THE SOCIALIZED INFRASTRUCTURE OF THE INTERNET ON THE COMPUTING LEVEL

Bing Li

Department of Computer Engineering, Xi'An Technological University, XiAn, China

ABSTRACT

To share the huge amount of heterogeneous information to the large-scale of heterogeneous users, the Internet on the computing level should be reconstructed since such a crucial infrastructure was designed without proper understandings. To upgrade, it must consist of five layers from the bottom to the top, including the routing, the multicasting, the persisting, the presenting and the humans. The routing layer is responsible for establishing the fundamental substrate and finding resources in accordance with social disciplines. The multicasting layer disseminates data in a high performance and low cost way based on the routing. The persisting layer provides the services of storing and accessing persistent data efficiently with the minimum dedicated resources. The presenting layer absorbs users' interactions to guide the adjustments of the underlying layers other than shows connected local views to users. Completely different from the lower software layers, the topmost one is made up totally with humans, i.e., users, including individual persons or organizations, which are social capital dominating the Internet. Additionally, within the upgraded infrastructure, besides the situation that a lower layer supports its immediate upper one only, the humans layer influences the lower ones in terms of transferring its social resources to them. That is different from any other traditional layer-based systems. Those resources lead to adaptations and adjustments of all of the software layers since each of them needs to follow social rules. Eventually, the updated underlying layers return latest consequences to users upon those modifications.

KEYWORDS

Social Computing, World Wide Web, Peer-to-Peer computing, Software Architecture, Social Network Services, Information Sharing, Routing

1. INTRODUCTION

It is believed that the Internet on the computing level should be reconstructed since it has never been designed in a proper way [1]. For its existing downsides, it is tough to share the huge amount of heterogeneous information to the large-scale of heterogeneous users. To upgrade, a five-layered infrastructure is proposed, which includes the routing, the multicasting, the persisting, the presenting and the humans from the bottom to the top, respectively.

When the Internet was proposed in 1960s, one neither envisioned what it would become in the future nor knew the essence and its disciplines at that time. Although the Internet became succeeded in terms of businesses, almost all of the marvellous applications were invented arbitrarily. That is why Ray Tomlinson [2], th[1]e inventor of the email, said, "Don't tell anyone! This isn't what we're supposed to be working on." And, the Google [3] founders attempted to sell their search engines to Excite [4] under one million dollars [5]. The famous Internet protocol, HTTP [6], stands for Hypertext Transfer Protocol. However, it is well known that texts are not the dominant data [7] of the current Internet.

The term of the Internet on the computing level is defined as the Internet-based distributed environment concentrated on resource management and fundamental architectures to support high performance and low cost applications. Usually, it is also be called, the Internet on the application level [8][9], which discusses about the multicasting over a large-scale distributed environment. As a counterpart beneath it, the Internet on the communication level or the network level [10] aims to support raw data transmission between two nodes over the Internet. At present, TCP/IP [11] is its primary representative of the level. Since the communication level is far from specific distributed environments, it can hardly come up with a high quality application for end users upon its solutions only though it is also critical. This paper extends the definition of the computing level in the sense that all of the issues over TCP/IP are included in the scope. To simplify, when mentioning the term of the Internet in the paper, it specifies the one on the computing level.

The routing layer is responsible to manage and retrieve the tremendously huge resources over the Internet. As a common sense, any distributed systems have a mechanism [12] through which any distributed nodes are able to find others or any other shared computing resources to collaborate. The requirement becomes more critical in the case of the Internet because of its large-scale while such a system works in a different manner. Unfortunately, the early designers did not pay much attention to that. Until now, a proper routing mechanism is still not available over the Internet. DNS [12] is incompetent with the role since it provides a mapping between URLs [13] and IP addresses only. Although Google-like systems take this responsibility and win the huge successes in the industrial domain, they do not touch the essences. It is convinced that the routing over the Internet must follow the disciplines of social resources [14] rather than any other machine-based algorithms [14].

It is no doubt that the multicasting layer is mandatory for the Internet system. At this moment, multicasting over the Internet also works in a high cost and low performance manner because of the lack of appropriate solutions. Until now, the protocol, HTTP [6], is still one of the most important solutions to transfer data over the Internet. However, when it was proposed, it aimed to transmit text-based or lightweight data without the timing issue. For that, it missed a lot of designs, such as the associated topology [9], the transmission environment and the data characteristics, etc. When heavyweight [16][17][18] data or timing-sensitive [16][17][18] data are to be disseminated, the protocol becomes out of date. The situation promotes the popularity of some so-called downloading applications, such as BitTorrent [19] and eMule [20]. To improve that, it is required to design a bunch of new algorithms based on the routing layer to distribute heavyweight and timing-sensitive data in the high performance and low cost manner.

The third layer is the persisting that supports high performance huge data saving and accessing over the Internet scale with minimum dedicated resources investments. In traditional distributed systems, persistent data is handled by the techniques of DFS (Distributed File Systems) [12], in which all of storage resources are dedicated. Unfortunately, it is impossible to apply them for data over the Internet because of the issues of the amount of data, the performance and the cost. The dilemma in the layer is that the performance can hardly be raised even with the availability of more dedicated resources for the special characters of the Internet. To upgrade, it must be aware that the law of the socialized distributed environment is extremely different from that of traditional machine-based distributed systems. The utilization of undedicated resources [22] that are controlled by their proprietary owners is crucial to the ideal data persisting layer. In addition, it is also required to enhance the system with the strong assistance from the underlying layers, i.e., the routing and the multicasting.

The presenting layer provides a fundamental application to consumers of the Internet services. The layer attempts to ameliorate the traditional WWW in terms of raising the quality of data connections and providing upgraded publishing and reading approaches. Moreover, the layer is also an approach to bind its upper layer with the underlying ones together. Traditionally, the

World Wide Web (WWW) [23] plays the role, which aims to connect all of the relevant data, i.e., Web pages [23], over the huge system. However, it does not provide mandatory management systems to support the linked world and it has no any guarantee to the quality of the connections among pages. Therefore, with the rapid growth of the Internet world, the WWW becomes withered since no one cares about whether a page is linked properly such that it does not make sense to follow any links to keep accessing referenced data.

The topmost layer is the one of the humans, which does not exist in any other traditional software infrastructures. When talking about an infrastructure in a conventional system, all of its subsystems run independently without human interventions. Even though interactions between users and the system are needed, they are different from the dominant effects of the humans layer to its underlying ones. Briefly speaking, over the Internet, the humans layer not only determines the consequences of the underlying layers but also governs the performance and cost of them, which is obviously beyond traditional user inputs or interactions. Moreover, for the distinctions of the humans layer, its influences are transferred to each of the underlying layers although it owns only one direct link between itself and the routing layer, which is responsible for absorbing and managing social resources from the humans layer. Traditionally, each lower layer is visible to its immediate upper neighbor only, such as the Open Systems Interconnection (OSI) model [24] or the n-tier distributed system architecture [25]. It is also a design principle of software engineering, such as the concept of encapsulation [26]. In another word, it means that a directed connection exists between two immediate neighboring layers, which play the roles of the source and the destination respectively from the lower to the upper. In contrast, the influences of the humans layer of the Internet span all of the layers, i.e., starting from itself, passing through the routing layer, the multicasting layer, the persisting layer, to the presenting layer finally.

