# Ranking Nodes By Silentness

Soheil Ghanbari
Faculty of Engineering Science
College of Engineering
University of Tehran
Tehran, Iran
Email: soheil.ghanbari@ut.ac.ir

Hasan Heydari
Faculty of Engineering Science
College of Engineering
University of Tehran
Tehran, Iran
Email: h_heydari@ut.ac.ir

Ali Moeini
Faculty of Engineering Science
College of Engineering
University of Tehran
Tehran, Iran
Email: moeini@ut.ac.ir

***Abstract*—Silentness in networks refers to the behavior that a node receives lots of information from other nodes but share nothing or little information with them. We can rank people in social networks by silentness. In this paper we present an algorithm based on random walks for ranking nodes by silentness. The time complexity of the proposed algorithm in a network with $n$ nodes is $O(log^2 n)$ with high probability, while the state-of-the-art algorithm does not specified time complexity and runs until holds convergence conditions and we show it does not converge in all cases by a counterexample. We assess the proposed algorithm with Fagin's intersection metric and Bperef methods and compare the implementation results of the algorithm on GPlus and Twitter datasets with PageRank and I/O ranking methods. We implement our algorithm on Hadoop framework, as well and in compare of the state-of-the-art algorithm reduces $64.48\%$ of the disk I/O.**

## I. Introduction

In recent years networks are inseparable parts of real-life communication systems. Networks are composed of individuals units, called nodes, connected together in some ways. Each unit in a network contains or generates information that can be sent by links. Technological Networks (e.g. The Internet, Power grids, Transportation Networks), Social Networks (e.g. Facebook, Google plus), Networks of Information (e.g. the World Wide Web, Citation Network) and Biological Networks (e.g. Biochemical, Neural and Ecological Networks) are four major classes of networks [1]. With this expanding definition, scientists are interested in analyzing networks in different aspects.

One of the well studied problems on networks is proposing a measure for calculating importance of nodes. Several measures like degree [2], closeness [3], betweenness [4] and PageRank [5] have been introduced. The PageRank originally used to rank webpages in Web, but it can be used in other networks area such as ranking journals [6], measuring the importance of components in mechatronic systems [7], endorsement deduction in social networks [8] and manufacturing service recommendations [9]. The PageRank can be computed by power method [5] but after that some works applied Monte Carlo method on PageRank. Earlier works like [10] used Monte Carlo method to proposing of reducing number of iterations and determined PageRank after only one iteration with high accuracy. But in recent years, Das Sarma et al. used Monte Carlo method to make PageRank more appropriate for distributed networks [11].

978-1-5090-5795-5 /17/$31.00 © 

Dealing with dangling nodes is a challenge for computing PageRank that even the original paper for PageRank suggested to omit dangling nodes from network but other works tried to handle them [12], [13], [14]. A dangling node in a network is a node that has one or more in-coming edges and no out-going edge. Dangling nodes is also referred as hanging nodes [14], zero-out nodes [15], dead end or sink nodes [13] and in the context of Web is called frontier pages [12].

More general types of dangling nodes are silent nodes which the in-degree to out-degree ratio is greater than one. In the different kind of networks the majority of nodes can be considered as silent nodes. For example, in social network *lurkers* consist majority members of community. According to "90-9-1" principle 90% of an online community only read and use information that community contains, 9% of members edit and share some information and only 1% of network members have an active role in producing information an sharing it with other members [16].

Mining the lurkers on social networks is an important problem [17]. Therefore, many efforts has been done for understanding of lurkers, definition of lurkers and discussing the reasons for emerging lurkers as mentioned in [16] and the references therein. Study the problem of encouraging the lurkers to do more activity on social networks leads to present some delurking approaches [18], [16], [19]. In addition other work consider the prediction of lurkers reading topics in social networks [20].

Definition of silent nodes is based on the ratio of in-degree and out-degree of a node. But lurkers in a network do not exactly coincident with this site of view. Lurkers have a behavior that we call it silentness. That means the lurker receives lots of information from network and share nothing or very little information. It should be noted that in-degree and out-degree of a node is not an only effective factor on silentness, there are also non-negligible factors like importance of in-neighbors and the amount of produced information that should be considered. Hence, ranking nodes by silentness can help us to recognize and manage lurkers on networks.

