# Massively Parallel Single-Source SimRanks in $o(\log n)$ Rounds

**Siqiang Luo**[1*] , **Zulun Zhu**[1*]

[1]Nanyang Technological University

{siqiang.luo, ZULUN001}@ntu.edu.sg

## Abstract

SimRank is one of the most fundamental measures that evaluate the structural similarity between two nodes in a graph and has been applied in a plethora of data mining and machine learning tasks. These tasks often involve single-source SimRank computation that evaluates the SimRank values between a source node $u$ and all other nodes. Due to its high computation complexity, single-source SimRank computation for large graphs is notoriously challenging, and hence recent studies resort to distributed processing. To our surprise, although SimRank has been widely adopted for two decades, theoretical aspects of distributed SimRanks with provable results have rarely been studied. In this paper, we conduct a theoretical study on single-source SimRank computation in the Massive Parallel Computation (MPC) model, which is the standard theoretical framework modeling distributed systems. Existing distributed SimRank algorithms enforce either $\Omega(\log n)$ communication round complexity or $\Omega(n)$ machine space for a graph of $n$ nodes. We overcome this barrier. Particularly, given a graph of $n$ nodes, for any query node $v$ and constant error $\epsilon > \frac{3}{n}$, we show that using $O(\log^2 \log n)$ rounds of communication among machines is enough to compute single-source SimRank values with at most $\epsilon$ absolute errors, while each machine only needs a space sub-linear to $n$. To the best of our knowledge, this is the first single-source SimRank algorithm in MPC that can overcome the $\Theta(\log n)$ round complexity barrier with provable result accuracy.

## 1 Introduction

Evaluating the structural similarity between two nodes in a graph is fundamental in plenty of data mining and machine learning tasks. Examples include recommendation systems [Chen *et al.*, 2022; Mo and Luo, 2021], avoiding customer churn [Luo *et al.*, 2019], spam detection [Badola and Gupta, 2021], link prediction [Zhou *et al.*, 2019; Xu *et al.*,

2017; Hours *et al.*, 2016] and graph mining [Zhu *et al.*, 2024; Jin *et al.*, 2011; Liao *et al.*, 2022; Zhu *et al.*, 2022]. Among many similarity measures, SimRank [Jeh and Widom, 2002] is one of the most widely adopted measures over graphs. SimRank is defined based on the intuition that two nodes are similar only when their neighboring nodes are similar. Formally, it uses the following recursive equation to compute the SimRank between two nodes $u$ and $v$, where $c \in (0, 1)$ and $\mathcal{I}(u)$ denotes the in-neighbor set of node $u$.

$$s(u,v) = \begin{cases} 1, & u = v \\ \frac{c}{|\mathcal{I}(u)||\mathcal{I}(v)|} \sum_{u' \in \mathcal{I}(u)} \sum_{v' \in \mathcal{I}(v)} s(u',v'), & u \neq v \end{cases} \quad (1)$$

Since it was proposed by Jeh and Widom [Jeh and Widom, 2002], SimRank has gained increasing popularity in various application domains, e.g., social analysis [Zheng *et al.*, 2013], nearest neighbor search [Lee *et al.*, 2012], as well as recommender systems [Antonellis *et al.*, 2008; Chuanyan and Xiaoguang, 2021; Symeonidis *et al.*, 2021].

**Distributed SimRank Computation.** We focus on *single-source SimRank computation*, whose importance has been uncovered in a plethora of recent studies [Wang *et al.*, 2020b; Li *et al.*, 2015; Wang *et al.*, 2020a; Shi *et al.*, 2020; Wang *et al.*, 2021; Maehara *et al.*, 2014; Tian and Xiao, 2016; Kusumoto *et al.*, 2014]. Given a graph $G$, a single-source SimRank for source node $u$ evaluates the SimRank values between $u$ and all other nodes in the graph. Single-source SimRank computation is widely used in applications where a ranking of the other objects with respect to an object is required. For example, it can be applied in search engines to locate the most similar web pages to a given one [Fogaras and Rácz, 2005], or in social network services to recommend new friends to a given user [He *et al.*, 2010; Nguyen *et al.*, 2015], or act as a ranking measurement to cluster objects [Cai *et al.*, 2008].

Given a graph of $n$ nodes, computing single-source SimRank is challenging for large graphs because it inherently involves $O(n)$ times of pairwise SimRank evaluations, each of which can already be too costly. Particularly, following the recursive form in Equation 1, computing $s(u,v)$ requires accessing many pairs of nodes in the graph, leading to $O(n^2)$ complexity. To address the efficiency issue, recent works [Jiang *et al.*, 2017; Zhang *et al.*, 2017; Shao *et al.*, 2015] employ a random-walk-based approach to

---

approximate the SimRank values. The main idea is to translate the SimRank computation into estimating the meeting probability of two decay-based random walks from the two source nodes (See Section 2.2 for more details). The computational complexity is then dependent on the number of random walks sampled, achieving a significant speed-up. However, even with the improved approach, it is still challenging to compute single-source SimRanks when $n$ is large.