It should emphasize that the paramount value is raised from users of the Internet rather than any other traditional algorithms. For that, a new computing model is set up for large-scale distributed systems over the Internet. That is, while the resources of the systems must include traditional ones, such as CPU cycles, storage and bandwidth, it is more important to introduce a new kernel to control those resources. The new kernel is the humans layer which provides the most precious value, i.e., the socialized capital [14][15], to the Internet.

The main contributions of this work can be summarized as follows.

It proposes a new infrastructure to construct the Internet on the computing level, which consists of five layers, the routing, the multicasting, the persisting, the presenting and the human layer. It also emphasizes the other layers are dominated by the humans layer, which is different from supporting one's immediate upper layer from the bottom to the top in traditional layer-based architectures.

It indicates that social capital from the humans layer becomes the primary resource not only to raise performance and lower costs but also to fulfill users' publishing and reading requirements. Thus, it is required to take into account it when upgrading the Internet besides traditional computing resources.

It puts forward that connected local views are the summary of users' requirement. The five layers aim to provide users with such an environment to access the huge amount of heterogeneous data over the Internet.

The paper consists of the following sections to present its primary ideas. Section 2 discusses the previous related researches in the domain and makes concise comparisons. Section 3 summarizes the generic requirements of users over the Internet. Section 4 explains in details about the layers

of the Internet on the computing level. Section 5 articulates the collaboration of the layers. Section 6 gives a brief summary about the current and future work.

2. RELATED WORK

The Internet has emerged for more than forty years. During the days, its drawbacks on design also came into being for such a long time. As Barabasi Albert-Laszlo [1] said researchers would not design the Internet like present if they knew its characters as deeply as nowadays [1].

2.1. The Routing

For any distributed systems, the first question is how to design a routing algorithm that helps each node to find colleagues and computing resources [12]. The issue is extraordinarily critical for a large-scale system over the Internet since such a system is open, huge, heterogeneous and even unrevealed such that no any dedicated coordinators are competent to handle the problem for the rest world. For the Internet, publishing and reading are two primary requirements of users. However, no efficient routing is designed for them.

2.1.1. Routing for Publishing

As the fundamental application of the Internet, the Web [23] aims to share data in the form of pages to users, which is defined as the publishing. For any authors, the goal of routing for publishing is to win as many readers as possible. However, when publishing a page, no any routing answers which users would like to be the potential readers. Therefore, the only solution is to post the page on a certain web site according to the author's experiences. The phenomenon still exists widely until now.

The problem is caused primarily by the fact that all of the users over the Internet are not registered uniquely. Moreover, it is tough to identify their interests even though their profiles are available. Thus, it can hardly find appropriate readers for publishing. Some later applications over the Internet intend to overcome the problem, such as Napster [27], Freenet [28] and Gnutella [29]. At least, almost of all of those systems require users to have an account. Even some traditional web sites would invite users to access them after registering. Recently, as the social network systems like Facebook [30], Twitter [31] and WeChat [32] became dominant, it was popular to have an identification to enter them. More different, any node in such a system always has a number of followers or friends. Therefore, when publishing, those users become the potential readers. This is a rough solution to the routing for publishing.

While those systems make some progresses, the issue of routing for publishing still remains the old states in the overall infrastructure of the Internet on the computing level. It is still required to construct a more open environment to attract potential readers than the closed one. The exact routing algorithms should be refined deeply by taking into account underlying social networks.

2.1.2. Routing for Reading

The questions on the issue of routing for reading are what is the appropriate way to present the huge amount of data to users and how to accomplish it, respectively.

Compared with its counterpart, the issue of routing for reading obtains many solutions during the past days. When the Web was popular in 1990s, for the lack of routing solutions for pages, a directory service, Yahoo! [33], took the role to categorize URLs [13] of web sites to help users navigate the Web. It believes that data is exhibited in a hierarchical manner. Therefore, before

reading real data, users made a series of choices about the categories of their requirements. Even though doing that correctly, it does not guarantee that the obtained data meet their interests. Additionally, it is difficult to maintain such a huge amount of heterogeneous web sites manually by a limited number of hired staffs. Even worse, the huge Web can hardly be abstracted to the static hierarchy of URLs upon a limited count of categories [34].

Then, the keyword searching, such as Google [3], takes the place of directory services. It aims to show data that matches users' keywords exactly or as approximately as possible. The service fulfills one of users' requirements. However, it is far from to answer the core question comprehensively. Different from traditional data retrieval approaches [35], it is expected that retrieved results form a particular structure upon their inherent senses rather than one series of equivalent pages or a list of irrelevant ones. Moreover, the consequences of such a retrieval system rely heavily on the form of keywords from users. Over the Web, the quantity of pages is extremely larger than that of words of natural languages. Thus, Google-like systems work better when searching by a sequence of keywords existing in pages rather than by a fuzzy term or a sequence of ones that should be split or analyzed. A sequence of keywords existing in pages maps uniquely to one or a limited number of copies whereas the fuzzy one leads to a large amount of results for ambiguous matching since it is tough to understand users' inputs and then match them within such a huge system. Another problem is that the systems have to match all of indexed pages crawled from the Web for any keywords and most of matching is a waste of cost. But it has no efficient solutions to avoid it with the current methodologies.

After that, some social network [36] based systems, such as Gnutella [29], Freenet [28], Twitter [31], Facebook [30] and WeChat [32], emerge to overcome the above problems. No matter whether their inventors are aware of them or not, those systems conduct their services in a more reasonable way. That is, data is presented in a context formed by a reader's friends or followees. While it can hardly claim that the data presentation is the ideal one, it makes an obvious progress such that users keep highly active to read information from the context. The topic is tough since it associates with the difficult question, i.e., what on earth a large-scale of heterogeneous and homogenous data is interweaved [37][38]. If it was answered, the ultimate solution to present data could be proposed. That is the goal of the routing for reading. The paper puts forward some disciplines to do that. Eventually, the potential solution is applied to the underlying infrastructure of the Internet on the computing level comprehensively in order to support the multicasting, the persisting and the presenting other than users.

2.2. The Multicasting

Multicasting is an old topic in a distributed system. However, over the Internet, more novel solutions are taken into account for its unique natures. The multicasting is divided into two categories, the messaging and the streaming, according to the data transmission character. If data is transferred in a request/response manner, it belongs to the messaging protocol. Instead, if data is transmitted among nodes in the form of flowing from a source to a destination continuously, it is believed to be a streaming protocol.

