VizResearch: Linking the Knowledge of People and the People with Knowledge

Sarah Masud\textsuperscript{a}, Monjura Afrin\textsuperscript{a}, Farhana Murtaza Choudhury\textsuperscript{a}, Syed Ishtiaque Ahmed\textsuperscript{b}

\textsuperscript{a}Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh
\textsuperscript{b}Department of Information Science, Cornell University, Ithaca, NY, Zip: 14850, USA.

Abstract

Access to information and interaction with peers are dominant demands of collaborative research. To provide academicians information regarding research papers, peers and events there are some research information repositories overlaping the need of interaction and collaboration. Although the booming growth of social networking sites has already captured huge audience from academic field, research shows existent Social Networking Services (SNSs) are mostly hedonic thus having a superficial contribution to research based interaction. To address this unmet demand, some systems have tailored the concept of SNS to fit research oriented social networking. Though they provide an interactive platform for researchers they lack in information visualization, binding location and time with research and evaluation of research work and related bodies. Hence our solution VizResearch unifies fragmented functionalities across existing platforms and brings all research beneficent features under a single hood.

Keywords: Scientific Collaboration, Social Network, Online Community, Location Based Service.

1. Introduction

Social networking reflects the social relations among people who share common interests, activities, beliefs or knowledge. It operates on many levels removing the boundary of time, space and language. SNS reveals the links between real world people. Whether it is making new friends, finding a new job, promoting business, or locating a destination these links between people come useful. By providing people a sense of ubiquitous connectedness and a platform to share multiple point of views social networking sites are getting popular rapidly.

As the harbinger of social-networking hype, facebook \cite{1} now has more than 800 million active users \cite{1}, and growing at a rate of around 83 percent per year on average over the last 5 years \cite{1}. Among other SNSs twitter \cite{2}, hi5 \cite{3}, myspace \cite{4}, orkut \cite{5} and rapidly emerging Google+ \cite{6} contribute to connect people from anywhere anytime.

Scientific research demands such connectedness among researchers to ensure collaboration and interaction. Research collaboration is any research activity that is carried out by multiple individuals overcoming institutional, disciplinary, and geographic boundaries. By collaboration, researchers can combine their strengths thereby producing more reliable and powerful results. It can enlarge the creation of innovation and save resources too.
There are situations when a researcher searches for publications and references, when a freshman looks for research funding and a supervisor to work with, or when an academian wants up-to-date information of any particular research field, or looks for upcoming events, i.e., seminar, conferences on a specific topic. Only dealing with factual information makes the task difficult and time consuming. Rather the urge is to connect researchers and their work across time and space only using a mouse-click. Also each research field gets deeper every day with increased amount of research works and the number of people involved. Given this inherently social and dynamical nature of research, an interactive network platform is urgently needed to fuel cooperation and information exchange globally.

As social networking services allows identity and network management, information management, and communication with peers [7] researchers can harness it to boost up research collaboration. Though existing SNSs can support interaction via message posting, chatting, file sharing, limited group based activity among researchers, they can only superficially benefitted from these virtual communities. Currently these SNSs do not support managing academic profile with updated publication list, co-author list, visualization of important research relationships, sharing publications or specialized search for publications, peers or events. Most importantly, these hedonic social networking sites can distract a researcher in many ways i.e., streams of non-academic news feeds, message, status updates as everyone in the network is not opting for research [8].

Even if we consider only research based interaction in current SNSs dropping information sharing and management, still it is not convincing. Being facebook friend for work discussion has some privacy issues [8]. It shows that lack of threaded discussion hampers the growth of appropriate interactive environment [9].

While some researchers have explored use of SNS in education, very few have delved into the methodology of how it should be integrated in research. Current studies include drawing out essential requirements for social networking in academic environments comprising of students, teachers, administrators and researchers [10], investigating present status opportunities and challenges of educational applications in SNS [11], integrating SNS in a formal learning environment [12]. Another research explores intersection of online social networking with medical professionalism [7]. A comparison of faculty and student responses indicates that students are more likely than faculty to use facebook and are significantly more open to the possibility of using facebook and similar technologies to support classroom work.