Therefore, it is increasingly popular to apply *distributed computation* [Wang *et al.*, 2020b; Li *et al.*, 2015; Song *et al.*, 2018] to SimRank computation, which involves multiple machines to compute SimRank values in a collaborative manner, ultimately scaling up the computation to large graphs. In modeling distributed computation, the local computation cost within each machine is typically dominated by the synchronization cost among the machines. Hence, by convention, the local computation cost is omitted in MPC analysis and the focus is on reducing communication rounds. Furthermore, there is an intrinsic trade-off between the number of communication rounds and space-per-machine.

**Open Problem.** Existing distributed SimRank algorithms (e.g., [Li *et al.*, 2015; Song *et al.*, 2018; Wang *et al.*, 2020b]) mostly focus on empirical evaluation, and non-trivial theoretical analysis with provable approximation result guarantees is rarely given. We analyze existing representative distributed SimRank algorithms and summarize their requirements regarding communication rounds and machine space in Table 1*. In a nutshell, they require either $\Omega(n)$ per-machine space (UniWalk [Song *et al.*, 2018]) or $\Omega(\log n)$ communication rounds (DISK [Wang *et al.*, 2020b] and CloudWalker [Li *et al.*, 2015]). We remark that overcoming the $\Theta(\log n)$ round complexity barrier in MPC for natural problems is usually challenging and it has attracted tremendous interest [Luo *et al.*, 2022; Lacki *et al.*, 2020; Luo, 2019; Behnezhad *et al.*, 2019; Czumaj *et al.*, 2021; Luo, 2020] to improve the distributed computation down to $O(\textbf{poly}(\log \log n))$ rounds. Hence, a natural *open theoretical problem* is raised:

*Given an error threshold $\epsilon \in (0,1)$, let an approximate single-source SimRank algorithm be an algorithm that outputs SimRank values with $\epsilon$ absolute errors. Is there a distributed approximate single-source SimRank algorithm over a graph of $n$ nodes that can be finished in $o(\log n)$ rounds, while each machine only needs $o(n)$ space?*

**Our Main Results.** In this paper, we give a positive answer to the aforementioned open problem and present a distributed single-source SimRank algorithm that suits the Massively Parallel Computing (MPC) model [Karloff *et al.*, 2010]. Particularly, for a graph of $n$ nodes and $m$ edges, we focus on the MPC model that involves a set of machines, each having a sub-linear space $S = n^\alpha$ for some $\alpha \in (0,1)$. Our main results can be stated by the following theorem. To the best of our knowledge, this is the first distributed single-source SimRank algorithm that achieves sub-$\log n$ communication rounds while only requiring strongly sub-linear per-machine

space. We leave the proofs of lemmas and theorems in the Appendix of our technique report [Luo and Zhu, 2024].

**Theorem 1.** *Given a source node $u$ in a graph of $n$ nodes and $m$ edges, there is an algorithm that computes $\epsilon$-absolute-error guaranteed SimRank values between $u$ and all the other nodes using $M$ machines in $O\left(\log^2 \log n\right)$ communication rounds with high probability* [†]. *This algorithm only requires the space per machine is $S = n^\alpha$ for some $\alpha < 1$, and $M$ is some value of $\tilde{O}((m+n)/S)$.*

## 2 Preliminaries

### 2.1 Distributed Computation Model

Distributed computation is one of the most important techniques to address various computation tasks [Luo *et al.*, 2023; Wang *et al.*, 2020b; Li *et al.*, 2015; Song *et al.*, 2018; Luo *et al.*, 2022; Luo *et al.*, 2014; Luo *et al.*, 2012; Klauck *et al.*, 2014]. In recent years, Massively Parallel Computation (MPC) [Karloff *et al.*, 2010] becomes a popular theoretical framework in modeling the complexity of a distributed algorithm, because it closely simulates the situation of general distributed computation. An MPC model has three important parameters: the input data size $F$, the number of involved machines $M$, and the space capacity (words) $S$ on each machine. The frequently used notations are listed in Table 2 in the Appendix. It is required that when given $F$ and $S$, the number of machines $M$ should be of $\tilde{O}(\frac{F}{S})$, where $\tilde{O}(\cdot)$ hides a poly-logarithmic factor. Our main focus on the MPC model derives from several perspectives as follows:

**Space.** Consider an input graph of $n$ nodes. MPC for graph algorithms can be categorized into three types: strongly **super-linear** space ($S = n^{1+\omega}$) for some constant $\omega > 0$, near **linear** space ($S = \Theta(n)$), and strongly **sub-linear** space ($S = n^\alpha$) for some constant $\alpha \in (0,1)$. A super-linear model can typically be employed with a local algorithm, which loses the generality to be employed on a large scale of data. Hence, many studies focus on sub-linear settings that better capture the scalability of a distributed system. In this paper, we focus on the strongly sub-linear setting where each machine cannot even store the whole set of graph nodes.

**Communication Rounds.** The computation in the MPC model is based on *communication rounds*. Initially, each edge is randomly assigned to a machine. We assume each node has an integer ID from 1 to $n$, and each machine has an integer ID from 1 to $M$. The computation proceeds in synchronized rounds. At the beginning of a round, each machine may receive the messages sent from some other machines in the previous rounds. During a round, every machine conducts local computation based on its local data or messages received. Then each machine will send the computed results, packed as messages, to target machines. Each machine creates message packages to be routed onto the network and hence the size of messages sent/received per machine in one round shall not exceed the space capacity $S$. A new round starts only after the end of the previous round. The term *round complexity* [Ghaffari *et al.*, 2019] refers to the number of rounds required

---

*For the limitation of space, we leave the related work analysis in Appendix C. As the authors did not give an analysis based on the MPC model, we analyze them on our own.