Ranking lurkers in a social network can aid to understanding the lurking behaviors and revealing the relationship between lurkers and other users. Interdonato and Tagarelli presented an iterative algorithm for ranking silent nodes in an information networks, that is the first ranking method in this context with the best of our knowledge. They considered the amount of the clamorous behavior of in-neighbors and amount of the silence behavior of out-neighbors in addition of in-degree/out-degree ratio of a node [21].

In this paper we present a random walk based algorithm for ranking nodes by silentness in a network. We try to introduce a ranking method that consider relative importance of in-neighbors in addition of in-degree/out-degree radio. Our algorithm will be terminated in $O(\log^2 n)$ rounds with high probability (w.h.p.). We assess the proposed algorithm with two popular assessment methods (Fagin's intersection metric and Bperef) and compare the implementation results of the algorithm on GPlus and Twitter datasets with two basic ranking methods (I/O and PageRank). Massive scale of social networks persuade us to implement our algorithm and the state-of-the-art algorithm on Hadoop framework and compare their performances which our algorithm reduce $64.48\%$ of disk I/O.

The rest of the paper is organized as follows: Section II provides some preliminaries, terminologies and notations that will be used throughout the rest of the paper. After that, our algorithm is explained in Section III. We proved that the time complexity of the proposed algorithm is $O(\log^2 n)$ w.h.p. In section IV we presented implementation results on Twitter and GPlus datasets and we assessed our results with methods like Fagins intersection metric and Bpref. Implementation of the proposed algorithm and the state-of-the-art algorithm on Hadoop framework and their comparison result presented, as well. We will talk about conclusions and some future works in section V. At the end, in Appendix we prove the state-of-the-art algorithm diverges in some cases.

## II. PRELIMANARIES

Let $G = (V, E)$ be a directed graph representing a network where $V$ and $E$ are the set of nodes and edges respectively. $G$ can be called graph of a social network in which $V$ is a set of peoples and $E$ is a set of connections between peoples. We denote the number of nodes in the network by $n$. For any particular edge $e \in E$ and $e = (v, u)$, $v$ and $u$ represent the initial and terminal vertices, respectively. The in-degree of $v \in V$ is denoted by $deg^-(v)$ and its out-degree is denoted by $deg^+(v)$, such that $deg^-(v) = \big|\{e|\forall e \in E \wedge e = (u, v)\}\big|$ and $deg^+(v) = \big|\{e|\forall e \in E \wedge e = (v, u)\}\big|$. The in-neighbors and out-neighbors of $v \in V$ is denoted by $N^-(v)$ and $N^+(v)$, respectively.

Node $v$ with $deg^-(v) = 0$ is called a *source*. Similarly, node $u$ with $deg^+(u) = 0$ is called a *sink*, *dangling* or *trivial-silent*. If a node is neither a source nor a sink, it is called an *internal*. A node $v \in V$ is called silent if $\frac{deg^-(v)}{deg^+(v)} > 1$ and clamorous if $\frac{deg^-(v)}{deg^+(v)} \leq 1$.

As standard, we use the phrase with high probability to mention an event has probability at least $1 - \frac{1}{n^c}$, where $n$ and $c$ are the number of nodes in the network and an arbitrary constant that greater than one, respectively.

Here we state a useful fact that we use it in the next section. Most of social networks such as Facebook and Twitter have small-world properties and are a scale-free network [22], [23], [24]. In small-world and scale-free networks, the diameter of the network is $O(\log n)$ or less [25].

## III. ALGORITHM

In this section we will explain our approach for ranking nodes by silentness based on random walks. At the first we try to separate dangling nodes and rank them. In next step we change the algorithm to rank all silent nodes (trivial and non-trivial silent nodes).

The main Idea of the algorithm is as follows. Let $G = (V, E)$ be a directed graph. We set that $k = \log n$ is the number of random walks starting from each node that obtained from our simulation studies. In each round, each random walk will be absorbed with probability of $\epsilon$ (we will talk about its value later) and with probability of $(1 - \epsilon)$ the random walk will choose one of the out-neighbors and transport to that node. If there is no out-going link the random walk absorbs any way. Each node counts the number of random walks absorbed at that node per each round. When all the random walks absorbed the algorithm will be finished. At this point each node can compute rank of itself by dividing number of absorptions in that node to the number of random walks.