2.2.1. The Messaging

The most popular solution to data distribution over the Internet is the protocol of HTTP [6], which transfers data in the messaging way. Through it, data resides on a centralized server unless it is requested by a client. For that, the server usually encounters the difficulty to respond clients when receiving a large number of concurrent requests. To resolve the problem, it has to prepare for a powerful server to endure the potential high pressure. Even worse, because of the richness of heavyweight data, such as video, the cost of the server becomes higher. In brief, as a messaging

protocol, the HTTP is suitable to the environment of lightweight data distribution with restricted concurrency. To distribute heavyweight data with intensive concurrency, it is required to pay for high performance servers or even a cluster of them.

Since the HTTP is supported by a hidden protocol, the TCP [11], which is solved by the communication level and no computing level routing is available, the entire transmission load has to be taken by a centralized server only. It is expected to upgrade the distributed environment and provide high quality routing services to accelerate the messaging.

2.2.2. The Streaming

As the HTTP is not a proper solution to transmit heavyweight data, a bunch of competitors are proposed, such as BitTorrent [19], BiToS [17] and Tribler [39], etc. All of those protocols attempt to form a decentralized topology to transmit heavyweight data in a streaming way with low cost and high performance.

With BitTorrent, all of the nodes perform two concurrent tasks, i.e., receiving and sending data. For its tit-for-tat algorithms [19], each node is willing to contribute in order to consume more. Its decentralized structure makes the distribution scalable and economic. The problems of BitTorrent include the random routing to find collaborators before a transmission and the low-quality routing during the procedure of the transmission. All of those are caused by the fact that BitTorrent is not aware of the preferences of multicasting participants, especially the social relationships among them.

BitTorrent is a protocol to transmit high volume timing-insensitive data like a large size of software package rather than timing-sensitive data like video because of its rarest-first algorithm [19]. To multicast timing-sensitive heavyweight data, BiToS splits data into two sets, a high priority one and a low priority one [17]. The size of the high priority set is apparently smaller than that of the low priority one. Then, it performs BitTorrent algorithms on both of them concurrently, but the high priority set obtains more CPU cycles than the low priority one. Therefore, although the rarest-first algorithm is still applied, the playback buffered data from the high-priority set is filled fast enough almost in a consecutive sequence. Meanwhile, the low priority set gains some data which is turned into the high priority set with the progress of video playback. It also lowers the load to transfer the high priority one. The problem of BiToS is the same as that of BitTorrent in terms of the lack of the support of high quality routing.

It is appropriate to regard Tribler as Facebook or Twitter for multicasting. To overcome the problem of classic BitTorrent, it finds potential collaborators from a buddy list [39] before performing the multicasting with BitTorrent. It is believed that the nodes in the buddy list must be willing to form a high efficiency multicasting topology to speed up heavyweight data transmission. However, the drawback of TRIBLER is also caused by the buddy list. While it is better than the random routing of BitTorrent, it cannot always guarantee the high performance since it does not make sense that a user always has the same preference as its friends.

It should consider the solution to put users into a wider context other than the group of friends. Then, it provides a node with more opportunities to find potential collaborators based on either social relations or common interests. Moreover, it is believed that the routing for multicasting is an interactive and adaptive procedure on the fly rather than a one-step task performed before data transmission. Thus, it is also required to improve the multicasting from that aspect.

2.3. The Persisting

The principal problem to persist the huge amount of data over the Internet is to confirm the crucial disciplines that determine the performance and the cost. Such an environment must follow the distinguished rules. Thereafter, reasonable solutions can be put forward.

The solutions of cloud computing [40] answer the above questions by providing sufficient dedicated resources with strengthened fundamental distributed frameworks [41][42] to establish the persisting system like data centers [43]. It is believed to be a traditional solution that does not consider the special characters of the Internet. Thus, the approach is a proper solution to enterprises rather than a high quality Internet foundation environment.

OceanStore [22] is a system which was proposed before the concept of cloud computing. However, distinguished from the later one, the technique utilizes the undedicated resources from proprietary users like cafeteria owners by attracting them with business incentives. The approach lowers the cost to persist a huge amount of data with the participation of business-oriented resources if the business plan is reasonable. In particular, with the support of the non-technical design, the performance even raises to a high degree, especially for the popular data.

It is required to apply the experiences of OceanStore widely to the persisting layer of the Internet on the computing level with in-depth understanding of the Internet. Social incentives and social capital other than business ones must exist pervasively in the design of the Internet infrastructure.

2.4. The Presenting

The World Wide Web [23] is the fundamental application of the Internet on the computing level. Without the support of an essential underlying infrastructure, the WWW becomes withered such that it is losing the value to present and manage the huge amount of data over the Internet.

Upon the World Wide Web, when creating a page, an author is authorized the privilege to connect one single term or one phrase inside the page to any other existing one over the Internet through the techniques of hyperlinks [44]. Tim-Berners Lee [45] believed that relevant data was weaved together in this way such that readers can follow the links to access the Internet in an efficient and significant manner.

Unfortunately, after a short period success, WWW became faded such that neither publishers nor readers took care about its original motivation. It is seldom for an author to spend sufficient time adding links inside a page since it brings additional workload for publishing. A reader loses interests to navigate along the routes formed by the links since it does not offer obvious benefits. As a matter of fact, WWW is being degenerated into a huge graph, which lacks meaningful connections from a user's point of view. While the claim is invalid when considering WWW as a complex network [46], a hyperlink is degraded to be a one-step pointer from an index-like or hub page [47] to ordinary or authority pages [47]. And, advertisements replace many high quality pages.

2.5. Conclusions

In brief, the current Internet on the computing level was full of drawbacks in terms of its infrastructure. It results in high cost and low performance applications. Moreover, it is also a barrier to fulfill users' functional requirements when they access the huge amount of heterogeneous data. The Internet on the computing level must be redesigned upon the nature of

the Internet to provide publishers and readers with an appropriate environment to access the large information world.

3. THE REQUIREMENTS TO THE INTERNET ON THE COMPUTING LEVEL

For the distinguished characters of the Internet, it needs to clarify users' requirements when accessing it. In the procedure of the development, for the lack of sufficient knowledge about its natures, no proper designs were available for this environment in the early stage. A couple of applications [2][32][3][48][49] emerge as the giants of the Internet out of imagination of all of the experts and even the inventers themselves. Those applications fulfill users' requirements over the Internet to some extent. That is why they are widely accepted eventually.

3.1. The Requirements

The principal goal of the Internet aims to share data in a large-scale heterogeneous environment. Because of the distinctions of the Internet, it is necessary to indicate that the requirements of sharing data within the environment become different from those of other traditional information systems. The requirements influence the design of the Internet on the computing level.

3.1.1. Finding Readers

When publishing in a large-scale system, the first question to answer is who are the potential readers. In any traditional systems, it is not necessary to consider the issue since the scale is small and the potential readers are predefined implicitly, such as a bulletin board system [50] for a particular organization. However, when doing that over the Internet, it becomes a difficult question since it is tough to learn whether a particular user is willing to accept or read the data to be published such that it is impossible to transfer the data to any users directly.