[†]We say an event happens with high probability, if there exists a constant $\tau > 0$ such that the event happens with probability at least $1 - \frac{1}{n^\tau}$, where $n$ is the number of graph nodes.

| Algorithm | Rounds | Total space cost per round | Accuracy error | Space cost per-machine |
|---|---|---|---|---|
| CloudWalker[1] [Li et al., 2015] | $O(\frac{nl(I+l)}{b})$ | $O\left(\frac{bl^2\log n}{\epsilon_p^2}+n+m\right)$ | no guarantees | $O\left(\frac{bl^2\log n}{\epsilon_p^2}+\frac{m+n}{M}\right)$ |
| UniWalk[2] [Song et al., 2018] | $O(l)$ | $O\left(\frac{n^2 l\log n}{\epsilon^2}\right)$ | $\epsilon+\epsilon_b$ | $O\left(\frac{n^2 l\log n}{\epsilon^2}\right)$ |
| DISK[3] [Wang et al., 2020b] | $O\left(\frac{\log^2 n}{\epsilon_d^2}+K\right)$ | $O(m\log n+Kn)$ | $\frac{c\left(1-c^K\right)\epsilon_d}{1-c}+c^{K+1}$ | $O\left(\frac{m\log n+Kn}{M}\right)$ |
| **Ours**[4] | $O\left(\log^2\log n\right)$ | $O\left(m\log n+n^{1+o(1)}\log^{4.5}n+\frac{n\log^{3+o(1)}n}{2\left(\epsilon-\frac{3}{n}\right)^2}\right)$ | $\epsilon$ | $O(n^\alpha)$ |

[1] $\epsilon_p$ is the error when estimating the random walk distribution, $b$ is the parameter that controls the number of nodes to be handled in a single machine, and $I$ is the number of iterations in the Jacobi method. $l$ is a user-defined walk length or number of steps. [2] $l$ is a user-defined walk length or number of steps, and $\epsilon_b$ is the absolute error of SimRank introduced by BiWalk [Fogaras and Rácz, 2005]. [3] $K$ is the number of truncated terms of linearized SimRank and $\epsilon_d$ is an internal estimation error threshold. [4] $\alpha < 1$ and $n^\alpha > \log^5 n/(\log\frac{1}{\sqrt{c}})$.

Table 1: Communication rounds of distributed single-source SimRank Computation.

for program execution and serves as a measure of distributed computation cost, primarily driven by the expensive communication among machines. The main principle of designing algorithms in the MPC model is to achieve low round complexity subject to the space constraint on each machine.

**Existence of an MPC Algorithm.** With a per-machine space of $S = n^\alpha$, we acknowledge the existence of an MPC algorithm such that when $n \geq n_0$ (with $n_0$ being a constant), a distributed algorithm can operate within this space constraint using $M$ machines, where $M = \tilde{O}(\frac{m+n}{S})$. For ease of analyzing the MPC algorithms, we first give a preparation lemma regarding machine space expansion, as follows.

**Lemma 1.** *If there exists an $O(f(n))$-round distributed graph algorithm that works for per-machine space $S = \Theta(n^\alpha)$ for any $0 < \alpha < 1$ using $M = \tilde{O}(\frac{m+n}{n^\alpha})$ machines, and $f(n)$ is a function not related to $\alpha$ (i.e., $\alpha$ only contributes a constant factor to the round complexity), then there exists an MPC algorithm that works for per-machine space $S = n^\alpha$ with the same round complexity.*

This lemma eliminates the obstacle of analyzing round complexity when there is a constant factor expansion of the machine space. For example, consider that we have an algorithm with a certain distributed algorithm such that (1) its round complexity hides $\alpha$ factors; (2) applies to any $\alpha \in (0, 1)$ with $M = \tilde{O}(\frac{m+n}{n^\alpha})$, and (3) the space per machine is $Cn^\alpha$ for some constant $C$. Then Lemma 1 guarantees the existence of an MPC algorithm with $n^\alpha$ machine space.

## 2.2 Approximate SimRank Computation

Given a directed graph $G = (V, E)$, and let $n = |V|$ and $m = |E|$. Following [Wang et al., 2020a; Shi et al., 2020], we aim to compute approximate SimRank values with constant errors. In particular, given any source node $u \in G$ and a constant error $\epsilon$, we aim to compute the SimRank values between $u$ and any other node $v \in G$, such that $|s(u,v) - \tilde{s}(u,v)| \leq \epsilon$, where $s(u,v)$ denotes the true SimRank value and $\tilde{s}(u,v)$ denotes the estimated value.

