

SNS Search in Developing Countries: Linking the People to End Digital Division in Information Retrieval

Hasan Shahid Ferdous, Saif Ahmed, Md. Tanvir Alam Anik, Mashrura Tasnim

Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology,

Email: hasan.shahid@csebuet.org, iconicsaif@gmail.com, tanviranik@gmail.com, mashrura_cse@yahoo.com

Abstract—Using Social Network Sites (SNS) as an information source has drawn the attention of the researchers for a while now. There has been many works that analyzed the types and topics of questions people ask in these networks and why. Topics like what motivate people to answer such queries, how to integrate the traditional search engines and SNS together are also well investigated. In this paper, we focus on a relevant but different issue - how SNS search varies in developed and developing regions of the world and why. With established statistics of Internet usage, e-Governance, and our experimental data collection, we have tried to emphasize the differences among them and provided insight that one might require to consider while developing any application for SNS based searching.

Keywords-Social network sites; search engine; query; developing countries.

I. INTRODUCTION

Being able to think and ask questions has been an integral part of the human race for supremacy, and for a long part of our history, helping one another in this quest was the only way. Then we learnt to preserve, convey, and spread our knowledge through written and printed medium. The digital revolution over the past three decades has provided us with new power to store and maintain large collection of data in a tiny amount of space. Specially the inception of search engines (SE) has enabled us to look into tremendous amount of information within seconds, a feat our ancestors could hardly imagine about. These achievements lead many of us to believe that we are at the pinnacle of information search and retrieval, but that was hardly the truth. As the past decade has seen the emergence of Social Networking Sites (SNS), many researchers are now wondering if history is repeating itself to bring us back to human intervention in information retrieval. We are somewhat back to the state when *everyone knows something* and when connected, *together knows everything*.

In this paper, we will use Facebook as an example SNS, without losing any generality. With one billion monthly active users and more than half a billion daily active users [1], currently (as of 22nd February, 2013) it is the number two site in the world considering Internet traffic, according to Alexa ranking [2]. On an average, the users spend 10.5 billion minutes per day on it, make 421 million status

message posts, 3.2 billion likes and comments, and have 140.3 friends in their network [3]. In this triumph of SNS, we are more connected with people around the world than ever before. Nowadays, it is no longer a source of entertainment and social connectivity only, it has paved a new way for information searching. Apart from using the search engines that can merely use the already available information in the public sites crawled in its memory and some algorithm to search and index the results without much personalization, we can simply ask the members of our social network and get useful information that the researchers found quick, useful, and in many cases, more robust.

In this paper, we emphasize on this phenomenon with special focus to developing regions of the world and see how SNS search has made significant changes in way people access information here. We discuss about seminal works in this area in Section II. The problem of ‘digital division’ is explained in Section III along with the concept of less biased SNS world in Section IV. Our experimental data along with methodology, interviews and findings are explained in Section V.

II. RELATED WORKS

Lampe *et. al.* [4] analyzed how the use of Facebook has changed over time, using three consecutive years of survey data and thorough interviews with a few of the survey people. They reported that though the uses of the site remain relatively constant over time, but the perceived audience for user profiles and attitudes about the site showed differences over the study period. They find that patterns of use, perception, and attitude somewhat changed over the time. Their study, consistent with others, found that the number of friends and time spent on Facebook increased at first and then leveled off, which from interviews, suggested that new users spend time adding people as friends and getting used to the site. After a while, this behavior lessens as time is spent more seeing what is happening to friends instead of expanding their friend base. Also, new users are more likely to use Facebook to “find people to date” or “meet new people” than long-term users.

One of the important studies in SNS based information search is done by Efron *et. al.* [5], who identified that micro-

blogging services like *www.twitter.com* are gradually becoming a popular venue for informal information interaction. They showed that question asking in micro-blogs is strongly tied to peoples' naturalistic interactions, which helped them to offer a taxonomy of questions in micro-blogs. They also showed that the act of asking questions in Twitter is not analogous to information seeking in more traditional information retrieval environments, which contextualize these articulations through analysis of a large body of tweets.

Teevan *et. al.* [6] discussed the types of information people used twitter to find, for example, breaking news, real-time content, popular trends, etc. This paper presented the systematic overview of search behavior on Twitter and differences with web search using questionnaire data along with analysis on query logs. They found that Twitter results included more social content and events, while web results contained more facts and navigation. Based on their study, they recommended that search engines can use trending Twitter queries to discover additional queries that have strong temporal components.