3.1.2. Transmitting Data

Even though the potential readers are found, what is the best approach to deliver data to them? Besides the issue of network bandwidth, it usually depends on the characters of data and the count of potential readers. If data is lightweight and the scale of readers is small, the transmission approach is simple while this is not a popular case. Usually, it is required to disseminate data to a large-scale of readers. Specially, if data is heavyweight and timing-sensitive, the load becomes high. Therefore, the transmission must be redefined to lower cost and raise performance.

3.1.3. Persisting Data

Because of the existence of a huge amount of data in a heterogeneous environment around the globe, it is required to be persisted with an efficient solution in terms of high performance and low cost. As a global scale system, data must be replicated to ensure the performance and availability for each corner of the world. It is required to take into account geographical locations to store replicas and the count to replicate for each place. More difficult, it is mandatory to consider the issues of heterogeneous tastes of users for the globe distribution.

3.1.4. Presenting Data

Even though the total amount of data is huge, any user has to access data through a small screen as slowly as reading news from a blackboard. The approach to present data must solve the contradiction. First of all, it is necessary to place data in a certain form to assist users to enter,

browse and navigate the information world. In addition, it is required to put forward a new solution to fulfill users' preferences by looking for a needle in a haystack. However, the needle should be attached with a pile of hay rather than separated from others. This is usually more important.

3.2. The Connected Local View

All of the requirements of users can be summarized as a simple phrase, the connected local view. It is defined as a user's perspective for data within a large-scale heterogeneous socialized environment. To understand all of the details, it is necessary to retrieve multiple domains of knowledge, such as social cognition [51], social networks [36], information and library science [34], interpersonal communication [52], economics [53][54] and so forth. The way to construct the connected local view, as illustrated in Figure 1, for each user is identical to the solution to fulfill the requirements of users of the Internet on the computing level.
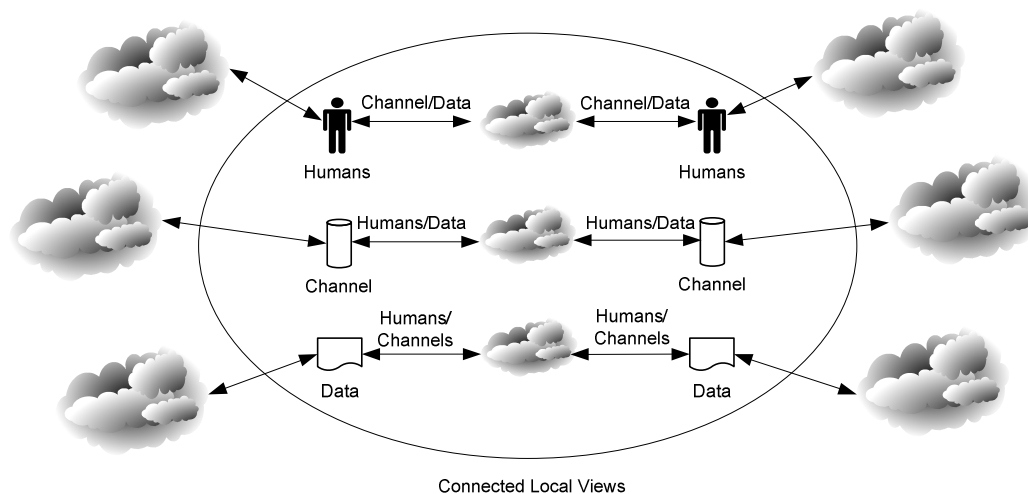

Figure 1. The Connected Local View

3.2.1. The Connected.

To share the huge amount of data, the first job is to connect them in a certain manner such that it is possible to take into account them as a whole. Fortunately, it is known that the huge count of nodes form large-scale graphs over the Internet. While they are also complex networks, the graphs are different from traditional ones [46] in terms of their nodes and connections.

The types of the graphs are divided by their node natures, i.e., data, channels and humans. Data is the information published over the Internet. The concept of channels represents the information categories, a virtual location, into which data is published. Humans are the large number of users who either publish or read over the Internet. More interesting, any one of the nodes plays the role of connections in the graphs of other nodes. That is, data is connected by either channels or humans in the data graph, channels are connected by either data or humans in the channel graph and humans are connected by either data or channels in the human graph. As a matter of fact, they are connected under the drivers of their inherent senses and associations. Except those innate powers, no any other algorithms could determine the relations among those nodes.

3.2.2. The Cluster

Another apparent character of the large-scale graphs is that it contains clusters [46], which are made up with the three types of nodes, respectively. Inside any one of the clusters, the connections are tight whereas they become loose outside them [55].

The same as that of connections, the clusters are also formed under the control of the inherent force inside and among those nodes rather than any other energy outside them. Each of the clusters represents a certain sense, which can hardly be articulated in a few words during the life cycle of the cluster with any machine algorithms. In short, a cluster is made up with the nodes which are interweaved together upon similar social preferences, requirements or goals.

3.2.3. The View

A view is a perspective of the graphs for a specific user. It is believed to be a tailored graph that is suitable for a user's individual sight. Since it is impossible to present the whole graphs over the Internet to users, a view is constructed for a particular user on the assumption that an individual always navigates a limited number of clusters and the neighbors of the user are also limited within the clusters for a relatively long period. The approach eliminates the scale of the graphs to be presented to a user obviously.

Even though the graphs are customized, it is possible that the total amount of nodes and connections in the cluster is too large to be presented within a person's outlook. Therefore, a series of management approaches, such as ranking, are required to filter the views further. Fortunately, the refined algorithms are easier to be found once if a view is found. Even some trivial ones, such as the count of clicks, are good enough indicators to select better nodes. Moreover, the customized cluster must achieve the goal that users can tolerate a certain degree of low quality data. Since views are defined restrictedly upon users' interactions, it is a feasible objective. And, users' tolerances also lower the burden of the management.

In short, the concept of views is a solid foundation to present users with high quality data, channels and humans within the ocean of the Internet.

3.2.4. The Local View

The local view is identical to the current one presented to a user. As a cluster represents certain significance, it is reasonable to present the unique view primarily for users at a specific moment. A local view assumes that a user usually accesses information nearby rather than around a distant area according to the small world phenomenon [56]. This concept must be related to the thinking approaches of human brains [51].

A user always enters the Internet following her own preferences. If the global information is presented properly, a user can make a decision to navigate conveniently. It is infeasible to bootstrap and lead users with any algorithms that intend to learn users' interests. When considering the Internet as a number of connected views for an individual, it is convenient to support most of the user's data accessing activities with high performance and low cost over the Internet.

3.3. The Solutions

Once if connected local views are determined for each user, it is easy to propose the solutions to publishing, multicasting, persisting and presenting. Over the current Internet world, each user

controls one real or virtual computing device, which ranges from a single mobile phone to a large-scale cluster. If a device is powerful, it dedicates its computing resources, such as CPU cycles, main memory, mass storage and bandwidth to applications. Otherwise, it contributes while consuming or it consumes only. The degrees of the contributions depend on their respective social roles over the Internet. In essence, it would believe users contribute their social capital [14] other than computing resources to the Internet.

