

Determining the Social Network Worth: Analysis, Quantitative Approach, and Application

Sujata Jindal

Computer Science Engineering Department,
SGT Institute of Engg. & Technology,
Gurgaon, India

Dr. Ritu Sindhu

Associate Professor, CSE Deptt.,
SGT Institute of Engg. & Technology,
Gurgaon, India

Abstract: Social networks are growing day by day. Users of the social networks are generating values for these networks. All the users can't be considered equal as they have different social network impact value. In this paper we analyze the social impact of a user and propose a method to estimate an individual's worth to a social network in terms of impact. The mathematical evaluations show the effectiveness of our method. Based on the proposed method many applications can be built taking into consideration the impact any individual's social profile has. We have tried to make various social data attributes more valuable and meaningful.

Keywords: Social networks, Social impact, Social data, Network computing, User engagement, Social media analytics

I. INTRODUCTION

Leading social networking companies like Google, Facebook, Twitter, and LinkedIn are making millions [2] by using user data. Imagine Google+[4], or Facebook[10] without users (that is us!) and Orkut's story is well known to all of us, Google announced closure of Orkut on September 30th 2014, [1] and the reason was growth of new social networking sites and users started moving to Facebook, so due to user base shift, Google had to shut down Orkut finally.

Such instances has given birth to a new thought in today's world where range of social networking sites are making use of users' growing interest in online networking, and they are using social data to their benefit and making millions of dollars by various means like selling it to other companies, behavioural analysis of masses, research to launch new products/services, online marketing, and many more countless ways. The biggest problem with this entire social networking arrangement is such that we, the biggest contributors, are hardly aware of this or may be just ignorant and happy with service or platform being provided to us to connect to the world. Each action, be it just a like, comment, tweet, +1, photo share, or location check-in; all of it is churning enormous amount of useful data behind the scenes every second for companies to use at will.

The amount of money these companies are making from this data is huge; a Forbes article on what a user might be worth is illustrated below and really worth looking at -

| Company name | Facebook | LinkedIn | Yahoo | Google |
|---------------------------------|----------|----------|---------|----------|
| Market cap (in billions) | \$100.56 | \$31.31 | \$27.67 | \$282.20 |
| Number of users (in millions) | 1,110 | 225 | 627 | 1,300 |
| Revenue (in billions) | \$1.813 | \$0.366 | \$1.135 | \$13.110 |
| Per user valuation | \$90.59 | \$131.55 | \$44.13 | \$217.08 |
| Average Revenue per User (ARPU) | \$1.63 | \$1.53 | \$1.81 | \$10.09 |

Figure 1: source: Forbes.com [2]

Total of average revenue per user from each of the sites is roughly USD \$15 (approx ₹1000), that's worth giving a serious thought as we users of these social networking sites are generating this real-time valuable data in return of one key service that is 'online networking' but the real questions that we all need to ask are -

- What is my real worth and how do I calculate it?
- If companies are making money off my data, should I not be paid a part of this - basis my worth?

Before we jump into math behind one's online worth, let's look at list of social metrics, data elements that are vital for online networking and their related actions and factors that turn all of this into valuable big data. For this research paper, we have looked at top four social networking sites – Facebook[10], Twitter[11], LinkedIn[12] and Google+[4].

Let's look at data elements, user actions and related impact factors of these four sites –

Table 1: Data elements, User actions and related Impact factors

| | Facebook | Twitter | LinkedIn | Google+ |
|-------------------------------------|--|--|--|--|
| Key Data Elements/Attributes | Common user info, User demographics info | Common user info, User demographics info | Common user info, User demographics info | Common user info, User demographics info |
| Network | Friends, Followers | Fans | Friends | Friends, Followers, Circles |
| Key Actions | Post, Like, Comment, Share | Tweet, Re-Tweet, Comment/Reply | Post, Like, Comment, Share | +1, Share, Comment |

While key social elements are inputs and social networking won't exist without it – imagine an online profile without your picture, location, education, job details etc. Similarly network is next important aspect of social networking, without friends or followers your profile has no real meaning in the online world. If the first two exist, actions would follow and that's what really takes your data to next level that's friends, friends of friends, and even the entire world if you are not careful with security settings of social networking sites.