Lampe *et. al.* investigated the Facebook user characteristics based on a survey of 614 people who used it to ask something [7]. They identified the perception of the relationships within network members as significant predictors of information seeking approach. They did not show any comparison between SNS and SE regarding obtaining any particular type of information. This question is addressed by Morris *et. al.* [8], where they explored the pros and cons of using SNS as information source and compared user interaction when they search anything either on SNS or SE, involving 12 participants on their study. They find that 53% of the users received quick responses from SNS and 83% received responses eventually as well.

The type of questions and answers in SNS are investigated by Morris *et. al.* [9] using a study of 624 people about their Facebook usage experience. They also explored the relationships between answer speed and quality, properties of participants (age, gender, and social network usage habits) and their questions (type, topic, and phrasing). Their study complies with the findings of many other researchers that while traditional SE is good for objective queries, SNS shows better results and interactions for subjective queries. There are many motivations for asking questions in SNS - among them the most important reason was that people in our social network knows our *context* better and thus may provide more relevant answers. Often people turn to SNS regarding objective questions if knowing the answer is not urgent, in the hope that some other friend in his network already knows the answer and will share his knowledge with him in due time.

Panovich *et. al.* [10] evaluated the role of *tie strength* in question-response behavior as an indication of how close the relationship is - close friends are strong ties, while acquaintances are weak ties. In their study, they asked

19 participants to ask some technological recommendation questions through status messages. After the participants rated the received answers' quality, they compared that with a tie strength metric, and found that stronger tie provides better answers than weaker ties, in general. Also, they find that friends who have expertise in the question topic provide more trustworthy answer irrespective of strong or weak ties.

Farnham *et. al.* [11] studied the suitability of *So.cl*: a web application that combines web browsing, search, and social networking, designed for the purposes of sharing and learning around topics of interest by taking feedback from 32 college students. Their findings present the importance of social media for inspiring learning around new topics through social connections. They found the easy, lightweight integration of sharing around search in *So.cl* effectively fostered serendipitous, informal learning online.

Naaman *et. al.* [18] examined the 350 users message and some system data to understand the individual's activity using their own developed content based categorization. Their analysis showed two common types of user behavior in terms of the content of the posted messages, and exposed differences between users in respect to these activities. But they did not address in this work the relationship between social network structure and social influence to the type of content posted by users.

A controlled study conducted on 282 persons by Teevan *et. al.* [12] analyzed effect of the factors: punctuation of status, scoping of audience, and precision on the response time, quantity and quantity of response. Their key findings are that a higher portion of questions with a question mark received responses (88.1% v. 76.3%, $p < .01$) and two-sentence questions received fewer and slower responses. They also noted that explicitly scoped questions resulted in better response.

Hecht *et. al.* tried to combine the benefits of SE and SNS searching in their system named *SearchBuddies* [13], a system that responds to Facebook status message questions with algorithmic search results. They proposed two agents - Investigator (search on SE), that connects people with information, and Social Butterfly (Search on SNS), that connects people with other people who may have the desired information. After deploying their 'Socially Embedded Search Engine' on 122 users for three months, they believed that it provides highly relevant information in a social context.

None of these researches investigate the difference in question-answer behavior in different parts of the world. Yang *et. al.* [14] addressed this issue and identified some key differences between SNS search in the Western and Eastern cultural hemisphere. Their survey included people from US and UK representing the Western culture and people from China and India representing the Eastern culture. They concluded that people in the Eastern culture are somewhat more likely to use SNS for getting objective information

than their counterpart and use it more often for the purpose. They explained this phenomenon using existing and established knowledge from sociology study that Western cultures are associated with an analytic and low-context cognitive pattern, along with individualism, while Asian cultures are associated with a holistic, high-context cognitive pattern, along with interdependence and collectivist social orientation. Our initial findings matches with them, except they did not include another possible explanation of this behavior - the existing web infrastructure deficit in the developing and undeveloped countries, commonly known as the *Digital Divide*. In our work, we will elaborate on this explanation.

III. THE DIVIDED WORLD

The term ‘Digital Division’ indicates the difference in technological advancement between the developed and underdeveloped/undeveloped parts of the world. Computers and other computing devices are essential commodities for the people in the developed region for the past two/three decades and their web presence is ubiquitous nowadays. Recent explosion in the smart-phone usage has enabled virtually everyone to remain connected to Internet round the clock. Nearly all the governmental and business services have their information published and updated in the web. Traditional search engines in that respect are very effective in capturing the required information as it is already there in Internet.