Now we want to discuss about the value of $\epsilon$. At the first step let $\epsilon$ be a constant (lower that 0.5). We will call it $\epsilon$-constant Alg. In this case in nodes with out-going links, a random walk is more probably to continue its path than terminate. But in case of dangling nodes the random walk will be terminated any way. Mostly, the number of termination in dangling nodes will be more than other nodes. In comparison between dangling nodes, a dangling node with more in-coming links and with in-neighbors that have more visitors is more likely to trap a random walk. But in some cases a silent node or a clamorous one may rank higher than a dangling node. Because of relative importance and the number of in-neighbors and it states the difference between definition of silent nodes and concept of silentness. That also is the key point of our algorithm.

With this approach we can separate dangling nodes from other nodes and rank them in addition of ranking other nodes. But in this case we cannot sift not-trivial silent nodes from other nodes. For this purpose in second step we do not set $\epsilon$ as a constant. Let $\epsilon$ be variable with the upper bound and the lower bound. We will call it $\epsilon$-variable Alg. On the other hand, as mentioned before, the diameter of a social network is $O(\log n)$ or less. We are supposed to let the random walk has a chance for traversing the diameter of the network. This helps us to consider the effection of nodes that located to far from given node $v$ in addition of it's neighbors on nodes ranking. Since, the number of steps of the random walk $(X)$ has a geometry distribution with parameter $\epsilon$. Thus, in order to the expectation of $X$ becomes $O(\log n)$, we choose $\epsilon = c/\log n$ where $c$ is a constant. We set $0.2/\log n$ for lower bound and $0.8/\log n$ for upper bound by experiments.

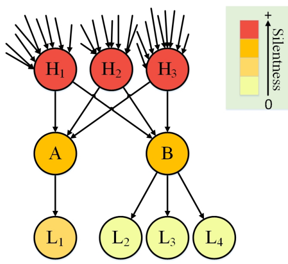



Fig. 1: Constant $\epsilon$ leads to same rank value for $A$ and $B$. But variable $\epsilon$ ranks $A$ higher than $B$.

---

**Algorithm 1** Ranking nodes by silentness

---

**Input (for every node)**: The number of nodes ($n$), the constant value of $\epsilon$ and the upper and lower bound of it, maximum out-going degree.
**Output**: A number between 0 and 1 representing rank of each node.

For each $v \in V$, $A_v$ and $NA_v$ shows the number of the absorbed and non-absorbed random walks, respectively.
$A_v = 0, NA_v = \log n$
**while** all random walks absorbed **do**
  **for** $i = 1, 2, ..., NA_v$ **do**
    With probability of $\epsilon$, $A_v = A_v + 1$ in otherwise select an out-going neighbor uniformly and transport to it.
  **end for**
**end while**
Each node computes its rank as $\frac{A_v}{n \log n}$.

---

The value of $\epsilon$ will be different for different nodes depending on number of out-going links and will be computed initially. For nodes that have no out-going degree or have one, the $\epsilon$ will be equal to the lower bound, for nodes with maximum out-going links $\epsilon$ equals to the upper bound and for others will be between the upper bound and the lower bound with proportion of number of out-going links to maximum of out-going degree.

In the other hand it means more silent the nodes become probability of a random walk termination will be increase. In this way we consider in/out ratio of a node in addition to strength of non-silent behavior of in-neighbors. To clarifying the difference of $\epsilon$-constant Alg. and $\epsilon$-variable Alg. in Fig. 1 $\epsilon$-constant Alg. leads to same rank value for $A$ and $B$. But $\epsilon$-variable Alg. ranks $A$ higher than $B$.

### A. Time complexity analysis

*Lemma 1:* Let a simple random walk absorbed with probability of $\epsilon = \frac{c}{\log n}$ (which $c$ is a constant) and $X$ be a random variable that counts the number of steps. Then, we have

$$\Pr[X \geq \frac{\log^2 n}{c}] \leq \frac{1}{n^{(\frac{\log n - 1}{3c})}}$$

*Proof:* $X$ has geometry distribution with parameter $\epsilon$. Thus, $E(X) = \frac{\log n}{c}$. By applying a chernoff bound [26], we have

$$\Pr[X \geq \frac{\log^2 n}{c}] \leq exp(-\frac{\log n(\log n - 1)}{3c}) = \frac{1}{n^{(\frac{\log n - 1}{3c})}}$$

■

*Theorem 1:* Let $G = (V, E)$ be a graph and $|V| = n$. If we start $\log n$ random walks from each vertex with absorbing probability $\epsilon = \frac{c}{\log n}$ (which $c$ is a constant) Then, all random walks absorbed in $\frac{\log^2 n}{c}$ rounds w.h.p.

