, such as using four or more reports from different peers (Jurca, Faltings, et al., 2009), asking each peer for both prior and posterior beliefs (Witkowski & Parkes, 2012), and having peers predict other reports (Radanovic & Faltings, 2013). Recent studies generalize DG13 from binary to multiple signals (Shnayder et al., 2016a) and simulate its convergence to an equilibrium (Shnayder et al., 2016b). See Faltings and Radanovic (2017), for a more comprehensive review on peer prediction and other methods for information elicitation.

For these studies, the methodology of this thesis has an academic contribution in that it leverages graphs to make peer prediction practical. Another limitation is that peer prediction methods need to grasp the correlation structure of signals (e.g., Assumption 2.3.3 for DG13)[34], which is unrealistic, especially for multiple signals with many correlation patterns. To overcome the limitation, Shnayder et al. (2016a) was the first to propose estimating the correlation structure of signals from accumulated reports. Such an estimation-based approach is further extended to cover heterogeneous tasks (Mandal et al., 2016) and heterogeneous peers (Agarwal et al., 2017). In this context, the methodology of this thesis—combining DG13 with random walks

[33] Dasgupta and Ghosh (2013) designated such Nash equilibria by uninformative strategies as *blind agreements*.

[34] Some earlier peer prediction methods (e.g., Jurca, Faltings, et al., 2009; Miller et al., 2005) were more limited in their utility because they required full knowledge of the signal distribution as well as the correlation structure.

on graphs—is an alternative graph-based approach that ensures the (positive) correlation structure of (binary) signals through the appropriate tasks-to-peers assignment (as reviewer assignment), thereby making peer prediction practical.

In this section, we confirmed the details of the multi-task peer prediction (DG13; Dasgupta & Ghosh, 2013). To summarize, DG13 is an extension of peer prediction methods originated from Miller et al. (2005), which achieved the strong truthfulness (Definition 2.3.1) under binary signals. The methodology of this thesis—combining DG13 with random walks on graphs—has an academic contribution in that it leverages graphs to ensure the correlation structure of signals (Assumption 2.3.3), thereby making peer prediction practical.

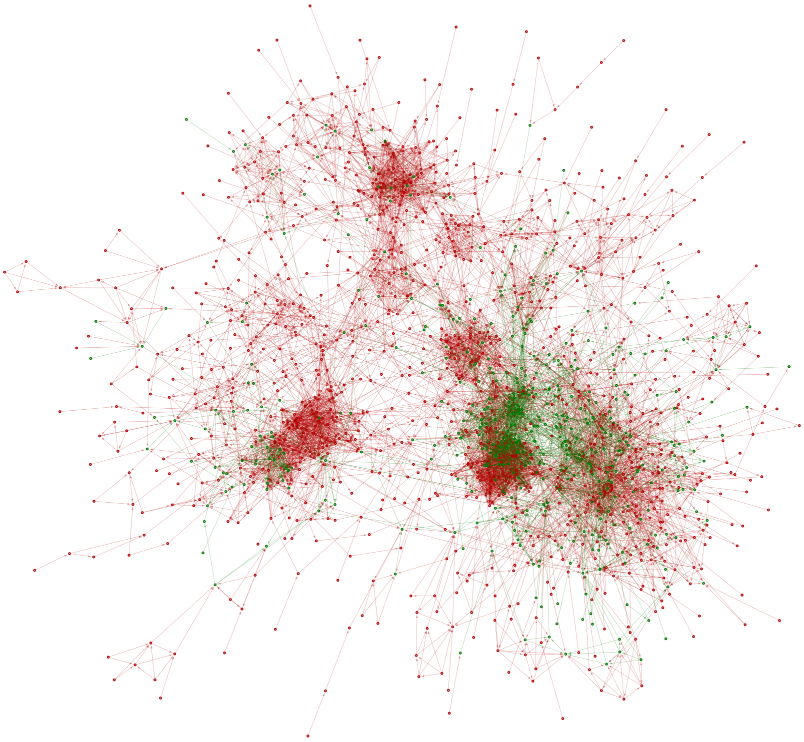
## 2.4 Experimental Datasets

As examples of the growing DAGs, the author retrieved datasets from three real-world citations, corresponding to scientific publications (from arXiv), patents (from USPTO), and web pages (from Google). All datasets are available in *Stanford Network Analysis Project* (SNAP) repository[35], and their visualization (Figure 2.3–2.4) is powered by *Cytoscape* (Shannon et al., 2003)[36].

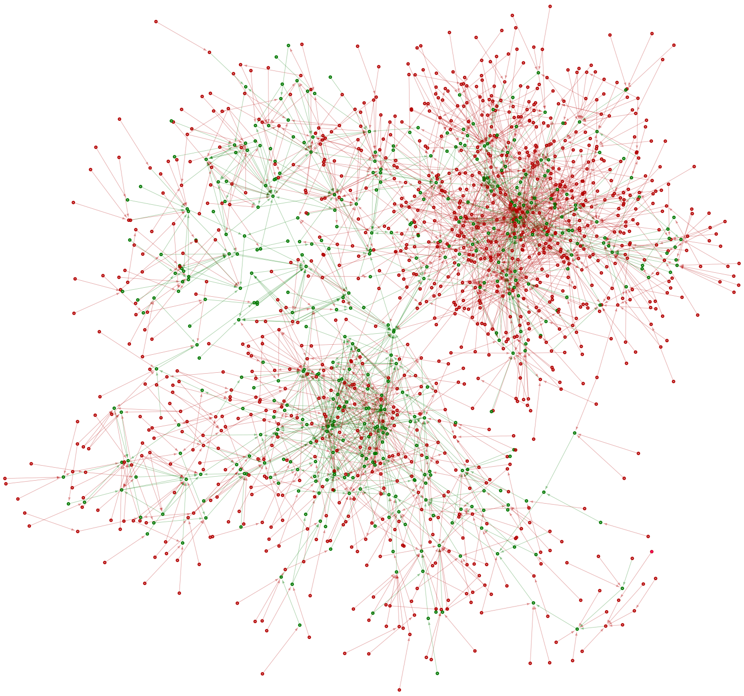
The dataset for scientific publications is *arXiv high-energy physics theory citation network* (Gehrke et al., 2003; Leskovec et al., 2005)[37], which collected citations for $27,770$ papers submitted to high-energy physics theory (HEP-TH) category in arXiv from January 1993 to April 2003. From this dataset, the author extracted a DAG structure with $1,421$ time-ordered vertices and $7,753$ edges (Figure 2.3a). Here, the green represents the citation relationship of the first 421 vertices, while the red represents that of the last $1,000$ vertices.

The dataset for patents is *the NBER U.S. patent citations data file* (Hall et al., 2001; Leskovec et al., 2005)[38], which collected citations for $3,774,768$ patents submitted (and accepted) to USPTO from 1975 to 1999[39]. From this dataset, the author extracted

[35] https://snap.stanford.edu/data/index.html, accessed August 27, 2020.
[36] https://cytoscape.org/, accessed August 27, 2020.
[37] https://snap.stanford.edu/data/cit-HepTh.html, accessed August 27, 2020.
[38] https://snap.stanford.edu/data/cit-Patents.html, accessed August 27, 2020.
[39] This dataset was originally proposed and managed in *National Bureau of Economic Research* (NBER; http://data.nber.org/patents/, accessed August 27, 2020).



(a) Scientific Publications



(b) Patents

**FIGURE 2.3: Citations on Scientific Publications and Patents.** Figure 2.3a depicts citations on scientific publications (1,421 time-ordered vertices and 7,753 edges) extracted from *arXiv high-energy physics theory citation network*, and Figure 2.3b depicts those on patents (1,500 time-ordered vertices and 2,243 edges) extracted from *the NBER U.S. patent citations data file*. Our experiments focus on their state transitions $(G_t)_{t=0}^{1000}$ where the 1,000 red vertices are sequentially added to the green initial state $G_0$.

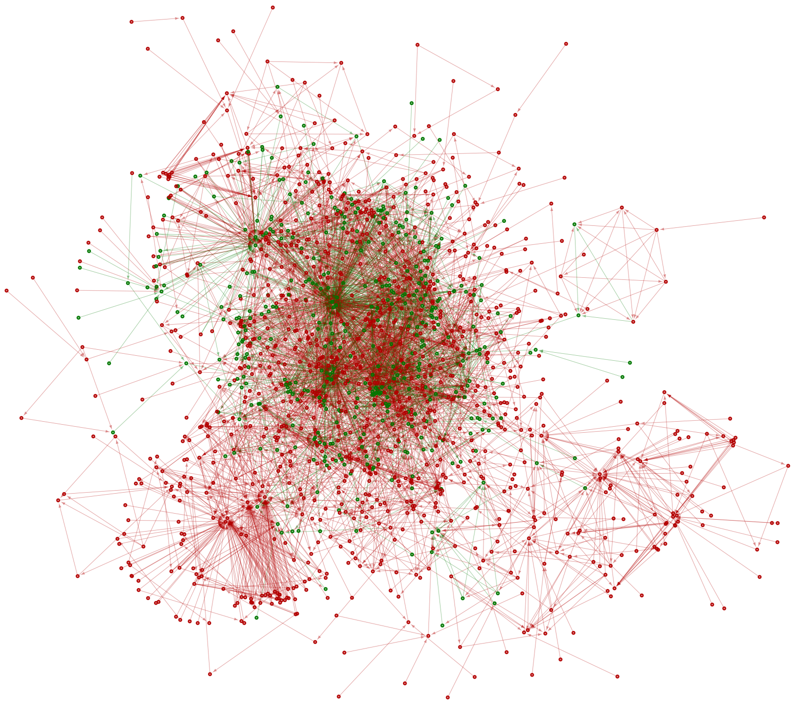


**FIGURE 2.4: Citations on Web Pages.** This thesis uses citations on web pages (1,439 time-ordered vertices and 3,862 edges) extracted from *Google web graph*. Our experiments focus on its state transition $(G_t)_{t=0}^{1000}$ where the 1,000 red vertices are sequentially added to the green initial state $G_0$.

a DAG structure with 1,500 time-ordered vertices and 2,243 edges (Figure 2.3b). Here, the green represents the citation relationship of the first 500 vertices, while the red represents that of the last 1,000 vertices.

The dataset for web pages is *Google web graph* (Leskovec et al., 2009)[40], which collected citations for 875,713 web pages until 2002. From this dataset, the author extracted a DAG structure with 1,439 time-ordered vertices and 3,862 edges (Figure 2.4)[41]. Here, the green represents the citation relationship of the first 439 vertices, while the red represents that of the last 1,000 vertices.

Our experiments in Chapters 3 and 4 use these three DAG structures, especially focusing on their state transitions $(G_t)_{t=0}^{1000}$ where the 1,000 red vertices are sequentially added to the green initial state $G_0$.

[40] https://snap.stanford.edu/data/web-Google.html, accessed August 27, 2020.

[41] Note that, as the dataset for web pages, the Google web graph (Leskovec et al., 2009) has a number of cycles. The author, therefore, removed all old-to-new edges for constructing a DAG structure.

## 2.5 Summary of This Chapter

This chapter covered the methodology, which examines the RQ through two incentive mechanisms (ex-ante or ex-post consensus) consisting of the same research fields—random walks on graphs and peer prediction methods. We can summarize all arguments in Chapter 2 as the description of mechanism components (2.2–2.3), experimental datasets (2.4), and answers to the questions presented at the beginning:

- *Why are the two incentive mechanisms important?* (2.1.1) — The two incentive mechanisms are important because they allow us to cover both citations with peer-review and citations without peer-review,
- *Why are random walks on graphs important?* (2.1.2) — Random walks on graphs are important (for the two incentive mechanisms) because they are useful to address reviewer assignment in citations,
- *Why are peer prediction methods important?* (2.1.2) — Peer prediction methods are important (for the two incentive mechanisms) because they are useful to address free-riding and strategic misreporting in P2P systems,
- *What are academic contributions of the methodology?* (2.2–2.3) — Academic contributions of the methodology are (i) providing strong truthfulness for random walks on graphs and (ii) leveraging graphs to make peer prediction practical.

Based on the above, the next Chapter 3 introduces the incentive mechanism with ex-ante consensus, which (i) covers citations with peer-review (e.g., those on scientific publications and patents) and (ii) consists of the PPR algorithm (Haveliwala, 2002) and DG13 (Dasgupta & Ghosh, 2013).

# Chapter 3

# Incentive Mechanism With Ex-Ante Consensus

Of the two proposals, Chapter 3 introduces the *incentive mechanism with ex-ante consensus*, which (i) covers citations with peer-review (e.g., those on scientific publications and patents) and (ii) consists of the PPR algorithm (Haveliwala, 2002) and DG13 (Dasgupta & Ghosh, 2013)[1]. The first half of this chapter (3.1) details its algorithms, while answering the following questions:

- *Why can the PPR algorithm solve reviewer assignment?* (3.1.3),
- *Why can DG13 solve free-riding and strategic misreporting?* (3.1.4).

In addition, the second half of this chapter (3.2) experimentally confirms that the algorithms ensure the aforementioned conditions for the reliable consensus-building (Assumption 1.4.3), by using real-world citation data on scientific publications (Figure 2.3a) and patents (Figure 2.3b).

## 3.1 Algorithms

As mentioned in Sections 1.5 and 2.1, to cover citations on scientific publications and patents, the incentive mechanism with ex-ante consensus first builds consensus on the validity of intellectual products (and their citations) and then accepts only peer-reviewed products into the system. This section details such algorithms in text,

[1]This chapter is based on the author's previous study Ito and Tanaka (2019). In this study, the incentive mechanism with ex-ante consensus—originally named *CitedTCR*—was a proposal to incorporate the expertise of anonymous peers (reviewers) into the consensus-building of Token Curated Registries (TCRs) by leveraging citations on posted contents.

**TABLE 3.1: Notations for Incentive Mechanism with Ex-ante Consensus.**

| Notations | Meanings |
|---|---|
| $G_t = (V_t, E_t)$ | Citations as a growing DAG in period $t$ |
| $V_t$ | Set of vertices as intellectual products in period $t$ |
| $E_t$ | Set of edges as citation relationships for $V_t$ |
| $\dot{G}_t = (\{k\} \cup V_k, E_k)$ | A proposal of new citations in period $t$ |
| $k$ | A new vertex in period $t$ |
| $V_k$ | Set of base vertices (references) for $k$ (i.e., $V_k \subseteq V_t$) |
| $E_k$ | Set of out-edges directed from $k$ to $V_k$ |
| $\dot{G}_{t'}$ | A proposal of new citations in a period other than $t$ |
| $\dot{G}_{t''}$ | A proposal of new citations in a period other than $t$ and $t'$ |
| $\dot{C}_t$ | Set of reviewers $\{1, 2, \cdots, \lambda\}$ for $\dot{G}_t$, selected from $V_t \setminus V_k$ |
| $\dot{R}_t$ | Set of reports $\{r_1^{\dot{G}_t}, r_2^{\dot{G}_t}, \cdots, r_\lambda^{\dot{G}_t}\}$ elicited from $\dot{C}_t$ |
| $\dot{X}_t$ | Set of rewards $\{x_1^{\dot{G}_t}, x_2^{\dot{G}_t}, \cdots, x_\lambda^{\dot{G}_t}\}$ for $\dot{R}_t$ |
| $\lambda$ | An exogenous parameter ($\geq 2$) for the number of reviewers |
| $\mu$ | An exogenous parameter ($\leq \lambda$) for the difficulty |
| $R_t$ | Set of the stock of all reports until period $t$ |

figures, and pseudocode, along with the role and notes on the PPR algorithm and DG13. See Table 3.1 above for the notations relevant to the incentive mechanism with ex-ante consensus.

### 3.1.1 Setup

Consider citations as a growing DAG $G_t = (V_t, E_t)$, where $V_t$ denotes the set of intellectual products that are synonymous with peers and individuals (Assumption 1.4.1), and $E_t \subseteq V_t \times V_t$ denotes their citation relationships in period $t$. Because $G_t$ involves peer-review, its state transition $(G_t)_{t=0}^{q}$ becomes an iterative process of reviewing whether to accept a proposal of new citations $\dot{G}_t = (\{k\} \cup V_k, E_k)$, where $k$ denotes a new vertex in period $t$; $V_k \subseteq V_t$ denotes the set of base vertices (references) for $k$; $E_k$ denotes the set of all out-edges from $k$ to $V_k$, respectively. For the sake of algorithmic description, we here assume that each period deals with only one $\dot{G}_t$, despite the importance of parallel consensus-building in P2P citation systems (Definition 1.2.2)[2].

[2]This assumption is just for the sake of convenience; the incentive mechanism with ex-ante consensus can handle multiple $\dot{G}_t$s in parallel.

Reviewer assignment proceeds soon after $\dot{G}_t$ arrives; the incentive mechanism selects a set of reviewers $\dot{C}_t = \{1, 2, \cdots, \lambda\}$ from $V_t \setminus V_k$[3], where $\lambda$ ($\geq 2$) is an exogenous parameter to determine the number of reviewers for one proposal[4]. Reviewers in $\dot{C}_t$ then send a set of reports $\dot{R}_t = \{r_1^{\dot{G}_t}, r_2^{\dot{G}_t}, \cdots, r_\lambda^{\dot{G}_t}\}$ which evaluates $\dot{G}_t$ (as a task) with binary signals $\{0, 1\}$[5]. We assume for convenience that reports 0 and 1 each designate *reject* and *accept*. After computing rewards $\dot{X}_t = \{x_1^{\dot{G}_t}, x_2^{\dot{G}_t}, \cdots, x_\lambda^{\dot{G}_t}\}$ for $\dot{R}_t$, $G_t$ is finally updated to $G_{t+1}$, which accepts the proposal $\dot{G}_t$ only if $\dot{R}_t$ includes $\mu$ ($\leq \lambda$) or more number of *accept*.

Accordingly, the state transition from $G_t$ to $G_{t+1}$ can be summarized as follows:

Step 1: A new vertex $k$ proposes $\dot{G}_t$ to $G_t$,

Step 2: Select $\lambda$ of vertices as $\dot{C}_t$ (*reviewer assignment*),

Step 3: Collect $\lambda$ of reports on $\dot{G}_t$ as $\dot{R}_t$,

Step 4: Compute $\dot{X}_t$ (*reward computation*),

Step 5: Update $G_t$ to $G_{t+1}$ which conditionally accepts $\dot{G}_t$.

Figure 3.1 graphically depicts an example of the state transition by focusing on a subgraph with vertices $\{a, b, c, d, e, f, g\}$, where Figures 3.1a–3.1e correspond to the Steps 1–5. Intuitively, this mechanism introduces rewards into the academic peer-review that delegates the evaluation of a newly submitted paper to multiple reviewers (who would have submitted preceding studies in the relevant field). As will be detailed in Sections 3.1.3 and 3.1.4, the incentive mechanism with ex-ante consensus uses the PPR algorithm for reviewer assignment (Step 2) and DG13 for reward computation (Step 4).

[3]We here exclude $V_k$ from the list of candidates in order to avoid a biased review. Section 3.1.4 will detail the biased review.

[4]Therefore, to manage the incentive mechanism with ex-ante consensus, we need an initial state $G_0$ with a sufficient number of vertices and edges.

[5]Note that the incentive mechanism with ex-ante consensus can be tasked with not only $\dot{G}_t$, but also $E_k$ and each element of $E_k$. In this Chapter 3, we assume $\dot{G}_t$ as the task for convenience, following the existing peer-review that also assesses the quality of $k$ (and perhaps $V_k$). On the other hand, we can let $\dot{C}_t$ review only the proposed citation relationships $E_k$, or even the elements of $E_k$ individually (i.e., the task becomes each out-edge from $k$). The incentive mechanism with ex-post consensus (Chapter 4) adopts this alternative setting to build a consensus for every single citation relationship.

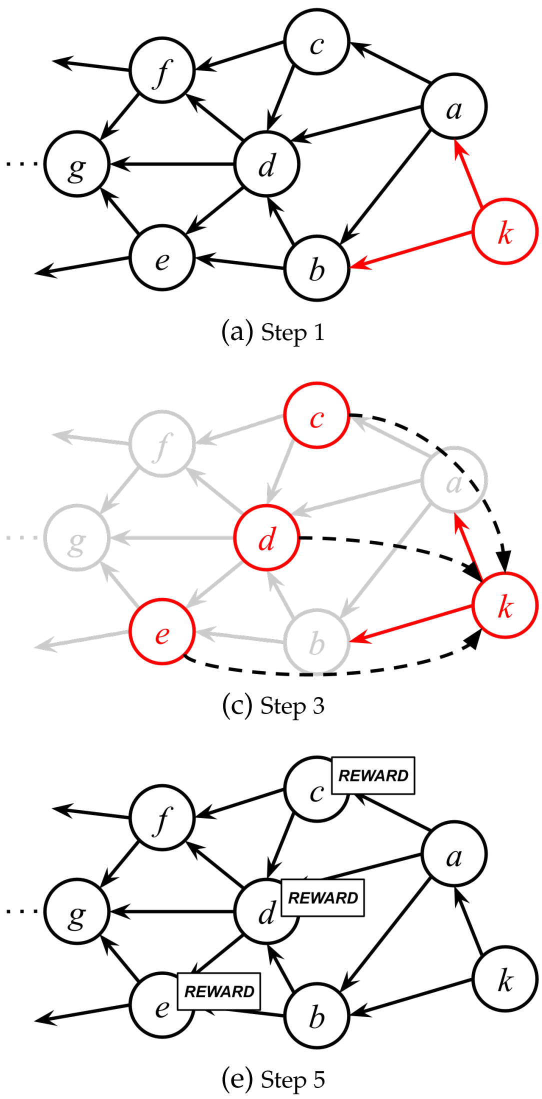



**FIGURE 3.1: Incentive Mechanism with Ex-ante Consensus.** If we focus on a subgraph with vertices $\{a, b, c, d, e, f, g\}$, an example of the Steps 1–5 can be depicted as Figures 3.1a–3.1e. These figures assume the following state transition: in Step 1, $k$ proposes $\dot{G}_t = (\{k, a, b\}, \{(k, a), (k, b)\})$; in Step 2, the mechanism selects $\dot{C}_t = \{c, d, e\}$ as reviewers of $\dot{G}_t$ (i.e., $\lambda = 3$); in Step 3, $\dot{C}_t$ evaluates $\dot{G}_t$ with binary reports $\{0, 1\}$; in Step 4, $\dot{C}_t$ can receive rewards whose amount was computed from their reports; finally, in Step 5, $\dot{G}_t$ is accepted because it obtained the sufficient (equal or more than $\mu$) number of report 1.