The scenario is quite opposite in the other parts of the world where the web culture has not flourished yet. If we focus on the South-East Asia region as an example of the developing part of the world, we can see from UN survey 2010 [15] that the average e-Governance ranking of the 8 countries in this region is 134, way beyond the developed regions. According to [16], about 8–10 percent people in this region has access to Internet. Even that is after the growth of Internet users in recent years, and the overall web presence is not good yet. Many important governmental and non-governmental institutions do not have their information in the web and often do not update their information regularly, if there is any.

The problem is twofold. People in this developing region cannot find the required information from web using traditional search engines as it is beyond its capacity to show any result that is not already in the web. Again, as the Internet culture has not flourished yet, many people are not used to search information in the web, or do not know how to find the right information if there is a lot of different search results. Though the Governments in these countries are trying to eradicate this digital division, it is proved as not easy. The world remains ‘divided’ and probably will remain so for a long time from now.

Table I
INTERNET AND FACEBOOK USAGE ANALYSIS (ALL FIGURES ARE IN MILLIONS OR PERCENTAGE)

Country	Population	Internet User	%	Facebook User	% of Internet User
Australia	22.8	17.9	78.3	11.7	65.7
USA	314.8	243.8	77.4	168.6	69.2
UK	62.3	51.2	82.2	33.8	66.0
Nepal	26.6	2.7	10.3	1.9	69.2
India	1210.2	125.0	10.3	60.6	48.5
Pakistan	181.3	15.9	8.8	7.6	47.7
Srilanka	20.3	3.2	15.6	1.5	46.2
Bangladesh	152.5	7.5	4.9	3.2	42.6

IV. THE UNIFIED SNS WORLD

In this section we will investigate the interaction of people of these undeveloped countries in the Internet. We consider ‘Bangladesh’ as representative country from the South East-Asia to provide some data on this. Bangladesh is ranked 3rd among the 8 countries in this region in the e-Government ranking. Despite the efforts of the Government to provide e-services to its citizens, the web presence of different Government and non-Government institutions is quite low. Internet access is available to only 5 percent of her citizens and many of those who have access to Internet use it seldom. But if we consider the SNS presence of the people in Bangladesh, they are not far behind [16], [17].

There has been dispute regarding the total number of Internet users in Bangladesh. But despite the dissimilarity about the total number of Internet users from different online sources, it is noticeable that the ratio of the total number of Facebook users to Internet users from all the sources are close and roughly 43% of the Internet users in Bangladesh use Facebook. If we compare this ratio with other countries in the world (Table I), we can see that the ratio is good enough. A significant part of our Internet users are SNS user too.

This connectivity among the users has paved a new way for information gathering and sharing for the people of developing countries like Bangladesh. SE cannot give them the data that is not there in the web, but through SNS their query can reach hundreds of the people of their acquaintance, and as Yang *et. al.* [14] has already mentioned, they are traditionally encouraged to share the experience with others. This is not the end of ‘digital division’ mentioned earlier, but we are getting a bit closer to unify the world in terms of information searching and retrieval capacity.

V. EXPERIMENTAL DATA

Our data collection process had two phases. In the first part, we made a proposal for volunteers through our research group from which we selected 10 enthusiastic participants from two universities. All our participants had more than

150 friends in their Facebook profile (average 270) and uses Facebook regularly in their day-to-day life. They were instructed about what type of data we need to collect and how. They monitored the data stream in their Facebook home pages passively for questions asked through status messages and recorded those status messages with responses after about one hour and after about 5 hours of posting. We collected data for about 4 weeks and received 257 of such queries. Then we analyzed those questions according to the categories mentioned by Morris *et. al.* [9]. We tried to search answers for those questions using traditional search engines and compared them with the answers obtained from Facebook. We are still gathering more data, so the explanation provided in this section are not claimed as complete. But it should give some indication, emphasize our logic, and provide future directions for work.

In our second phase, we choose 10 participants using our already collected data. 5 of them has asked at least one question in the past one month while the rest has responded to at least one query made in Facebook. We tried to investigate the motivation behind using social network as an information source and the inspiration that worked behind answering it. Our interview data strongly supported our previous findings and also supported the findings made by Morris *et. al.* [9].