3.3.1. Publishing

When publishing, the potential readers must be the humans within the local view where an author is located in. It is also reasonable to include additional readers that are nearby from the local view. Therefore, the problem to find potential readers is resolved and an author is able to prepare data to post.

3.3.2. Multicasting

To lower the burden and raise the performance of publishing data, it is convenient to form a socialized topology and transmit data between the author and its potential readers. No matter whether the scale of the readers is large or not and whether the volume of data is heavy or light, the solution to the topic becomes straightforward if the way to find readers is available.
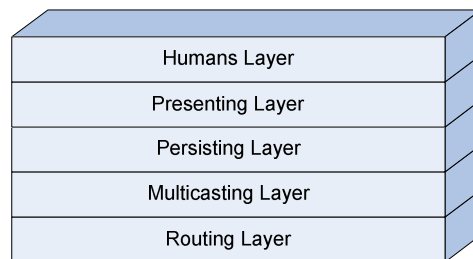
Humans Layer
Presenting Layer
Persisting Layer
Multicasting Layer
Routing Layer

Figure 2. The Five Layers

3.3.3. Persisting

The connected local view is also regarded as the virtual location to persist data. Or in anther word, while multicasting data to readers, the persisting is also being performed since the tough questions [57] are answered during the procedure. Those questions include how many storage resources are required, how to deploy them physically and what the exact number of replicas is.

3.3.4. Presenting

It is reasonable to present the three types of graphs within the connected local views to users. The solution to the topic is simplified to an easier problem how to sift the graphs of data, channels and humans within a specific connected local view of a user rather than forcing millions of machines to understand interests of users and sense of data.

4. THE INFRASTRUCTURE OF THE INTERNET ON THE COMPUTING LEVEL

As the connected local view is the requirements to the Internet on the computing level, it is indispensable to establish the infrastructure to constitute the views for users. As a foundation, the system consists of five layers, the routing, the multicasting, the persisting, the presenting and the humans, which are shown in Figure 2.

The term, the computing level, is used to differentiate the work of this paper from the one that supports fundamental communications between any two nodes over the Internet. As the protocol of TCP/IP [11] is the primary approach to connect nodes and exchange data between them, it defines that all of the tasks beyond the TCP/IP belong to the domain of the Internet on the computing level whereas its counterpart, the Internet on the communication level, solves the data transmission between two nodes only. The levels are illustrated in Figure 3.

4.1. The Differences

Different from traditional layer-based architectures that work in the way of a lower layer supporting its immediately upper layer only from the bottom to the top, the infrastructure of the Internet on the computing level is an interactive layered system. In the architecture, similar to traditional ones, a lower layer supports its immediately upper layer. In addition, the routing layer is also visible to the humans one such that it has to adjust for crucial updates of the topmost layer. Within the fundamental infrastructure, it indicates that all of the lower ones are under the domination of the topmost, the humans. The influences of the topmost one are transferred directly or indirectly to the underlying ones. For the distinctness of the system, it claims that the Internet should be involved with social resources other than computing resources. That is, such a system relies mainly on humans rather than computers. In another word, the social capital from the humans layer is the guarantee of the quality of the Internet applications.
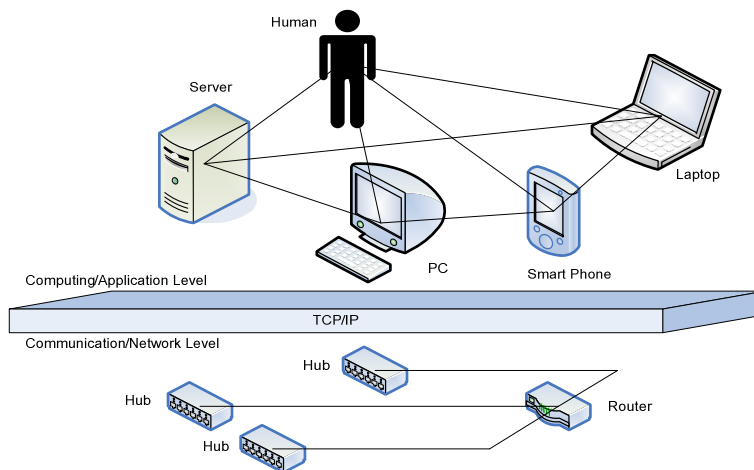

Figure 3. The Computing Level and the Communication Level

NIICL

Different from classic layer-based software architectures that work only in the way of a lower layer supporting its immediately upper layer from the bottom to the top, the infrastructure of the Internet on the computing level is an interactive layered system. In the architecture, similar to traditional ones, a lower layer supports its immediately upper one. In addition, the routing layer is also visible to the humans one such that it has to adjust for crucial updates of the topmost layer. Within the fundamental infrastructure, all of the lower ones are under the domination of the topmost, the humans. The influences of the topmost one are transferred directly or indirectly to the underlying ones. For the distinctness of the system, it claims that the Internet should be involved with social resources other than computing resources. That is, such a system relies mainly on humans rather than computers. In another word, the social capital from the humans layer guarantees the quality of the Internet applications.

4.2. The Humans Layer

The layer of the humans is made up with all of the users of the Internet on the computing level. According to the scales, the users are divided into individuals and organizations. According to the roles, the users consist of publishers and readers. All of them behave online only in accordance with their inherences, preferences, experiences, knowledge and common senses, i.e., human capital [14][15][36], which is accumulated during a long term.

It is known that the ways to capture those human capital are unavailable yet. For that, it is impossible to know a user's tastes and intents with their actions directly. As a system to share huge information, it becomes tough to publish and read since it is difficult to find relevant targets, the readers and the data. Then, it represents it is impractical to collaborate each other and perform the Internet computing efficiently.

Fortunately, rather than human capital, social capital [14][36][15] or social resources are convenient to be caught. No matter whether to publish or read, users interact with others. It is feasible to retain the interactions by the concept of connections or neighbors. With the growth of the system, a huge graph is constructed.

In short, human capital from the layer dominates the Internet through users' interactions to access data. Although they are tough to be identified, it is convenient to map human capital to social capital through interactions among users.

4.3. The Routing Layer

The routing layer is defined as the mechanism to discern and find sufficient resources for any other layers to support their online activities in a high performance and low cost manner over the Internet on the computing level. When finding resources, it is equivalent to exploring routes between consumers and providers. That is why the layer is named the routing. For a large-scale system like the Internet, it is expected that the routes be as short as possible to minimize network latencies.

One of the primary tasks of the layer is to discern social resources from the topmost, the humans layer. To do that, the concept of a connection is defined as the interaction between a publisher and a reader. As all of the users of the Internet perform publishing and reading following their preferences, it is time to retain those connections and construct a graph upon those interactions. It indicates that the topmost layer is associated with the bottom one directly, which results in the three graphs distinguished by their node types, data, channels and humans, respectively.