### 3.1.2 Pseudocode

Pseudocode allows for a detailed description of the state transition as Algorithms 1 and 2 below, where the former is the whole process, and the latter is the part related to peer review (i.e., Steps 2 and 3). For these algorithms, two properties should be noted. First, Algorithm 1 returns not only $G_{t+1}$ and $\dot{X}_t$ but also the stock of reports $R_{t+1}$. This property is specific to DG13, whose reward computation leverages both

**Algorithm 1** State transition in the incentive mechanism with ex-ante consensus

$G_t \leftarrow (V_t, E_t)$
$\dot{G}_t \leftarrow (\{k\} \cup V_k, E_k)$
$\{\lambda, \mu\} \leftarrow$ exogenous parameters
$R_t \leftarrow$ stock of reports until period $t$ ▷ Specific to DG13
$\dot{R}_t \leftarrow \textsc{PeerRev}(\mu, V_t \setminus V_k, \{\emptyset\}, G_t)$ ▷ See Algorithm 2
Compute rewards $\dot{X}_t$ with $R_t$ and $\dot{R}_t$ ▷ Use DG13
**return** $\dot{X}_t$
$R_{t+1} \leftarrow R_t \cup \dot{R}_t$ ▷ Specific to DG13
**return** $R_{t+1}$ ▷ Specific to DG13
**if** $\mu \geq |\{r \in \dot{R}_t | r = 1\}|$ **then**
    $G_{t+1} \leftarrow G_t$
**else**
    $G_{t+1} \leftarrow G_t \cup \dot{G}_t$
**end if**
**return** $G_{t+1}$

**Algorithm 2** Peer-review

**function** $\textsc{PeerRev}(n, C, R, G)$
    $C' \leftarrow n$ reviewers selected from $C$ in $G$ ▷ Use the PPR algorithm
    $R' \leftarrow$ reports collected from $C'$ within a given period of time
    $R \leftarrow R \cup R'$
    **if** $|R'| = n$ **then**
        **return** $R$
    **else**
        $n \leftarrow n - |R'|$
        $C \leftarrow C \setminus C'$
        $\textsc{PeerRev}(n, C, R, G)$
    **end if**
**end function**

the flow and stock of elicited reports ($\dot{R}_t$ and $R_t$) as one of the multi-task peer predictions. Algorithm 1 could be simpler if we adopt other intratemporal mechanisms (e.g., the token-staking scheme). Second, Algorithm 2 integrates Steps 2 and 3 as $\textsc{PeerRev}(n, C, R, G)$ function, which returns a set of reports $R$ for the following four arguments: $n$, the number of reports; $C$, the set of vertices that are candidates for the reviewer; $R$, the initial value of the set of reports; and $G$, the graph containing $C$[6]. This property aims to handle the case in which (some or all) assigned reviewers do not provide their reports within a given time. In this case, the $\textsc{PeerRev}(n, C, R, G)$ continues to reselect new vertices as replacements for unresponsive reviewers until it collects $n$ reports.

[6] As in Algorithm 1, four arguments $(n, C, R, G)$ corresponds to $(\mu, V_t \setminus V_k, \{\emptyset\}, G_t)$ in the incentive mechanism with ex-ante consensus.

### 3.1.3 Role and Notes on the PPR Algorithm

In these algorithms, the PPR algorithm can solve reviewer assignment because it assigns appropriate reviewers $\dot{C}_t$ (with similarity to $\dot{G}_t$), while ensuring condition (i) of Assumption 1.4.3—*peers can be reviewers more often as they get higher PR scores*. To confirm its specific usage, let us now turn to equation 2.2 on the PPR algorithm, presented in Section 2.2.1:

$$\boldsymbol{P}_{PPR} = (1-\alpha)\boldsymbol{P} + \alpha \frac{1}{|V_b|}\boldsymbol{B}, \tag{3.1}$$

where the incentive mechanism substitutes $V_k$ into $V_b$, thereby computing the PPR score for $V_t \setminus V_k$ from the viewpoint of $k$. $\dot{C}_t$ (the $\lambda$ number of vertices) is stochastically selected, according to the PPR score[7]. As we will confirm experimentally in Section 3.2.1, such PPR-based reviewer assignment can maintain a positive correlation with PR scores that are not biased towards $k$.

On the other hand, there exist two notes on the application of the PPR algorithm. First, similar to preceding studies (e.g., Gori & Pucci, 2006; Küçüktunç et al., 2012), the PPR algorithm considers $G_t$ to be undirected in its computation[8]. This is important because if the PPR algorithm were on a DAG structure, the score would be concentrated on the peer with no out-edges (i.e., the oldest vertex in the growing DAG), making it a less useful measure for $k$'s importance. Second, as already mentioned, the PPR algorithm excludes $V_k$ from the candidates of $\dot{C}_t$. Although peers in $V_k$ can obtain high PPR scores, the mechanism does not select them to reduce the bias such that assigned reviewers accept $k$ to increase their number of citations (thereby increasing their opportunity of becoming a reviewer again in the future)[9].

[7] Considering the correlation with the PR score, reviewer assignment should be deterministic (i.e., selecting the top-$\lambda$ vertices with the highest PPR score) rather than stochastic. However, this incentive mechanism employs stochastic assignment because it can both diversify reviewers and prevent $k$ from identifying $\dot{G}_t$'s reviewers in advance.

[8] Strictly speaking, preceding studies add an edge with the opposite direction for each edge, instead of considering advantages as undirected. In each case, $\boldsymbol{P}(G)$ has the same content.

[9] Note also that, even though we exclude $V_k$ from the candidates of $\dot{C}_t$, the incentive mechanism cannot completely remove this type of bias because (i) the PPR algorithm computes its score from the entire graph structure and (ii) the next proposal $\dot{G}_{t+1}$ would change depending on the shape of $G_{t+1}$. In other words, a reviewer's report in period $t$ will inevitably affect their future PPR scores. We therefore implicitly assume that this bias is not so large enough for reviewers to change their binary reports {*accept*, *reject*}. It is one of the future tasks to eradicate this bias from the incentive mechanism with ex-ante consensus. (As will be explained in Chapter 4, thanks to the two-path mechanism, there is no such bias in the incentive mechanism with ex-post consensus.)

### 3.1.4 Role and Notes on DG13

In these algorithms, DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers $\dot{C}_t$, while ensuring condition (ii) of Assumption 1.4.3—*peers can maximize the amount of expected rewards per review by always reporting true beliefs*. To confirm its specific usage, let us now turn to equation 2.3 on DG13, presented in Section 2.3.1.

$$x_i^{m^*} = \delta(r_i^{m^*}, r_j^{m^*}) - \delta(r_i^n, r_j^{n'}), \tag{3.2}$$

where the incentive mechanism substitutes $\dot{G}_t$ into the overlapping task $m^*$, two different proposals $\dot{G}_{t'}$ and $\dot{G}_{t''}$ into $n$ and $n'$ ($t'$ and $t''$ are randomly selected from periods other than $t$), respectively. DG13 then computes rewards for all reviewers by randomly picking the peer $j \in \dot{C}_t$ for reference. As we will confirm experimentally in Section 3.2.2, such DG13-based reward computation can elicit true beliefs, even under the reviewer assignment with the PPR algorithm.

On the other hand, note that DG13 cannot work until both peers $i$ and $j$ finish reporting multiple tasks. Specifically, DG13 requires at least one overlapping task $m^*$ ($\dot{G}_t$) and two other tasks $n$ and $n'$ ($\dot{G}_{t'}$ and $\dot{G}_{t''}$) for $i$ and $j$, which implies that peers cannot get rewards soon after their initial reports. A practical reward computation would therefore be once a given number of periods (e.g., 100 periods), rather than every period[10]. In this case, DG13 picks up all reports (to which equation 2.3 is newly applicable) from the stock $R_t$ and computes their rewards simultaneously[11].

We detailed algorithms for the incentive mechanism with ex-ante consensus in this section and the role and notes on the PPR algorithm and DG13. To summarize, in the algorithms (Figure 3.1, Algorithms 1 and 2), the PPR algorithm can solve reviewer assignment because it assigns appropriate reviewers (with similarity) while ensuring condition (i) of Assumption 1.4.3, and DG13 can solve free-riding and

[10] The experiment in Section 3.2.2 also computes rewards together, after all, scheduled periods are completed.

[11] Here, when both peers $i$ and $j$ have already done three tasks, DG13 always works because we can ensure $m^*$, $n$, and $n'$ even if the three are all overlapping tasks. Accordingly, the incentive mechanism with ex-ante consensus can run DG13 every period, if its review targets not the entire $\dot{G}_t$ but elements of $E_t$, and the constraint $|E_t| \geq 3$ holds for all $t$.

strategic misreporting because it computes rewards for reviewers while ensuring condition (ii) of Assumption 1.4.3, respectively.

## 3.2 Experimental Studies

This section experimentally confirms that the incentive mechanism with ex-ante consensus ensures the two conditions of Assumption 1.4.3, thereby supporting the claims made in the previous section. For real-world citation data on scientific publications (Figure 2.3a) and patents (Figure 2.3b), we had two-step experiments which first use only the PPR algorithm to examine condition (i) for reviewer assignment, then incorporate DG13 to examine condition (ii) for free-riding and strategic misreporting. All materials for the experiments are available in the Github repository[12].

### 3.2.1 Experiments for Reviewer Assignment

The first experiment computes the correlation between the frequency distribution of reviewer assignment and the PR score for all vertices. This is important because the PPR-based reviewer assignment differs from the PR score in that it is (i) biased towards $k$, (ii) done period-by-period, and (iii) stochastic rather than deterministic. For the aforementioned real-world citation data (Figure 2.3), the former can be derived by applying the PPR algorithm $1,000$ times along with the state transition $(G_t)_{t=0}^{1000}$[13], while the latter can be derived by applying the PR algorithm to the last state $G_{1000}$[14]. This experiment derived 200 patterns—10 times for every $\lambda = \{1, 2, \cdots, 20\}$ cases—of reviewer assignment from both scientific publications and patents, then computed Spearman's rank correlation coefficients between their frequency distributions and the (deterministic) PR score[15].

Figure 3.2 represents the experimental results, where Figure 3.2a is the trend of 200 correlation coefficients computed from scientific publications, and Figure 3.2b is

[12] https://github.com/knskito/materials_thesis

[13] For simplicity, this state transition accepts all $1,000$ (red) vertices to $G_t$, i.e., we assume $\mu = 0$.

[14] We set the dumping factor $\alpha = 0.15$ in both PR and PPR algorithms; furthermore, the PR algorithm (as well as the PPR algorithm) considers $G_t$ to be undirected in its computation.

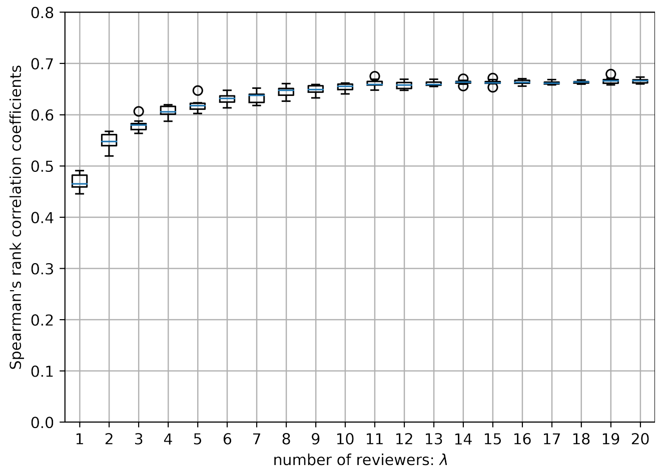
[15] This experiment cannot use Pearson correlation coefficients because neither the frequency distribution of reviewer assignment nor the PR score for all vertices follows normal distributions. We confirmed this by using Shapiro-Wilk test (Shapiro & Wilk, 1965) for both scientific publications and patents data.

the trend of those from patents. Box plots show the median value as blue or orange lines, 25/75 percentile as boxes, pseudo-maximum/minimum value as bars, and outliers as circles. We can see that all $200 * 2$ correlation coefficients are within the range of 0.4 to 0.7, indicating that they are moderately correlated. This result—the moderate positive correlation between the PPR-based reviewer assignment and the PR score—supports condition (i) of Assumption 1.4.3[16]. Moreover, there exist two other implications. First, correlation coefficients begin to converge between 0.6 and 0.7 when $\lambda$ exceeds 10. Second, patents have consistently lower correlation coefficients than scientific publications, which may be related to the density of DAG structures.

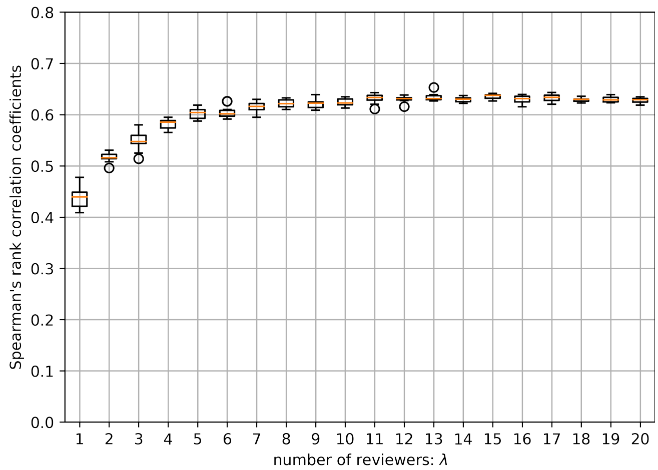
### 3.2.2 Experiments for Free-riding and Strategic Misreporting

The second experiment computes $\mathbb{E}[X_i]$—the amount of $i$'s expected rewards per review—by incorporating DG13 into the first experiment. This is important because we need to confirm whether the DG13-based reward computation can ensure strong truthfulness (Definition 2.3.1) even in conjunction with the PPR-based reviewer assignment. To compare various patterns, this experiment stochastically allocated both signals and strategies to all vertices in advance. For signals, we assume that all vertices, as $k$, emit either signal 0 or 1 to indicate the quality of their proposed task $\dot{G}_t$. This experiment allocated such signals according to nine rules with different randomness $Pr(S=0) = \{0.1, 0.2, \cdots, 0.9\}$, where $S$ is a random variable denoting the signal allocated to each vertex. Note that, from the viewpoint of vertices, $S$ covers sending-signals, while $S_i$ covers received-signals. For strategies, we assume that all vertices, as reviewer, take either the truthful strategy $\boldsymbol{P}(\sigma_i^*) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ or an uninformative strategy $\boldsymbol{P}(\sigma_i) = \left(\begin{smallmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{smallmatrix}\right)$. This experiment allocated such strategies according to 11 rules with different randomness $\epsilon = \{0.0, 0.1, \cdots, 1.0\}$, where $\epsilon$ is an exogenous parameter denoting the probability of allocating the uninformative strategy. That is, all vertices take the truthful strategy if $\epsilon = 0.0$ and the uninformative strategy if $\epsilon = 1.0$. Finally, we computed $\mathbb{E}[X_i]$ resulting from the state transition $(G_t)_{t=0}^{1000}$, for each of the 99 patterns with different signal-strategy allocation pairs

[16]Let us recall that this reviewer assignment does not select the base vertices $V_k$ as reviewers, which may have reduced correlation coefficients. We found that, in an additional experiment modified to include $V_k$ in reviewer candidates, all correlation coefficients have increased by approximately 0.1.




(a) Scientific Publications

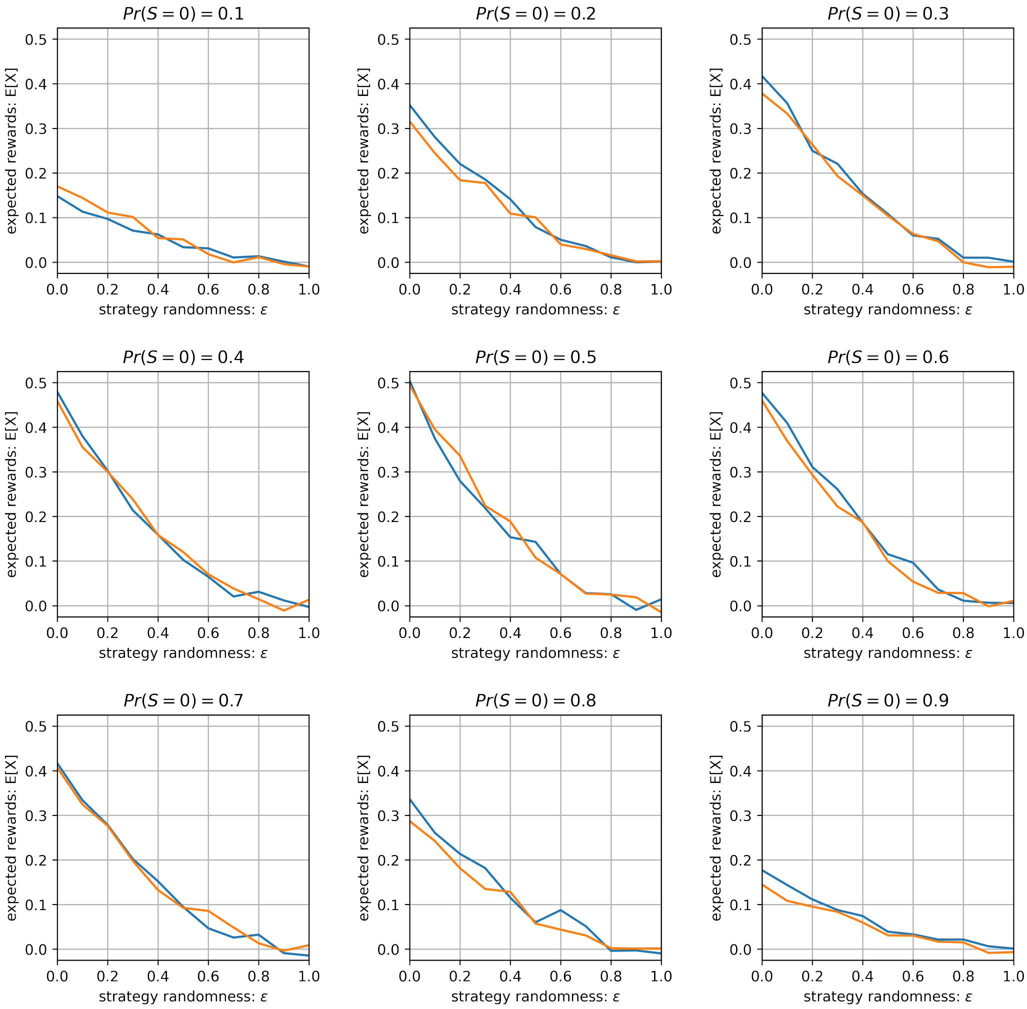



(b) Patents

**FIGURE 3.2: Reviewer Assignment in the Incentive Mechanism with Ex-ante Consensus.** The first experiment computed Spearman's rank correlation coefficients between the frequency distribution of reviewer assignment in $(G_t)_{t=0}^{1000}$ and the PR score in $G_{1000}$. The two box plots for scientific publications (Figure 3.2a) and patents (Figure 3.2b) represent that all 200 (10 times for every $\lambda = \{1, 2, \cdots, 20\}$ cases) $*2$ coefficients are moderately correlated. This result—the moderate positive correlation between the PPR-based reviewer assignment and the PR score—supports condition (i) of Assumption 1.4.3.

$\{0.1, 0.2, \cdots, 0.9\} \times \{0.0, 0.1, \cdots, 1.0\}$. Here, remaining exogenous parameters are fixed as $\lambda = 10$ and $\mu = 0$ (i.e., the state transition always accepts $\dot{G}_t$ into $G_t$).

Figure 3.3 represents the experimental results, where each of the nine graphs depicts the results for a different $Pr(S = 0)$; blue and orange lines show the computed $\mathbb{E}[X_i]$ on scientific publications and patents, respectively. We can see that, for all $Pr(S = 0)$ rules, $\mathbb{E}[X_i]$ is maximized when all vertices take the truthful strategy (i.e., $\epsilon = 0.0$). This result—the maximized expected rewards under truthful strategies—supports condition (ii) of Assumption 1.4.3. Moreover, there exists another




*Note*: blue is the result of scientific papers, and orange is that of patents.

**FIGURE 3.3: Free-riding and Strategic Misreporting in the Incentive Mechanism with Ex-ante Consensus.** The second experiment computed $\mathbb{E}[X_i]$ resulting from the state transition $(G_t)_{t=0}^{1000}$, for each of the 99 patterns with different signal-strategy allocation pairs. The nine graphs covering both scientific publication (blue line) and patents (orange line) represent that, for all $Pr(S = 0) = \{0.1, 0.2, \cdots, 0.9\}$ signal allocations, $\mathbb{E}[X_i]$ is maximized when all vertices take the truthful strategy (i.e., $\epsilon = 0.0$). This result—the maximized expected rewards under truthful strategies—supports condition (ii) of Assumption 1.4.3.

implication that the amount of maximized $\mathbb{E}[X_i]$ decreases as $Pr(S = 0)$ deviates from 0.5. This trend can be inferred from the fact that the penalty term (Equation 2.3) would always be 1 if $Pr(S = 0)$ were 0.0 or 1.0, even though both vertices $i$ and $j$ take the truthful strategy[17].

[17] This is the reason why this experiment does not deal with the case of $Pr(S = 0) = 0.0$ or 1.0. Note that these cases are outside the scope of DG13 because we cannot apply the positively correlated signals (Assumption 2.3.3) to the environment where all signals are identical.

In this section, we confirmed that the incentive mechanism with ex-ante consensus could ensure the two conditions of Assumption 1.4.3, through two-step experiments. To summarize, the first experiment (on the PPR-based reviewer assignment) yielded results supporting condition (i) for reviewer assignment, and the second experiment (on the DG13-based reward computation) yielded results supporting condition (ii) for free-riding and strategic misreporting, respectively.

## 3.3 Summary of This Chapter

This chapter introduced the incentive mechanism with *ex-ante* consensus, which (i) covers citations with peer-review (e.g., those on scientific publications and patents) and (ii) consists of the PPR algorithm (Haveliwala, 2002) and DG13 (Dasgupta & Ghosh, 2013). We can summarize all arguments in Chapter 3 as answers to the questions presented at the beginning:

- *Why can the PPR algorithm solve reviewer assignment?* (3.1.3) — The PPR algorithm can solve reviewer assignment because it assigns appropriate reviewers (with similarity), while ensuring condition (i) of Assumption 1.4.3,
- *Why can DG13 solve free-riding and strategic misreporting?* (3.1.4) — DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers, while ensuring condition (ii) of Assumption 1.4.3.

These answers are supported by the two-step experiments (3.2) as well. Similarly, the next Chapter 4 introduces incentive mechanism with ex-post consensus, which (i) covers citations without peer-review (e.g., those on web pages) and (ii) consists of the two-path mechanism (Babichenko et al., 2018) and DG13 (Dasgupta & Ghosh, 2013).