Calculating the SimRank value iteratively according to Equation 1 may occupy large memory space and incur a high computation cost. Therefore, state-of-the-art approaches employ the following $\sqrt{c}$-decay walk-based computation, first proposed in [Tian and Xiao, 2016]:

**Lemma 2.** *For any two nodes $u, v \in G$, the SimRank between $u$ and $v$ is equal to the meeting probability of two $\sqrt{c}$-decay walks starting at $u$ and $v$ on the reverse graph $\bar{G}$, where two walks meet if there exists an integer $i \geq 0$, such that the $i$-th step of the two walks visit the same node.*

Here, a $\sqrt{c}$-decay walk on $\bar{G}$ from a node $u$ is a traversal on $\bar{G}$ such that at each step of the walk, it has $1 - \sqrt{c}$ probability to stop at the current node, and otherwise jumps to the next node that is a uniformly chosen out-neighbor of the current node. Here we define a length-$l$ walk as a path that includes $l + 1$ nodes and $l$ edges. We also say a length-$l$ walk has $l$ steps. Particularly, we have $(\sqrt{c})^i \cdot (1 - \sqrt{c})$ probability to generate a length-$i$ walk from a given starting node.

**Monte Carlo Method.** By sampling $N$ pairs of $\sqrt{c}$-walks from $u$ and $v$, and if $H$ pairs of the walks meet, then $\frac{H}{N}$ is an estimator of $s(u,v)$. The sampling number $N$ controls the estimation accuracy since more random walks approximate the value of $s(u,v)$ more accurately.

## 3 $O(\log^2\log n)$-**Round Algorithm**

It is not too challenging to design $O(\log n)$-round MPC algorithm (See Appendix E for details), but further improving the round complexity can be difficult. To explain, we outline challenges of direct adaptation of $\sqrt{c}$-walk based approach. First, the maximum length of a $\sqrt{c}$-decay walk can be infinite; a straightforward method to compute length-$l$ in a distributed environment easily entails $l$ communication rounds because one step may require one communication round when some neighbors are located in a different machine. Such design easily leads to $\Omega(\log n)$-round algorithms. Second, evaluating single-source SimRank values for node $u$ requires running random walks from both $u$ and all the other nodes, resulting in a large number of random walks being conducted. This can lead to a high space cost per machine because some hub nodes are prone to be passed by many random walks. Therefore, a careful design of the MPC algorithm is required to guarantee a small number of rounds and a low space cost for each machine. Third, the random walks, once computed in the MPC model, are stored in different machines. As such, detecting whether two walks meet may overload the machine regarding space cost.

**Algorithm 1:** Overall algorithm
___
**Input** : Graph $G = (V, E)$; source node $u$; decay factor $c$;
Montel-Carlo failure probability $\delta$; accuracy error $\epsilon$
**Output:** SimRank scores $\tilde{s}(u, v)$ for each $v \in V$
___
**1** *Store $G$ across multiple machines*;
**2** *Reverse the edges of $G$ within each machine to form $\bar{G}$*;
**3** $l \leftarrow \log_{\frac{1}{\sqrt{c}}} n$;
**4** **for** $i = 0$ **to** $l$ in parallel **do**
**5** $\quad \left| \quad N_i = \left\lceil \frac{\log 2n}{2(\epsilon - \frac{3}{n})^2} \cdot (\sqrt{c})^i \cdot (1 - \sqrt{c}) \right\rceil \right.$;
**6** $\quad \left| \quad$ **foreach** *node $v \in V$ in parallel* **do**
**7** $\quad \quad \left| \quad \right.$ Generate $N_i$ random walks from node $v$ in
$\quad \quad \quad \quad O\left(\log^2 \log n\right)$ rounds;
**8** *Shuffle and decompose random walks (See Sections 4.2 and 4.3);*
**9** *Run **Algorithm 2** to calculate $\tilde{s}(u, v)$ for each $v \in V$;*
___

## 3.1 New Interpretation of Walk-based Approach

To address the aforementioned challenges, we reinterpret the SimRank computation between node $u$ and node $v$ in a batch manner using the following three operations, which are more MPC-friendly:

**(a) Random walk generation.** Instead of generating $\sqrt{c}$-decay random walks whose lengths are non-deterministic, we generate random walks with deterministic length distribution. Particularly, to generate $N$ $\sqrt{c}$-walks from $u$, we will generate $N \cdot \sqrt{c}^i (1 - \sqrt{c})$ length-$i$ walks for $i \geq 0$, following the corresponding geometric distribution of walk lengths. Similar operations are conducted for random walks starting from $v$. Fixing the random walk lengths is more MPC-friendly. As we will show in Section 4.1, we can generate sufficient such random walks both from the source node $u$ and each node $v \in V$ in only $O\left(\log^2 \log n\right)$ MPC rounds.

**(b) Random walk shuffling and matching.** Since the random walks from $u$ (or $v$) are generated in the order of increasing lengths (due to Step (a)), pairing up the random walks from $u$ and $v$ directly for meeting detection is not valid for estimating SimRanks as the $i$-th walks from $u$ and $v$ are correlated (they have the same length). In order to maintain the randomness, we shuffle the random walks above.

**(c) SimRank computation.** Last, we compute the SimRanks by detecting how many pairs of walks meet and estimating the SimRank value $s(u, v)$ by Lemma 2.

## 3.2 Overview of Main Algorithmic Steps

Based on the new interpretation, we give our main algorithmic steps in Algorithm 1, which consists of the following five main stages. For ease of presentation, we first outline the main idea in this section and defer the detailed MPC-related operations in Section 4.