*Proof:* We define event $E_i^j$ $(1 \leq i \leq n, 1 \leq j \leq \log n)$ that indicates the $j^{th}$ random walk started from $i^{th}$ vertex, absorbed in less than $\frac{\log^2 n}{c}$ steps. Then, we have

$$\begin{aligned}
&\Pr(E_1^1 \cap E_1^2 \cap ... \cap E_1^{\log n} \cap E_2^1 \cap ... \cap E_n^{\log n}) \\
&= \big(\Pr(E_1)\big)^{n \log n} \\
&= \big(1 - \Pr(\bar{E}_1)\big)^{n \log n} \\
&\geq (1 - \frac{1}{n^{(\frac{\log n - 1}{3c})}})^{n \log n} \qquad \text{(By Lem. 1)} \\
&\geq exp\big(-\frac{2n \log n}{n^{(\frac{\log n - 1}{3c})}}\big) > e^{-\frac{1}{n^{\Omega(1)}}} \geq 1 - \frac{1}{n^{\Omega(1)}}
\end{aligned}$$

■

## IV. Experiments

### A. Assessment

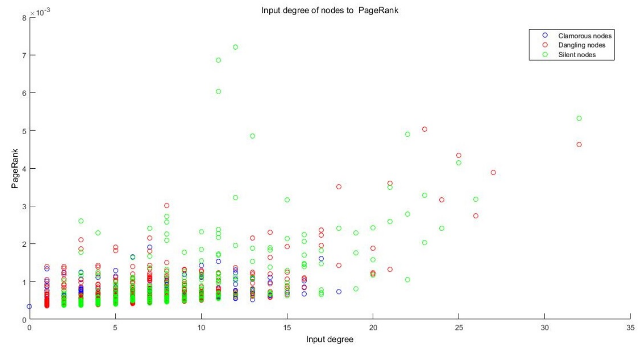
For assessment of the algorithm ($\epsilon$-variable Alg.) we use GPlus dataset with over 107,000 nodes and 13,000,000 edges, and Twitter dataset with 81,000 nodes and 1,700,000 edges, which both of them are directed graphs. In the graphs each edges between two nodes means one of the nodes follow another one. A silent node in such graphs represents a person who has followers but does not follow much people.

At the first, the direction of links in GPlus and Twitter networks are from follower-to-followee. Running the proposed algorithm on these networks leads to rank foe influence of celebrities. In other hand, by inversing the direction of a network from follower-to-followee to followee-to-follower we can show the data flow in the social network and in this new graph a silent node is a person whos followee is more than his/her followers and actually that is more fit with silence definition. We applied our algorithm on inversed datasets.

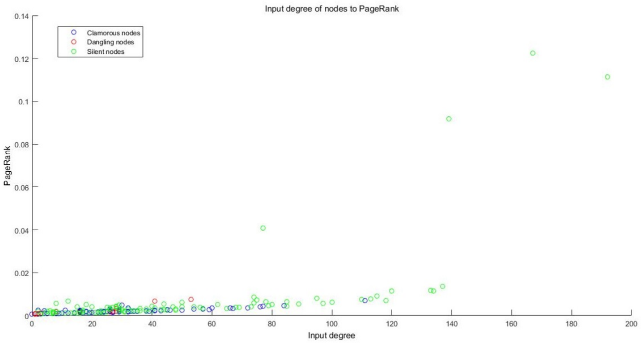
For assessment of our algorithm we compared our rank list against PageRank as most popular ranking method and I/O method which includes in-degree/out-degree nodes in a network.

Hear we used two popular measures Bpref [27] and Fagins intersection metric [28] in order to compare our method with others (PageRank and I/O) and measure the effectiveness of our ranking algorithms.