# Chapter 4

# Incentive Mechanism With Ex-Post Consensus

Of the two proposals, Chapter 4 introduces the *incentive mechanism with ex-post consensus*, which (i) covers citations without peer-review (e.g., those on web pages) and (ii) consists of the two-path mechanism (Babichenko et al., 2018) and DG13 (Dasgupta & Ghosh, 2013)[1]. The first half of this chapter (4.1) details its algorithm, while answering the following questions:

- *Why can the two-path mechanism solve reviewer assignment?* (4.1.3),
- *Why can DG13 solve free-riding and strategic misreporting?* (4.1.4).

In addition, the second half of this chapter (4.2) experimentally confirms that the algorithms ensure the aforementioned conditions for the reliable consensus-building (Assumption 1.4.3) by using real-world citation data on web pages (Figure 2.4).

## 4.1 Algorithms

As mentioned in Sections 1.5 and 2.1, to cover citations on web pages, the incentive mechanism with ex-post consensus first accepts all products coming into the system. It then builds consensus on the validity of intellectual products (and their citations). This section details such algorithms in text, figures, and pseudocode, along with the

[1]This chapter is based on the author's previous studies Ito et al. (2018, 2019). In these studies, the incentive mechanism with ex-post consensus—was originally named *strongly-truthful two-path mechanism*—was a proposal to strengthen the truthfulness of the two-path mechanism (that aims to find an influential vertex in non-cooperative DAGs) by leveraging DG13. On the other hand, this thesis modified the way of task and reviewer-assignment in Ito et al. (2018, 2019), to make the mechanism more practical.

**TABLE 4.1: Notations for Incentive Mechanism with Ex-post Consensus.**

| Notations | Meanings |
|---|---|
| $G_t = (V_t, E_t)$ | Citations as a growing DAG in period $t$. |
| $V_t$ | Set of vertices as intellectual products in period $t$. |
| $E_t$ | Set of edges as citation relationships for $V_t$. |
| $\dot{G}_t$ | A proposal of new citations in period $t$. |
| $\dot{C}_t$ | Set of reviewers $\{1, 2, \cdots\}$ in the random walk. |
| $\dot{R}_t$ | Set of reports $\{r_1^{(1,i)}, r_1^{(1,j)}, \cdots, r_2^{(2,i)}, r_2^{(2,j)}, \cdots\}$ elicited from $\dot{C}_t$. |
| $\dot{X}_t$ | Set of rewards $\{x_1, x_2, \cdots\}$ for $\dot{R}_t$. |
| $P_1$ | Totally-ordered set of vertices, as one random path. |
| $P_2$ | Totally-ordered set of vertices, as the other random path. |
| $U$ | Set of vertices as mark by the two-path mechanism. |
| $z$ | The first intersection of $P_1$ and $P_2$. |
| $R_1$ | Set of reports on $P_1$'s out-edges elicited from $P_1$. |
| $R_2$ | Set of reports on $P_2$'s out-edges elicited from $P_2$. |
| $R_3$ | Set of reports on $P_1$ (or $P_2$)'s out-edges elicited from $P_2$ (or $P_1$). |

role and notes on the two-path mechanism and DG13. See Table 4.1 above for the notations relevant to the incentive mechanism with ex-post consensus.

### 4.1.1 Setup

Consider citations as a growing DAG $G_t = (V_t, E_t)$, where $V_t$ denotes the set of intellectual products that are synonymous with peers and individuals (Assumption 1.4.1), and $E_t \subseteq V_t \times V_t$ denotes their citation relationships in period $t$. Unlike the previous Chapter 3, this time $G_t$ may not be the "true" structure because anyone can add new intellectual products without the peer-review process (as with WWW). Accordingly, its state transition $(G_t)_{t=0}^{q}$ becomes an iterative process of accepting a proposal of new citations $\dot{G}_t = (\{k\} \cup V_k, E_k)$ and reviewing the true structure of $G_t$ as ex-post consensus-building, where $k$, $V_k$, and $E_k$ are the same as those in Chapter 3. For the sake of algorithmic description, we again assume that each period deals with only one $\dot{G}_t$, despite the importance of parallel consensus-building in P2P citation systems (Definition 1.2.2)[2].

[2]This assumption is just for the sake of convenience; the incentive mechanism with ex-post consensus can handle multiple $\dot{G}_t$s in parallel.

Review in this mechanism is thus not for $\dot{G}_t$ but for $G_t$; this is actually the process of random walk on $G_t$ itself, which runs once $\dot{G}_t$ is proposed and proceeds while asking each vertex about its own out-edges[3]. Here, the set of reviewers $\dot{C}_t = \{1, 2, \cdots\}$ is a group of all vertices in the (elicited) random walk[4], and the set of their reports $\dot{R}_t = \{r_1^{(1,i)}, r_1^{(1,j)}, \cdots, r_2^{(2,i)}, r_2^{(2,j)}, \cdots\}$ designates the existence of all out-edges of $\dot{C}_t$ (as a task) with binary signals $\{0, 1\}$[5]. We assume that reviewers first report their out-edges and then, to generate overlapping tasks, report the out-edges of other reviewers as well[6]. Namely, $\dot{R}_t$ may include the reports such as $r_1^{(2,i)}, r_1^{(2,j)}, \cdots$ and $r_2^{(1,i)}, r_2^{(1,j)}, \cdots$. After computing rewards $\dot{X}_t = \{x_1, x_2, \cdots\}$ based on $\dot{R}_t$, $G_t$ is finally updated to $G_{t+1}$, which always accepts the proposal $\dot{G}_t$ regardless of $\dot{R}_t$.

Therefore, the state transition from $G_t$ to $G_{t+1}$ can be summarized as follows:

Step 1: A new vertex $k$ proposes $\dot{G}_t$ to $G_t$,

Step 2: Select all vertices in the random walk as $\dot{C}_t$ (*reviewer assignment*),

Step 3: Collect reports on $\dot{C}_t$'s out-edges as $\dot{R}_t$,

Step 4: Compute $\dot{X}_t$ (*reward computation*),

Step 5: Update $G_t$ to $G_{t+1}$ which always accepts $\dot{G}_t$.

Figure 4.1 graphically depicts an example of the state transition by focusing on a subgraph with vertices $\{a, b, c, d, e, f, g\}$, where Figures 4.1a–4.1e correspond to the Steps 1–5; Figure 4.1b illustrates a random walk by the two-path mechanism that draws two independent random paths $c, d, e$ and $b, d, f$[7]. Intuitively, this mechanism introduces rewards into the ranking algorithm that performs a random walk while eliciting out-edges from each web page (who may strategically misreport its links). As will be detailed in Sections 4.1.3 and 4.1.4, the incentive mechanism with ex-post consensus uses the two-path mechanism for reviewer assignment (Step 2 and 3) and DG13 for reward computation (Step 4).

[3] The frequency of the random walk, once for every $\dot{G}_t$, is also just for the sake of convenience. We can run the random walk any number of times for $\dot{G}_t$ (see Section 4.2.1), or independently of $\dot{G}_t$.

[4] Due to this rule, we cannot determine the number of reviewers by the exogenous parameter $\lambda$.

[5] In other words, vertices report some out of all tasks $V_t \times V_t$, and it is equivalent to reporting the $i$-th column of the adjacency matrix $A(G_t)$ when a reviewer $i$ reports its own out-edges. However, there is another assumption on the treatment of 0 reports, as we will discuss in Section 4.1.4.

[6] See Sections 4.1.2 and 4.1.3 for details on how this incentive mechanism assign reviewers to the out-edges of other reviewers.

[7] Note that random path (Definition 2.2.3) is a specific type of random walk (Definition 2.2.1).

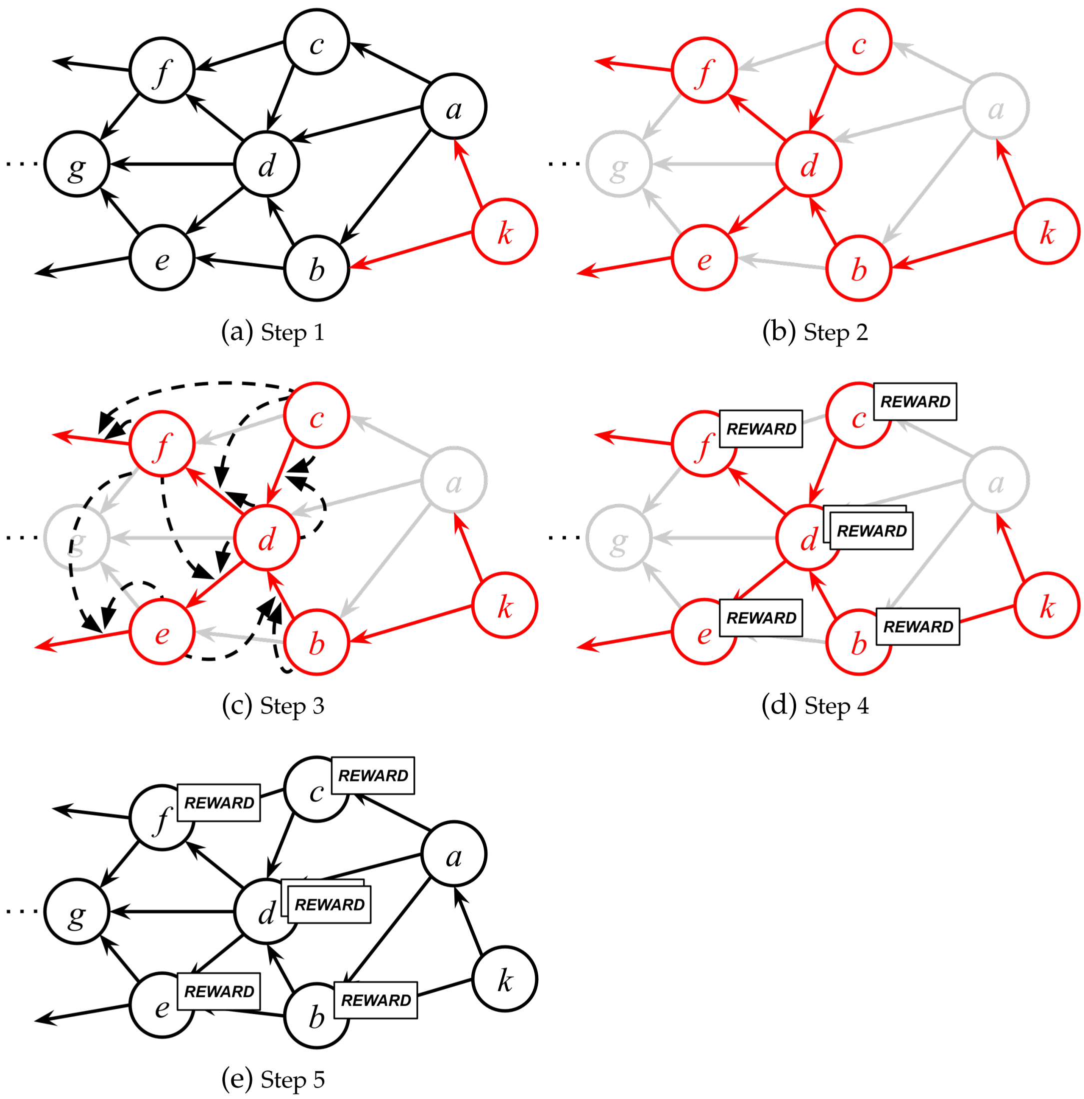



**FIGURE 4.1: Incentive Mechanism with Ex-post Consensus.** If we focus on a subgraph with vertices $\{a, b, c, d, e, f, g\}$, an example of the Steps 1–5 can be depicted as Figures 4.1a–4.1e. These figures assume the following state transition: in Step 1, $k$ proposes $\dot{G}_t = (\{k, a, b\}, \{(k, a), (k, b)\})$; in Step 2, the mechanism selects $\dot{C}_t = \{b, c, d, e, f\}$ as reviewers of their own (potential) out-edges; in Step 3, $\dot{C}_t$ reciprocally evaluates their out-edges with binary reports $\{0, 1\}$; in Step 4, $\dot{C}_t$ can receive rewards whose amount was computed from their reports; finally, in Step 5, $\dot{G}_t$ is accepted regardless of $\dot{C}_t$'s reports.

### 4.1.2 Pseudocode

Pseudocode allows for a detailed description of the state transition. Let us first confirm the two-path mechanism that the author quotes from Babichenko et al. (2018) as Algorithm 3 below. The gist of the two-path mechanism, as already mentioned in Section 2.2.2, is to regard an important vertex as the first intersection of the two independent random paths, which provides weak truthfulness (Definition 2.2.4) to

**Algorithm 3** The two-path mechanism (Babichenko et al., 2018)

$U \leftarrow \emptyset$
**while** $U \neq V$ **do**
  Pick $x \in V$ uniformly at random
  $P_1 \leftarrow$ random path starting at $x$
  Pick $y \in V$ uniformly at random
  $P_2 \leftarrow$ random path starting at $y$
  **if** $P_1 \cap P_2 = \emptyset$ **then**
    $U \leftarrow U \cup P_1 \cup P_2$
  **else**
    $z \leftarrow$ the first vertex in $P_1 \cup P_2$
    **if** $z \in U$ **then**
      **return** $\emptyset$
    **else**
      **return** $z$
    **end if**
  **end if**
**end while**

the selection process. In Algorithm 3, the two random paths $P_1$ and $P_2$ are the totally-ordered sets of vertices, which are drawn iteratively until they intersect or until all vertices in the network are marked. This mark, denoted by the set $U$, is attached to all vertices on which the two paths have passed when they do not intersect. The marked vertices will never be selected as an important vertex $z$[8]. This algorithm satisfies weak truthfulness because any vertex can no longer manipulate its probability to be the first intersection, at the point of reporting out-edges for drawing $P_1$ and $P_2$ (Proposition 2.2.1).

State transition in the incentive mechanism with ex-post consensus consists of Algorithms 4 and 5, which are an extension of Algorithm 3. Algorithm 4 represents the whole process of state transition, which covers graphs $\dot{G}_t, G_t, G_{t+1}$, reports $\dot{R}_t, R_1, R_2, R_3$, and rewards $\dot{X}_t$, by adding several new lines (gray highlights) to Algorithm 3. On the other hand, Algorithm 5 represents the specific part of report collection (i.e., Step 3) as $\textsc{Overlap}(P_1, P_2)$ function. To generate overlapping tasks, Algorithm 5 randomly pairs all vertices in $P_1 \cup P_2$ with a vertex in the other path, then lets all pairs $(i, j)$ review each other's out-edges. As will be explained in Section 4.1.3, this property aims to make the out-edge reporting indifferent to reviewer

[8]This is important because, if it were not for the setting, vertices could increase their probability of being the first intersection by strategically misreporting they have no out-edges and resetting the two-path drawing.

**Algorithm 4** State transition in the incentive mechanism with ex-post consensus

$G_t \leftarrow (V_t, E_t)$ ▷ newly added
$\dot{G}_t \leftarrow$ proposal of new citations ▷ newly added
$\dot{R}_t \leftarrow \emptyset$ ▷ newly added
$U \leftarrow \emptyset$
**while** $U \neq V_t$ **do**
    Pick $x \in V_t$ uniformly at random
    $P_1 \leftarrow$ random path starting at $x$
    $R_1 \leftarrow$ out-edge reports for $P_1$ ▷ newly added
    Pick $y \in V_t$ uniformly at random
    $P_2 \leftarrow$ random path starting at $y$
    $R_2 \leftarrow$ out-edge reports for $P_2$ ▷ newly added
    $R_3 \leftarrow \text{OVERLAP}(P_1, P_2)$ ▷ newly added
    $\dot{R}_t \leftarrow \dot{R}_t \cup R_1 \cup R_2 \cup R_3$ ▷ newly added
    **if** $P_1 \cap P_2 = \emptyset$ **then**
        $U \leftarrow U \cup P_1 \cup P_2$
    **else**
        $z \leftarrow$ the first vertex in $P_1 \cup P_2$
        **if** $z \in U$ **then**
            **return** $\emptyset$
        **else**
            **return** $z$
        **end if**
    **end if**
**end while**
Compute rewards $\dot{X}_t$ with $\dot{R}_t$ ▷ Use DG13, newly added
**return** $\dot{X}_t$ ▷ newly added
$G_{t+1} \leftarrow G_t \cup \dot{G}_t$ ▷ newly added
**return** $G_{t+1}$ ▷ newly added

**Algorithm 5** Generating Overlapping Tasks

**function** $\text{OVERLAP}(P_1, P_2)$
    $R_3 \leftarrow \emptyset$
    **if** $|P_1| \geq |P_2|$ **then**
        $(A, B) \leftarrow (P_1, P_2)$
    **else**
        $(A, B) \leftarrow (P_2, P_1)$
    **end if**
    Make surjection $f : A \twoheadrightarrow B$ at random, subject to $i \neq f(i)$ for all $i \in A$
    **for** all $i \in A$ **do**
        $R_i \leftarrow$ reports by $f(i) = j$ for the out-edges of $i$
        $R_j \leftarrow$ reports by $i$ for the out-edges of $f(i) = j$
        $R_3 \leftarrow R_3 \cup R_i \cup R_j$
    **end for**
    **return** $R_3$
**end function**

assignment, thereby removing bias from random-path drawings.

### 4.1.3 Role and Notes on the Two-path Mechanism

In these algorithms, the two-path mechanism can solve reviewer assignment because it assigns reviewers $\dot{C}_t$ (under the weak truthfulness) while ensuring condition (i) of Assumption 1.4.3—*peers can be reviewers more often as they get higher PR scores*. Algorithm 3 implies that $\dot{C}_t$ is the set of vertices on all random paths drawn in period $t$[9]. As we will confirm experimentally in Section 4.2.1, such a two-path reviewer assignment can maintain a positive correlation with PR scores that are derived from random walks in general, rather than random paths.

Furthermore, the two-path mechanism has an important role in removing bias from the reviewer assignment. If we selected $\dot{C}_t$ from a simple random walk, for example, out-edge reports would affect reviewer candidates of the reporters themselves (i.e., vertices can manipulate their reviewers to some extent, through out-edge reports). The two-path mechanism is effective for this bias because, under the additional setting (Algorithm 5) of selecting reviewers from the other path, vertices can no longer manipulate their reviewers at the point of reporting out-edges. Thus, the two-path mechanism has a synergy with DG13 in terms of removing bias.

On the other hand, there exist two notes on the application of the two-path mechanism. First, the two-path mechanism implicitly assumes rewards other than $\dot{X}_t$. As mentioned in Section 2.2.2, we assumed that vertices, who aim to be selected by the mechanism, can obtain (some fixed amount of) rewards if they become $z$. Nevertheless, Chapter 4 can focus only on $\dot{X}_t$ because the weak truthfulness for $z$ still remains. In other words, out-edge reports do not affect the expected amount of rewards for $z$, but only affect that of $\dot{X}_t$ by DG13. Second, the two-path mechanism, including Algorithm 5, leads to an asymmetry in the number of reviews per reward computation. For example, if $|P_1| = 1,000$ and $|P_2| = 1$, the vertex in $P_2$ has to review $1,000$ times for one opportunity of reward computation[10]. Resolving this asymmetry is one of the future studies; however, the author believes that it can be mitigated by using more than two random-paths for Algorithm 5[11].

[9] $\dot{C}_t$ is not always $P_1 \cup P_2$, given that two random paths may be drawn more than once in period $t$.
[10] See the next Section 4.1.4 for the detail of reward computation by DG13.
[11] This would be possible by running both Algorithm 5 and the reward computation, once a given number of periods (e.g., 100 periods) rather than every period.

### 4.1.4 Role and Notes on DG13

In these algorithms, DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers $\dot{C}_t$, while ensuring condition (ii) of Assumption 1.4.3—*peers can maximize the amount of expected rewards per review by always reporting true beliefs*. To confirm its specific usage, let us now turn to equation 2.3 on DG13, presented in Section 2.3.1.

$$x_i^{m^*} = \delta(r_i^{m^*}, r_j^{m^*}) - \delta(r_i^n, r_j^{n'}), \tag{4.1}$$

where, for a reviewer $i \in \dot{C}_t$, the incentive mechanism randomly selects one of its out-edge reports (including those $i$ did for other reviewers) as $r_i^{m^*}$, another reviewer sharing the overlapping task $m^*$ as $j \in \dot{C}_t$, and two other out-edge reports corresponding to $i$ and $j$ as $r_i^n$ and $r_j^{n'}$, respectively[12]. DG13 then computes rewards for all reviewers by making such random selection for each vertex in $\dot{C}_t$. As we will confirm experimentally in Section 4.2.2, such DG13-based reward computation can elicit true beliefs, even under the two-path reviewer assignment.

Furthermore, DG13 has an important role in enhancing the solution concept of the two-path mechanism from weak to strong truthfulness. Strictly speaking, weak truthfulness does not solve free-riding and strategic misreporting because it makes the truthful strategy and other strategies indifferent (i.e., vertices have no clear incentive to report true out-edges). DG13 is effective for this problem as an additional layer, where reporting itself generates $\dot{X}_t$, and the truthful strategy maximizes their expected amount. Thus, DG13 has a synergy with the two-path mechanism in terms of enhancing the solution concept.

On the other hand, there exist two notes on the application of DG13. First, DG13 leverages Algorithm 5 to extract 0 reports from $\dot{C}_t$. One of the remaining problems is what to report as 0 in the process of random walk on $G_t$ where vertices report their out-edges as 1. We address this problem by assuming that, whenever a randomly-selected $j \in P_2$ (or $P_1$) reviews out-edges of $i \in P_1$ (or $P_2$) in Algorithm 5, $i$ and $j$ also make $r_i^{(i,x)}$ for $\{x \mid x = j \vee r_j^{(j,x)} = 1 \vee r_j^{(i,x)} = 1\}$ and $r_j^{(i,x)}$

[12]Thus, Algorithm 5 implies (i) $i$ and $j$ are always in different random paths, and (ii) $m^*$ is an out-edge from either $i$ or $j$.

for $\{x \mid r_i^{(i,x)} = 1 \vee r_i^{(i,x)} = 0\}$, respectively[13]. For example, if $c \in \{c, d, e\}$ and $d \in \{b, d, f\}$ in Figure 4.1c are $i$ and $j$, out-edge reports on their own $r_c^{(c,d)} = r_c^{(c,f)} = 1$ and $r_d^{(d,e)} = r_d^{(d,f)} = r_d^{(d,g)} = 1$ make $r_c^{(c,e)} = r_c^{(c,g)} = 0$; furthermore, if $d$ reports $r_d^{(c,b)} = 1$ for $c$'s out-edges, it makes $r_c^{(c,b)} = 0$ and $r_d^{(c,d)} = r_d^{(c,e)} = r_d^{(c,f)} = r_d^{(c,g)} = 0$. Note here that, all reports, except $r_d^{(d,e)} = r_d^{(d,f)} = r_d^{(d,g)} = 1$, are for overlapping tasks, whether they are 0 or 1. Given that Algorithm 5 considers all reviewers (including $j$) in $P_1 \cup P_2$ to be $i$, this implies that all out-edge reports by $\dot{C}_t$ will be candidates of $r_i^{m^*}$. Second, DG13 computes $\dot{X}_t$ per the iteration of two-path drawings (i.e., per the two-path mechanism). This is important because $\dot{C}_t$ could manipulate their (expected) opportunity of reward computation, if we computed rewards per the two-path drawing (i.e., compute $\dot{X}_t$ in the while statement of Algorithm 4) or per the reviewer assignment to generate overlapping tasks (i.e., compute $\dot{X}_t$ in Algorithm 5). Specifically, the former is vulnerable to the strategic misreporting of $P_1 \cap P_2 = \emptyset$ (to repeat the two-path drawing)[14]. At the same time, the latter is vulnerable to that of $|P_1| < |P_2|$ or $|P_1| > |P_2|$ (to make asymmetry in the frequency of reviewer assignment)[15]. Reward computation is thus for the iteration of two-path drawings, not for each two-path drawing or reviewer assignment[16].