**Initial State.** The SimRank algorithms are based on the topology information of the input graph. Initially, the whole graph should be partitioned across different machines, as assumed by the MPC model. We assume that each machine holds a random partition of edges of the graph (Line 1 in Algorithm 1), and hence each machine holds roughly $\frac{m}{M}$ edges. We also reverse the edges so that the graph represents $\bar{G}$ (Line

**Algorithm 2:** Detect meeting and calculate SimRank
___
**Input** : Tuple set $D_G$ from Section 4.3; source node $u$
**Output:** Estimated SimRank score $\tilde{s}(u, v)$ for each $v \in V$
___
**1** *Sort all elements in $D_G$;*
**2** *Calculate $z = \frac{\log 2n}{2(\epsilon - \frac{3}{n})^2}$, $|Z_{u,l}| = z \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c})$ and*
$\quad |\hat{Z}_{u,l}| = \lceil |Z_{u,l}| \rceil$ *for $l = 1, ..., \log_{\frac{1}{\sqrt{c}}} n$ and $u \in V$ ;*
**3** **for** *all tuple $(i, j, v_j, l, v) \in D_G$ in parallel* **do**
**4** $\quad \left| \quad \right.$ **if** $v = u$ **then**
**5** $\quad \quad \left| \quad \right.$ Label $(i, j, v_j, l, v)$ as 1
**6** $\quad \left| \quad \right.$ **else**
**7** $\quad \quad \left| \quad \right.$ Label $(i, j, v_j, l, v)$ as 0
**8** **for** *all tuple $(i, j, v_j, l_2, v)$ labeled by 0 in parallel* **do**
**9** $\quad \left| \quad \right.$ Link it to the closest tuple $(i, k, v_k, l_1, u)$ labeled by 1
$\quad \quad$ using PREDECESSOR;
**10** $\quad \left| \quad \right.$ **if** $(j = k) \wedge (v_j = v_k)$ **then**
**11** $\quad \quad \left| \quad \right.$ Send message $(j, u, l_1, v, l_2)$ to $\lfloor \frac{vM}{n} \rfloor$-th machine.
**12** **for** *each unique message $(j, u, l_1, v, l_2)$ in each machine in parallel* **do**
**13** $\quad \left| \quad \right.$ $\tilde{s}(u, v) \leftarrow \tilde{s}(u, v) + \frac{|Z_{u,l_1}| \cdot |Z_{v,l_2}|}{|\hat{Z}_{u,l_1}| \cdot |\hat{Z}_{v,l_2}|} \cdot \frac{1}{z}$
___

2 of Algorithm 1).

**Parallel Random Walks Generation.** This stage corresponds to Operation (a) mentioned earlier. As we need to compute single-source SimRank values from any source node $u$ to all other nodes, we generate $N$ $\sqrt{c}$-walks from each node. Further, we truncate them at $\log_{1/\sqrt{c}} n$ length (Line 3 of Algorithm 1). By setting $N = \frac{\log 2n}{2(\epsilon - 3/n)^2}$, we show that the truncated walks only cause negligible influence on the final accuracy (see Section 4), and we can guarantee the $\epsilon$ error bound of SimRank values. Another problem is when generating $N_i = N \cdot (\sqrt{c})^i (1 - \sqrt{c})$ length-$i$ walks, $N_i$ may not be an integer. In Section 5 we give a rounding technique and prove that the rounding still guarantees the error bound. We also prove that these random walks can be obtained in the MPC model using $O(\log^2 \log n)$ communication rounds (Lines 4-7 of Algorithm 1 and details in Section 4.1).

**Shuffling Random Walks.** It's important to note that in Lemma 2, the lengths of the random walks originating from $u$ and $v$ may differ, as they are independently and randomly generated. To guarantee the randomness of each pair of random walks in our Monte Carlo simulation, the generated random walks have to be shuffled (Line 8 of Algorithm 1). The shuffling operation in the MPC model is not as trivial as in a single machine because the generated walks from the same node can be located in different machines. Shuffling these walks incurs communication between machines, which may violate the space bound in each machine.

**Decomposing Random Walks.** After shuffling, we pair the $j$-th walk from the source node $u$ with the $j$-th walks from other nodes respectively, and detect whether two walks in each pair meet. Detecting whether two walks meet is relatively simple in the single-machine setting because all information can be loaded locally. However, in a distributed environment, we need to detect whether two paired-up walks meet while the walks are stored at different machines. This
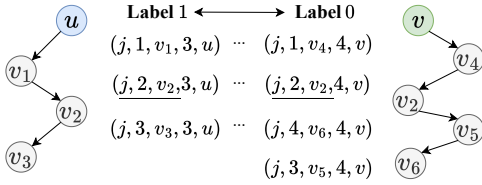
Figure 1: Detect meeting using tuple association.

poses drastically different challenges. To address the challenges, we decompose the generated random walks into tuples, each containing one visited node in the walk (except the starting node). Each tuple contains five elements to convey all the information needed to detect whether the walk intersects another walk in the specified node. As shown in Figure 1, there is a walk $\{u, v_1, v_2, v_3\}$ decomposed into three tuples $(j, 1, v_1, 3, u)$, $(j, 2, v_2, 3, u)$, $(j, 3, v_3, 3, u)$, where $j$ implies that the tuple is decomposed from the $j$-th walk of those starting from $u$. The remaining elements in the tuple are the *walk step*, *visited node at the particular walk step*, *the total steps in the walk*, and *source node*, respectively.