Fagin's intersection metric is a measure to show how two Ranking list are identical. Assume L1 and L2 are identical then the Fagin's intersection metric measure will be equal to 1. And if L1 and L2 be completely different the measure is actually

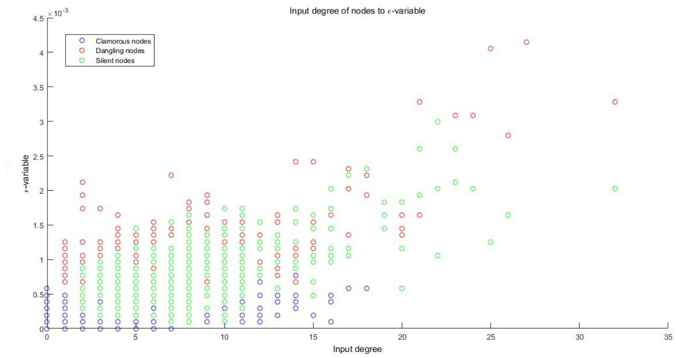



(a) GPlus

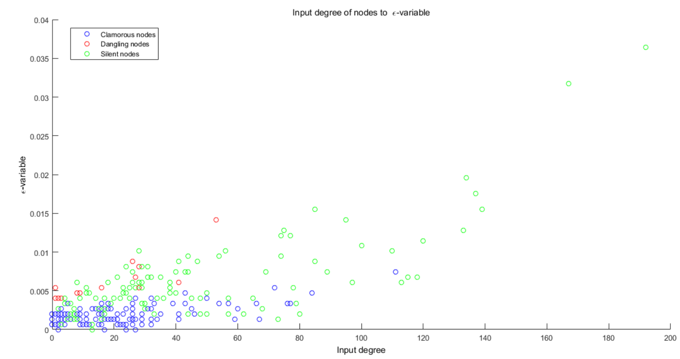



(b) Twitter

Fig. 2: PageRank on GPlus and Twitter




(a) GPlus




(b) Twitter

Fig. 3: $\epsilon$-variable Alg

equals to 0. This measure helps us to show even though we present a new ranking method it is related with PageRank and I/O its not completely irrelevant. Fagin's intersection metric measure also weights the lists in order to top of the lists gets higher weights from the tail.

Bpref is a measure for evaluating performance of ranking method for this reason we judged some nodes as relevant nodes ($R$) and judged some nodes as non-relevant ($N$) and we let others be unjudged, then Bpref computed by Bpref $(R, N) = \frac{1}{R}\sum_r \left(1 - \frac{|n \text{ ranked higher than } r|}{max(R,N)}\right)$ where $r$ is relevant nodes retrieved and $n$ is judged non-relevant nodes retrieved. According to the Bpref score is based on number of non-relevant nodes that ranked higher than relevant nods and does not get account unjudged nodes.

### *B. Results*

In this section we will show the result of applied $\epsilon$-variable Alg. on datasets. Now we will see the result of PageRank and $\epsilon$-variable Alg. separately, after that we will assess result using two assessment criteria.

In Fig. 2 we show the result of PageRank on GPlus and Twitter datasets, red points represents trivial-silent nodes or dangling nodes, green points show non-trivial silent nodes and blue ones are clamorous nodes with in-degree/out-degree $< 1$, the vertical axis shows ratio of ranks and horizontal axis is in-degree of a node, obviously PageRank do not rank nodes by silence.

Fig. 3 includes result of $\epsilon$-variable Alg. As we can see $\epsilon$-variable Alg. can rank non-trivial silent nodes and non-silent nodes as well as trivial-silent nodes. If we consider nodes with in-degree, this method can separate nodes with respect of amount of silence.

Now let assess our result using Bpref and Fagin's intersection metric. For using Bpref we judged dangling nodes as related and others as non-related nodes, with this assumption Bpref measure for I/O methods equaled to one, because for dangling nodes in-degree/out-degree is infinite. So we will talk about other methods.

As we can see in table I, $\epsilon$-variable Alg. have near one Bpref value. It shows that this method are suitable for ranking silent nodes, this happened because the number of non-silent nodes and nontrivial-silent nodes that ranked higher than trivial-silent nodes are really low. But Bpref value for PageRank is about $0.5$ that means PageRank is not proper Ranking method for ranking silent nodes. I/O is not appropriate Rank method for Ranking silent nodes because I/O ratio for each node was not influenced by neighbors and I/O methods can not make difference between nodes with same I/O ration. As mentioned before I/O method and our method both have Bpref measure value about one, but is it meaning that they are identical? The answer is negative.