We detailed algorithms for the incentive mechanism with ex-post consensus in this section and the role and notes on the two-path and DG13. To summarize, in the algorithms (Figure 4.1, Algorithms 4 and 5), the two-path mechanism can solve reviewer assignment because it assigns appropriate reviewers (under the weak truthfulness) while ensuring condition (i) of Assumption 1.4.3, and DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers while ensuring condition (ii) of Assumption 1.4.3, respectively.

[13] Although $i$'s own out-edge reports are equivalent to reporting $i$-th column of $A(G_t)$, we need this assumption because (i) the incentive mechanism cannot capture the precise number of all vertices $|V_t(G_t)|$ for a large network such as WWW, and (ii) too many 0 reports would lower $\mathbb{E}[X_i]$ to an impractical level.

[14] This type of strategic manipulation can occur when the two-path mechanism selects $z$ as well, which is the reason why Babichenko et al. (2018) introduces the mark $U$ in Algorithm 3.

[15] This type of strategic manipulation and the possibility of having to review in many times for one opportunity of reward computation (mentioned in Section 4.1.3) are two sides of the same coin.

[16] Since this is the matter of timing, we can compute rewards multiple times (e.g., 10 or 100 times) for each reviewer to obtain $\dot{X}_t$, as long as the process is per the iteration of two-path drawings. In addition, although not treated here for simplicity, DG13 could cover the original incentive of the two-path mechanism as well by making the number of reward computations for $z$ twice that of other reviewers.

## 4.2 Experimental Studies

This section experimentally confirms that the incentive mechanism with ex-post consensus ensures the two conditions of Assumption 1.4.3, thereby supporting the claims made in the previous section. For real-world citation data on web pages (Figure 2.4), we had two-step experiments which first use only the two-path mechanism to examine condition (i) for reviewer assignment, then incorporate DG13 to examine condition (ii) for free-riding and strategic misreporting. All materials for the experiments are available in the Github repository[17].

### 4.2.1 Experiments for Reviewer Assignment

The first experiment computes the correlation between the frequency distribution of reviewer assignment and the PR score for all vertices. This is important because the two-path reviewer assignment differs from the PR score in that it is (i) based on two random-paths (without dumping factor), (ii) done period-by-period, and (iii) stochastic rather than deterministic[18]. For the aforementioned real-world citation data (Figure 2.4), the former can be derived by applying the two-path mechanism along with the state transition $(G_t)_{t=0}^{1000}$, while the latter can be derived by applying the PR algorithm to the last state $G_{1000}$[19]. To compare the different number of reviewer assignments, we here generalized the number of two-path mechanisms executed in one period from 1 to an exogenous parameter $\beta$[20]. This experiment derived 200 patterns—10 times for every $\beta = \{1, 2, \cdots, 20\}$ cases—of reviewer assignment from web pages, then computed Spearman's rank correlation coefficients between their frequency distributions and the (deterministic) PR score[21].

Figure 4.2 represents the experimental results regarding the trend of 200 correlation coefficients computed from web pages. Box plots show the median value as blue lines, 25/75 percentile as boxes, pseudo-maximum/minimum value as bars,

[17] https://github.com/knskito/materials_thesis

[18] Note that this experiment is different from that of Babichenko et al. (2018), which confirms the relationship between the frequency distribution to be $z$ and an importance score based on the PR score (see footnote 17 of Section 2.2.2 for its details).

[19] We set the dumping factor $\alpha = 0.15$ in the PR algorithm.

[20] For example, if $\beta = 10$, the state transition executes the two-path mechanism a total of $1,000 * 10$ times; this extension does not affect the gist of the incentive mechanism with ex-post consensus.

[21] This experiment cannot use Pearson correlation coefficients because neither the frequency distribution of reviewer assignment nor the PR score for all vertices follows normal distributions. We confirmed this by using Shapiro-Wilk test (Shapiro & Wilk, 1965).

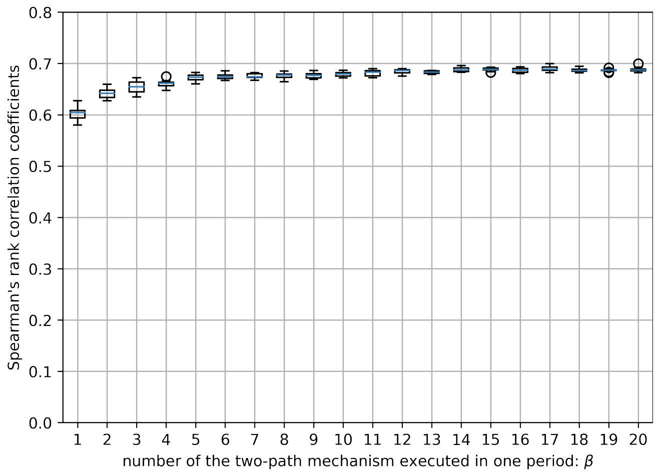



**FIGURE 4.2: Reviewer Assignment in the Incentive Mechanism with Ex-post Consensus.** The first experiment computed Spearman's rank correlation coefficients between the frequency distribution of reviewer assignment in $(G_t)_{t=0}^{1000}$ and the PR score in $G_{1000}$. The box plot represents that all 200 (10 times for every $\beta = \{1, 2, \cdots, 20\}$ cases) coefficients are moderately correlated. This result—the moderate positive correlation between the two-path reviewer assignment and the PR score—supports condition (i) of Assumption 1.4.3.

and outliers as circles. We can see that most of the 200 correlation coefficients are within the range of 0.6 to 0.7, indicating that they are moderately correlated. This result—the moderate positive correlation between the two-path reviewer assignment and the PR score—supports condition (i) of Assumption 1.4.3[22]. Moreover, this figure shows that correlation coefficients begin to converge around a little below 0.7 when $\beta$ exceeds 12.

### 4.2.2 Experiments for Free-riding and Strategic Misreporting

The second experiment computes $\mathbb{E}[X_i]$—the amount of $i$'s expected rewards per review—by incorporating DG13 into the first experiment. This is important because we need to confirm whether the DG13-based reward computation can ensure strong truthfulness (Definition 2.3.1) even in conjunction with the two-path reviewer assignment. To compare various patterns, this experiment stochastically allocated strategies to all vertices in advance while keeping their graph structure (i.e., signals emitted from $V_t \times V_t$) as the original. We assume that all vertices, as reviewer, take either the truthful strategy $\boldsymbol{P}(\sigma_i^*) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ or uninformative strategies $\boldsymbol{P}(\sigma_i) = \left(\begin{smallmatrix} \gamma & \gamma \\ 1-\gamma & 1-\gamma \end{smallmatrix}\right)$. Given the low average out-degree of Figure 2.4 (approx. 2.7

[22]Let us recall that PR scores in this Section 4.2.1 consider citations as DAGs, while those in Section 3.2.1 consider citations as undirected graphs. We found that, in an additional experiment computing the PR score from web pages as undirected graphs, all correlation coefficients are around 0.3; thus, there is a weak positive correlation even in this case.

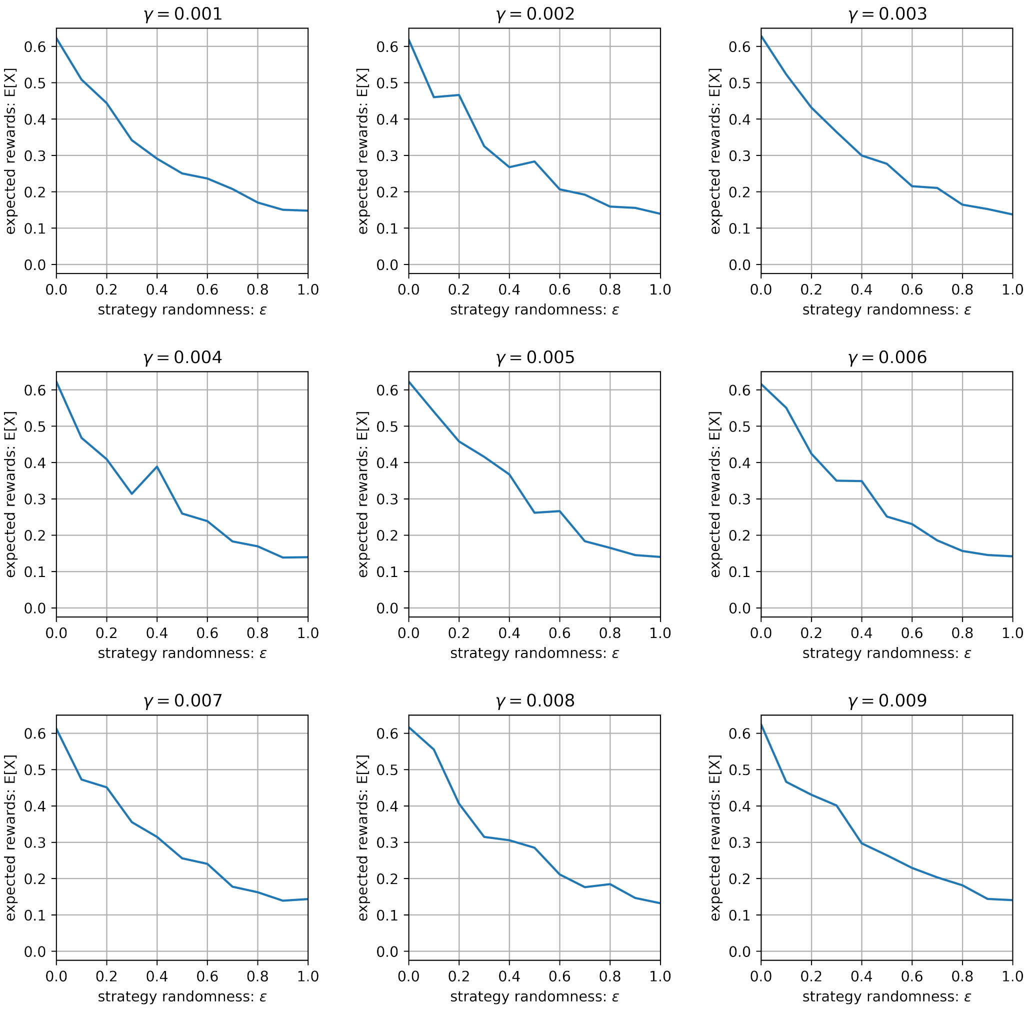



**FIGURE 4.3: Free-riding and Strategic Misreporting in the Incentive Mechanism with Ex-post Consensus.** The second experiment computed $\mathbb{E}[X_i]$ resulting from the state transition $(G_t)_{t=0}^{1000}$, for each of the 99 patterns with different strategy-strategy allocation pairs. The nine graphs covering web pages represent that, for all uninformative strategies with $\gamma = \{0.001, 0.002, \cdots, 0.009\}$, $\mathbb{E}[X_i]$ is maximized when all vertices take the truthful strategy (i.e., $\epsilon = 0.0$). This result—the maximized expected rewards under truthful strategies—supports condition (ii) of Assumption 1.4.3.

for 1,439 vertices), we considered $\gamma = \{0.001, 0.002, \cdots, 0.009\}$ cases[23]. This experiment allocated one of the nine uninformative strategies according to 11 rules with different randomness $\epsilon = \{0.0, 0.1, \cdots, 1.0\}$, where $\epsilon$ is an exogenous parameter denoting the probability of allocating the strategy. That is, all vertices take the truthful strategy if $\epsilon = 0.0$ and one of the uninformative strategies if $\epsilon = 1.0$. Finally, we computed $\mathbb{E}[X_i]$ resulting from the state transition $(G_t)_{t=0}^{1000}$, for each of the 99 patterns with different strategy-strategy allocation pairs $\{0.001, 0.002, \cdots, 0.009\} \times \{0.0, 0.1, \cdots, 1.0\}$. Here, $\beta$ is fixed as 1.

[23]Note again that tasks for a vertex $i$ are $i$-th and its peer $j$-th columns of $\boldsymbol{A}(G_t)$.

Figure 4.3 represents the experimental results, where each of the nine graphs depicts the results for a different $\gamma$; the blue line shows the computed $\mathbb{E}[X_i]$ on web pages. We can see that, for all uninformative strategies with different $\gamma$, $\mathbb{E}[X_i]$ is maximized when all vertices take the truthful strategy (i.e., $\epsilon = 0.0$). This result—the maximized expected rewards under truthful strategies—supports condition (ii) of Assumption 1.4.3.

In this section, we confirmed that the incentive mechanism with ex-post consensus could ensure the two conditions of Assumption 1.4.3, through two-step experiments. To summarize, the first experiment (on the two-path reviewer assignment) yielded results supporting condition (i) for reviewer assignment, and the second experiment (on the DG13-based reward computation) yielded results supporting condition (ii) for free-riding and strategic misreporting, respectively.

## 4.3 Summary of This Chapter

This chapter introduced the incentive mechanism with *ex-post* consensus, which (i) covers citations without peer-review (e.g., those on web pages) and (ii) consists of the two-path mechanisms (Babichenko et al., 2018) and DG13 (Dasgupta & Ghosh, 2013). We can summarize all arguments in Chapter 4 as answers to the questions presented at the beginning:

- *Why can the two-path mechanism solve reviewer assignment?* (4.1.3) — The two-path mechanism can solve reviewer assignment because it assigns appropriate reviewers (under the weak truthfulness), while ensuring condition (i) of Assumption 1.4.3,
- *Why can DG13 solve free-riding and strategic misreporting?* (4.1.4) — DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers, while ensuring condition (ii) of Assumption 1.4.3.

These answers are supported by the two-step experiments (4.2) as well. Based on the above, the next Chapter 5 discusses potential applications and future studies of the two incentive mechanisms we have confirmed in Chapters 3 and 4.

# Chapter 5

# Discussion

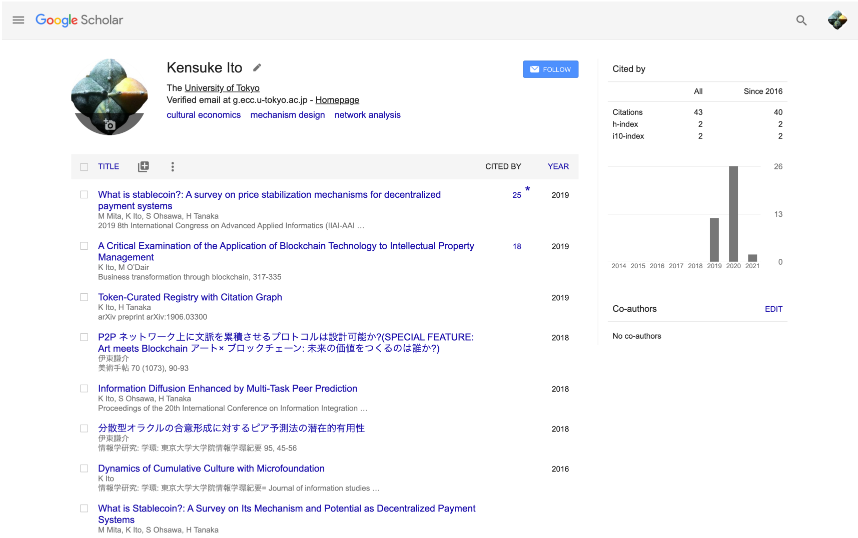
We have thus far confirmed the details of *incentive mechanism with ex-ante consensus* (Chapter 3) and *incentive mechanism with ex-post consensus* (Chapter 4). In this Chapter 5, the author discusses potential applications and future studies of these two incentive mechanisms, while answering the following questions:

- *What is the potential application of the two incentive mechanisms?* (5.1)
- *What is the social contribution of the RQ?* (5.1)
- *What is the future study of the two incentive mechanisms?* (5.2)

Here, potential applications lead to the social contribution of the RQ as well, and future studies are summarized as open questions (5.2.3), except for those related to Assumptions 1.4.1 and 1.4.2.

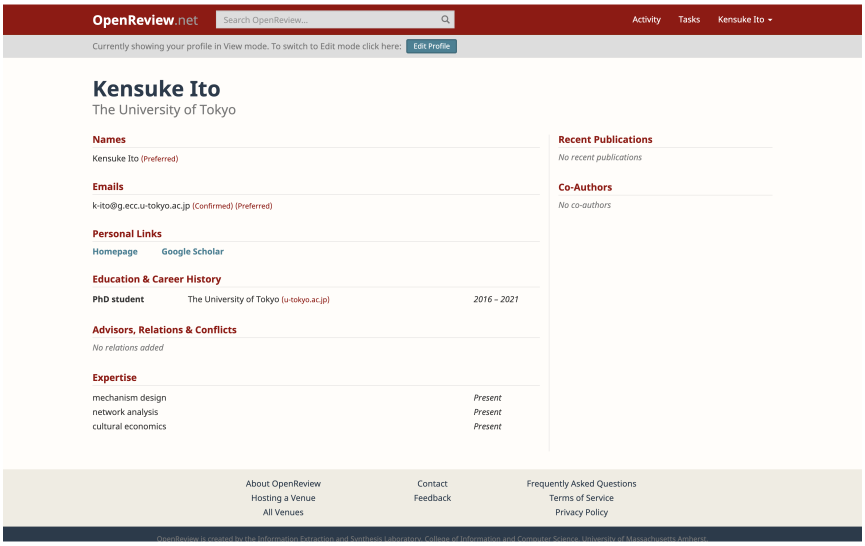
## 5.1 Potential Applications

Because of their simplicity, the two incentive mechanisms have a variety of potential applications for intellectual products (as will be noted in Section 5.1.2, we can apply the two incentive mechanisms to almost all intellectual products). This section discusses such potential applications and the (resulting) social contribution of the RQ, according to the three categories we have consistently used: scientific publications, patents, and web pages.




(a) Google Scholar

*Source*: bQTxPTMAAAAJ-profile (https://scholar.google.com/citations?user=bQTxPTMAAAAJ&hl=en, accessed February 15, 2021)




(b) OpenReview.net

*Source*: Profile (https://openreview.net/profile?id=~Kensuke_Ito1, accessed February 15, 2021)

**FIGURE 5.1: Systems Related to the Potential Application to Scientific Publications.** *Google Scholar* (Figure 5.1a) is an article search system that automatically collects publications and citations of registered authors, and *OpenReview.net* (Figure 5.1b) is one of the crowd-sourced reviews for efficient conference managements. The potential application to scientific publications would functionally be like a hybrid of these (non-P2P) systems in that peers can check not only their publications and citations but also their reviews and rewards.

### 5.1.1 The Potential Application to Scientific Publications

For scientific publications, the potential application is to make their crowd-sourcing robust and productive. As mentioned in Section 1.2.2, scientific publications have leveraged *crowd-sourced review* (Ford, 2013) to shorting the time between submission to publication, but the centralized authority (i.e, editorial board) for consensus-building has restricted its robustness and productivity. The two incentive mechanisms, especially the one with ex-ante consensus, can improve this situation as a component of P2P citation systems (Definition 1.2.2). Figure 5.1 depicts related systems—*Google Scholar*[1] (Figure 5.1a, e.g., Noruzi, 2005) is an article search system that automatically collects publications and citations of registered authors, and *OpenReview.net*[2] (Figure 5.1b, Soergel et al., 2013) is one of the crowd-sourced reviews for efficient conference managements. The potential application to scientific publications would functionally be like a hybrid of these (non-P2P) systems in that peers can check their publications and citations and their reviews and rewards.

In addition to the robustness and productivity as a P2P citation system, this application leads to the social contribution of the RQ: developing a new reward source for open-access intellectual products[3]. Scientific publications are open-access when their authors use some pre-print server (e.g., arXiv) or open-access journal[4], but this is implicitly on the premise that authors earn grants from external sources (e.g., universities, research institutes). In this context, our proposal—rewards not directly from intellectual products but indirectly from the consensus-building on their citation relationships—can be interpreted as an alternative scheme to existing grants. Furthermore, it is worth noting that the incentive mechanism with ex-ante consensus has a synergy with scientific communities because (i) rewards can be a sufficient incentive (even without Assumption 1.4.2) as an evaluation measure for researchers, like *h-index* (Hirsch, 2005), and (ii) peer-review as an unpaid task has already been

[1] https://scholar.google.com/, accessed February 15, 2021.

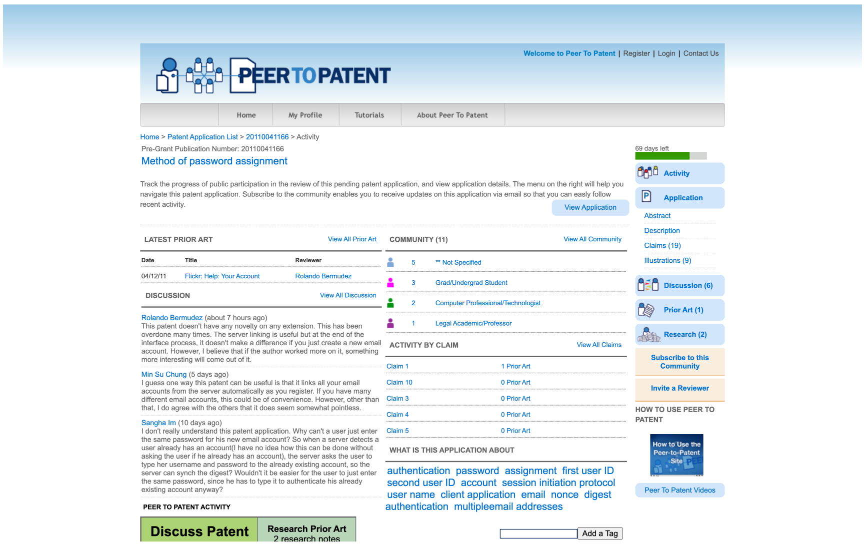
[2] https://openreview.net/, accessed February 15, 2021.