II. PERSONAL BIG DATA AND ITS REACH

Personal data is the new class of data that is emerging due to this digital world [7-8].

Each of the element can be configured to be either private or non-private (in varied degree from friends, friends of friends, the world etc), as soon as you mark it as non-private (any degree), it's reach and impact increases exponentially. Same is with any of the posts (or tweets etc), depending upon the visibility its reach and impact increases manifold and further depending upon the size of network and user engagement, it can really be quite impactful in overall network.

III. CALCULATING SOCIAL IMPACT

Social impact assessment is a methodology to review the social influence of an individual on his or her network. There are 4 key contributors to one's social impact:

- **Data Elements (D)** - User information, Demographic information, e.g. Name, Location, Work Details, Academic Information etc.

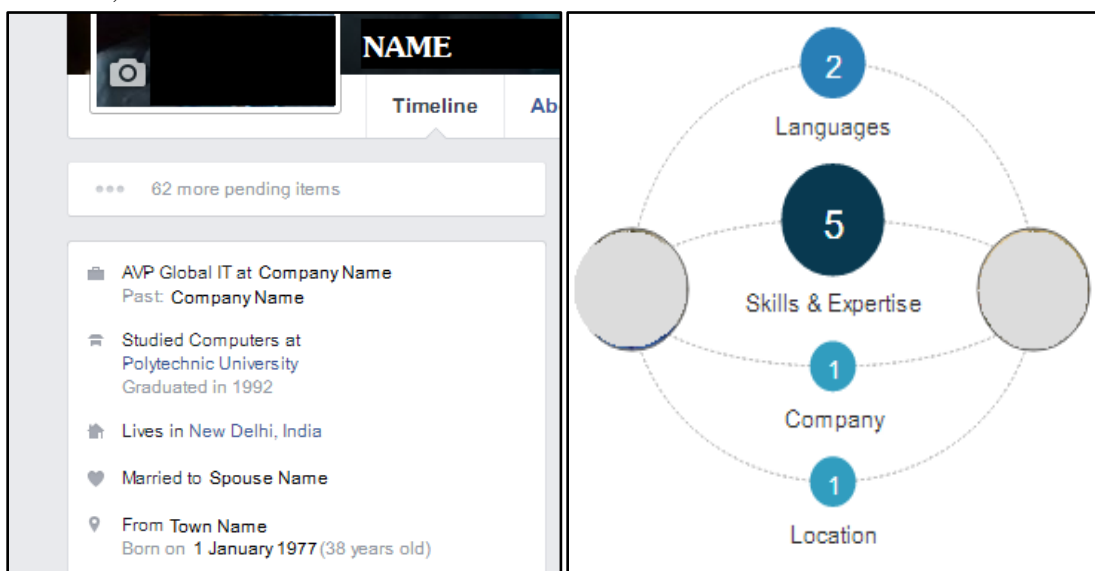


Figure 2: Sample Facebook profile data and LinkedIn examples. Such information can be used to find match with other profiles

- **Activities (A)** – Any user action is an activity, e.g. Post (update status, upload photo/video, tag someone), Like, Share, Comment etc. A network has almost no impact without regular activities. Activities play most important role in creating an impactful social network. This is the reason why companies are encouraging users to use mobile apps, which enable them capture more than what a user may like to share – like geolocation, place information etc, apart from giving users a way to stay connected more often by removing dependency on PC.