After the graphs are available, it is time to maintain, rank and find relevant resources upon them whenever publishing and reading are performed. The primary principle to do that is to identify those nearby nodes of a publisher or a reader since it is believed that they are the potential resources to fulfill the requirements of the publisher or the reader. That is, they are willing to accept data from the publisher or they hold data that the reader needs.

4.4. The Multicasting Layer

The multicasting layer is responsible for transmitting data to potential computing devices controlled by humans, which are found by the routing layer before the transmission is started and during the procedure of the transmission.

With respect to the characteristics of social networks [27][29][58], it is known that a large portion of data is published by an extremely small number of users, i.e., super nodes [27]. Most users play the role of readers as ordinary nodes [27], which seldom publish or win few attentions. Therefore, it brings heavy load to publishers. To lower the burden, it is required to seek additional resources, such as network bandwidth, to assist the transmission. The seeking is performed initially before the transmission is started. Once if those resources are available, an appropriate topology is constructed between a publisher and its routed resources to balance the load and raise the performance.

In addition, it needs to emphasize that it is also critical to perform the routing on the fly since the potential resources are not reliable. Different from traditional multicasting, all of the devices participating in the transmission belong to proprietary users such that it is possible for them to dislike the data being transmitted and consumed. Then, they leave the topology casually during the multicasting procedure. To lower the affects of the departure, it is required to bind those devices with a fault-tolerant graph rather than a tree. It is also necessary to keep seeking high quality resources by interacting with the underlying routing layer during the entire procedure. Once if higher quality resources are obtained, they can be used to replace lower quality ones immediately.

4.5. The Persisting Layer

The layer of persisting aims to save the huge amount of data into mass storage devices. It needs to indicate that it does not consider the issue to persist private data, which is restrictedly visible to a particular scope of people or the owner only. Only the data public to all of users is the task for the layer to persist.

The first problem to answer is whether storage resources are dedicated or not. For traditional systems, all of resources are dedicated since data belongs to a particular organization only. However, as a storage system for the Internet, it is impossible to do that in the same way because of the large scale. A reasonable system is made up with both dedicated and undedicated storage resources. While dedicated storage space comes from a service provider, it is critical to find undedicated resources from proprietary users. The willingness of those users depends on two cases. First, original publishers would rather share their storage resources. In addition, replicas are spread by secondary publishers, i.e., as forwarders. They are also potential users to contribute their resources.

The next issue is where to persist. As a global system, the locations of data are critical to the performance. Since consumers of public data reside each corner of the globe, it is tough to answer the question. Fortunately, it is unnecessary to care about the issue because of the remarkable routing layer. Since readers might become potential resource providers such that their physical locations are identical to where the data should be persisted. For the nature of the routing layer, new users gain high performance from the previously persistent data. In the solution, it is interesting that physical locations are hidden by social positions [Borgatti et al. 1994], i.e., the social capital, to a great extent.

The final problem is how many replicas to persist. The number of data replicas is another critical factor to determine the performance for a distributed storage system. It is also difficult to estimate the number for the Internet-based storage system with traditional approaches. To resolve the problem, data is replicated and persisted to the owners who are willing to share with the assistance of the routing layer. During the procedure, the number of replicas is controlled by its original publisher and secondary publishers. The count of replicas is raised with the popularity of

data, and vice versa. In essence, the quality of data determines its quantity of replicas. Thus, it is unnecessary to take into account the issue either when persisting.

4.6. The Presenting Layer

The presenting layer is responsible for interacting with users and showing connected local views to corresponding users. The layer aims not only to fulfill users' requirements directly but also to achieve the goal to cohere users with the system. Through the interactions, users become the dominators to guide the underlying layers to work, adjust and optimize. It is the existence of the layer that converts human and social capital to a special type of computing resources that govern the entire system. Its goal is different from traditional presentations which attempt to separate users from computing systems and leave all of tasks to machines.

While the infrastructure is governed by the underlying graphs of data, channels and humans, they are updated on the fly following users' behaviors through the layer. Those behaviors consist of two types, i.e., the writing and the reading. The writing one includes publishing, forwarding and commenting and the reading one is divided into browsing, navigating and searching. All of those behaviors result in creating or updating connections for data, channels and humans. Thus, they are collected by the presenting layer and transferred to the underlying ones.

5. THE COLLABORATION OF THE LAYERS

The five layers form the infrastructure of the Internet on the computing level. They works in the same way as traditional ones that the lower layer supporting its immediately upper one from the bottom to the top. In addition, they also interact with one another under the principle that social capital [15] determines the effects and the performance of the underlying layers. The collaboration among the layer is illustrated in Figure 4.
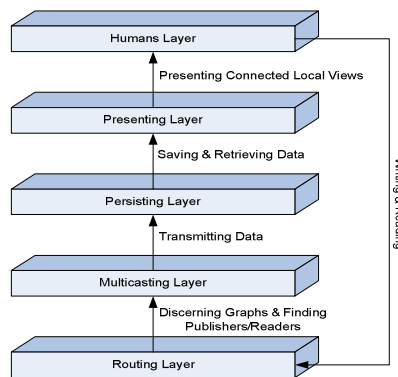

Figure 4. The Collaboration of the Five Layers

5.1. The Supporting From the Bottom to the Top

The routing layer supports its immediately upper layer, the multicasting, by providing sufficient numbers of potential readers and publishers. With its support, a huge amount of data can be published to a large number of potential readers easily. Moreover, a reader obtains accurate and comprehensive responses from relevant publishers and other readers who are willing to work collaboratively to send data to the reader.

The multicasting layer plays the foundation role to keep data transmission in a high quality way for the persisting layer. Only if data is disseminated efficiently enough, the services of persisting

data, such as saving, copying, moving, retrieving and removing, are performed efficiently in a large scale and heterogeneous environment.

The persisting layer supports the presenting layer by keeping published data in a reasonable manner in terms of proper locations and rational replicas. As a global level system, that is the bedrock to present data to users to access from all of the corners of the world in a high performance and low cost way. Although they are located in different places, users are hidden from physical positions.

The presenting layer shows connected local views to users of the humans layer. It covers the underlying mechanism such that users gain a comfortable data writing and reading environment in the sense to match users' personal preferences and social cognition [51].

5.2. The Interactions Among the Layers

The amazing character of the infrastructure is that those layers interact with each other instead of following the discipline of traditional layer-based architectures to support immediately upper layers only. The primary cause for that is the core resources of the Internet on the computing level are human and social resources rather than computing resources [14][15]. The graphs has to be updated, adjusted and ameliorated according to behaviors from the humans layer in order to provide users with better services. In brief, the layer of humans is the dominant component of the infrastructure regardless of being users split from machines in traditional ways.

When users enter the system initially, they become members of the humans layer. With the support of the presenting layer, it is convenient to observe and locate relevant connected views they are interested in. The local views consist of three types of connected nodes, i.e., data, humans and channels. Whatever they intend to do, writing or reading, their behaviors are supported efficiently and effectively through those views. More important, it leads to updates of the connections surrounding them. It indicates that the updates of the potential consequences from the routing layer. Thereafter, the relevant multicasting and the persisting are improved for both of them depending on the routing. It is no doubt that it also results in updated connected local views for users.