[3] The term open-access means "free and unrestricted online availability" (Cuplinskas et al., 2002, para. 2), which was originally for scientific publications (e.g., Albert, 2006; Xia et al., 2012). On the other hand, this thesis uses the following broader concept provided by *United Nations Educational, Scientific and Cultural Organization* (UNESCO):

> Open access (OA) means free access to information and unrestricted use of electronic resources for everyone. Any kind of digital content can be OA, from texts and data to software, audio, video, and multi-media. While most of these are related to text only, a growing number are integrating text with images, data, and executable code. OA can also apply to non-scholarly content, like music, movies, and novels. (UNESCO, n.d. para. 2)

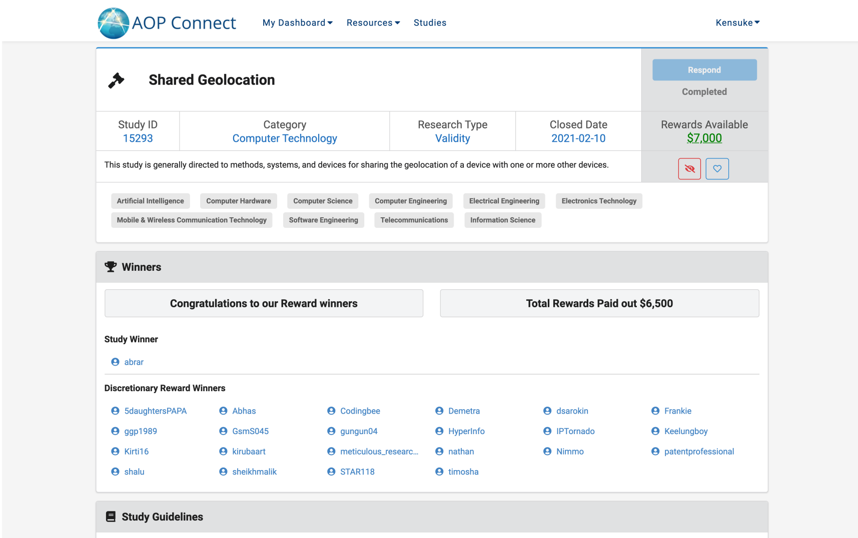
See, for example, Bailey Jr (2007) for the history and other definitions of open-access.

[4] We can search for specific examples of open-access journals at *Directory of Open Access Journals* (DOAJ), https://doaj.org/, accessed February 5, 2021.




(a) Peer To Patent

*Source*: Pre-Grant Publication Number: 20110041166 (https://web.archive.org/web/20110412153728/http://www.peertopatent.org/patent/20110041166/activity, accessed February 15, 2021).




(b) Article One Partners

*Source*: 15293-shared-geolocation in AOP connect (https://app.articleonepartners.com/study/index/15293-shared-geolocation, accessed February 15, 2021)

**FIGURE 5.2: Systems Related to the Potential Application to Patents.** *Peer To Patent* (Figure 5.2a) is one of the crowd-sourced prior art searches that accepts voluntary review from anyone as a social experiment, and *Article One Partners* (Figure 5.2b) is another example that distributes monetary rewards from the client patent office to the validated-reviewers as a business. The potential application to patents would functionally be like a hybrid of these (non-P2P) systems in that peers can earn rewards via patent review without any validation from centralized authority.

called into question (e.g., Engers & Gans, 1998; Gasparyan et al., 2015; Smith, 2006).

### 5.1.2 The Potential Application to Patents

For patents, as with scientific publications, the potential application is to make their crowd-sourcing robust and productive. As mentioned in Section 1.2.2, patents have leveraged *crowd-sourced prior art search* (Ghafele et al., 2011) to facilitate the review of submitted inventions, but the centralized authority (i.e., patent examiner) for consensus-building has restricted its robustness and productivity. The two incentive mechanisms, especially the one with ex-ante consensus, can improve this situation as a component of P2P citation systems. Figure 5.2 depicts related systems—*Peer To Patent*[5] (Figure 5.2a, Bestor & Hamp, 2010; Noveck, 2006) is one of the crowd-sourced prior art searches that accepts voluntary review from anyone as a social experiment, and *Article One Partners*[6] (Figure 5.2b, Malone, 2011) is another example that distributes monetary rewards from the client patent office to the validated-reviewers as a business. The potential application to patents would functionally be like a hybrid of these (non-P2P) systems in that peers can earn rewards via patent review without any validation from centralized authorities.

In addition to the robustness and productivity as a P2P citation system, this application again leads to the social contribution of the RQ: developing a new reward source for open-access intellectual products. Patented inventions are open-access once accepted by examiners, but this is on the premise that inventors can earn royalties from everyone who uses their patents for new inventions[7]. In this context, our proposal—rewards not directly from intellectual products but indirectly from the consensus-building on their citation relationships—can be interpreted as an alternative scheme to existing royalties. Furthermore, it is worth noting that *Creative Commons* (CC; Lessig, 2004) extends this new reward source to other open-access intellectual products that do not use citations. CC is "a set of various licenses that allow people to share their copyrighted work to be copied, edited, built upon, etc., while retaining the copyright to the original work (often used attributively)" (Dictionary.com, n.d.). It has been attached to a variety of intellectual products, such as

[5] http://www.peertopatent.org/, accessed February 15, 2021.

[6] https://app.articleonepartners.com/index, accessed February 15, 2021.

[7] The impact of patents on innovation has been studied extensively, especially from an economic perspective. See, for example, Cimoli et al. (2011), Dosi, Stiglitz, et al. (2014) for preceding studies.

*Source*: Wikidata (https://www.wikidata.org/wiki/Q6905942, accessed February 12, 2021).

**FIGURE 5.3: CC BY-SA License for Two Incentive Mechanisms.** CC BY-SA license is one of the licenses CC currently offers that allows everyone to modify and distribute the intellectual product, on condition that they attach both an appropriate credit (Attribution) and the same CC BY-SA license (ShareAlike). If we treated this credit as a citation, the two incentive mechanisms could be applied to almost all intellectual products.

music, movies, novels, and source codes[8]. Figure 5.3 depicts CC BY-SA license—, one of the licenses CC currently offers—that allows everyone to modify and distribute the intellectual product, on condition that they attach both an appropriate credit (Attribution) and the same CC BY-SA license (ShareAlike). If we treated this credit as a citation[9], CC BY-SA licenses would have the opportunity to earn rewards from the consensus-building on their citation relationships. This implies we can apply the two incentive mechanisms to almost all intellectual products.

### 5.1.3 The Potential Application to Web Pages

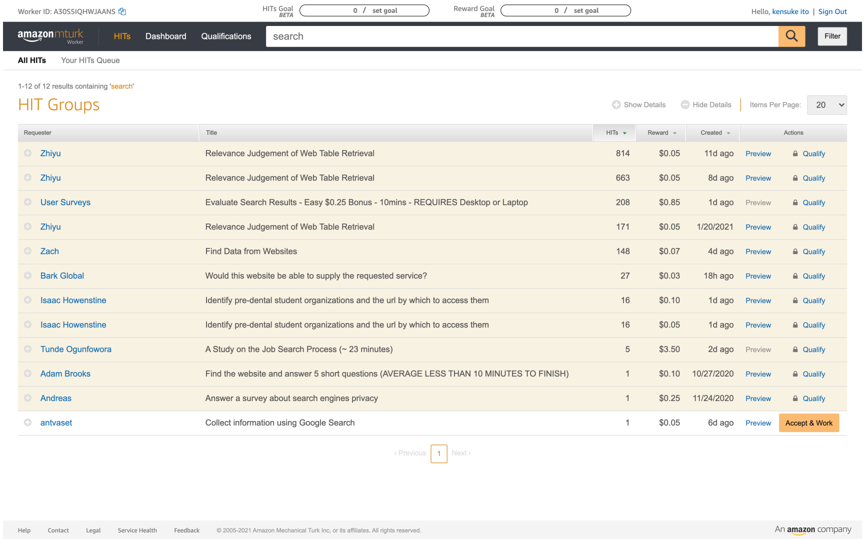
For web pages, as with scientific papers and patents, the potential application is to make their crowd-sourcing robust and productive. As mentioned in Section 1.2.2, web pages have leveraged *crowd-sourced human-based computing* (Wightman, 2010) to enhance the information retrieval from WWW, but the centralized authority (i.e, search engine) for consensus-building has restricted its robustness and productivity. The two incentive mechanisms, especially the one with ex-post consensus, can improve this situation as a component of P2P citation systems. Figure 5.4 depicts related systems—*Amazon Mechanical Turk*[10] (Figure 5.4a, Fort et al., 2011) is a platform for crowd-sourced human-based computing that allows users to outsource a variety of human intelligence tasks (e.g., assessing the relevance between two web pages) for a small fee, and *Brave*[11] (Figure 5.4b, Brave, 2021) is a web browser with

[8] https://creativecommons.org/, accessed February 7, 2021.

[9] This analogy is natural because CC assumes that the appropriate credit at least includes a link to the license and a statement of whether changes were made, according to the original explanation (https://creativecommons.org/licenses/by-sa/3.0/deed.en, accessed February 20, 2021).

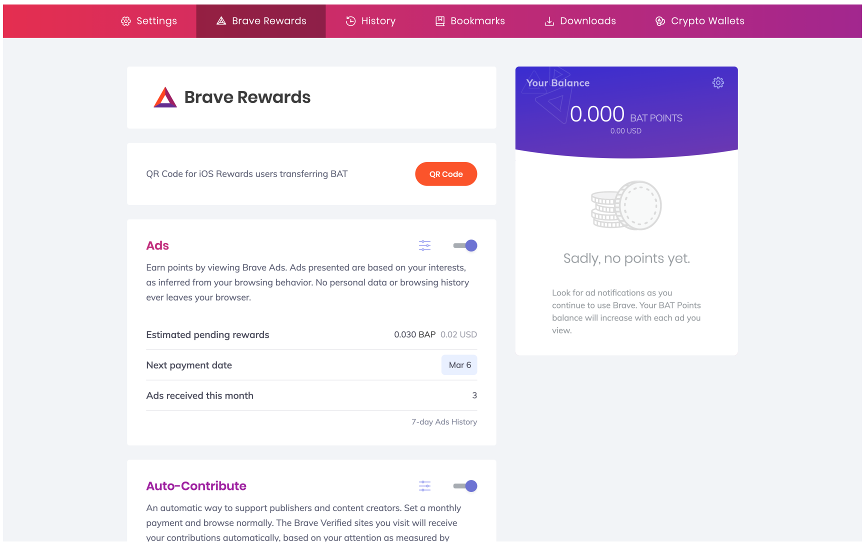
[10] https://www.mturk.com/, accessed February 15, 2021.

[11] https://brave.com/, accessed February 15, 2021.




(a) Amazon Mechanical Turk

*Source*: HIT-groups (https://worker.mturk.com/?filters%5Bsearch_term%5D=search&page_size=20&page_number=1&sort=num_hits_desc&filters%5Bmin_reward%5D=0.01, accessed February 15, 2021)




(b) Brave

*Source*: Brave Rewards (brave://rewards/ in Brave web browser retrieved from https://brave.com/, accessed February 15, 2021)

**FIGURE 5.4: Systems Related to the Potential Application to Web Pages.** *Amazon Mechanical Turk* (Figure 5.4a) is a platform for crowd-sourced human-based computing that allows users to outsource a variety of human intelligence tasks for a small fee, and *Brave* (Figure 5.4b) is a web browser with a built-in wallet that allows users to earn reward tokens in exchange for replacing existing advertisements with different ones. The potential application to web pages would functionally be like a hybrid of these (non-P2P) systems in that peers can earn reward tokens from human intelligence tasks.

a built-in wallet that allows users to earn reward tokens in exchange for replacing existing advertisements with different ones. The potential application to web pages would functionally be like a hybrid of these (non-P2P) systems in that peers can earn reward tokens from the consensus-building on citation relationships as a human intelligence task.

In addition to the robustness and productivity as a P2P citation system, this application again leads to the social contribution of the RQ: developing a new reward source for open-access intellectual products. Web pages are open-access unless administrators impose some restriction on their viewing, but this is often on the premise that administrators can earn advertising revenue from web pages[12]. In this context, our proposal—rewards not directly from intellectual products but indirectly from the consensus-building on their citation relationships—can be interpreted as an alternative scheme to existing advertisements. Furthermore, it is worth noting that this new reward source may contribute to attempts to make information on web pages machine-readable, such as *Semantic Web* (Berners-Lee et al., 2001) and *Linked Open Data* (Berners-Lee, 2006)[13]. Despite the promise of better information retrieval, these attempts are currently not very popular because there is little incentive (relative to the cost) for administrators to add high-quality machine-readable descriptions to web pages and hyperlinks (e.g., Cuel et al., 2011; Hendler & Berners-Lee, 2010; Simperl et al., 2013). In contrast, the new reward source could facilitate their widespread adoption if the consensus-building got to cover not only citation relationships but also their semantics (e.g., does a hyperlink from one lab's web page to a personal web page mean "student" or "professor"?). How to reflect citation semantics on consensus-buildings of the two incentive mechanisms is one of the future studies for this thesis (see Section 5.2.3)[14].

[12]See Liu-Thompkins (2019), Ratliff and Rubinfeld (2010) for the history and preceding studies on online advertising.

[13]Specifically, "The Semantic Web is not a separate Web but an extension of current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee et al., 2001, p. 36), and "Linked Open Data (LOD) is Linked Data which is released under an open licence, which does not impede its reuse for free" (Berners-Lee, 2006, para. 44, added in 2010), where Linked Data represents a concept that extends the Semantic Web from documents to data in general. Semantic Web and LOD are often categorized as *Web 3.0* (e.g., Naik & Shivalingaiah, 2008).

[14]It is also worth noting that the aforementioned CC licenses indicate their requirements (e.g., Attribution, ShareAlike) not only with visual icons but also with machine-readable descriptions. In other words, there is a close relationship between the use of CC to expand the scope of two incentive mechanisms and the widespread adoption of the Semantic Web.

In this section, we discussed the potential applications of the two incentive mechanisms and the (resulting) social contribution of the RQ. To summarize, the potential application is to make crowd-sourcing (for scientific publications, patents, and web pages) robust and productive as a component of the P2P citation system. This attempt has the social contribution of developing a new reward source for open-access intellectual products as an alternative scheme to grants, royalties, and advertisements.

## 5.2 Future Studies

On the other hand, we need further studies to implement such potential applications, mainly due to strong Assumptions 1.4.1 and 1.4.2. This section first considers how to relax these two assumptions, then summarizes other remaining studies as open questions.

### 5.2.1 Relaxing One-to-one Correspondence (Assumption 1.4.1)

One of the main future studies is to relax Assumption 1.4.1—*In P2P citation systems (Definition 1.2.2), individuals-to-peers, peers-to-products, and thus individuals-to-products are all one-to-one correspondence*. This is important because *spamming* and *Sybil attack* are critical risks for potential applications. For example, suppose a large portion of $V_t(G_t)$ were occupied by one individual as a result of these attacks. In that case, the individual could manipulate subsequent consensus-buildings as they like, which implies the failure of P2P citation systems. Accordingly, the two incentive mechanisms need some additional features to prevent spamming and Sybil attacks.

Spamming is "the act of spreading unsolicited and unrelated content" (Hayati et al., 2010, p. 1). For the aforementioned three layers (Figures 1.4 and 1.6), this means the (extreme) one-to-many correspondences in individual-to-products. While spamming may cause system failure, P2P systems are more convenient when an individual can register (or spread) multiple intellectual products. Preceding studies have mitigated spamming by imposing some small cost on registration, thereby ensuring that individuals have no incentive to register too many products. For example, *Hashcash* (Back et al., 2002)—an anti-spam system mainly for email—requires a small amount of computational resources to peers whenever they use networks[15]; the Bitcoin protocol (Nakamoto et al., 2008) requires transaction fee (in the form of bitcoin) to peers whenever they make a transaction. The two incentive mechanisms can leverage both preceding studies for their registration of intellectual products. Still, the latter approach (i.e., let individuals pay a reward token as a registration fee) would be particularly practical because it can also provide a use for the reward as a token.

Sybil attack is "the forging of multiple identities" (Douceur, 2002, p. 251). For the aforementioned three layers (Figures 1.4 and 1.6), this means the (extreme) one-to-many correspondences in individual-to-peers. While Sybil attack may cause system failure, P2P systems, by their nature, need to accept peers without any validation from centralized authorities. Preceding studies have therefore mitigated Sybil attack with a variety of decentralized approaches (e.g., Levine et al., 2006; Mohaisen & Kim, 2013). For example, *resource testing* imposes some small cost on creating new peers, such as computational resources (Borisov, 2006; Li et al., 2012), human resources (Von Ahn et al., 2003) and IP address (Freedman & Morris, 2002); *economic incentive* gives monetary rewards to peers who report malicious peers (Margolin & Levine, 2007); *reputation system* detects Sybil attack by constructing a graph structure (referred to as social network or trust graph) from the activities of each peer (Cheng & Friedman, 2005; Yu, 2011)[16]. In addition, we can interpret the proof-of-work in the Bitcoin protocol as an application of the resource testing in that rewards depend not on the number of peers but on computational resources[17]. The two incentive mechanisms can leverage any preceding study for their creation of peers. Still, it is a topic for future study to assess which one (or combination of ones) is the best to mitigate Sybil attacks.

Thus, to relax one-to-one correspondence (Assumption 1.4.1), the two incentive

[15]The approach of imposing computational resources as a cost has been generalized as *proof-of-work* (Dwork & Naor, 1992; Jakobsson & Juels, 1999). Note here that the Bitcoin protocol adopts a proof-of-work similar to Hashcash, but this is not for anti-spamming but consensus-building, as described in Section 1.3.2.

[16]Note that, when Douceur (2002) introduces the concept of Sybil attack, he also discusses potential solutions including the resource testing, but concludes that the validation by centralized authority is the only effective means. Subsequent decentralized approaches are mainly aimed at mitigating rather than completely solving Sybil attack.

[17]Considering that preceding approaches have only increased the cost of Sybil attacks, the proof-of-work in the Bitcoin protocol is very elegant because it eliminates the incentive for individuals to create multiple peers.

mechanisms need some additional features that can cover spamming and Sybil attacks. Preceding studies, such as Nakamoto et al. (2008) and Margolin and Levine (2007), would be useful for their design.

### 5.2.2 Relaxing Expected Rewards as Objective (Assumption 1.4.2)

Another future study is to relax Assumption 1.4.2—*peers aim to maximize the total amount of their expected rewards*. This is important because rewards, which may take the (exchangeable) token form, are not always sufficient incentives for peers. For example, suppose their market price were too low (e.g., $ 0.1 per review). In that case, peers would not commit consensus-building or even register their intellectual products, which implies the failure of P2P citation systems. Accordingly, the two incentive mechanisms need some additional features to make rewards incentives.

First of all, rewards as incentives need to ensure their market price. From the viewpoint of economics, the market price of goods is determined as the intersection of the value (from the supply-side) depending on the marginal cost and the value (from the demand-side) depending on the marginal utility (Marshall, 1890). In this framework, the Bitcoin protocol has features that ensure the value of both. On the supply-side, the block-reward halving—cutting the amount of rewards (bitcoin) for miners in half for every $210,000$ blocks[18]—contributes to gradually increasing the marginal cost of rewards (i.e., the amount of computational resources to obtain a unit of bitcoin); the effect of this feature on the market price is studied by Meynkhard (2019), Pagnotta and Buraschi (2018)[19]. On the demand-side, the above-mentioned transaction fee contributes to increasing the marginal utility of rewards, in addition to their use-value as electronic cash; the effect of this feature on the market price is studied by Easley et al. (2019), Lehar and Parlour (2019). The two incentive mechanisms can leverage these features. Specifically, the concept of block-reward halving can be applied by halving the amount of rewards per review for every certain number of intellectual products accumulated in $G_t$ (instead of blocks)[20]. The transaction

[18]https://en.bitcoin.it/wiki/Controlled_supply, accessed February 10, 2021.

[19]Pagnotta and Buraschi (2018) argues that the block reward-halving may have both positive and negative effects on the bitcoin price, as it reduces the increasing rate of total supply, not the total supply (i.e., disinflation, not deflation).

[20]Here, the marginal cost of rewards is the human resource devoted per review. It is also worth noting that this reward-halving can contribute to user acquisition in the early stages since it provides

fee corresponds to the above-mentioned registration fee.

Secondly, rewards as incentives need to stabilize their market price. Despite the contribution of reward halving and registration fees, rewards do not work well unless their market price is stable. The Bitcoin protocol addresses this problem through the *difficulty adjustment*. This feature changes the difficulty of proof-of-work for every $2,016$ block to keep the block interval 10 minutes, which contributes to stabilizing the marginal cost of rewards in the short term[21]. For the difficulty adjustment, Saito and Iwamura (2019) and Iwamura et al. (2019) proposed a modification for further stability; Tiutiun et al. (2018) designed a model for *Stablecoin*[22]; Noda et al. (2020) analyzed the behavior of strategic peers under the difficulty adjustment. The two incentive mechanisms can leverage this feature by replacing the difficulty and the block interval with the number of reward computations per state and the interval between state transitions $(G_t)_{t=0}^{q}$, respectively. In other words, we could stabilize the market price of rewards through the adjustment, which increases the number of reward computations per state if the interval between state transitions increases (and vice versa), although Algorithms 1 and 4 assumed for convenience that reward computation is once per state[23].

Finally, rewards as incentives need to take into account the incentive outside P2P citation systems. Particularly in the presence of negative incentives (e.g., bribes to encourage misreporting, short-selling on reward-token exchanges), individuals may not follow the truthful strategy or may act to decrease the market price of rewards (i.e., *Goldfinger attack*; Kroll et al., 2013). Even though this problem remains to be discussed so much in the context of consensus-building in P2P systems, several

the two incentive mechanisms with a kind of first-mover advantage in that those who register their intellectual products first will earn a larger amount of rewards per review.

[21]From this difficulty and the number of blocks generated in a given period, we can estimate how many computational resources are currently being put into the Bitcoin protocol per second (i.e., the total hash rate). The total hash rate is an important metric for estimating the marginal cost of bitcoin.

[22]Stablecoin is an approach to address this problem of high price volatility, which is defined such as "a digital currency that is pegged to another stable asset like gold, or to major fiat currencies like Euros, Pounds or the US dollar" (Lund, 2018), "cryptocurrency that has price stable characteristics" (Tomaino, 2017), and "a digital token that will have low price volatility as a result of being pegged to some underlying fiat currency, thereby acting as a store of value, a medium of exchange and unit of accounting for blockchain payments" (Hassani et al., 2018). See survey paper Mita et al. (2019), Mita et al. (2019) for the detail of stablecoin.