Figure 3: Sample Facebook activities list

- **Reach (R)** - Follower base (friends, fans, followers) defines one's reach in network, and reach grows exponentially as the network grows (1st level network, 2nd level network, and so on)

| What they studied 🔍 | Count | What they're skilled at 🔍 | Count | How you are connected | Count |
|--|-------|---------------------------|-------|-----------------------|-------|
| Electrical and Electronics Engineering | 8 | Testing | 6 | 1st Connections | 1 |
| Engineering | 6 | Troubleshooting | 4 | 2nd Connections | 1 |
| Computer Science | 3 | Project Engineering | 4 | Group Members | 2 |
| Computer and Information Sciences an... | 2 | Automation | 3 | 3rd + Everyone Else | 23 |
| Electrical, Electronics and Communica... | 2 | Agile Methodologies | 3 | | |
| Engineering Technologies and Enginee... | 2 | Engineering | 3 | | |

Figure 4: LinkedIn example showing reach categorized by study area, skills, network

- **Engagement Rate (E)** - Facebook defines Engagement Rate as “Engagement rate is the percentage of people who saw a post that liked, shared, clicked or commented on it” [3]

Engagement is a user action that can be anything like - views, likes, shares, retweet, clicks etc - which is directly proportional to user activities and overall Reach of an individual. Activities contribute to engagement with your users/viewers/visitors etc. An engaged user is key for the success of any content, so this is one of the most important metrics in online media.

- **Impact (I)** - Your overall impact on your network can be calculated basis above four factors - Data Element, Activity, Reach, and Engagement. Relationship among these five can be expressed as below -

$$\text{Impact} = \mathbf{D} \cdot \mathbf{A} \cdot \mathbf{R} \cdot \mathbf{E}$$

I call this expression as 'I DARE'

IV. FORMULATION AND CALCULATIONS

Impact of D (Data Elements)

To accurately measure the impact of D (data elements), we need to consider the V (visibility) exposure factor of those respective items. Impact is directly proportional to visibility, and same can be represented in the following expression -

$$\mathbf{D} = \sum_{i=0}^n (D_i \times V_i)$$

Here D is data element and V is visibility factor of that data element; which is fixed at 0.25 for private sharing, 0.5 for friends circle sharing, 0.75 for 'friends of friends' sharing and 1 for data element that is shared to public. The reason for assigning a non-zero coefficient for private sharing is the fact that it is still exposed to the social networking host/site and carries some value.

Impact of Activities (A)

Similar to impact of data elements, the impact of network activities can be calculated, and the visibility factor has same impact here as well, the only difference being the consolidation of sets of activities that share same visibility factor.

The impact of activities can be represented by the following expression.

$$\mathbf{A} = \sum_v \left(V_j \times \sum_a A(j) \right)$$

Reach Factor (R)

Let's look at R (reach) factor, which can be represented by the following expression.

$$R = \frac{\sum_{k=0}^m \frac{Fr_k}{N_k}}{\sum Fr}$$

Here N is network depth (1st level, 2nd level and so on) of kth friend (Fr), so closer the contact (or friend) in a network, more the overall reach. To find the average reach coefficient, the resultant is divided by total users in a network.

Engagement Rate (E)

Engagement Rate (E) is defined by engaged users and total reach of any content being shared. Anyone's network is his/her total potential reach, which comprises of fans, friends, followers etc. Engagement rate can be calculated for every single data element, or for a profile or page/post as a whole, here we shall look at individual data element as that the level we are trying to calculate the overall Impact value.

Overall Impact Factor

So, this is what the final *IDARE* formula looks like -

Impact = Impact of D x Impact Factor of Activities x Network Reach x Engagement Rate

$$I = \sum_{i=0}^n (D_i \times V_i) \times \sum_v \left(V_j \times \sum_a A(j) \right) \times \frac{\sum_{k=0}^m \frac{Fr_k}{N_k}}{\sum Fr} \times E$$

V. METHOD EVALUATION (MEASUREMENT AND RESULT ANALYSIS)

Let's see how IDARE formula works and can be used to calculate the social impact index. We shall look at two examples; first evaluation is based on sample numbers while the second one is based on a Facebook profile of one of the public figure in India.