Table II shows some analysis of our obtained data. The data has good similarities with the data obtained by others, specially, like Yang *et. al.* [14], our data also indicates that people in the eastern culture asks less subjective queries than people in the western countries. However, unlike many other works, our study finds that significant part of the queries are related to finding factual information. When we analyzed the queries of such kind, we could understand the reason. Though these questions are objective and have definite answers, the users could not find the information in the web, and thus turning to SNS was the only option, aside contacting specific persons for it. As indicated by Morris *et. al.* [8], people often do a Google search before asking anything through SNS, probably this was the case with our queries too. But the ratio of such queries is astonishingly high in this region and considering the fact that web culture here has not expanded that much, it was somewhat expected.

There are many queries on different topics that we could not find specific answer in the web. Some examples were like “When is the next performance by Shironamhin/James?” (two popular bands in Bangladesh). In the developed countries, we can expect that the music providers keep record of their future events and update it frequently. But here in Bangladesh, we could not find any specific site maintained by them. But when people asked it in Facebook, they got the information almost instantly (within 5 minutes).

Another of the interesting queries and responses was about the traffic situation in a particular day. A person was on a very tight schedule to attend a workshop in Dhaka,



Figure 1. Example of question-answer in Facebook.

Bangladesh. He was supposed to land in Dhaka Airport at 8 AM, and his speech was scheduled at 10 AM in front of Governmental dignitaries. So he was asking people in this locality about possible real life traffic scenario during that time, describing the challenge he has to face. This kind of traffic information for Bangladesh is not available through Google map or any other service. But his friends could make valuable comments (Fig. 3), including an effective suggestion to get a front row sit while taking boarding pass so that the queue in front of him in the immigration remains small. During interview, he pointed that this is one of the reason he prefers to ask such questions in Facebook as it may show unexpectedly unorthodox but useful solutions.

Another interesting query we find was about “Does xx University publish any journal?”. Using Internet, we were able to find 3 journals published from that university. But in Facebook, the comments contained information about 6 journals. We contacted the relevant departments to verify that the information from those Facebook comments were accurate indeed. Those journals being local hard-copy only had no online presence, and thus quite hard to find using search engines.

Local information is another kind of information that people seemed to seek through SNS. Queries like “Has there been any accident in xx Road?”, “Do we have class test tomorrow?”, “What movies are now showing in xx cinema

Table II
QUESTION TYPES AND RESPONSE ANALYSIS

Question Type	Percent	Avg. First Response	Avg. Total Response	Appropriate Answer	Time Required to Search through SE
Recommendation	21.1	8.5 min.	6.2	Seemed somewhat appropriate	About 30 minutes searching to obtain reasonable information
Opinion	15.8	4.7 min.	9.5	No defined answer	No defined answer
Factual Knowledge	40.4	7 min.	6.9	Accurate in the 91.3 percent cases, the rest are unanswered	No information for 56.5 percent queries, about 5 min. for others.
Rhetorical	7	5 min.	12	Not applicable	Not applicable
Invitation	5.3	4.2 min.	15.5	Each got min. 1 positive reply	Not applicable
Favor	5.3	5.1 min.	7.1	Each got min. 1 positive reply	Not applicable
Soc. Conn.	3.5	5 min.	15	Yes	SE were not suitable
Offer	1.8	4.3 min.	8.2	Yes	Not applicable



Figure 2. Another example of question-answer in Facebook.



Figure 3. Another example of question-answer in Facebook.

hall?“, etc. are such examples. These queries are answered promptly by friends in the SNS, but we could not find answers to them through searching the web.

When we asked our participants about why they have chosen Facebook to ask the questions, all of them agreed to the fact that many of their queries are not satisfied by traditional search engines. While dealing with objective questions, they usually go through Google first. Only when they cannot find the information there, or are uncertain about the validity of it, they turn to Facebook to get the answer. However, there are some other cases too.

A person made a query about the location of the service centre of a particular mobile operator. He got prompt reply from his friends. That information was available in the web and could be found easily. In fact, his friends has Googled it

for him and gave him the answer. When we asked him about it, he agreed that this information was not urgent for him, and as he pass a lot of time using Facebook, he just made a post in the hope that someone may know it personally. He did not expect that this information in already available in the web and can be searched for. This shows that there is a gap in understanding the flourish of web technology in this region and often people are not aware what have changed around them in the past decade.

Our face-to-face interview session was focused on two things - why do we ask in Facebook and why do we care to answer those queries. Answer of the first topic could strongly support our primary hypothesis of living in a digitally in advanced world, while the second one shows significant correlation with other researchers' findings about motivation in

replying.