In another study while students have found class social networking more user-friendly and comfortable giving them a sense of belonging [13], using existing platforms i.e., facebook often leads to some drawbacks. Concerns related to privacy and anxiety in interacting with professors in this environment [14], a belief that it does not serve an academic purpose [15] and the opinion that faculty should simply avoid educationally appropriating these backstage social spaces [16] have been expressed [17]. Another work shows that where students use facebook for formal or informal learning purpose involving student-to-student interaction they are less likely to use it for teaching purpose involving instructor-student interaction [18].

Current research indicates while people welcome integrating concept of SNS with educational or academic purpose [17], [9], [18] using existing SNS is not indisputable. In this regard we propose niche social networking which is tailored to meet specific demands of research. Such concept is not unprecedented as blackboard, moodle, sakai, teachstret.com are educational SNSs, Academia.edu [19], mendely [20] and ResearchGate [21] are research oriented SNSs. Our uniqueness lies in identifying the complete set of needs encompassing interaction, collaboration and information management and binding them into a single platform.

2. Related Works

As stated earlier, currently there are some platforms to fuel research activities. Some of them provide researchers with various research-oriented data regarding publication, events, or researchers while others emphasize on interaction and collaboration of researchers. Hence they can be categorized into two types:

- Research Information Management Tools
- Social Networks for Research Communities

2.1. Research Information Management Tools

Underscoring the significance of information management, currently, Google Scholar [22], CiteSeerX [23], Pubzone [24], ArXiv [25] along with professional bodies like IEEE [26], ACM [27] satisfy academicians demand on research information substantially. They provide information for publications based on keywords, authors, conferences,
and journals. Microsoft Research [28], a research division of Microsoft, works for developing various computer science ideas and integrating them into Microsoft products [29]. This site shares publications, project updates, videos and information on researchers. But it focalizes only the domain of Computer Science and deprives a huge majority of academicians from other faculties.

Microsoft Academic Search [30], a free academic search engine of Microsoft Research [28], covers more than 27 million publications and over 16 million authors across a variety of domains with weekly updates [29]. It allows users to innovatively explore and visualize academic papers, authors and events using co-author graph, citation graph, call for paper CFP calendar and domain trend. They are the first to merge the concept of location with research. Here users can visualize the geographical location of researchers and events on map. ArnetMiner [31] is another impressive tool like Microsoft Academic Search [30] for visualization of research related information. It comes up with comprehensive search and mining services [31].

Although the above mentioned platforms meet up the demand of information impressively, there is no support for collaboration and interaction between researchers.

2.2. Social Networks for Research Communities

With the rapid growth of web 2.0 technologies, academicians along with the rest of the online population increasingly use niche SNSs like, LinkedIn [32], Academia.edu [19], ResearchGate [21], mendeley [20], Researchr.org [33], Epernicus [34], ResearcherID [35], SciSpace [36], Dropbox [37]. Here we are focalizing on the major sites.

LinkedIn [32], officially launching at 2003 and connecting more than 120 million people by August 2011, is the fastest growing professional networking site [32]. It allows members to create business contacts, search for jobs, and find potential clients. Individuals have the ability to create their own professional profile. Ultimately, this is business-related social networking and has very little to contribute to research.

More than half a million [19] academician users update and share information about their work and interests in Academia.edu [19] making it probably the largest social network for academicians. Features like friend finding, status update, maintaining online academic profile, searching jobs, news feedback contributed to its buzzing demand. Though people can interact through following and exchanging personal messages, there is a lack of support for group-based interactions.

ResearchGate [21], currently involving over 1.2 million users, came up with the idea of crowd-sourced research [21]. Unlike Academia.edu [19] it supports workgroup besides small scale personal interaction. Presenting information on people, publications, events and jobs with only a single search query fuels to its popularity. Its powerful semantic search capacity scouring internal resources and some external research databases has outdone that of Academia.edu [19].

Mendeley Desktop is a cross platform tool for organizing research paper collection and citations. It automatically extracts references from documents, generates bibliographies from your PDFs and turns them into a searchable full-text database. It also syncs continuously with other reference managers like Zotero or CiteULike. Mendeley Web lets you access your research paper library from anywhere, share documents in closed groups, and collaborate on research projects online [10].