Example 1: Sample Data

Table 2: Calculation of impact of data elements (D)

| Data Element(s) | Count (D) | Visibility | Visibility Factor (V) | Impact of Data Element (D x V) |
|-----------------|-----------|--------------------|-----------------------|--------------------------------|
| Name | 1 | Public | 1.00 | 1.00 |
| Work | 1 | Friends | 0.50 | 0.50 |
| Location | 1 | Friends of Friends | 0.75 | 0.75 |
| Education | 1 | Public | 1.00 | 1.00 |

| | | | | |
|-----|---|---------|---------------------|-------------|
| DOB | 1 | Private | 0.25 | 0.25 |
| | | | Total I(D) = | 3.50 |

Table 3: Calculation of impact of activities (A) per day

| Activities | Public (x1.0) | Friends (x0.75) | Friends of Friends (x0.50) | Private (x0.25) | Total Impact |
|------------|---------------|-----------------|----------------------------|--------------------|--------------|
| Posts | 30.00 | 10 | 15 | 8 | 47.00 |
| Comments | 70.00 | 0 | 0 | 0 | 70.00 |
| Likes | 75.00 | 0 | 0 | 0 | 75.00 |
| Shares | 50.00 | 0 | 0 | 0 | 50.00 |
| | | | | A (per day) | 8.07 |

Table 4: Calculation of reach factor (R)

| Network | Count (Fr) | Network Depth (N) | Reach (R=Fr/N) |
|--------------------|------------|--------------------------|----------------|
| Friends | 100 | 1.00 | 100.00 |
| Friends of Friends | 1000 | 1.25 | 800.00 |
| Followers | 50 | 1.50 | 33.33 |
| | | Total Reach (R) = | 933.33 |
| | | Avg. of R = | 0.81 |

For this formula evaluation exercise, let us assume the total engaged users in this given social network paradigm are 75 (max number of users across all activities - 70 comments, 75 likes, and 50 shares), and engagement rate would be -

$$E = 75 / (100+1000+50) = 0.07$$

With all given values, let us find out social impact of this user using IDARE equation -

$$I = 3.50 \times 8.07 \times 0.81 \times 0.07 = 1.49$$

Higher the value; bigger the impact would be.

Example 2: Data taken from official page of an Indian political leader as on 25th March 2015[5]

Table 5: Calculation of impact of data elements (D)

| Data Element(s) | Count (D) | Visibility | Visibility Factor (V) | Impact of Data Element (D x V) |
|-----------------|-----------|------------|-----------------------|--------------------------------|
|-----------------|-----------|------------|-----------------------|--------------------------------|

| | | | | |
|--|---|--------|---------------------|-------------|
| Name, Country, Description, Interests, Gender, Personal Info, Email, Website | 8 | Public | 1.00 | 8.00 |
| | | | Total I(D) = | 8.00 |

Table 6: Calculation of impact of activities (A) per day

| Activities | Public (x1.0) | Friends (x0.75) | Friends of Friends (x0.50) | Private (x0.25) | Total Impact |
|------------|---------------|-----------------|----------------------------|--------------------------------|-----------------|
| Posts | 6.00 | 0 | 0 | 0 | 6.00 |
| Comments | 8737.00 | 0 | 0 | 0 | 8737.00 |
| Likes | 540789.00 | 0 | 0 | 0 | 540789.00 |
| Shares | 7243.00 | 0 | 0 | 0 | 7243.00 |
| | | | | N_a (per day) | 79539.29 |

Table 7: Calculation of reach factor (R)

| Network | Count (Fr) | Network Depth (N) | Reach (R=Fr/N) |
|--------------------|------------|--------------------------|-------------------|
| Friends | 0 | 1.00 | 0.00 |
| Friends of Friends | 0 | 1.25 | 0.00 |
| Followers | 6349000 | 1.50 | 4232666.67 |
| | | Total Reach (R) = | 4232666.67 |
| | | Avg. of R = | 0.67 |

For this formula evaluation exercise, let us assume the total engaged users in this given social network paradigm are 540789 (max number of users across all activities - comments, likes, and shares), and engagement rate would be -

$$E = 540789 / 6349000 = \mathbf{0.09}$$

With all given values, let us find out social impact of this user using IDARE equation -

$$I = 8 \times 79539.29 \times 0.67 \times 0.09 = \mathbf{40649.52}$$

Now if we compare this profile's social impact with sample profile used in the 1st example, it is clearly evident that the social impact of the second profile is about 27300 times higher on IDARE scale.