With respect to result in table II it can be concluded that I/O method and our method have completely different rank for $k = 50$, and Fagin's intersection metric measure value for comparing our methods with PageRank is about $0.5$ that means $\epsilon$-variable Alg. is not identical with PageRank too but

| | GPlus | | | | Twitter | | | |
|---|---|---|---|---|---|---|---|---|
| n | 49380 | | 100470 | | 2470 | | 78603 | |
| k | 50 | 100 | 50 | 100 | 50 | 100 | 50 | 100 |
| Our Alg. vs. I/O | 0.147 | 0.154 | 0.054 | 0.083 | 0.329 | 0.545 | 0.000 | 0.001 |
| Our Alg. vs. PageRank | 0.762 | 0.763 | 0.776 | 0.754 | 0.632 | 0.641 | 0.579 | 0.624 |

TABLE II: Fagin's Result

| | GPlus | | Twitter | |
|---|---|---|---|---|
| n | 49380 | 100470 | 2470 | 78603 |
| PageRank | 0.468 | 0.482 | 0.396 | 0.337 |
| Our Alg. | 0.829 | 0.878 | 0.771 | 0.841 |
| I/O | 1 | 1 | 1 | 1 |

TABLE I: BPref Result

have some relations.

Massive scale of social networks persuade us to implement our algorithm and the state-of-the-art algorithm on Hadoop framework and compare their performances which $\epsilon$-variable Alg. reduce $64.48\%$ of disk I/O. The result on Twitter and GPlus datasets will be shown in Fig. 4. The flow chart of MapReduce will be shown in Fig. 5, as well.

## V. Conclusion and Future Works

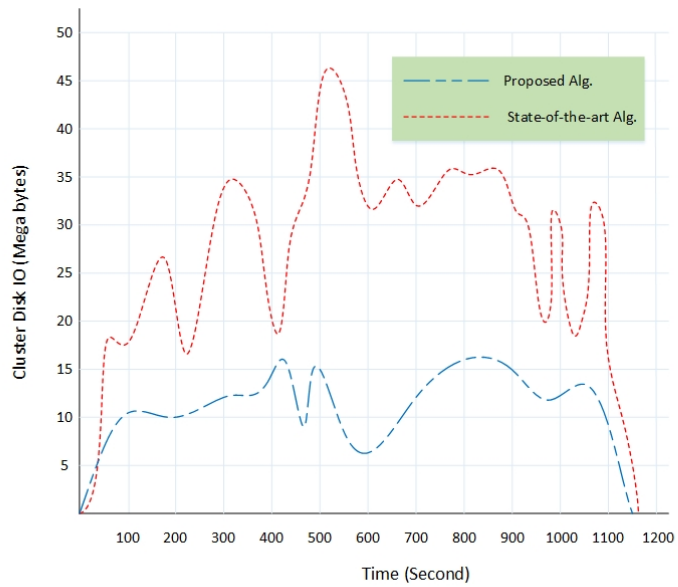
In this paper we have presented an algorithm based on random walks for ranking nodes by silentness. The time complexity of the proposed algorithm in a network with $n$ nodes is $O(log^2 n)$ w.h.p. We have assessed the proposed algorithm with Fagin's intersection metric and Bperef methods and compare the implementation results of the algorithm on GPlus and Twitter datasets with I/O and PageRank ranking methods. We have implemented our algorithm on Hadoop framework, as well and in compare of the state-of-the-art algorithm reduces $64.48\%$ of the disk I/O. As the future work, its a good idea to present a distributed version of the proposed algorithm. Moreover, we can work to propose an approach in order to use this algorithm in recommender systems.

## Appendix

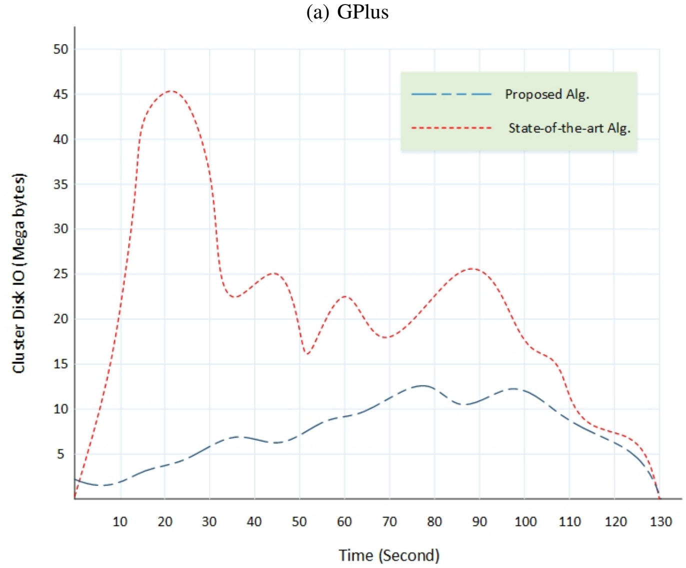
The state-of-the-art algorithm for ranking silent nodes proposed the following approach. In graph $G = (V, E)$ the rank of node $v \in V$, iteratively computed by following relation until the rank of nodes converge.