[23]If we adopted the difficulty adjustment of the Bitcoin protocol more directly, the two incentive mechanisms would adjust the amount of human resources devoted per review rather than the number of reward computations per state. It is a topic for future study to investigate the feasibility of such an adjustment, with a survey of peer prediction methods. Moreover, it may be possible for the incentive mechanism with ex-ante consensus to use the exogenous parameter $\mu$ for the difficulty adjustment.

peer-prediction methods have explored ways to address it. For example, Jurca and Faltings (2006) computed the threshold of negative incentives that Miller et al. (2005) can handle under (given) budget constraints; Goel et al. (2019, 2020a), Goel et al. (2020b) developed a peer prediction method for *decentralized oracles*[24], while assuming such negative incentives; moreover, even Dasgupta and Ghosh (2013) mentioned the robustness of DG13 to reporting costs[25]. Therefore, it is a topic for future study to assess how well DG13 in the two incentive mechanisms can handle the incentive outside P2P citation systems and which other peer-prediction methods are a better solution.

Thus, to relax expected rewards as objective (Assumption 1.4.2), the two incentive mechanisms need some additional features that can cover the market price of rewards, the stability of market price, and the incentive outside P2P citation systems. Preceding studies, such as Nakamoto et al. (2008), Iwamura et al. (2019), and Goel et al. (2020a), would be useful for their design.

### 5.2.3 Other Open Questions

Aside from relaxing Assumptions 1.4.1 and 1.4.2, the two incentive mechanisms leave the following (interdisciplinary) open questions:

*How the reviewer assignment should be?* — The PR score, a quantitative measure this thesis uses in Assumption 1.4.3, is one of the simplest criteria for the RAP. Preceding studies on the RAP have proposed systems that focus on not only centrality (e.g., the PR score), but also similarity (Küçüktunç et al., 2012)[26], diversity (Liu et al., 2014), and CoI (Long et al., 2013). We, therefore, have to consider how the reviewer assignment should be. The answer would vary according to intellectual products, with different citation practices (Meyer, 2000),

[24] See Section 1.3.2 and its footnotes for the detail of decentralized oracles.

[25] However, the discussion here assumes that we can scale the amount of rewards as much as desired by an exogenous parameter, and does not go into the robustness of net rewards (i.e., rewards minus reporting costs). See Section 4 of Dasgupta and Ghosh (2013) for more details.

[26] As we confirmed in Chapter 3, the incentive mechanism with ex-ante consensus ensures similarity by the PPR algorithm, just like Küçüktunç et al. (2012).

*How to reflect differentiated citation practices?* — Furthermore, citation practices are differentiated even in the same intellectual product, especially as the growth of citations. For example, Hurt (1987) compared three groups of scientific publications (physics, engineering, sociology) and revealed their different citation practices[27]; similar results were pointed out for patents (e.g., Jaffe & Trajtenberg, 1999) and web pages (e.g., Barnett & Sung, 2005). It is therefore a future study to reflect such differentiated citation practices on our consensus-building. This may be possible by partitioning $G_t$ into multiple components with *graph clustering* (e.g., Malliaros & Vazirgiannis, 2013; Schaeffer, 2007), then applying different (but compatible) methods of reviewer assignment and reward computation to each cluster[28],

*How to reflect citation semantics?* — Citation does not always have positive semantics in practice. For example, Moravcsik and Murugesan (1975) estimated 14 percent of citations are to dispute the research, from 30 scientific publications on high-energy physics theory[29]. It is, therefore, another future study to reflect such citation semantics on our consensus-building. This is relevant to the discussion of the potential application to web pages in terms of considering Semantic Web and Linked Open Data and to the discussion of using other peer prediction methods (e.g., Miller et al., 2005; Shnayder et al., 2016a) in terms of considering not binary- but multiple-signals emitted from citation edges,

*How should consensus-building be?* — Even if the two incentive mechanisms succeed in eliciting truthful reports from appropriate reviewers, there remains the discussion of how the consensus-building should be done. Given the explanations in Chapters 3 and 4, an ex-ante consensus is a vote of confidence in that $\dot{G}_t$ is accepted when more than $\mu$ reviewers report 1; an ex-post consensus is a dictatorship in that each citation relationship is determined from the out-edge report by one reviewer; however, are these appropriate? This is the topic closely related to *the theory of social choice* (Arrow, 1951). The axiomatic approach of this theory would be useful for the

[27]See also Braun et al. (1995a, 1995b) and Hargens (2000).

[28]In addition to graph clustering, peer prediction methods dealing with heterogeneous tasks (Mandal et al., 2016) and heterogeneous peers (Agarwal et al., 2017) would be useful for this problem.

[29]This is one of the earliest studies on the research topic nowadays referred to as *citation context analysis*. See Bornmann and Daniel (2008), Hernández-Alvarez and Gomez (2016), McCain and Salvucci (2006) for the detail and subsequent studies on the citation context analysis.

normative analysis of consensus-building, as a future study,

*How to exclude misappropriated intellectual products?* — Misappropriated intellectual products (i.e., registering the intellectual products created by others without their permission) is another risk in the management of P2P citation systems. Since we cannot completely prevent it ex-ante, a practical solution would be to exclude already-registered intellectual products (vertices) from the system or reviewer candidates ex-post[30], once they are regarded as misappropriated ones. Considering the specific process—from selecting suspicious products to the consensus-building on their misappropriation—of this feature is also a future study for potential applications. Preceding studies, such as Suryanarayana and Taylor (2004) and Hoffman et al. (2009), are useful for this problem, which survey detection and exclusion of malicious peers in existing P2P systems[31],

*How should governance be?* — To be sustainable P2P citation systems, the two incentive mechanisms have to consider their governance (i.e., the consensus-building for changing the incentive mechanisms themselves) as well. For example, the Bitcoin protocol has governance where community developers propose, discuss, and update the protocol according to *Bitcoin Improvement Proposal* (BIP)[32]. An important feature of this process is that anyone who opposes the BIP can split the blockchain and run another incompatible protocol, which leaves room for other stakeholders to participate in the governance (e.g., De Filippi & Loveluck, 2016; DiRose & Mansouri, 2018). If the two incentive mechanisms adopt similar governance, their protocol will also split during the growth of citations. This would result in incompatible clustered citations, where intellectual products can no longer cite each other between clusters in $G_t$,

[30] Peers cannot get rewards from the two incentive mechanisms once their (misappropriated) intellectual products are excluded from reviewer candidates.

[31] See also Ito and O'Dair (2019) that survey problems of intellectual property management with the blockchain. Note that the two incentive mechanisms do not need to consider such misappropriation if they construct citation relationships of the public domain (e.g., classical art, classical music) or metadata (e.g., novel titles, movie titles). This is the reason why the author originally proposed the incentive mechanism with ex-ante consensus (Ito & Tanaka, 2019) as a TCR that aims to curate a high-quality, reliable list of any content (e.g., restaurants, universities, web pages) in a decentralized manner.

[32] https://github.com/bitcoin/bips, accessed January 27, 2021. Note that the BIP-based governance is spontaneous, and Nakamoto et al. (2008) does not mention the governance of Bitcoin protocol.

In this section, we discussed future studies of the two incentive mechanisms, which first considered how to relax Assumptions 1.4.1 and 1.4.2, then summarized other remaining studies as open questions. To summarize, future study is to design additional features that can (i) prevent spamming, and Sybil attack, (ii) make rewards a sufficient incentive, and (iii) address other open questions.

## 5.3 Summary of This Chapter

This chapter discussed potential applications and future studies of the two incentive mechanisms detailed in the previous Chapters 3 and 4. We can summarize all arguments in Chapter 5 as answers to the questions presented at the beginning:

- *What is the potential application of the two incentive mechanisms?* (5.1) — Potential application is to make crowd-sourcing (for scientific publications, patents, and web pages) robust and productive, as a component of P2P citation systems,
- *What is the social contribution of the RQ?* (5.1) — Social contribution of the RQ is to develop a new reward source for open-access intellectual products, as an alternative scheme to grants, royalties, and advertisements,
- *What is the future study of the two incentive mechanisms?* (5.2) — Future study is to design additional features that can (i) prevent spamming, and Sybil attack, (ii) make rewards sufficient incentives, and (iii) address other open questions.

Based on the above, the next Chapter 6 concludes this thesis with a summary of each chapter and the answer to the RQ.

# Chapter 6

# Conclusion

Chapter 6 finally concludes this thesis with a summary of each chapter and the answer to the RQ. This thesis clarified that subject to several assumptions (including Assumptions 1.4.1–1.4.3), the answer to the RQ is *Yes*.

## 6.1 Summary of Each Chapter

**Chapter 1** introduced some backgrounds and preliminaries of this thesis. We summarized its all arguments as follows:

- *Why are citations important?* (1.1) — Citations are important because their structure as a growing DAG can be a source of quantitative measures to evaluate the ever-increasing intellectual products efficiently,
- *Why are citations in a P2P system important?* (1.2) — Citations in a P2P system are important because they are independent of possibly misbehaved centralized authorities and can obtain both robustness and productivity,
- *Why are citations in a P2P system difficult?* (1.3) — Citations in a P2P system are difficult because their consensus-building has remained unreliable due to three problems: free-riding, strategic misreporting, and reviewer assignment.
- *What is the Research Question (RQ) of this thesis?* (1.4) — The RQ of this thesis is: *Can we design some incentive mechanism to solve the unreliable consensus-building in P2P citation systems?*,
- *What are academic contributions of the RQ?* (1.3–1.4) — Academic contributions of the RQ are (i) incorporating expertise into the consensus-building in P2P

systems, (ii) extending the RAP to a group of strategic peers, (iii) bridging How and Why approaches on network formation.

In addition, this chapter provided a thesis outline (1.5) as well.

**Chapter 2** covered the methodology, which examines the RQ through two incentive mechanisms (with ex-ante or ex-post consensus) consisting of the same research fields—random walks on graphs and peer prediction methods. We summarized its all arguments as follows:

- *Why are the two incentive mechanisms important?* (2.1.1) — The two incentive mechanisms are important because they allow us to cover both citations with peer-review and citations without peer-review,
- *Why are random walks on graphs important?* (2.1.2) — Random walks on graphs are important (for the two incentive mechanisms) because they are useful to address reviewer assignment in citations,
- *Why are peer prediction methods important?* (2.1.2) — Peer prediction methods are important (for the two incentive mechanisms) because they are useful to address free-riding and strategic misreporting in P2P systems,
- *What are academic contributions of the methodology?* (2.2–2.3) — Academic contributions of the methodology are (i) providing strong truthfulness for random walks on graphs and (ii) leveraging graphs to make peer prediction practical.

In addition, this chapter provided the description of mechanism components (2.2–2.3) and experimental datasets (2.4) as well.

**Chapter 3** introduced the incentive mechanism with *ex-ante* consensus, which (i) covers citations with peer-review (e.g., those on scientific publications and patents) and (ii) consists of the PPR algorithm (Haveliwala, 2002) and DG13 (Dasgupta & Ghosh, 2013). We summarized its all arguments as follows:

- *Why can the PPR algorithm solve reviewer assignment?* (3.1.3) — The PPR algorithm can solve reviewer assignment because it assigns appropriate reviewers (with similarity), while ensuring condition (i) of Assumption 1.4.3,

- *Why can DG13 solve free-riding and strategic misreporting?* (3.1.4) — DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers, while ensuring condition (ii) of Assumption 1.4.3.

In addition, this chapter provided experimental studies (3.2) as well.

**Chapter 4** introduced the incentive mechanism with *ex-post* consensus, which (i) covers citations without peer-review (e.g., those on web pages) and (ii) consists of the two-path mechanisms (Babichenko et al., 2018) and DG13 (Dasgupta & Ghosh, 2013). We summarized its all arguments as follows:

- *Why can the two-path mechanism solve reviewer assignment?* (4.1.3) — The two-path mechanism can solve reviewer assignment because it assigns appropriate reviewers (under the weak truthfulness), while ensuring condition (i) of Assumption 1.4.3,

- *Why can DG13 solve free-riding and strategic misreporting?* (4.1.4) — DG13 can solve free-riding and strategic misreporting because it computes rewards for reviewers, while ensuring condition (ii) of Assumption 1.4.3.

In addition, this chapter provided experimental studies (4.2) as well.

**Chapter 5** discussed potential applications and future studies of the two incentive mechanisms detailed in the previous Chapters 3 and 4. We summarized its all arguments as follows:

- *What is the potential application of the two incentive mechanisms?* (5.1) — Potential application is to make crowdsourcing (for scientific publications, patents, and web pages) robust and productive, as a component of P2P citation systems,

- *What is the social contribution of the RQ?* (5.1) — Social contribution of the RQ is to provide a new reward source for open-access intellectual products, as an alternative scheme to grants, royalties, and advertisements,

- *What is the future study of the two incentive mechanisms?* (5.2) — Future study is to design additional features that can (i) prevent spamming, and Sybil attack, (ii) make rewards sufficient incentives, and (iii) address other open questions.

The above are summaries of the entire thesis.

## 6.2 Answer to the RQ

As Definition 1.4.1 indicates, the RQ was: *Can we design some incentive mechanism to solve the unreliable consensus-building in P2P citation systems?*. Based on the results of Chapters 3 and 4, this thesis concludes that the answer to the RQ is ***Yes***.

On the other hand, it should be noted that this conclusion is subject to the following three assumptions about the RQ:

**Assumption 1.4.1** (One-to-one correspondence). In P2P citation systems (Definition 1.2.2), individuals-to-peers, peers-to-products, and thus individuals-to-products are all one-to-one correspondence,

**Assumption 1.4.2** (Expected rewards as objective). Peers aim to maximize the total amount of their expected rewards,

**Assumption 1.4.3** (Reliable consensus-building). Consensus-building is reliable if peers can (i) be reviewers more often as they get higher PR scores and (ii) maximize the amount of expected rewards per review by always reporting true beliefs,

and other four assumptions about the peer prediction method:

**Assumption 2.3.1** (Signal Reporting). Peers $i$ and $j$ each report what *signals* $S_i$ and $S_j$ were, which are discrete random variables emitted from the task,

**Assumption 2.3.2** (Binary Signals). $s_i, s_j \in \{0, 1\}$,

**Assumption 2.3.3** (Positive Correlation). Binary signals $\{0, 1\}$ to peers $i$ and $j$ are positively correlated; namely, $Pr(S_i = 0 \mid S_j = 0) > Pr(S_i = 0)$ and $Pr(S_i = 1 \mid S_j = 1) > Pr(S_i = 1)$.

**Assumption 2.3.4** (Strategies as probability matrices). Peers $i$ and $j$ follow mixed strategies $\sigma_i$ and $\sigma_j$ that have probability matrices $\boldsymbol{P}(\sigma_i)$ and $\boldsymbol{P}(\sigma_j)$, respectively.

This clarification is the outcome of this thesis.

As discussed in Chapter 5, one of the future studies is relaxing these assumptions, especially those on the RQ.

# Appendix A

# Expected Rewards in a Simple Token-Staking Scheme

Consider a simple token-staking example in which $n$ individuals stake a fixed $q$ number of tokens on one of the options. Let $k$ be the amount of (net) rewards that individuals can obtain when their selections become the consensus, and let $p$ be the individuals' subjective probability of realizing this event. Then, the expected reward of this example is $\mathbb{E}\left[\,k\,\right] = pk - (1-p)q$.

Here, $k$ is the redistribution of the total staked tokens $nq$ among the individuals who have staked on the consensus except for one's stake, $q$. Accordingly, if we let $n^*$ be the number of individuals who have staked on the consensus, $k = \frac{n}{n^*}q - q = \frac{n-n^*}{n^*}q$. By substituting this into the equation of $\mathbb{E}\left[\,k\,\right]$, we can derive the following condition:

$$\mathbb{E}\left[\,k\,\right] \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 0, \text{ if } \frac{p/(1-p)}{n^*/(n-n^*)} \left\{ \begin{array}{c} > \\ = \\ < \end{array} \right\} 1,$$

where $\frac{p/(1-p)}{n^*/(n-n^*)}$ represents the odds ratio between the expected and actual value of the probability of one's choice becoming the consensus; i.e., the anticipated reward of the model takes a positive value only when we estimate the odds to be higher than their actual value and is zero as long as our estimation is precise (as a result of the zero-sum game). Furthermore, the expected reward under precise odds estimation is negative if we take the cost of staking into account[1].

These results reveal that the token-staking scheme does not have sufficient incentive to engage individuals in consensus-building. Providing new reward tokens to individuals in proportion to the score of the peer-prediction mechanism is one possible approach to this problem.

[1]If we assume the cost of staking as $c$, the expected rewards in this example become $\mathbb{E}\left[k\right] = p(k - c) - (1-p)(q+c)$. This extension shifts the condition for $\mathbb{E}\left[k\right] = 0$, from $\frac{p/(1-p)}{n^*/(n-n^*)} = 1$ to $\frac{p/(1-p)}{n^*/(n-n^*)} = \frac{q+c}{q - \frac{n^*}{n-n^*}c}$, where the right-hand side of the new condition must be greater than one.

# Appendix B

# Proof of the Strong Truthfulness in DG13

For the proof of the strong truthfulness (Definition 2.3.1) in DG13, we need to derive $\mathbb{E}\left[X_i\right]$ from the following realization of $X_i$ (equation 2.3):

$$x_i^{m^*} = \delta(r_i^{m^*}, r_j^{m^*}) - \delta(r_i^n, r_j^{n'}). \tag{B.1}$$

The expected value corresponding to the reward term $\delta(r_i^{m^*}, r_j^{m^*})$ depends not only on strategies $\sigma_i, \sigma_j$ but also on the probability distribution of signals, which can be described as follows:

$$\sum_{s_i=0}^{1} \sum_{s_j=0}^{1} Pr(S_i = s_i, S_j = s_j) \cdot (p_{0,s_i}\, p_{0,s_j} + p_{1,s_i}\, p_{1,s_j}), \tag{B.2}$$

where $Pr(S_i = s_i, S_j = s_j)$ is the joint distribution; moreover, $p_{0,s_i}\, p_{0,s_j} + p_{1,s_i}\, p_{1,s_j}$ denotes the probability that two reports will match for given $s_i$ and $s_j$. $p_{0,s_i}, p_{1,s_i}$ and $p_{0,s_j}, p_{1,s_j}$ are elements of $\boldsymbol{P}(\sigma_i)$ and $\boldsymbol{P}(\sigma_j)$, respectively. Note that B.2 does not require superscript $m^*$ because Assumption 2.3.3 holds across the tasks.

The expected value corresponding to the penalty term $\delta(r_i^n, r_j^{n'})$, on the other hand, is different from B.2 in that the term uses $n \in M_i \setminus \{m^*\}$ and $n' \in M_j \setminus \{m^*\}$ instead of $m^*$, which can be described as follows:

$$\sum_{s_i=0}^{1} \sum_{s_j=0}^{1} Pr(S_i = s_i) Pr(S_j = s_j) \cdot (p_{0,s_i}\, p_{0,s_j} + p_{1,s_i}\, p_{1,s_j}), \tag{B.3}$$

where $Pr(S_i = s_i) Pr(S_j = s_j)$ is the product of marginal distributions. B.3 does not use the joint distribution $Pr(S_i = s_i, S_j = s_j)$ because the penalty term covers two different tasks, $n$ and $n'$.

Consequently, using B.2 and B.3, $\mathbb{E}[X_i]$ can be described as follows:

$$\mathbb{E}[X_i] = \sum_{s_i=0}^{1}\sum_{s_j=0}^{1} \left[Pr(S_i = s_i, S_j = s_j) - Pr(S_i = s_i)Pr(S_j = s_j)\right] \cdot (p_{0,s_i}\, p_{0,s_j} + p_{1,s_i}\, p_{1,s_j}). \tag{B.4}$$

The terms in square brackets clearly represent the correlation between $S_i$ and $S_j$; specifically, if realizations $s_i$ and $s_j$ are positively correlated, then $Pr(S_i = s_i, S_j = s_j) - Pr(S_i = s_i)Pr(S_j = s_j) > 0$ holds (and vice versa), because $Pr(S_i = s_i, S_j = s_j) = Pr(S_i = s_i \mid S_j = s_j)Pr(S_j = s_j) = Pr(S_j = s_j \mid S_i = s_i)Pr(S_i = s_i)$.

Accordingly, given the Assumption 2.3.3, the following condition holds for the expansion of B.4:

$$\begin{aligned}
\mathbb{E}[X_i] = {} & \left[Pr(S_i = 0, S_j = 0) - Pr(S_i = 0)Pr(S_j = 0)\right]_{>0} \cdot (p_{0,0}\, p_{0,0} + p_{1,0}\, p_{1,0}) \\
& + \left[Pr(S_i = 0, S_j = 1) - Pr(S_i = 0)Pr(S_j = 1)\right]_{<0} \cdot (p_{0,0}\, p_{0,1} + p_{1,0}\, p_{1,1}) \\
& + \left[Pr(S_i = 1, S_j = 0) - Pr(S_i = 1)Pr(S_j = 0)\right]_{<0} \cdot (p_{0,1}\, p_{0,0} + p_{1,1}\, p_{1,0}) \\
& + \left[Pr(S_i = 1, S_j = 1) - Pr(S_i = 1)Pr(S_j = 1)\right]_{>0} \cdot (p_{0,1}\, p_{0,1} + p_{1,1}\, p_{1,1}),
\end{aligned}$$

where $[a]_{>0}$ and $[a]_{<0}$ indicate that $a$ is positive and negative, respectively.

It is apparent that $\mathbb{E}[X_i]$ is maximized when both $i$ and $j$ adopt truthtelling strategy or perverse strategy, i.e., $\boldsymbol{P}(\sigma_i^*) = \boldsymbol{P}(\sigma_j^*) = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$ or $\boldsymbol{P}(\sigma_i) = \boldsymbol{P}(\sigma_j) = \left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$, because only these strategy pairs can achieve $p_{0,0}\, p_{0,0} + p_{1,0}\, p_{1,0} = p_{0,1}\, p_{0,1} + p_{1,1}\, p_{1,1} = 1$ and $p_{0,0}\, p_{0,1} + p_{1,0}\, p_{1,1} = p_{0,1}\, p_{0,0} + p_{1,1}\, p_{1,0} = 0$ simultaneously[1]. Thus, DG13 satisfies strong truthfulness. □

[1] Here, as long as Assumption 2.3.3 holds, we do not need to know the full signal distribution (i.e., DG13 is detail-free).

# Bibliography


Abraham, R. H. (1997). Webometry: Measuring the complexity of the world wide web. *World Futures: Journal of General Evolution*, *50*(1-4), 785–791.

Adar, E., & Huberman, B. A. (2000). Free riding on gnutella. *First monday*, *5*(10).

Agarwal, A., Mandal, D., Parkes, D. C., & Shah, N. (2017). Peer prediction with heterogeneous users, In *Proceedings of the 2017 acm conference on economics and computation*. ACM.

Aggarwal, C., & Subbian, K. (2014). Evolutionary network analysis: A survey. *ACM Computing Surveys (CSUR)*, *47*(1), 1–36.

Aguillo, I. (1997). Cybermetrics' 97 (jerusalem, israel). *Retrieved May*, *12*, 2007.