VI. COMPARISON ANALYSIS

The chart below (Figure 5) depicts how three different profiles or pages can be compared using IDARE scale.

For this comparison analysis we have taken three real Facebook profiles' data – two of which are politicians and one is a sportsman. It is clearly evident that Impact factor has a positive correlation with Reach factor. Relation between Activity and Impact factor needs to be carefully examined for the profiles of Politician 2 and Sportsman; the Impact factor has dropped with Activity factor for Politician 2, but increase in Reach factor for Sportsman has resulted in increase of the Impact factor despite further drop in activities for him. Engagement rate has not been able to influence much for this case as the change in Engagement rate is overall negligible in comparison to Reach factor.

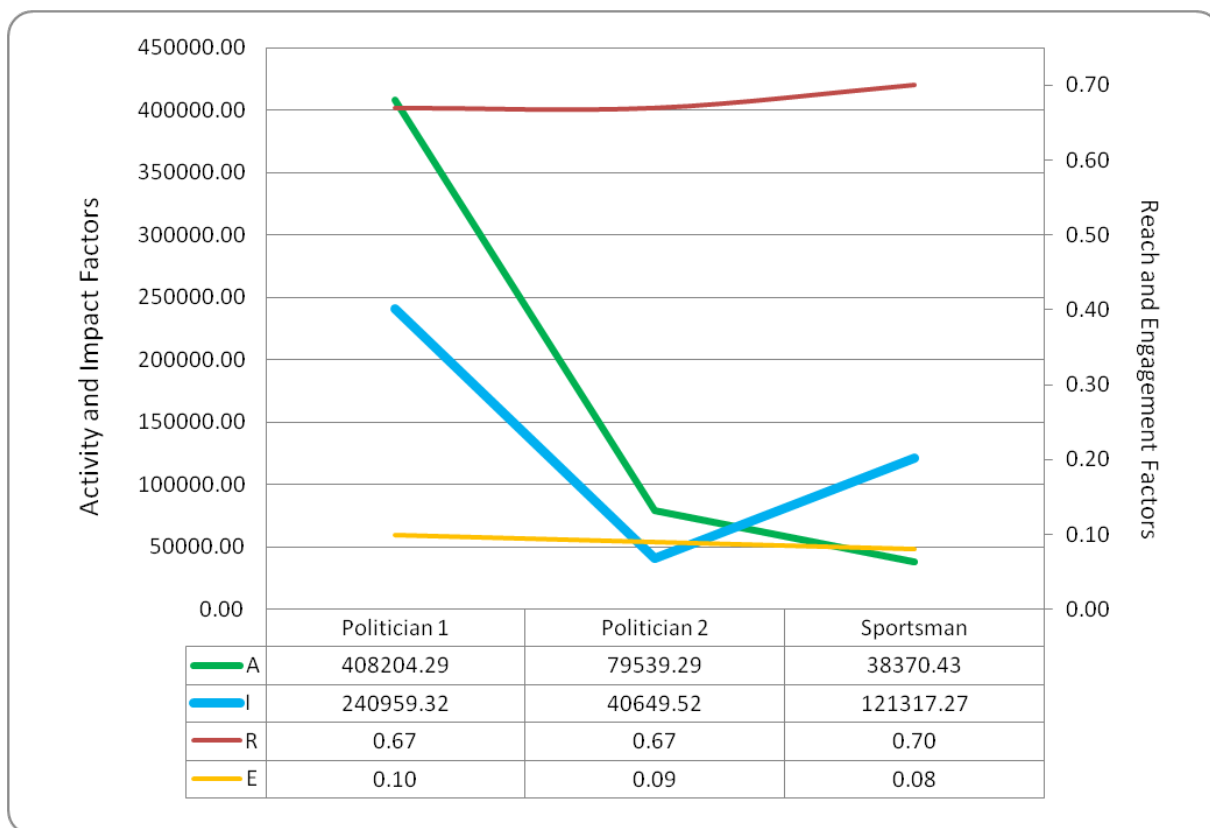


Figure 5: Comparison Chart of Three Profiles

VII. IDARE IMPLEMENTATION ON FACEBOOK

In this section, we discuss the effectiveness of the proposed method. For this purpose, we conducted several experiments to validate that our proposed method indeed measures the social network worth of social network users. To perform the experiments on Facebook real data, we developed a Facebook application (called IDARE).

The aim of this application is to:

- 1) Gather user's information (various inputs)
- 2) Apply this data to IDARE algorithm
- 3) Generate resulting impact factor