6. THE CURRENT AND FUTURE WORK

The research has been carried out for more than nine years since November 2005. A couple of demonstrations are available. For the rise of the mobile devices, a more novel system is being implemented over Android [60][61]. The system attempts to convert the existing WWW to the new infrastructure. The scale of existing WWW is huge such that it is impossible to do that with resources from a single organization. Fortunately, since the Internet on the computing level is inherently made up with undedicated proprietary resources, the current system is regarded as a seed that can be fostered to be mature. In particular, the resources to be invested are extremely less than those with traditional solutions.

During the procedure of the research, it is found that the critical problems over the Internet on the computing level are associated with multiple domains, such as social cognition [51], social network analysis [36], economy [53][54], statistical physics [62], library & information science [34], social computing [63] and statistics [64], etc. It is required to improve the current work while the primary disciplines and the overall architecture must be important progresses to the current Internet.

As a huge system to share information, it is mandatory to place data in an appropriate form for publishers and readers. From my point of view, all of the traditional approaches [34] do not fit in the complicated situation of the Internet. Unfortunately, since my system has not been deployed,

no reliable data is available to verify my solutions. However, during the past years of the Internet development, the trend is apparent. That is, traditional Web sites [2][23][50], directory sites [33] and even search engines [3] are not as significant as social networks oriented systems [27][28][29][30][31][32][48][49] even in the domain of data sharing [18][19][20][57][58]. A large number of hyperlinks are replicated through those socialized systems rather than traditional ways [65]. Through those phenomena, it is known that information structures have innate ties with social capital. It is required to be clarified in the future research.

Even though the relationship between the structure of large amount of heterogeneous data and the socialized systems is elucidated, it is still required to answer whether the relationship is strong enough to fulfill users' data accessing requirements. What are the approaches to evaluate the satisfaction degree? If the consequence of the evaluation is high, which types of users are satisfied and which ones are not? Or, which categories of data are maintained better or worse? If the consequence of the evaluation is low, any improvements can be performed? Since it seems no more efficient approaches are available in the current Internet, what are the natures the low evaluation tells us to share information over the Internet?

Finally, all of the above solutions to share information have to be bound with social networks. Is it the unique? They all have an assumption that the social structure maps to the inherence of knowledge or information to some extent. To the best of my knowledge, there are no research results to discuss about the structure when a huge amount of heterogeneous information is accumulated like the Internet. The information or library science [34] touches some topics. However, it is required to have sufficient staffs to manage books. That is another story since it is impossible to hire people to manage the Internet. In addition, even though the structure is known, is it feasible to implement it over the Internet? What are the best approaches for users to access it if it really exists? If the solution is negative, what is the conclusion?

REFERENCES

[1] Albert-laszlo Barabasi. 2014. Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. Publisher: Basic Books, ISBN-10: 0465085733, ISBN-13: 978-0465085736

[2] Ray Tomlinson. 1971. Email Home. http://openmap.bbn.com/~tomlinso/ray/home.html

[3] Google. 1998. http://en.wikipedia.org/wiki/Google

[4] Excite. 1995. http://en.wikipedia.org/wiki/Excite

[5] MG Siegler. 2010. When Google Wanted To Sell To Excite For Under $1 Million – And They Passed. http://techcrunch.com/2010/09/29/google-excite/

[6] HTTP. 1996. HTTP – Hypertext Transfer Protocol Overview. http://www.w3.org/Protocols/

[7] Sandvine. 2014. Global Internet Phenomena Report. https://www.sandvine.com/trends/global-internet-phenomena/

[8] Xuxian Jiang. 2003. GnuStream: a P2P Media Streaming System Prototype. Multimedia and Expo, 2003, Proceedings 2003 International Conference on Volume 2, Pages: 325-328

[9] Suman Banerjee, et al. 2002. Scalable Application Layer Multicast. SIGCOMM '02 Proceedings of the 2002 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, Pages: 205-217

[10] Miguel Castro and Peter Druschel, et al. 2002. SCRIBE: A Large-Scale and Decentralized Application-Level Multicast Infrastructure. IEEE Journal on Selected Areas in Communications, 2002, Volume: 20, Issue 8, Pages: 1489-1499

[11] Kevin R. Fall and W. Richard Stevens. 2011. TCP/IP Illustrated Volume 1, the Protocols Second Edition. Addison-Wesley Professional Computing Series, 2011

[12] George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair. 2011. Name Services, Chapter 13, Distributed Systems: Concepts and Design, the Fifth Edition. Addison-Wesley, 2011

[13] URL. 1994. Uniform Resource Locators (URL), http://www.w3.org/Addressing/URL/url-spec.txt

[14] Yilei Shao. 2007. Exploring Social Networks in Computing Systems. PhD Dissertation, Princeton University, 2007

[15] James S. Coleman. 1988. Social Capital in the Creation of Human Capital. American Journal of Sociology, 1988, Volume 94 Supplement, Pages: 95-120

[16] Dongyan Xu. 2002. On Peer-to-Peer Media Streaming. Distributed Computing Systems, 2002, Proceedings of 22nd International Conference, Pages: 363-371

[17] Aggeglos Vlavianos, Marios Iliofotou and Michalis Faloutsos. 2006. BiToS: Enhancing BitTorrent for Supporting Streaming Applications. Proceedings of 25th IEEE International Conference on Computer Communications, 2006, Pages: 1-6

[18] Chris Dana, Danjue Li, David Harrison and Chen-Nee Chuah. 2005. BASS: BitTorrent Assisted Streaming System for Video-on-Demand. 2005 IEEE 7th Workshop on Mulitmedia Signal Processing, Pages: 1-4

[19] Bram Cohen. 2003. Incentives Build Robustness In BitTorrent. 1st Workshop on the Economics of Peer-2-Peer System, 2003

[20] eMule. 2002. http://www.emule-project.net

[21] George Coulouris, Jean Dollimore, Tim Kindberg and Gordon Blair. 2011. Distributed File Systems, Chapter 12, Distributed Systems: Concepts and Design, the Fifth Edition. Addison-Wesley, 2011

[22] John Kubiatowicz, et al. 2000. OceanStore: An Architecture for Global-Scale Persistent Storage. ACM SIGPLAN Notices, Volume 35 Issue 11, Nov. 2000, Pages: 190-201

[23] W3C. 2000. World Wide Web Consortium. http://www.w3.org

[24] Hubert Zimmermann. 1980. OSI Reference Model – The ISO Model of Architecture for Open Systems Interconnection. IEEE Transaction on Communications 28 (4), 1980, Pages: 425-432

[25] Frank Buschmann, Regine Meunier, Hans Rohnert, Peter Sommerlad and Michael Stal. 1996. Pattern-Oriented Software Architecture. Volume 1, A System of Patterns, Wiley, August 1996, ISBN: 978-0-471-95869-7

[26] Michael Lee Scott. 2006. Programming Language Pragmatics, Edition 2. Morgan Kaufmann, 2006, ISBN 0-12-633951-1