Although Academia.edu [19], ResearchGate [21], and Mendeley provides useful news feed of papers, conferences, and job news they lack some effectual features. Namely, they neglect to bind location as a parameter for finding people or events, which academicians will find very helpful. None of them provide any visualization of various important research relationship, like, adviser-advisee hierarchy, co-author chain and citation chains. Users can not filter the news feed they provide which results into a constantly clogged feed with information about people and paper one is not interested in. There is also an opportunity to improve these platforms by utilizing user-feedback and number of citations of publications to develop a ranking of research literatures. This ranking can be further used to formulate ranking of authors and their corresponding institutions. Thus a dynamic evaluation of research work can be performed. Based on the over-all activities in various research fields i.e., number of interested researchers, publications, events a research trend curve can be also generated. The above discussion can be summarized in Table 1.

Presently, there is no single system to meet up the demand of information and demand of collaboration simultaneously. One need to explore too many sites which is time consuming. More importantly, traversing sparse platforms along with ones group and peers can baffle even the smartest one. Hence our solution VizResearch is to bring all vital research facilities under the same hood, namely, information sharing, information visualization, interaction and dynamic evaluation of research.
3. Our Work

The main four components of VizResearch includes information sharing, visualization, user interaction and dynamic evaluation of research. Here one can maintain his academic profile while viewing those of others. Researchers can interact via personal message and follow option. Like-minded people can form groups and share files, create events, and run threaded discussion. One can search for all research related essential information here, regarding publication, events, and peers. These searches are fortified using time and location as parameters. After accessing a publication one can post review or rate it. A ranking for papers is generated using number of citations and user rating. This ranking is further extended to rank corresponding authors and their institutions. Moreover, based on the overall participation of researchers from various field a research trend curve is generated. The four components are described here in detail.

3.1. Sharing Information

It involves mainly two types of information, regarding researchers and publications as shown in Figure 1. While signing-up for the site one needs to provide basic research related information like, academic background, research interest, supervisor and so on. Based on the provided information and our existing database, a supervisor tree is constructed representing the adviser-advisee hierarchy. In Figure 2, a typical supervisor hierarchy has been presented. The person on the top is the supervisor of the person just below him. Similarly, the person who is second from the top is actually supervisor of the person just below him. As the supervisor of the person on top has not enrolled in the system, his information is not available. In this way VizResearch offers us a visualization of the academic genealogy. Authors can share their publications partially or fully maintaining copyright issue. Upon sharing a paper, the publication list and co-author list are automatically updated in corresponding profiles of authors. Thus as in Figure 3, ones publication list, co-author list, research experience -everything appears in his profile, making his complete academic profile reachable within a mouse click.
While sharing a paper citation data is also extracted from paper. In the detail view of a paper one can find the forward links i.e., the papers that cited this paper as well as backward links i.e., the papers this paper cited. As a result publications of similar topic are much easier to get. Besides, as shown in Figure 4 the detail view of a paper aggregates information like, authors, bibliography, date of publication, publishers, keywords, venue of publication etc.

3.2. Information Visualization

Here users can search in following criteria:

- Paper: keyword(s) can be topic, publication year, publisher, author.
Figure 3: Profile of a user containing co-author list, paper list and research experience

Figure 4: Detail view of a paper containing forward and backward link

- People: keyword(s) can be research interest, institute, location.
- Events: keyword(s) can be topic and/or location.
• Group search.

To increase efficacy of searching, search results of researchers and events are visualized in Google Maps. Upon searching for an event one can find all the details of that event, i.e., organizers, venue, submission deadline, event scope, important dates etc. One can also join such an event as this virtual participation notifies him about that event via SMS and e-mail before due date. These features tie people, event, time and location together.

3.3. User Interaction

Users can explore profiles of others and initiate interaction by following or sending personal messages. In case of follow, one gets several updates from people he is following, such as, recent publications, recent events, groups or status updates. To enhance interaction, people of similar interest can create groups. Groups can be restricted or open for all. Thus there is a chance to share something confidential with a desired audience. Group members can share files, post messages, run threaded discussions on a topic and arrange or notify about events like, conference, seminar, workshop. The total interaction hierarchy is plotted in Figure 5.

3.4. Dynamic Evaluation of Research Work

As one can browse publications, he can rate and review them. Based on their rating and number of citations of the paper, its impact factor is calculated. While showing search result of paper, they are sorted according to this impact factor. An author is rated based on the impact factor of his publication(s). Eventually using the rating of member researcher(s) a ranking of institutions is generated.