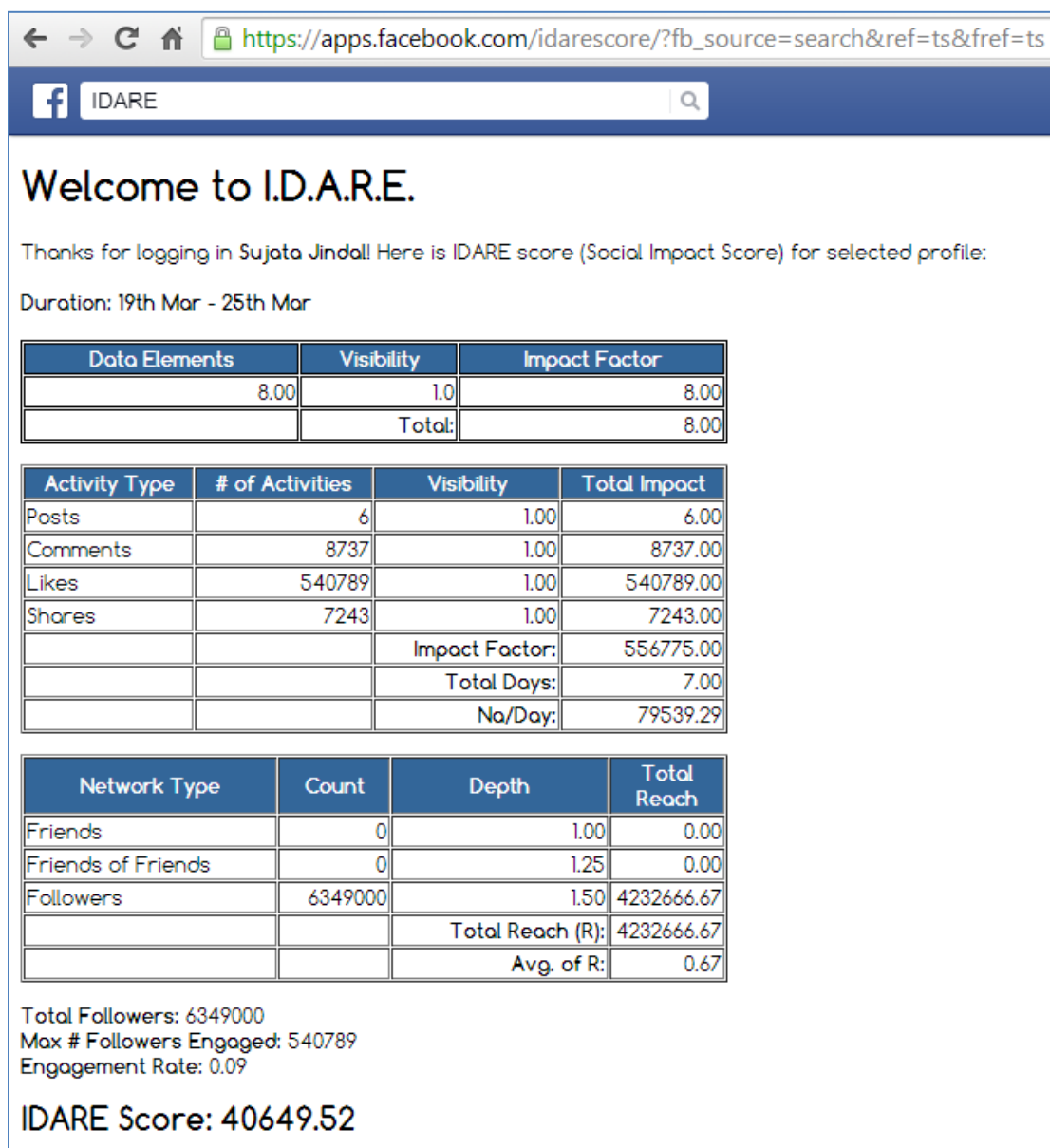


Figure 6: Facebook Application App – IDARE

VIII. CONCLUSION, DISCUSSION AND FUTURE WORK

The use of social media offers many advantages to both users and social networks. The point is whether the users are getting the real worth of their data in terms of benefits, various services. We need to have a more effective way to scientifically measure various social media activities including analytics and reporting, and their overall impact.

We present a systematic and scientific method to estimate an individual's impact value to a social network. The approach involves 3 steps - 1) Input various data elements that are exposed, 2) Average out the values, 3) Calculate the Impact value. The benefit of our method is twofold, first to a user (individual or company) and second to social networks to improve their services based on the impact of a particular user.

Various applications can be built based on this impact value e.g. to connect to high net worth individuals and last but not the least, it opens a new debate today if we as social network contributors should get the share of revenue or not?

ACKNOWLEDGEMENT

The work presented here could not have been possible without the contributions of many supportive and knowledgeable people.

First of all, I am grateful to my supervisor and guide Dr. Ritu Sindhu, Associate Professor (Department of CSE, SGT Institute of Engineering and Technology) for her support, encouragement and guidance throughout my work. Her feedback on my paper has been invaluable. Special thanks to Mr. Ashok Gupta [6], IT expert for providing valuable feedback and comments and helping me understand the field of big data and social networks. I would like to thank the faculty members and fellows at SGT Institute of Engineering and Technology for their inputs, valuable discussions and accessibility.