Ahmed, F., Dickerson, J. P., & Fuge, M. (2017). Diverse weighted bipartite b-matching. *arXiv preprint arXiv:1702.07134*.

Albert, K. M. (2006). Open access: Implications for scholarly publishing and medical libraries. *Journal of the Medical Library Association*, *94*(3), 253.

Albert, M. B., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research policy*, *20*(3), 251–259.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, *74*(1), 47.

Alcácer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, *88*(4), 774–779.

Alcácer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in us patents: An overview and analysis. *Research Policy*, *38*(2), 415–427.

Aldous, D., & Fill, J. (1995). Reversible markov chains and random walks on graphs. Berkeley.

Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: Methodological approaches to 'webometrics'. *Journal of documentation*, *53*(4), 404–426.

Alonso, O., Rose, D. E., & Stewart, B. (2008). Crowdsourcing for relevance evaluation, In *Sigir forum*.

Androutsellis-Theotokis, S., & Spinellis, D. (2004). A survey of peer-to-peer content distribution technologies. *ACM computing surveys (CSUR)*, *36*(4), 335–371.

Arrow, K. J. (1951). Social choice and individual values.(cowles commission mongr. no. 12.).

Asgaonkar, A., & Krishnamachari, B. (2018). Token curated registries-a game theoretic approach. *arXiv preprint arXiv:1809.01756*.

Aumann, R. J., & Myerson, R. B. (1988). Endogenous formation of links between players and of coalitions: An application of the shapley value. *The Shapley Value*, 175–191.

Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential attachment as a unique equilibrium, In *Proceedings of the 2018 world wide web conference*.

Axelsson, O. (1996). *Iterative solution methods*. Cambridge university press.

Babichenko, Y., Dean, O., & Tennenholtz, M. (2018). Incentive-compatible diffusion, In *Proceedings of the 2018 world wide web conference on world wide web*. International World Wide Web Conferences Steering Committee.

Back, A.  Et al. (2002). Hashcash-a denial of service counter-measure.

Backstrom, L., & Leskovec, J. (2011). Supervised random walks: Predicting and recommending links in social networks, In *Proceedings of the fourth acm international conference on web search and data mining*.

Bailey Jr, C. W. (2007). What is open access? http://digital-scholarship.org/cwb/WhatIsOA.htm. (accessed February 5, 2021).

Bala, V., & Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, *68*(5), 1181–1229.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, *286*(5439), 509–512.

Barnett, G. A., & Sung, E. (2005). Culture and the structure of the international hyperlink network. *Journal of Computer-Mediated Communication*, *11*(1), 217–238.

Bauwens, M. (2005a). Peer to peer and human evolution. *Integral Visioning*, *15*.

Bauwens, M. (2005b). The political economy of peer production. *CTheory*, 12–1.

Bauwens, M. (2009). Class and capital in peer production. *Capital & class*, *33*(1), 121–141.

Becchetti, L., Castillo, C., Donato, D., Baeza-Yates, R., & Leonardi, S. (2008). Link analysis for web spam detection. *ACM Transactions on the Web (TWEB)*, *2*(1), 1–42.

Beel, J., Gipp, B., Langer, S., & Breitinger, C. (2016). Research-paper recommender systems: A literature survey. *International Journal on Digital Libraries*, *17*(4), 305–338.

Benkler, Y. (2002a). Coase's penguin, or, linux and" the nature of the firm". *Yale law journal*, 369–446.

Benkler, Y. (2002b). Freedom in the commons: Towards a political economy of information. *Duke LJ*, *52*, 1245.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Bergstrom, C. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *College & research libraries news*, *68*(5), 314–316.

Berners-Lee, T. (2006). Linked data. https://www.w3.org/DesignIssues/LinkedData.html. (accessed February 8, 2021).

Berners-Lee, T., Hendler, J., Lassila, O., Et al. (2001). The semantic web. *Scientific american*, *284*(5), 28–37.

Berners-Lee, T. J., & Cailliau, R. (1990). Worldwideweb: Proposal for a hypertext project.

Berthaud, C., Capelli, L., Gustedt, J., Kirchner, C., Loiseau, K., Magron, A., Medves, M., Monteil, A., Rivrieux, G., & Romary, L. (2014). Episciences–an overlay publication platform. *Information Services & Use*, *34*(3-4), 269–277.

Bestor, D. R., & Hamp, E. (2010). Peer to patent: A cure for our ailing patent examination system. *Nw. J. Tech. & Intell. Prop.*, *9*, 16.

Biggs, N., Lloyd, E. K., & Wilson, R. J. (1986). *Graph theory, 1736-1936*. Oxford University Press.

Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American society for information science and technology*, *55*(14), 1216–1227.

Björneborn, L.  Et al. (2004). *Small-world link structures across an academic web space: A library and information science approach*. Citeseer.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, *424*(4-5), 175–308.

Bollen, J., Rodriquez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, *69*(3), 669–687.

Bondi, A. B. (2000). Characteristics of scalability and their impact on performance, In *Proceedings of the 2nd international workshop on software and performance*. ACM.

Borisov, N. (2006). Computational puzzles as sybil defenses, In *Sixth ieee international conference on peer-to-peer computing (p2p'06)*. IEEE.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? a review of studies on citing behavior. *Journal of documentation*, *64*(1), 45–80.

Bornmann, L., Schier, H., Marx, W., & Daniel, H.-D. (2011). Is interactive open access publishing able to identify high-impact submissions? a study on the predictive validity of atmospheric chemistry and physics by using percentile rank classes. *Journal of the American Society for Information Science and Technology*, *62*(1), 61–71.

Bott, E. (1957). Family and social network: Roles. *Norms, and External Relationships in Ordinary Urban Families (Tavistock, London, UK, 1957)*.

Bott, E., & Spillius, E. B. (2014). *Family and social network: Roles, norms and external relationships in ordinary urban families*. Routledge.

Bozzon, A., Brambilla, M., & Ceri, S. (2012). Answering search queries with crowdsearcher, In *Proceedings of the 21st international conference on world wide web*. ACM.

Braun, T., Glänzel, W., & Grupp, H. (1995a). The scientometric weight of 50 nations in 27 science areas, 1989–1993. part i. all fields combined, mathematics, engineering, chemistry and physics. *Scientometrics*, *33*(3), 263–293.

Braun, T., Glänzel, W., & Grupp, H. (1995b). The scientometric weight of 50 nations in 27 science areas, 1989–1993. part ii. life sciences. *Scientometrics*, *34*(2), 207–237.

Brave. (2021). Basic attention token (bat) blockchain based digital advertising. https://basicattentiontoken.org/static-assets/documents/BasicAttentionTokenWhitePaper-4.pdf. (accessed February 13, 2021).

Breitzman, A, & Narin, F. (1996). A case for patent citation analysis in litigation. *Law Works*, *3*(3).

Breitzman, A. F., & Narin, F. (2001). Method and apparatus for choosing a stock portfolio, based on patent indicators [US Patent 6,175,824]. Google Patents.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, *30*(1-7), 107–117.

Bruck, P., Réthy, I., Szente, J., Tobochnik, J., & Érdi, P. (2016). Recognition of emerging technology trends: Class-selective study of citations in the us patent citation network. *Scientometrics*, *107*(3), 1465–1475.

Buterin, V. (2015). Visions part i: The value of blockchain technology. *Ethereum Blog*, *23*. https://blog.ethereum.org/2015/04/13/visions-part-1-the-value-of-blockchain-technology/. (accessed February 27, 2020).

Buterin, V. Et al. (2014). Ethereum: A next-generation smart contract and decentralized application platform. *https://github.com/ethereum/wiki/wiki/%5BEnglish%5D-White-Paper*, *7*.

Butler, B., Joyce, E., & Pike, J. (2008). Don't look now, but we've created a bureaucracy: The nature and roles of policies and rules in wikipedia, In *Proceedings of the sigchi conference on human factors in computing systems*. ACM.

Carpenter, M. P., Narin, F., & Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information*, *3*(4), 160–163.

Chandra, A. K., Raghavan, P., Ruzzo, W. L., Smolensky, R., & Tiwari, P. (1996). The electrical resistance of a graph captures its commute and cover times. *Computational Complexity*, *6*(4), 312–340.

Charlin, L., & Zemel, R. (2013). The toronto paper matching system: An automated paper-reviewer assignment system.

Cheng, A., & Friedman, E. (2005). Sybilproof reputation mechanisms, In *Proceedings of the 2005 acm sigcomm workshop on economics of peer-to-peer systems*.

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Chung, F., & Zhao, W. (2010). Pagerank and random walks on graphs, In *Fete of combinatorics and computer science*. Springer.

Cimoli, M., Dosi, G., Mazzoleni, R., & Sampat, B. (2011). Innovation, technical change and patents in the development process: A long term view. *Intellectual Property Rights: Legal and Economic Challenges for Development*, *57*.

Clarke, I., Miller, S. G., Hong, T. W., Sandberg, O., & Wiley, B. (2002). Protecting free expression online with freenet. *IEEE Internet Computing*, *6*(1), 40–49.

Clarke, I., Sandberg, O., Wiley, B., & Hong, T. W. (2001). Freenet: A distributed anonymous information storage and retrieval system, In *Designing privacy enhancing technologies*. Springer.

Cohen, B. (2003). Incentives build robustness in bittorrent, In *Workshop on economics of peer-to-peer systems*.

Crespo, A., & Garcia-Molina, H. (2002). Routing indices for peer-to-peer systems, In *Proceedings 22nd international conference on distributed computing systems*. IEEE.

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? inventor vs. examiner citations in european patents. *Research policy*, *37*(10), 1892–1908.

Cuel, R., Morozova, O., Rohde, M., Simperl, E., Siorpaes, K., Tokarchuk, O., Wiedenhöfer, T., Yetim, F., & Zamarian, M. (2011). Motivation mechanisms for participation in human-driven semantic content creation. *International Journal of Knowledge Engineering and Data Mining*, *1*(4), 331–349.

Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., Hagemann, M., Johnson, R., La Manna, M., de Souza, S., Suber, P., & Velterop, J. (2002). Budapest open access initiative. https://www.budapestopenaccessinitiative.org/read. (accessed February 5, 2021).

Dasgupta, A., & Ghosh, A. (2013). Crowdsourced judgement elicitation with endogenous proficiency, In *Proceedings of the 22nd international conference on world wide web*. ACM.

Davidson, S., De Filippi, P., & Potts, J. (2016). Economics of blockchain. *Available at SSRN 2744751*.

De Filippi, P., & Loveluck, B. (2016). The invisible politics of bitcoin: Governance crisis of a decentralized infrastructure. *Internet Policy Review*, *5*(4).

De Paula, Á. (2020). Econometric models of network formation.

Dechezleprêtre, A., Martin, R., & Mohnen, M. (2013). *Knowledge spillovers from clean and dirty technologies: A patent citation analysis*. Grantham Research Institute on Climate Change; the Environment.

Deng, Z., Lev, B., & Narin, F. (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal*, *55*(3), 20–32.

Di Mauro, N., Basile, T. M., & Ferilli, S. (2005). Grape: An expert review assignment component for scientific conference management systems, In *International conference on industrial, engineering and other applications of applied intelligent systems*. Springer.

Dictionary.com. (n.d.). Creative commons. https://www.dictionary.com/browse/creative-commons?action=save. (accessed February 6, 2021).

Diestel, R. (2012). Graph theory, volume 173 of. *Graduate texts in mathematics*, 7.

DiRose, S., & Mansouri, M. (2018). Comparison and analysis of governance mechanisms employed by blockchain-based distributed autonomous organizations, In *2018 13th annual conference on system of systems engineering (sose)*. IEEE.

Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the world-wide web. *Communications of the ACM*, *54*(4), 86–96.

Dosi, G., Stiglitz, J. E. Et al. (2014). The role of intellectual property rights in the development process, with some lessons from developed countries: An introduction. *Intellectual property rights: Legal and economic challenges for development*, *1*, 1–55.

Douceur, J. R. (2002). The sybil attack, In *International workshop on peer-to-peer systems*. Springer.

Dumais, S. T., & Nielsen, J. (1992). Automating the assignment of submitted manuscripts to reviewers, In *Proceedings of the 15th annual international acm sigir conference on research and development in information retrieval*.

Dwork, C., & Naor, M. (1992). Pricing via processing or combatting junk mail, In *Annual international cryptology conference*. Springer.

Easley, D., O'Hara, M., & Basu, S. (2019). From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics*, *134*(1), 91–109.

Elliott, M. A. (2007). *Stigmergic collaboration: A theoretical framework for mass collaboration* (Doctoral dissertation).

Engers, M., & Gans, J. S. (1998). Why referees are not paid (enough). *The American Economic Review*, *88*(5), 1341–1349.

Erdős, P., & Rényi, A. (1959). On random graphs. *Publicationes mathematicae*, *6*(26), 290–297.

Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, *5*(1), 17–60.

Erdős, P., & Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, *12*(1-2), 261–267.

Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science*, *38*(2), 189–200.

Falk, B. H., & Tsoukalas, G. (2018). *Token weighted crowdsourcing* (tech. rep.). Working Paper.

Faltings, B., & Radanovic, G. (2017). Game theory for data science: Eliciting truthful information. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *11*(2), 1–151.

Feigenbaum, J., Papadimitriou, C., & Shenker, S. (2000). Sharing the cost of muliticast transmissions (preliminary version), In *Proceedings of the thirty-second annual acm symposium on theory of computing*.

Feigenbaum, J., & Shenker, S. (2004). Distributed algorithmic mechanism design: Recent results and future directions, In *Current trends in theoretical computer science: The challenge of the new century vol 1: Algorithms and complexity vol 2: Formal models and semantics*. World Scientific.

Feller, J., Fitzgerald, B. Et al. (2002). *Understanding open source software development*. Addison-Wesley Longman Publishing Co., Inc.

Fitzpatrick, K. (2010). Peer-to-peer review and the future of scholarly authority. *Social Epistemology*, *24*(3), 161–179.

Ford, E. (2013). Defining and characterizing open peer review: A review of the literature. *Journal of Scholarly Publishing*, *44*(4), 311–326.

Fort, K., Adda, G., & Cohen, K. B. (2011). Amazon mechanical turk: Gold mine or coal mine? *Computational Linguistics*, *37*(2), 413–420.

Freedman, M. J., & Morris, R. (2002). Tarzan: A peer-to-peer anonymizing network layer, In *Proceedings of the 9th acm conference on computer and communications security*.

Freeman, L. (2004). The development of social network analysis. *A Study in the Sociology of Science*, *1*, 687.

Garfield, E. (1955). Citation indexes for science. *Science*, *122*(3159), 108–111.

Garfield, E. (1957). Breaking the subject index barrier-a citation index for chemical patents. *J. Pat. Off. Soc'y*, *39*, 583.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, *178*(4060), 471–479.

Garfield, E., Malin, M. V., & Small, H. (1983). Citation data as science indicators.

Garfield, E., & Merton, R. K. (1979). *Citation indexing: Its theory and application in science, technology, and humanities* (Vol. 8). Wiley New York.

Garner, R., Lunin, L., & Baker, L. (1967). Three drexel information science research studies.

Gasparyan, A. Y., Gerasimov, A. N., Voronov, A. A., & Kitas, G. D. (2015). Rewarding peer reviewers: Maintaining the integrity of science communication. *Journal of Korean medical science*, *30*(4), 360.

Gehrke, J., Ginsparg, P., & Kleinberg, J. (2003). Overview of the 2003 kdd cup. *Acm Sigkdd Explorations Newsletter*, *5*(2), 149–151.

Ghafele, R., Gibert, B., & DiGiammarino, P. (2011). Crowdsourcing patent application review to capitalize on innovation.

Ginsparg, P. (1994). First steps towards electronic research communication. *Computers in physics*, *8*(4), 390–396.

Ginsparg, P. (2011). Arxiv at 20. *Nature*, *476*(7359), 145.

Gleich, D. F. (2015). Pagerank beyond the web. *Siam Review*, *57*(3), 321–363.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378.

Gnosis, T. (2017). Gnosis-whitepaper. https://gnosis.pm/resources/default/pdf/gnosis_whitepaper.pdf. (accessed May 3, 2018).

Göbel, F, & Jagers, A. (1974). Random walks on graphs. *Stochastic processes and their applications*, *2*(4), 311–336.

Goel, N., Filos-Ratsikas, A., & Faltings, B. (2019). Decentralized oracles via peer-prediction in the presence of lying incentives.

Goel, N., Filos-Ratsikas, A., & Faltings, B. (2020a). Peer-prediction in the presence of outcome dependent lying incentives, In *Proceedings of the 29th international joint conference on artificial intelligence (ijcai)*.

Goel, N., van Schreven, C., Filos-Ratsikas, A., & Faltings, B. (2020b). Infochain: A decentralized, trustless and transparent oracle on blockchain, In *Proceedings of the 29th international joint conference on artificial intelligence (ijcai)*.

Goldin, M. (2017a). Token-curated registries 1.0. *Medium, September 14*. https://medium.com/@ilovebagels/token-curated-registries-1-0-61a232f8dac7. (accessed April 3, 2019).

Goldin, M. (2017b). Token curated registries 1.1, 2.0 tcrs, new theory, and dev updates. *Medium, December 14*. https://medium.com/@ilovebagels/token-curated-registries-1-1-2-0-tcrs-new-theory-and-dev-updates-34c9f079f33d. (accessed April 3, 2019).

Goldin, M., Soleimani, A., & Young, J. (2017). The adchain registry.

González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The sjr indicator. *Journal of informetrics*, *4*(3), 379–391.

Gori, M., & Pucci, A. (2006). Research paper recommender systems: A random-walk based approach, In *2006 ieee/wic/acm international conference on web intelligence (wi 2006 main conference proceedings)(wi'06)*. IEEE.

Grady, C., & Lease, M. (2010). Crowdsourcing document relevance assessment with mechanical turk, In *Proceedings of the naacl hlt 2010 workshop on creating speech and language data with amazon's mechanical turk*. Association for Computational Linguistics.

Gross, P. L., & Gross, E. M. (1927). College libraries and chemical education. *Science*, *66*(1713), 385–389.

Guidera, J. (2009). Content and citation analysis of interdesciplinary humanities textbooks within a framework of curriculum theory.

Gupta, M., Judge, P., & Ammar, M. (2003). A reputation system for peer-to-peer networks, In *Proceedings of the 13th international workshop on network and operating systems support for digital audio and video*.

Gyongyi, Z., Garcia-Molina, H., & Pedersen, J. (2004). Combating web spam with trustrank, In *Proceedings of the 30th international conference on very large data bases (vldb)*.

Haase, P., Siebes, R., & van Harmelen, F. (2008). Expertise-based peer selection in peer-to-peer networks. *Knowledge and Information Systems*, *15*(1), 75–107.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The nber patent citation data file: Lessons, insights and methodological tools* (tech. rep.). National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of economics*, 16–38.

Hargens, L. L. (2000). Using the literature: Reference networks, reference contexts, and the social structure of scholarship. *American sociological review*, 846–865.

Hart, H. C., & Goldsmith, E. H. (1949). Re: Citation system for patent office/back to nature. *J. Pat. & Trademark Off. Soc'y*, *31*, 714.

Hassani, H., Huang, X., & Silva, E. (2018). Banking with blockchain-ed big data. *Journal of Management Analytics*, *5*(4), 256–275.

Haveliwala, T. H. (2002). Topic-sensitive pagerank, In *Proceedings of the 11th international conference on world wide web*. ACM.

Hayati, P., Potdar, V., Talevski, A., Firoozeh, N., Sarenche, S., & Yeganeh, E. A. (2010). Definition of spam 2.0: New spamming boom, In *4th ieee international conference on digital ecosystems and technologies*. IEEE.

Hendler, J., & Berners-Lee, T. (2010). From the semantic web to social machines: A research challenge for ai on the world wide web. *Artificial intelligence*, *174*(2), 156–161.

Henzinger, M. R., Motwani, R., & Silverstein, C. (2002). Challenges in web search engines, In *Acm sigir forum*. ACM New York, NY, USA.

Hernández-Alvarez, M., & Gomez, J. M. (2016). Survey about citation context analysis: Tasks, techniques, and resources. *Natural Language Engineering*, *22*(3), 327–349.

Hileman, G., & Rauchs, M. (2017). 2017 global blockchain benchmarking study.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, *102*(46), 16569–16572.

Hoffman, K., Zage, D., & Nita-Rotaru, C. (2009). A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys (CSUR)*, *42*(1), 1–31.

Hopcroft, J., & Sheldon, D. (2008). *Network reputation games* (tech. rep.).

Howe, J. (2006a). Crowdsourcing: A definition. https://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html. (accessed November 15, 2019).

Howe, J. (2006b). The rise of crowdsourcing. *Wired magazine*, *14*(6), 1–4.

Hurt, C. D. (1987). Conceptual citation differences in science, technology, and social sciences literature. *Information Processing & Management*, *23*(1), 1–6.

IEEE, S. C. Et al. (1990). Ieee standard glossary of software engineering terminology. *IEEE Std 610.12-1990*, 1–84. https://doi.org/10.1109/IEEESTD.1990.101064

Ingwersen, P. (1998). The calculation of web impact factors. *Journal of documentation*, *54*(2), 236–243.

Ito, K. (2018). Bunsan gata oracle no gouikeisei ni taisuru peer prediction no yuuyousei [a potential utility of peer prediction method to consensus building on decentralized oracle systems]. *Journal of Information Studies, Interfaculty Initiative in Information Studies, The University of Tokyo*, *95*, 45–56.

Ito, K., Ohsawa, S., & Tanaka, H. (2018). Information diffusion enhanced by multi-task peer prediction, In *Proceedings of the 20th international conference on information integration and web-based applications & services*. ACM.

Ito, K., Ohsawa, S., & Tanaka, H. (2019). Information diffusion enhanced by multi-task peer prediction. *Journal of Data Intelligence*, *1*(1), 018–035.

Ito, K., & O'Dair, M. (2019). A critical examination of the application of blockchain technology to intellectual property management, In *Business transformation through blockchain*. Springer.

Ito, K., & Tanaka, H. (2019). Token-curated registry with citation graph. *Ledger*, *4*.

Iwamura, M., Kitamura, Y., Matsumoto, T., & Saito, K. (2019). Can we stabilize the price of a cryptocurrency?: Understanding the design of bitcoin and its potential to compete with central bank money. *Hitotsubashi Journal of Economics*, 41–60.

Jackson, M. O. (2005a). A survey of network formation models: Stability and efficiency. *Group formation in economics: Networks, clubs, and coalitions*, *664*, 11–49.

Jackson, M. O. (2005b). The economics of social networks.

Jackson, M. O. (2010). *Social and economic networks*. Princeton university press.

Jackson, M. O., & Rogers, B. W. (2007). Meeting strangers and friends of friends: How random are social networks? *American Economic Review*, *97*(3), 890–915.

Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of economic theory*, *71*(1), 44–74.