**Computing SimRank Values.** Evaluating the SimRank value between $u$ and $v$ is based on the probability that two walks from $u$ and $v$ meet. This can be detected by sorting the decomposed tuples (Line 9 of Algorithm 1→ Line 1 of Algorithm 2). Figure 1 shows the $j$-th walks from $u$ and $v$ respectively. The two walks meet because they share the same node $v_2$ in their third walk steps. Corresponding to $v_2$, there are two decomposed tuples from the two walks, denoted by $(j, 2, v_2, 3, u)$ and $(j, 2, v_2, 4, v)$. We note that they *share the first three elements*. In general, it is easy to verify that if two walks meet, there must be two decomposed tuples, each from one walk, that share the first three elements in the tuple. More formally, suppose there are two $j$-th walks $\{v_0, v_1, \ldots, v_{l_1}\}$ and $\{u_0, u_1, \ldots, u_{l_2}\}(0 \leq l_1, l_2 \leq \log_{1/\sqrt{c}} n)$ starting from $v_0$ and $u_0$. Walk $\{v_0, v_1, \ldots, v_{l_1}\}$ is decomposed into $l_1$ tuples, denoted by $(j, i, v_i, l_1, v_0)$ $(1 \leq i \leq l_1)$; Walk $\{u_0, u_1, \ldots, u_{l_2}\}$ is decomposed into $l_2$ tuples, denoted by $(j, k, u_k, l_2, u_0)$ $(1 \leq k \leq l_2)$. Then, if the two walks meet, or equivalently, they share the same node $v_i$ in the same step $i$, there must be $v_i = u_i$ and the corresponding two tuples (for $v_i$ and $u_i$) share the first three elements. Again, the challenges are in a proper rearrangement of the tuples across the machines, and the collection and aggregation of the walk-meeting cases within each machine.

With the above five stages, we highlight the major differences between our approach and the classic $\sqrt{c}$-walk-based algorithms: **(a)** the lengths of the walks are regularized to geometric distributions, and longer walks are truncated and we show that this would not violate the error tolerance in Section 5; **(b)** most steps are redesigned non-trivially so that all the steps can be efficiently implemented in the MPC model.

## 4 Detailed MPC Operations

### 4.1 Parallel Random Walks Generation

Performing random walks in the MPC model will incur communication rounds because the next sampled neighbor for each node can be located in different machines. Let us consider the situation of a node $v_0$ sampling its neighboring node

$v_1$ in the walk. In the MPC model, as $v_0$ and $v_1$ can be in different machines, forming a walk of $v_0 \rightarrow v_1$ may need one communication round. In general, forming an $l$-step random walk in the MPC model typically incurs $l$ communication rounds. For the purpose of reducing communication rounds, we first introduce a result provided by WRME [Lacki *et al.*, 2020], and then give our extension in Theorem 3:

**Theorem 2.** *Let $G$ be a directed graph. Let $N$ and $l$ be positive integers such that $l = o(S)/\log^3 n$, where $S$ is the available space per machine. For the task that samples $N$ independent random walks of length $l$ starting from each node $v$ in $G$, there exists an MPC algorithm that runs in $O\left(\log^2 \log n + \log^2 l\right)$ rounds and uses $O\left(m + n^{1+o(1)}l^{3.5} + Nnl^{2+o(1)}\right)$ total space and strongly sub-linear space per machine $S = n^\alpha$ $(0 < \alpha < 1)$. The algorithm is an imperfect sampler that does not fail with probability $1 - O\left(n^{-1}\right)$.*

Essentially, Theorem 2 states that the generation of a length-$l$ random walk from *every* node can be round-efficient in the MPC model. Particularly, it takes only $O\left(\log^2 \log n + \log^2 l\right)$ communication rounds to generate $N$ length-$l$ random walks, each with different source node. However, even with Theorem 2, we cannot directly give a reasonable round-complexity for SimRank evaluations if we use the original $\sqrt{c}$-decay walk-based method. The reason is that the length $l$ of a $\sqrt{c}$-decay walk can be infinite. To address this issue, as we show in Algorithm 1 Line 3 and Line 5, we carefully design the truncated walk length (Line 3) and the number of random walks to be sampled for each walk length (Line 5). The truncated length guarantees a reasonable bound of $O\left(\log^2 \log n\right)$ communication rounds when $l = \log_{1/\sqrt{c}} n$, and the number of random-walk samples ensures the SimRank estimation accuracy (we will formally prove it in Section 5).

**Parallel Random Walks in $O(\log^2 \log n)$ Rounds.** To effectively apply Theorem 2, we let Task-$i$ be the generation of $N_i$ length-$i$ walks from every node. We let all Task-$i$ $(1 \leq i \leq \log_{1/\sqrt{c}} n)$ be conducted in MPC in parallel. We apply $\log_{1/\sqrt{c}} n$ WRME algorithm [Lacki *et al.*, 2020] instances *concurrently* in the MPC model, where each instance corresponds to one task.