REFERENCES

- [1] "Orkut." *Wikipedia*. Wikimedia Foundation. Web. 28 Mar. 2015. <<http://en.wikipedia.org/wiki/Orkut>>.
- [2] "How Much Is A User Worth?" *Forbes*. Forbes Magazine. Web. 28 Mar. 2015. <<http://www.forbes.com/sites/tristanlouis/2013/08/31/how-much-is-a-user-worth/>>.
- [3] "Facebook." *How Is Engagement Rate Defined?* Web. 28 Mar. 2015. <<https://www.facebook.com/help/178043462360087>>.
- [4] "One Account. All of Google." *Google*. Web. 03 Apr. 2015. <<http://plus.google.com/>>
- [5] "Arvind Kejriwal." *Facebook*. Web. 28 Mar. 2015. <<https://www.facebook.com/AAPkaArvind>>.
- [6] "Ashok Gupta, PMP | LinkedIn." 2014. 27 Mar. 2015 <<https://in.linkedin.com/in/ashokgupta>>
- [7] "Rethinking Personal Data | World Economic Forum ..." 2014. 9 Feb. 2015 <<http://www.weforum.org/projects/rethinking-personal-data>>
- [8] Jindal S, Sindhu R (2014) Personal Big Data Usage and Controls – Review, *International Journal of Advanced Research in Computer Science and Software Engineering* 4(10), October- 2014, pp. 351 – 354
- [9] "One Account. All of Google." *Google*. Web. 03 Apr. 2015. <<http://plus.google.com/>>
- [10] "Facebook Logo." *Facebook*. Web. 03 Apr. 2015. <http://facebook.com/>
- [11] "Welcome to Twitter - Login or Sign up." 2007. 3 Apr. 2015 <<https://twitter.com/>>
- [12] "LinkedIn: World's Largest Professional Network." 2003. 3 Apr. 2015 <<https://www.linkedin.com/>>