Jaffe, A. B., & Trajtenberg, M. (1999). International knowledge flows: Evidence from patent citations. *Economics of innovation and new technology*, *8*(1-2), 105–136.

Jakobsson, M., & Juels, A. (1999). Proofs of work and bread pudding protocols, In *Secure information networks*. Springer.

Jeh, G., & Widom, J. (2003). Scaling personalized web search, In *Proceedings of the 12th international conference on world wide web*.

Johnson, S., Miller, N. H., Pratt, J. W., & Zeckhauser, R. J. (2002). Efficient design with interdependent valuations and an informed center.

Jurca, R., & Faltings, B. (2005). Enforcing truthful strategies in incentive compatible reputation mechanisms, In *International workshop on internet and network economics*. Springer.

Jurca, R., & Faltings, B. (2006). Minimum payments that reward honest reputation feedback, In *Proceedings of the 7th acm conference on electronic commerce*. ACM.

Jurca, R., Faltings, B. Et al. (2009). Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, *34*(1), 209.

Kamvar, S. D., Schlosser, M. T., & Garcia-Molina, H. (2003). The eigentrust algorithm for reputation management in p2p networks, In *Proceedings of the 12th international conference on world wide web*.

Karakaya, M., Korpeoglu, I., & Ulusoy, Ö. (2009). Free riding in peer-to-peer networks. *IEEE Internet computing*, *13*(2), 92–98.

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American documentation*, *14*(1), 10–25.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, *46*(5), 604–632.

Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. S. (1999). The web as a graph: Measurements, models, and methods, In *International computing and combinatorics conference*. Springer.

Kolasa, T., & Krol, D. (2011). A survey of algorithms for paper-reviewer assignment problem. *IETE Technical Review*, *28*(2), 123–134.

Koskela, T., Kassinen, O., Harjula, E., & Ylianttila, M. (2013). P2p group management systems: A conceptual analysis. *ACM Computing Surveys (CSUR)*, *45*(2), 20.

Kreiss, D., Finn, M., & Turner, F. (2011). The limits of peer production: Some reminders from max weber for the network society. *New Media & Society*, *13*(2), 243–259.

Kroll, J. A., Davey, I. C., & Felten, E. W. (2013). The economics of bitcoin mining, or bitcoin in the presence of adversaries, In *Proceedings of weis*.

Küçüktunç, O., Saule, E., Kaya, K., & Çatalyürek, Ü. V. (2012). Recommendation on academic networks using direction aware citation analysis. *arXiv preprint arXiv:1205.1143*.

Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D, Tomkins, A., & Upfal, E. (2000). Stochastic models for the web graph, In *Proceedings 41st annual symposium on foundations of computer science*. IEEE.

Lampe, R. (2012). Strategic citation. *Review of Economics and Statistics*, *94*(1), 320–333.

Lamport, L., Shostak, R., & Pease, M. (1982). The byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, *4*(3), 382–401.

Larson, R. R. (1996). Bibliometrics of the world wide web: An exploratory analysis of the intellectual structure of cyberspace, In *Proceedings of the annual meeting-american society for information science*.

Lehar, A., & Parlour, C. A. (2019). *Liquidity demand and bitcoin transaction fees* (tech. rep.). Working paper.

Leibowitz, N., Ripeanu, M., & Wierzbicki, A. (2003). Deconstructing the kazaa network, In *Proceedings the third ieee workshop on internet applications. wiapp 2003*. IEEE.

Leskovec, J., Backstrom, L., Kumar, R., & Tomkins, A. (2008). Microscopic evolution of social networks, In *Proceedings of the 14th acm sigkdd international conference on knowledge discovery and data mining*.

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2005). Graphs over time: Densification laws, shrinking diameters and possible explanations, In *Proceedings of the eleventh acm sigkdd international conference on knowledge discovery in data mining*.

Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2009). Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, *6*(1), 29–123.

Lessig, L. (2004). The creative commons. *Mont. L. Rev.*, *65*, 1.

Levine, B. N., Shields, C., & Margolin, N. B. (2006). A survey of solutions to the sybil attack. *University of Massachusetts Amherst, Amherst, MA*, *7*, 224.

Li, F., Mittal, P., Caesar, M., & Borisov, N. (2012). Sybilcontrol: Practical sybil defense with computational puzzles, In *Proceedings of the seventh acm workshop on scalable trusted computing*.

Li, X., & Watanabe, T. (2013). Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. *Procedia Computer Science*, *22*, 633–642.

Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the American society for information science and technology*, *58*(7), 1019–1031.

Liu, X., Suel, T., & Memon, N. (2014). A robust model for paper reviewer assignment, In *Proceedings of the 8th acm conference on recommender systems*. ACM.

Liu, Z., Luong, N. C., Wang, W., Niyato, D., Wang, P., Liang, Y.-C., & Kim, D. I. (2019). A survey on applications of game theory in blockchain. *arXiv preprint arXiv:1902.10865*.

Liu-Thompkins, Y. (2019). A decade of online advertising research: What we learned and what we need to know. *Journal of advertising*, *48*(1), 1–13.

Lockyer, M. (2018). Token curated registry (tcr) design patterns. *Medium, May 21*. https://hackernoon.com/token-curated-registry-tcr-design-patterns-4de6d18efa15. (accessed April 3, 2019).

Long, C., Wong, R. C.-W., Peng, Y., & Ye, L. (2013). On good and fair paper-reviewer assignment, In *2013 ieee 13th international conference on data mining*. IEEE.

Lovász, L. Et al. (1993). Random walks on graphs: A survey. *Combinatorics, Paul erdos is eighty*, *2*(1), 1–46.

Lukach, R., & Lukach, M. (2007). Ranking uspto patent documents by importance using random surfer method (pagerank). *Available at SSRN 996595*.

Lund, J. (2018). Stable coins: Enabling payments on blockchain through alternative digital currencies. https://www.ibm.com/blogs/blockchain/2018/07/stable-coins-enabling-payments-on-blockchain-through-alternative-digital-currencies/ (accessed May 5, 2020).

Lv, Q., Cao, P., Cohen, E., Li, K., & Shenker, S. (2002). Search and replication in unstructured peer-to-peer networks, In *Proceedings of the 16th international conference on supercomputing*. ACM.

Ma, N., Guan, J., & Zhao, Y. (2008). Bringing pagerank to the citation analysis. *Information Processing & Management*, *44*(2), 800–810.

Mahapatra, G. (2009). *Bibliometric studies in the internet era*. Indiana Publishing House.

Malliaros, F. D., & Vazirgiannis, M. (2013). Clustering and community detection in directed networks: A survey. *Physics reports*, *533*(4), 95–142.

Malone, C. (2011). Method and system for the requesting receipt and exchange of information [US Patent 7,991,624]. Google Patents.

Mandal, D., Leifer, M., Parkes, D. C., Pickard, G., & Shnayder, V. (2016). Peer prediction with heterogeneous tasks. *arXiv preprint arXiv:1612.00928*.

Margolin, N. B., & Levine, B. N. (2007). Informant: Detecting sybils using incentives, In *International conference on financial cryptography and data security*. Springer.

Marshall, A. (1890). *The principles of economics* (tech. rep.). McMaster University Archive for the History of Economic Thought.

McCain, K. W., & Salvucci, L. J. (2006). How influential is brooks' law? a longitudinal citation context analysis of frederick brooks' the mythical man-month. *Journal of Information Science*, *32*(3), 277–295.

McCreadie, R., Macdonald, C., Ounis, I., Giles, J., & Jabr, F. (2012). An examination of content farms in web search using crowdsourcing, In *Proceedings of the 21st acm international conference on information and knowledge management*. ACM.

Meho, L. I. (2007). The rise and rise of citation analysis. *Physics World*, *20*(1), 32.

Meyer, M. (2000). What is special about patent citations? differences between scientific and patent citations. *Scientometrics*, *49*(1), 93–123.

Meynkhard, A. (2019). Fair market value of bitcoin: Halving effect. *Investment Management and Financial Innovations*, *16*(4), 72–85.

Middleton, C., & Baeza-Yates, R. (2007). A comparison of open source search engines.

Miller, N., Resnick, P., & Zeckhauser, R. (2005). Eliciting informative feedback: The peer-prediction method. *Management Science*, *51*(9), 1359–1373.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European journal of operational research*, *246*(1), 1–19.

Mita, M., Ito, K., Ohsawa, S., & Tanaka, H. (2019). What is stablecoin?: A survey on price stabilization mechanisms for decentralized payment systems, In *2019 8th international congress on advanced applied informatics (iiai-aai)*. https://doi.org/10.1109/IIAI-AAI.2019.00023

Mita, M., Ito, K., Ohsawa, S., & Tanaka, H. (2019). What is stablecoin?: A survey on its mechanism and potential as decentralized payment systems. *arXiv preprint arXiv:1906.06037*.

Mohaisen, A., & Kim, J. (2013). The sybil attacks and defenses: A survey. *arXiv preprint arXiv:1312.6349*.

Moravcsik, M. J., & Murugesan, P. (1975). Some results on the function and quality of citations. *Social studies of science*, *5*(1), 86–92.

Moulin, H. (1986). *Game theory for the social sciences*. NYU press.

Myerson, R. B. (1977). Graphs and cooperation in games. *Mathematics of operations research*, *2*(3), 225–229.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, *85*(5), 1313–1326.

Naik, U., & Shivalingaiah, D. (2008). Comparative study of web 1.0, web 2.0 and web 3.0.

Nakamoto, S. Et al. (2008). Bitcoin: A peer-to-peer electronic cash system.

Nalimov, V. V., & Mulchenko, B. (1969). Scientometrics. *Studies of science as a process of information. Moscow, Russia: Science*.

Napster. (1999). Napster. https://us.napster.com (accessed October 22, 2019).

Napster. (2000). Napster messages. http://opennap.sourceforge.net/napster.txt (accessed October 22, 2019).

Nash-Williams, C. S. J. (1959). Random walk and electric currents in networks, In *Mathematical proceedings of the cambridge philosophical society*. Cambridge University Press.

Newman, M. (2010). *Networks: An introduction*. Oxford University Press.

Nisan, N., & Ronen, A. (2001). Algorithmic mechanism design. *Games and Economic behavior*, *35*(1-2), 166–196.

Nisan, N., Roughgarden, T., Tardos, E., & Vazirani, V. V. (2007). *Algorithmic game theory*. Cambridge university press.

Noda, S., Okumura, K., & Hashimoto, Y. (2020). An economic analysis of difficulty adjustment algorithms in proof-of-work blockchain systems, In *Proceedings of the 21st acm conference on economics and computation*.

Noruzi, A. (2005). Google scholar: The new generation of citation indexes. *Libri*, *55*(4), 170–180.

Noveck, B. S. (2006). Peer to patent: Collective intelligence, open review, and patent reform. *Harv. JL & Tech.*, *20*, 123.

Ocean Protocol, F. (2019). Ocean protocol: A decentralized substrate for ai data & services technical whitepaper. https://oceanprotocol.com/tech-whitepaper.pdf. (accessed March 14, 2020).

O'Neil, M. (2011). Wikipedia and authority. *G. Lovink and N. Tkacz (eds)*, 309–324.

O'Neil, M. (2014). Hacking weber: Legitimacy, critique, and trust in peer production. *Information, Communication & Society*, *17*(7), 872–888.

Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature i. *Libri*, *46*(3), 149–158.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The pagerank citation ranking: Bringing order to the web.* (tech. rep.). Stanford InfoLab.

Pagnotta, E., & Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets. *Available at SSRN 3142022*.

Parthasarathy, S., Ruan, Y., & Satuluri, V. (2011). Community discovery in social networks: Applications, methods and emerging trends, In *Social network data analytics*. Springer.

Pease, M., Shostak, R., & Lamport, L. (1980). Reaching agreement in the presence of faults. *Journal of the ACM (JACM)*, *27*(2), 228–234.

Peterson, J., Krug, J., Zoltu, M., Williams, A. K., & Alexander, S. (2015). Augur: A decentralized oracle and prediction market platform. *arXiv preprint arXiv:1501.01042*.

Pillai, S. U., Suel, T., & Cha, S. (2005). The perron-frobenius theorem: Some of its applications. *IEEE Signal Processing Magazine*, *22*(2), 62–75.

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information processing & management*, *12*(5), 297–312.

Pöschl, U. (2004). Interactive journal concept for improved scientific publishing and quality assurance. *Learned Publishing*, *17*(2), 105–113.

Price, D. J. D. S. (1965). Networks of scientific papers. *Science*, 510–515.

Price, D. J. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American society for Information science*, *27*(5), 292–306.

Price, D. J. D. S. (1978). Editorial statements. *Scientometrics*, *1*(1), 3–8.

Pritchard, A.  Et al. (1969). Statistical bibliography or bibliometrics. *Journal of documentation*, *25*(4), 348–349.

Prug, T. (2010). Open-process academic publishing. *Ephemera: Theory & Politics in Organization*, *10*(1).

Radanovic, G., & Faltings, B. (2013). A robust bayesian truth serum for non-binary signals, In *Proceedings of the 27th aaai conference on artificial intelligence (aaai" 13)*.

Rahman, M. R. (2009). A survey of incentive mechanisms in peer-to-peer systems. *Cheriton School of Computer Science, University of Waterloo, Tech. Rep. CS-2009-22*.

Ramaswamy, L., & Liu, L. (2003). Free riding: A new challenge to peer-to-peer file sharing systems, In *36th annual hawaii international conference on system sciences, 2003. proceedings of the*. IEEE.

Ratliff, J. D., & Rubinfeld, D. L. (2010). Online advertising: Defining relevant markets. *Journal of Competition Law and Economics*, *6*(3), 653–686.

Ratnasamy, S., Francis, P., Handley, M., Karp, R., & Shenker, S. (2001). *A scalable content-addressable network* (Vol. 31). ACM.

Raval, S. (2016). *Decentralized applications: Harnessing bitcoin's blockchain technology*. " O'Reilly Media, Inc."

Resnik, D. B., & Elmore, S. A. (2018). Conflict of interest in journal peer review. SAGE Publications Sage CA: Los Angeles, CA.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social networks*, *29*(2), 173–191.

Rodriguez, M. A., & Bollen, J. (2006). Simulating network influence algorithms using particle-swarms: Pagerank and pagerank-priors. *arXiv preprint cs/0602002*.

Rodriguez, M. A., & Bollen, J. (2008). An algorithm to determine peer-reviewers, In *Proceedings of the 17th acm conference on information and knowledge management*.

Rowstron, A., & Druschel, P. (2001). Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems, In *Ifip/acm international conference on distributed systems platforms and open distributed processing*. Springer.

Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A bayesian approach to filtering junk e-mail, In *Learning for text categorization: Papers from the 1998 workshop*. Madison, Wisconsin.

Saito, K., & Iwamura, M. (2019). How to make a digital currency on a blockchain stable. *Future Generation Computer Systems*, *100*, 58–69.

Saroiu, S., Gummadi, P. K., & Gribble, S. D. (2001). Measurement study of peer-to-peer file sharing systems, In *Multimedia computing and networking 2002*. International Society for Optics and Photonics.

Schaeffer, S. E. (2007). Graph clustering. *Computer science review*, *1*(1), 27–64.

Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics Management*, (1), 93–107.

Schollmeier, R. (2001). A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications, In *Proceedings first international conference on peer-to-peer computing*. IEEE.

Seidel, A. H. (1949). Citation system for patent office. *Journal of the Patent Office Society*, *31*(5), 54.

Shaffer, M. J. (2011). Entrepreneurial innovation: Patent rank and marketing science.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*(11), 2498–2504.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Shnayder, V., Agarwal, A., Frongillo, R., & Parkes, D. C. (2016a). Informed truthfulness in multi-task peer prediction, In *Proceedings of the 2016 acm conference on economics and computation*. ACM.

Shnayder, V., Frongillo, R., & Parkes, D. C. (2016b). Measuring performance of peer prediction mechanisms using replicator dynamics.

Simperl, E., Cuel, R., & Stein, M. (2013). Incentive-centric semantic web application engineering. *Synthesis Lectures on the Semantic Web: Theory and Technology*, *3*(1), 1–117.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, *24*(4), 265–269.

Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the royal society of medicine*, *99*(4), 178–182.

Soergel, D., Saunders, A., & McCallum, A. (2013). Open scholarship and peer review: A time for experimentation.

Specification, G. P. (2003). V0. 4. *Available from World Wide Web: http://www9. limewire. com/developer/gnutella protocol 0.4. pdf, 155.*

Steinmetz, R., & Wehrle, K. (2005). *Peer-to-peer systems and applications* (Vol. 3485). Springer.

Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., & Balakrishnan, H. (2001). Chord: A scalable peer-to-peer lookup service for internet applications. *ACM SIGCOMM Computer Communication Review*, *31*(4), 149–160.

Stox. (2017). Stox platform for prediction markets white paper v03. https://resources.stox.com/stox-whitepaper.pdf. (accessed March 14, 2020).

Suryanarayana, G., & Taylor, R. N. (2004). A survey of trust management and resource discovery technologies in peer-to-peer applications.

Szabo, N. (1997). Formalizing and securing relationships on public networks. *First Monday*, *2*(9).

Tang, W., Tang, J., Lei, T., Tan, C., Gao, B., & Li, T. (2012). On optimization of expertise matching with various constraints. *Neurocomputing*, *76*(1), 71–83.

Tapscott, D., & Williams, A. D. (2008). *Wikinomics: How mass collaboration changes everything*. Penguin.

Thelwall, M. (2002). Conceptualizing documentation on the web: An evaluation of different heuristic-based models for counting links between university web sites. *Journal of the American Society for Information Science and Technology*, *53*(12), 995–1005.

Thomas, P. (1999). The effect of technological impact upon patent renewal decisions. *Technology Analysis & Strategic Management*, *11*(2), 181–197.

Thulasiraman, K, & Swamy, M. (1992). *Graphs: Theory and algorithms*. John Wiley & Sons.

Tiutiun, R., Porco, L., Gord, M., & Lee, D. S. (2018). Usdx: A decentralized monetary policy system.

Tomaino, N. (2017). Stablecoins: A holy grail in digital currency. https://thecontrol.co/stablecoins-a-holy-grail-in-digital-currency-b64f3371e111 (accessed May 5, 2020).

Trajtenberg, M. (1989). The welfare analysis of product innovations, with an application to computed tomography scanners. *Journal of political Economy*, *97*(2), 444–479.

Trajtenberg, M. (1990). A penny for your quotes: Patent citations and the value of innovations. *The Rand Journal of Economics*, 172–187.

Trotman, A., Clarke, C. L., Ounis, I., Culpepper, S., Cartright, M.-A., & Geva, S. (2012). Open source information retrieval: A report on the sigir 2012 workshop, In *Acm sigir forum*. ACM.

Tsoumakos, D., & Roussopoulos, N. (2003). Adaptive probabilistic search for peer-to-peer networks, In *Proceedings third international conference on peer-to-peer computing (p2p2003)*. IEEE.

Uhlir, A. (2017). Dagui: A dataflow graphical user interface.

UNESCO. (n.d.). What is open access? https://en.unesco.org/open-access/what-open-access. (accessed February 5, 2021).

Von Ahn, L., Blum, M., Hopper, N. J., & Langford, J. (2003). Captcha: Using hard ai problems for security, In *International conference on the theory and applications of cryptographic techniques*. Springer.

Voshmgir, S. (2019). *Token economy: How blockchains and smart contracts revolutionize the economy*. BlockchainHub.

Waggoner, B., & Chen, Y. (2014). Output agreement mechanisms and common knowledge., In *Hcomp*.

Wang, F., Chen, B., & Miao, Z. (2008). A survey on reviewer assignment problem, In *International conference on industrial, engineering and other applications of applied intelligent systems*. Springer.

Wang, F., Shi, N., & Chen, B. (2010). A comprehensive survey of the reviewer assignment problem. *International Journal of Information Technology & Decision Making*, *9*(04), 645–668.

Wang, F., & Sun, Y. (2008). Self-organizing peer-to-peer social networks. *Computational Intelligence*, *24*(3), 213–233.

Wang, Y. L., & Krishnamachari, B. (2018). Enhancing engagement in token-curated registries via an inflationary mechanism. *arXiv preprint arXiv:1811.09680*.

Wąs, T., Rahwan, T., & Skibski, O. (2019). Random walk decay centrality, In *Proceedings of the aaai conference on artificial intelligence*.

Watanabe, S., Ito, T., Ozono, T., & Shintani, T. (2005). A paper recommendation mechanism for the research support system papits, In *International workshop on data engineering issues in e-commerce*. IEEE.

Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, *34*(2), 331–341.

West, D. B. Et al. (1996). *Introduction to graph theory* (Vol. 2). Prentice hall Upper Saddle River, NJ.

Wightman, D. (2010). Crowdsourcing human-based computation, In *Proceedings of the 6th nordic conference on human-computer interaction: Extending boundaries*. ACM.

Wilcox-O'Hearn, B. (2002). Experiences deploying a large-scale emergent network, In *International workshop on peer-to-peer systems*. Springer.

Witkowski, J., & Parkes, D. C. (2012). Peer prediction without a common prior, In *Proceedings of the 13th acm conference on electronic commerce*. ACM.

Wood, G. Et al. (2014). Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, *151*(2014), 1–32.

Wu, B., & Davison, B. D. (2005). Identifying link farm spam pages, In *Special interest tracks and posters of the 14th international conference on world wide web*.

Xia, J., Gilchrist, S. B., Smith, N. X., Kingery, J. A., Radecki, J. R., Wilhelm, M. L., Harrison, K. C., Ashby, M. L., & Mahn, A. J. (2012). A review of open access self-archiving mandate policies. *portal: Libraries and the Academy*, *12*(1), 85–102.

Xu, Y., Zhao, H., Shi, X., & Shah, N. B. (2019). On strategyproof conference peer review, In *Proceedings of the 28th international joint conference on artificial intelligence*. AAAI Press.

Yarowsky, D., & Florian, R. (1999). Taking the load off the conference chairs-towards a digital paper-routing assistant, In *1999 joint sigdat conference on empirical methods in natural language processing and very large corpora*.

Yu, H. (2011). Sybil defenses via social networks: A tutorial and survey. *ACM SIGACT News*, *42*(3), 80–101.

Yuen, M.-C., King, I., & Leung, K.-S. (2011). A survey of crowdsourcing systems, In *2011 ieee third international conference on privacy, security, risk and trust and 2011 ieee third international conference on social computing*. IEEE.

Zamfir, V. (2015). What is cryptoeconomics? *Presentation at CryptoEconomicon 2015*. https://www.youtube.com/watch?v=9lw3s7iGUXQ. (accessed February 27, 2020), Crypto Technology Conference.

Zhao, B. Y., Huang, L., Stribling, J., Rhea, S. C., Joseph, A. D., & Kubiatowicz, J. D. (2004). Tapestry: A resilient global-scale overlay for service deployment. *IEEE Journal on selected areas in communications*, *22*(1), 41–53.