Unfortunately, concurrently conducting multiple WRME algorithms will incur a space cost higher than $n^\alpha$ in each machine because by default each algorithm instance can incur a local cost up to $n^\alpha$. To address this issue, we let $\beta$ satisfy that $n^\beta = n^\alpha / \log_{1/\sqrt{c}} n$, and we apply the WRME algorithm concurrently for each Task-$i$ with $S = n^\beta$. Particularly, each machine space $n^\alpha$ is divided into subspaces of size $n^\beta$ to compute Task-$i$. By careful analysis of the round complexity and total space cost, we give the following theorem.

**Theorem 3.** *The random walk generation in Algorithm 1 can be done in $O\left(\log^2 \log n\right)$ rounds with high probability, with machine space $S = n^\alpha$ such that $n^\alpha > \log^5 n/(\log \frac{1}{\sqrt{c}})$ and a total space $O\left(m \log n + n^{1+o(1)} \log^{4.5} n + \frac{n \log^{3+o(1)} n}{2(\epsilon - \frac{3}{n})^2}\right)$.*

We note that the WRME algorithm is an imperfect sampler that fails with at most a probability $O(n^{-1})$. We only apply $O(\log n)$ times the sampler, which can still easily guar-

antee that our algorithm is successful with high probability (because $O(\log n) = o(n^\tau)$ for any constant $\tau > 0$).

## 4.2 Shuffling Random Walks

We need to shuffle the generated $N$ random walks sourced at each node to fully mimic the behavior of generating a $\sqrt{c}$-walk. The reason for shuffling is that when we select two random walks that start from two nodes and calculate their meeting probability, the lengths of these two walks are not bound to be the same. This shuffling operation helps us guarantee the randomness in our Monte Carlo simulation of the generated $\sqrt{c}$-decay walks.

**Shuffling in $O(1)$ Rounds.** We note that sorting in MPC can be done in $O(1)$ rounds [Goodrich *et al.*, 2011] as long as the number of items is $O(n^\eta)$ for some constant $\eta > 0$. Observe that $N = O(\frac{\log n}{\epsilon^2})$, and hence there exists $\eta$ such that $nN = O(n^\eta)$. Hence, sorting the $nN$ walks all together can be finished in constant rounds in MPC, where the comparison of two walks is based on the comparison of the corresponding nodes in the walks. As such, all the walks sourced at the same node will be clustered.

For sufficiently large $n$ we have $n^\alpha > 2N \log_{1/\sqrt{c}} n$, and the walks sourced at the same node, occupying a space of $N \log_{1/\sqrt{c}} n$, can be held in one machine. Since all the walks are sorted and $\frac{n^\alpha}{N \log_{1/\sqrt{c}} n}$ may not be an integer, each set of $N$ walks sourced at the same node may go across two machines. Let $\boldsymbol{W}_u$ be the set of walks sourced at node $u$. We discuss two cases to shuffle the walks of the same source node.

**Case 1:** If the walks in $\boldsymbol{W}_u$ are fully located in a machine, shuffling is done locally. We shuffle all sets of walks, and for each set we assign numbers from 1 to $N$ to each walk in $\boldsymbol{W}_u$ after shuffling.

**Case 2:** If the walks sourced at the same node are located at two consecutive machines. Without loss of generality, we denote the source node as $u$, and there are $N_0$ walks in the first machine and $N - N_0$ in the second. We then swap those $N - N_0$ walks in the second machine with the last $N - N_0$ walks sourced at node $u - 1$, as illustrated in Figure 2 in the Appendix. Since $n^\alpha > 2N \log_{1/\sqrt{c}} n$ for sufficiently large $n$, it is guaranteed that those $N - N_0$ walks sourced at $u - 1$ are located in the first machine and have been shuffled locally in Case 1. Once swapped, the shuffling within $\boldsymbol{W}_u$ can be done locally. We note that by guaranteeing $n^\alpha > 2N \log_{1/\sqrt{c}} n$ we can make sure the swaps are not conflicting and thus can be done in one round.

Any set $\boldsymbol{W}_u$ can be shuffled in Case 1 or Case 2. During shuffling, each walk is labeled by 1 to $N$ sequentially, and the $j$-th walk refers to the walk assigned with a number $j$.

## 4.3 Decomposing Random Walks

**Decomposition in $O(1)$ Rounds.** Recall that each length-$t$ walk $\{v_0, v_1, \ldots, v_t\}$ is to be decomposed into $t$ tuples each of which has 5 elements: $(j, i, v_i, t, v_0)$, for $1 \le i \le t$. Tuple $(j, i, v_i, t, v_0)$ implies that in the $j$-th random walk of the walks starting from $v_0$, the $i$-th step of the walk visits node $v_i$ and the walk has overall $t$ steps. In the MPC model, the decomposition of a random walk can be done within the machine holding the walk. While each walk will generate mul-

tiple tuples, the total space is only amplified by a constant factor (i.e., 5 times) because each tuple corresponds to one visited node along the walk. Hence, this operation will not incur any communication among the machines, and only has a constant expansion of the space cost, which has a negligible effect due to Lemma 1. We denote the set of all the decomposed tuples by $D_G$.

## 4.4 Detecting Meeting-Walks

As the decomposed tuples are stored at different machines, challenges exist if we use these tuples to detect whether some pairs of random walks meet. Algorithm 2 shows the pseudo-code of detecting walk-meetings and computing SimRank values.