[27] Stefan Saroiu, Krishna P. Gummadi and Steven D. Gribble. 2003. Measuring and Analyzing the Characteristics of Napster and Gnutella Hosts. Multimedia System, 2003, Volume 9, Issue 2, Pages: 170-184

[28] Ian Clarke, Oskar Sandberg, Brandon Wiley and Theodore W. Hong. 2001. Freenet: A Distributed Anonymous Information Storage and Retrieval System. International Workshop on Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, 2001

[29] Matei Ripeanu. 2001. Peer-to-Peer Architecture Case Study: Gnutella Network. Proceedings of 1st International Conference on Peer-to-Peer Computing, 2001, Pages: 99-100

[30] Atif Nazir, Saqit Raza and Chen-nee Chuah. 2008. Unveiling Facebook: A Measurement Study of Social Network Based Applications. Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, 2008, Pages: 43-56

[31] Haewoon Kwak, Changhyun Lee, Hosung Park and Sue Moon. 2010. What is Twitter, a Social Network or a News Media? Proceedings of the 19th International Conference on World Wide Web, 2010, Pages: 591-600

[32] WeChat. 2011. http://www.wechat.com/en

[33] Yahoo!. 1994. https://www.yahoo.com

[34] Thomas Mann. 2005. The Oxford Guide to Library Research, the Third Edition. Oxford University Press, 2005, ISBN-13: 978-0195189988, ISBN-10: 0195189981

[35] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schutze. 2008. Introduction to Information Retrieval. Cambridge University Press, 2008, ISBN: 978-0-521-86571-5

[36] John Scott and Peter J. Carrington. 2011. The Sage Handbook of Social Network Analysis. Sage Publications Ltd., 2011, ISBN: 978-1-84787-395-8

[37] Mark Granovetter. 1983. The Strength of Weak Ties. American Journal of Sociology 78, 1983, Pages: 1360-1380

[38] Ronald S. Burt. 1992. Structural Holes: the Social Structure of Competition. Harvard University Press, 1992, ISBN: 0-674-84371-1

[39] J. A. Pouwelse, et al. 2008. Tribler: A Social-Based Peer-to-Peer System. Journal of Concurrency and Computation: Practice & Experience – Recent Advances in Peer-to-Peer Systems and Security (P2P 2006), Volume 20, Issue 2, February 2008, Pages: 127-138

[40] Qusay Hassan. 2011. Demystifying Cloud Computing. Journal of Defense Software Engineering, 2011 (January/February), Pages: 16-21

[41] Apache Hadoop. 2008. http://hadoop.apache.org

[42] OpenStack. 2010. Open Source Cloud Computing Software. http://www.openstack.org

[43] Roxanne E. Burkey and Charles V. Breakfield. 2000. Designing a Total Data Solution: Technology, Implementation and Deployment. Auerbach Best Practices, CRC Press, 2000, ISBN: 0-8493-0893-3

[44] Harald Weinreich, Hartmut Obendorf and Winfried Lamersdorf. 2001. The Look of the Link – Concepts for the User Interface of Extended Hyperlinks. Proceedings of the 12th ACM Conference on Hypertext and Hypermedia, 2001, Pages: 19-28

[45] Tim-Berners Lee. 1989. http://www.w3.org/People/Berners-Lee/

[46] Albert-Laszlo Barabasi and Eric Bonabeau. 2003. Scale-Free Networks. Scientific American, May 2003, Pages: 50-59

[47] Jon Kleinberg. 1999. Authoritative Sources in a Hyperlinked Environment. Journal of the ACM Volume 46 Issue 5, September 1999, Pages: 604-632

[48] Facebook. 2004. https://www.facebook.com

[49] Twitter. 2006. https://twitter.com

[50] The BBS organization. 1999. http://thebbs.org

[51] Susan T. Fiske. 2013. Social Cognition, Sage Library in Social Psychology. Sage Publications, 2013, ISBN: 978-1446254738

[52] Mark L. Knapp and John A. Daly. 2011. The SAGE Handbook of Interpersonal Communication, the Fourth Edition. Sage Publications, 2011, ISBN: 978-1412974745

[53] Walter Nicholson and Christopher Snyder. 2011. Microeconomic Theory: Basic Principles and Extensions, the Eleventh Edition. South-Western Cengage Learning, 2011, ISBN-10: 1111525536, ISBN-13: 978-1111525538

[54] Richard T. Froyen. 2012. Macroeconomics: Theories and Policies, the Tenth Edition. Prentice Hall, 2012, ISBN-10: 013283152X, ISBN-13: 978-0132831529

[55] Ronald S. Burt. 2004. Structural Holes and Good Ideas. American Journal of Sociology, Volume 110, Number 2, September 2004, Pages: 349-399

[56] Jon M. Kleinberg. 2000. Navigation In a Small World. Nature 406:845, 24 August 2000

[57] Ragib Hasan, et al. 2005. A Survey of Peer-to-Peer Storage Techniques for Distributed File Systems. ITCC '05 Proceedings of the International Conference on Information Technology: Coding and Computing, Volume 02, 2005, Pages: 205-213

[58] Matei Ripeanu, Adriana Iamnitchi and Ian Foster. 2002. Mapping the Gnutella Network. Journal of IEEE Internet Computing, Volume 6 Issue 1, 2002, Pages: 50-57

[59] Stephen P. Borgatti, Ajay Mehra, Daniel J. Brass and Gluseppe Labianca. 1994. Network Analysis in the Social Sciences. Science, 1994, 323(5916): 892-895

[60] Android. 2007. https://www.android.com

[61] Bing Li. Blog, 2015. http://greatfree.lofter.com

[62] Boris M. Smirnov. 2010. Principles of Statistical Physics. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2010, ISBN-13: 978-3-527-40613-5, ISBN-10: 3-527-40613-1

[63] Nitin Agarwal. 2009. Social Computing in Blogosphere. PhD Dissertation, Arizona State University, August 2009

[64] Morris H. DeGroot and Mark J. Schervish. 2011. Probability and Statistics, 4th Edition. Pearson, 2011, ISBN-10: 0321500466, ISBN: 978-0321500465

[65] Josh Constine. 2013. Why Facebook Needs Trending Links.
http:// http://techcrunch.com/2013/06/30/facebook-trends/

ACKNOWLEDGEMENTS

AUTHORS

Bing Li is a faculty of Xi'An Technological University. I have been working here for four years since 2013. Before that, I worked as an associate professor for Peking University, the most famous university in China between 2005 and 2013. I got my PhD degree in the area of software engineering from Arizona State University in 2004. Moreover, I have some industrial working experiences in Lucent and IBM. He focused on the research in the area of the system-level software for Internet. He proposed the upgraded WWW and tried to program the entire system until now. Because of his rich experiences of programming on the system-level of the Internet, he summarizes a series of APIs and design patterns for large-scale distributed system programming, which is called the GreatFree Cloud Programming Environment. His personal blog is located at http://greatfree.lofter.com.