Based on the overall research activities of users, a research domain trend is presented plotting research impact versus year. We have considered number of researchers, publications and related events to calculate research impact. In Figure 6 such a curve is presented showing trend in 4 major branches of computer science namely Algorithms, System, Software Engineering (SE), and Artificial Intelligence (AI). It shows more people are inclining to AI and Systems rather than Algorithm. It should be noted that we have used our system database to formulate such trend. This database is developed from researchers who enrolled in the system. The categorization of a subject is done by researchers when they create an account and add research details. One can add new sub-branch to an existing branch or can start a whole new branch as his research interest. This categorization is visible to all account holders. As in social networks we are more inclined towards collective intelligence, we preferred researchers themselves to create and enlarge such research interest categorization.
4. Discussion

This system is built with ASP.Net on Microsoft .Net framework 4.0 using C#. As the backend support we used Oracle 11g.

4.1. Usability

This system supports Unicode(AL32UTF8) based character set. So it can store multiple language groups. Using this feature the system can be enhanced to provide a multi-lingual interactive web service like LinkedIn [32]. Academic institutions can use this system to maintain their local repository as well as conduct interactive communication between expert and novice researchers. It can contribute to connect the world’s best scientific thinkers with researchers in isolated and remote areas overleaping institutional, disciplinary and geographical boundary. It can reveal a network of like-minded academics and keep them updated about the latest research.

4.2. Business Feasibility

Various professional and research oriented sites like LinkedIn [32], Academia.edu [19], ResearchGate [21] are not only welcomed by academicians but they also caught the attention of venture capitalists. In the later part of last year investors from the US, UK and Germany converged to offer a first round of capital to ResearchGate [21]. Investors included the co-founder of generalist SNS Bebo, Michael Birch and the co-founder of global price comparison platform Idealo.com, Martin Sinner [38]. On the other hand LinkedIn [32] is publicly held and has a diversified business model with revenues coming from hiring solutions, marketing solutions and premium subscriptions [32]. This system can follow any of the ways. To make a self sustaining platform we can maintain a job board.

4.3. User Friendliness

To make the system more cohesive and user friendly we considered page load speed, accessibility, navigation and most importantly information. It is easy to search and access information regarding paper, peers and events. Google maps provide better visualization of location based search result. It is designed to be browser compatible and excludes unnecessary flash contents and lumbering graphics. Almost all pages are only two click away from the home page. When running as a real time application, it can easily incorporate user-developed applications. Thus the dynamic demand of users can be met.
5. Future Work

Though our system covers prominent features for fueling research based activities there are few things to promote it. Currently the text miner can extract data only from .pdf files. Better data mining tools can be used for data extraction from publications in various file formats. A lighter version of the system can be developed to ensure easy access from smart phone. For more financial robustness we intend to maintain a job board.

As the idea to get all researchers in a single platform is quite ambitious, improved algorithm can be embedded to extract information from distributed web about people who are not in the community presently. To enrich the repository, this system can be a client to a subset of data provided in a research information system. The repository can be further extended by exchanging data between various institutions on mutual agreement. Location based organization, researcher and event finding facility with GPS data can be implemented. Moreover we plan to generate a recommendation system for researchers to accelerate collaborative research based on interest and location of peers. Currently we are working on these updates.

6. Conclusion

In essence VizResearch brings all vital research facilities namely, information sharing, visualization, peer interaction, and dynamic evaluation of research and researchers under a single hood. It aims to become a platform to link the knowledge of people and people with the knowledge and thereby fueling research collaboration and information exchange globally.

References

URL http://www.springerlink.com/content/b2525286784440h0/fulltext.pdf
URL http://dl.acm.org/citation.cfm?id=1381300.1381451
[22] Google Scholar, http://scholar.google.com/, a freely accessible web search engine that indexes the full text of scholarly literature across an array of publishing formats and disciplines.

[23] CiteSeerX, http://citeseerx.ist.psu.edu/index, a public search engine and digital library and repository for scientific and academic papers with a focus on computer and information science.


[28] Microsoft Research, http://research.microsoft.com/en-us/default.aspx, the research division of Microsoft for developing various computer science ideas and integrating them into Microsoft products.


[34] Epernicus, http://www.epernicus.com/, a social networking website and professional networking platform.


[37] Dropbox, https://www.dropbox.com/, a web-based file hosting service that uses cloud storage to enable users to store and share files and folders with others across the Internet using file synchronization.