$$r_v = \alpha\Bigg[\Big(\frac{1}{deg^+(v)}\sum_{u\in N^-(v)} w(u,v)\frac{deg^+(u)}{deg^-(u)}r_u\Big)\Big(1+ \\ \big(\frac{deg^-(v)}{\sum_{u\in N^+(v)} deg^-(u)}\sum_{u\in N^+(v)} w(v,u)\frac{deg^-(u)}{deg^+(u)}r_u\big)\Big)\Bigg] + \frac{1-\alpha}{|V|}$$

We want to show the above relation does not converge in some cases. For this purpose, let the graph $G = (V, E)$ be a directed $k$-regular graph. In other words, $\forall v \in V, deg^-(v) = deg^+(v) = k$. We set $\alpha = 0.85$ and $\forall i, j \wedge i \neq j : w(i, j) = 1$. We denote to the rank of node $v \in V$ in $i^{th}$ iteration by $r_v^i$ and




(a) GPlus




(b) Twitter

Fig. 4: Cluster disk I/O usage

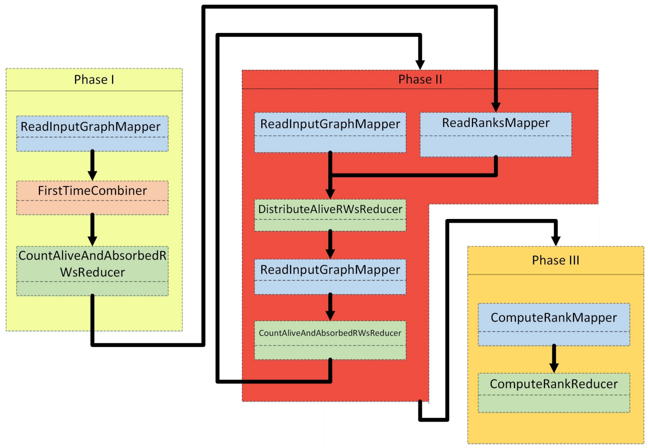



Fig. 5: Hadoop implementation flow chart of $\epsilon$-variable Alg.

in initialization we set $\forall v \in V : r_v^0 = c$, which $c \geq 1$. In the first iteration, we have $\forall v \in V, r_v^1 = 0.85 \times c \times (1+c) + \frac{0.15}{|V|}$ in which $r_v^1 > c$. In the second iteration, $\forall v \in V : r_v^2 > r_v^1 > c$. Since, $\forall v \in V$ the sequence $\sigma_v = r_v^0, r_v^1, r_v^2, ...$ is an increasing sequence and $\lim_{n\to\infty} r_v^n = \infty$, the sequence diverges.