Strong ties like close friends, work peers, neighbors are more encouraged to reply to queries in Facebook, supporting the earlier research works. Another important motivation for replying is to make a positive introduction of oneself to the asker. People are often more motivated to answer the queries made by their seniors, or someone with which they want a more positive relation with. And of course people often do it selflessly, to friend, or to show others about his expertise in the relevant topic. It is in the nature of human beings to help others, and that will remain as the driving force behind the success of SNS search.

VI. CONCLUSION

In this work, we have focused on differences of the SNS searching habits in different regions of the world based on their economic context. We showed that the motivation for SNS research in developing regions could be quite different than in the developed parts. The lack of information availability has played a major role in peoples' turning to SNS to get answer than from traditional search engines. Whether other factors like culture, religion, etc. plays a significant role alongside these factors remain as a major research challenge.

ACKNOWLEDGEMENT

This research is conducted by the students and Faculty members of the Human Technology Interaction (HTI) research group at the Department of Computer Science and Engineering, Bangladesh University of Engineering and Technology (BUET).

REFERENCES

- [1] "Key Facts - Facebook Newsroom," <http://newsroom.fb.com/Key-Facts>.
- [2] "Alexa Ranking of Facebook," <http://www.alexa.com/siteinfo/facebook.com>.
- [3] "Facebook Usage Statistics," <http://www.thesocialskinny.com/100-social-media-statistics-for-2012/>.
- [4] C. Lampe, N. B. Ellison, and C. Steinfield, "Changes in use and perception of facebook," in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, ser. CSCW '08. New York, NY, USA: ACM, 2008, pp. 721–730. [Online]. Available: <http://doi.acm.org/10.1145/1460563.1460675>
- [5] M. Efron and M. Winget, "Questions are content: a taxonomy of questions in a microblogging environment," in *Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47*, ser. ASIS&T '10. Silver Springs, MD, USA: American Society for Information Science, 2010, pp. 27:1–27:10. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1920331.1920371>
- [6] J. Teevan, D. Ramage, and M. R. Morris, "#twittersearch: a comparison of microblog search and web search," in *Proceedings of the fourth ACM international conference on Web search and data mining*, ser. WSDM '11. New York, NY, USA: ACM, 2011. [Online]. Available: <http://doi.acm.org/10.1145/1935826.1935842>
- [7] C. Lampe, J. Vitak, R. Gray, and N. Ellison, "Perceptions of facebook's value as an information source," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12. New York, NY, USA: ACM, 2012, pp. 3195–3204. [Online]. Available: <http://doi.acm.org/10.1145/2207676.2208739>
- [8] M. R. Morris, J. Teevan, and K. Panovich, "A comparison of information seeking using search engines and social networks," in *ICWSM*, 2010.
- [9] —, "What do people ask their social networks, and why?: a survey study of status message q&a behavior," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '10. New York, NY, USA: ACM, 2010, pp. 1739–1748. [Online]. Available: <http://doi.acm.org/10.1145/1753326.1753587>
- [10] K. Panovich, R. Miller, and D. Karger, "Tie strength in question & answer on social network sites," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ser. CSCW '12. New York, NY, USA: ACM, 2012, pp. 1057–1066. [Online]. Available: <http://doi.acm.org/10.1145/2145204.2145361>
- [11] S. Farnham, M. Lahav, D. Raskino, L. Cheng, and T. Laird-McConnell, "So.cl: An interest network for informal learning," in *International AAAI Conference on Weblogs and Social Media*, 2012.
- [12] J. Teevan, M. R. Morris, and K. Panovich, "Factors affecting response quantity, quality, and speed for questions asked via social network status messages," in *ICWSM*, L. A. Adamic, R. A. Baeza-Yates, and S. Counts, Eds. The AAAI Press, 2011.
- [13] B. Hecht, J. Teevan, M. R. Morris, and D. J. Liebling, "Searchbuddies: Bringing search engines into the conversation," in *ICWSM*, 2012.
- [14] J. Yang, M. R. Morris, J. Teevan, L. A. Adamic, and M. S. Ackerman, "Culture matters: A survey study of social q&a behavior," in *ICWSM*, 2011.
- [15] "United Nations E-Government Survey 2010," http://www2.unpan.org/egovkb/documents/2010/E_Gov_2010_Complete.pdf.
- [16] "Internet User Statistics," <http://data.worldbank.org/indicator/IT.NET.USER/countries>.
- [17] "Facebook User Statistics," <http://www.socialbakers.com/facebook-statistics/page-2/?interval=last-6-months>.
- [18] M. Naaman, J. Boase, and C. hui Lai, "Is it really about me? message content in social awareness streams," in *In Proc. CSCW 2010*, 2010.