## References

[1] M., Newman, Networks: An Introduction, Oxford University Press, 2010
[2] S., Wasserman, Social Network Analysis: Methods and Applications, Cambridge University Press ,1994
[3] A., Bavelas, "Communication patterns in task-oriented groups", Journal of the Acoustical Society of America, vol. 22, pp. 725–730, 1950
[4] L. C., Freeman, "A set of measures of centrality based on betweenness", Sociometry, vol. 40, pp. 35–41, 1977
[5] S., Brin, L., Page, "The anatomy of a large-scale hypertextual Web search engine", Computer Networks and ISDN Systems, vol. 30, pp. 107–117, 1998
[6] D., Yu, W., Wang, S., Zhang, W., Zhang, R., Liu, "A multiple-link, mutually reinforced journal-ranking model to measure the prestige of journals", Scientometrics, doi:10.1007/s11192-017-2262-9, 2017
[7] Y., Wang, L., Bi, S., Lin, M., Li, H., Shia, "A complex network-based importance measure for mechatronics systems", Physica A: Statistical Mechanics and its Applications, vol. 466, pp. 180–198, 2017
[8] H., Prez-Ross, F., Seb, J., Maria Rib, "Endorsement deduction and ranking in social networks", Computer Communications, vol. 73, pp. 200–210, 2016
[9] W. Y., Zhang, S., Zhang, S. S., Guo, "A PageRank-based reputation model for personalised manufacturing service recommendation", Enterprise Information Systems, vol. 11, pp. 672-693, 2017
[10] K., Avrachenkov, N., Litvak, D., Nemirovsky, N., Osipova, "Monte Carlo methods in PageRank computation: When one iteration is sufficient", SIAM Journal on Numerical Analysis, vol. 45, pp. 890–904, 2007
[11] A., Das Sarma, A. R., Molla, G., Pandurangan, E., Upfal, "Fast distributed PageRank computation", Theoretical Computer Science, vol. 561, pp. 113–121, 2015
[12] N., Eiron, K. S., McCurley, J. A., Tomlin, "Ranking the web frontier", in Proceedings of the 13th International Conference on World Wide Web, pp. 309–318, New York, 2004
[13] P. R., Kumar, A. G. K., Leng, A. K., Singh, A., Mohan, "Efficient methodologies to determine the relevancy of hanging pages using stability analysis", Cybernetics and Systems, vol. 47, pp. 376–391, 2016
[14] S. A., Kumar, P. R., Kumar, L. A. G., Kwang, "Efficient methodologies to handle hanging pages using virtual node", Cybernetics and Systems, vol. 42, pp. 621–635, 2011
[15] X., Wang, T., Tao, J. T., Sun, A., Shakery, C., Zhai, "DirichletRank: Solving the zeroone-gap problem of PageRank," in Proceedings of the ACM Transactions on Information Systems, doi:10.1145/1344411.1344416, New York, 2008
[16] N., Sun, P. P., Rau, L., Ma, "Understanding lurkers in online communities: A literature review", Computers in Human Behavior, vol. 38, pp. 110–117, 2014
[17] A., Tagarelli, R., Interdonato, "Time-aware analysis and ranking of lurkers in social networks", Social Network Analysis and Mining, pp. 5–46, 2015
[18] R., Interdonato, C., Pulice, A., Tagarelli, "Community-based delurking in social networks", Advances in Social Networks Analysis and Mining (ASONAM), pp. 263–270, San Francisco, 2016
[19] M.A., Javarone, R., Interdonato, A., Tagarelli, "Modeling evolutionary dynamics of lurking in social networks", in The 2016 International Conference on Complex Networks, pp. 227–239, Dijon, 2016
[20] W., Gong, E., Lim, F., Zhu, "Posting topics $\neq$ reading topics: on discovering posting and reading topics in social media", in Proceedings of the 12th International Conference and School on Advances in Network Science, pp. 14–28, Wroclaw, 2016
[21] R., Interdonato, A., Tagarelli, "Ranking silent nodes in information networks: a quantitative approach and applications", in 3rd International Conference Frontiers in Diagnostic Technologies, pp. 3641, Frascati, 2015
[22] S., Aparicio, J., Villazn-Terrazas, G., lvarez, "A model for scale-free networks: application to twitter", Entropy, vol. 17, pp. 5845–5867, 2015
[23] T., Suzumura, M., Watanabe, "How social network is evolving?: a preliminary study on billion-scale twitter network" in Proceedings of the 22nd International Conference on World Wide Web, pp. 531–534, Rio de Janeiro, 2013
[24] J., Wohlgemuth, M. T., Matache, "Small-world properties of Facebook group", Complex Systems, vol. 23, pp. 197–225, 2014
[25] R., Cohen, S., Havlin, "Scale-free networks are ultrasmall", Phys. Rev. Lett., vol. 90, doi: 10.1103/PhysRevLett.90.058701, 2003
[26] D., Dubhashi, A., Panconesi, "Concentration of Measure for the Analysis of Randomized Algorithms", Cambridge University Press, 2012
[27] C., Buckley, E. M., Voorhees, "Retrieval evaluation with incomplete information", in Proc. ACM SIGIR, pp. 25–32, Sheffield, 2004
[28] R., Fagin, R., Kumar, D., Sivakumar, "Comparing Top k Lists", SIAM Journal on Discrete Mathematics. vol. 17, pp. 134–160, 2003