**Sorting in $O(1)$ Rounds.** Our first step is to sort all the tuples across the machines (Line 1). The sorting is based on the *multi-dimensional* sorting because the tuple contains five elements, i.e., sorting is first done based on the first element, and then for the elements that share the same first element, sorting is based on the second element, and so on. Elements are sorted based on their node IDs. We also intentionally let the node ID of source node $u$ be **smaller** than other nodes to guarantee that walks sharing the same first three elements start with the walk sourced at $u$.

Recall in Section 3 we show that if two walks meet, then there must be two decomposed tuples, each from one walk, sharing the first three elements. After sorting the tuples, the tuples sharing the first three elements will be clustered together and the tuple whose fifth element is $u$, if exists, will be placed in the first position in the cluster. Particularly, suppose we have tuples $(j, i, v, l_1, u)$, $(j, i, v, l_2, v_1)$, $(j, i, v, l_3, v_2)$ that share the first three elements, implying that the $j$-th walk started from $u$ meets the $j$-th walks started from $v_1$ and $v_2$. These three tuples will be clustered in the order of $(j, i, v, l_1, u)$, $(j, i, v, l_2, v_1)$, $(j, i, v, l_3, v_2)$ after sorting. The formal result is as follows.

**Lemma 3.** *If the $j$-th walk $W$ started from the query source node $u$ meets the $j$-th walk $W'$ started from $v$, then there must be a decomposed tuple $T$ of $W$ and a decomposed tuple $T'$ of $W'$ that share the first three elements. Furthermore, after sorting the tuples, $T$ is the closest tuple that is decomposed from $W$ and comes before $T'$.*

**Walk Paired-Up in $O(1)$ Rounds.** To compute the SimRank between $u$ and $v_1$, we need to pair up $(j, i, v, l_1, u)$ and $(j, i, v, l_2, v_1)$; similarly, we also need to pair up $(j, i, v, l_1, u)$ and $(j, i, v, l_3, v_2)$ to compute the SimRank between $u$ and $v_2$. Pairing-up can be challenging because the tuples are stored at different machines, and hence communication between machines is unavoidable. In the MPC model, a machine that communicates with other machines would require the same space as the size of communication messages. To reasonably bound the communication cost between the machines, we employ the following *PREDECESSOR* procedure [Behnezhad *et al.*, 2019] to couple the tuples sharing the same first three elements, which can be finished in $O(1)$ communication rounds.

**PREDECESSOR**: *Considered an ordered list of tuples such that each tuple is labeled by 0 or by 1. Then, for each tuple $T'$ labeled by 0, PREDECESSOR associates the closest tuple $T$ labeled by 1 such that $T$ comes before $T'$ in the or-*

*dering. PREDECESSOR can be implemented in $O(1)$ MPC rounds with $n^\alpha$ space per machine, for any constant $\alpha > 0$.*

To apply *PREDECESSOR* for pairing up the tuples, each machine can run a local algorithm to distinguish the tuples that are generated from the random walks starting at node $u$ and others. As shown in Figure 1, all machines assign label 1 to the tuples whose fifth elements are $u$ in parallel. Otherwise, the tuples are labeled by 0. By Lemma 3, we can apply the *PREDECESSOR* procedure to associate the closest Label-1 tuple (i.e., the tuple generated from the walk sourced at $u$) to each of its following Label-0 tuples. Here, the *association* means that the machine holding the Label-0 tuple is aware of its closest Label-1 tuple before it. For each such associated pair $T$ and $T'$, if their first three elements are the same, then the corresponding walks of $T'$ and $T$ meet.

### 4.5 Computing SimRanks

The paired-up tuples in the previous step will contribute to the corresponding SimRank value. For example, suppose two walks from $u$ and $v$ meet, then this pair of walks contribute to the SimRank value $s(u, v)$, because $s(u, v)$ is estimated by the meeting probability of the walks from node $u$ and node $v$. The challenge is how to concurrently distribute and aggregate these *SimRank contributions* across multiple machines. Note that it is crucial to let those *SimRank contributions* corresponding to the same pair of nodes (e.g., $(u, v)$) be processed in the same machine, to avoid repeat counting of walk-meets. For example, walks $\{u, u_1, u_2, w\}$ and $\{v, u_1, u_2, k\}$ meet at both node $u_1$ and node $u_2$. To ensure accuracy, the event of each walk-meet should be counted only once. For this purpose, we let each machine specifically handle SimRank evaluations of $\frac{n}{M}$ nodes with respect to $u$. Without loss of generality, we assume the node IDs are from 1 to $n$. Then, the $i$-th machine handles the node set with node IDs in $\left[\frac{n(i-1)}{M} + 1, \frac{ni}{M}\right]$.

**Computing SimRanks in $O(1)$ Rounds.** The detailed MPC operations of computing SimRanks are performed as follows (also illustrated in Algorithm 2 Lines 8 - 11). Using the PREDECESSOR procedure, for each tuple $(i, j, v_j, l_2, v)$ with label 0, it can be linked to the closest tuple $(i, k, v_k, l_1, u)$; if $j = k$ and $v_j = v_k$, then that indicates the meet of two walks represented by the two tuples. Then, a *walk-meet message* $(j, u, l_1, v, l_2)$ will be sent (by the machine holding tuple $(j, i, v_i, l_1, u)$) to the target machine that is responsible to computing SimRank $s(u, v)$. Here, the walk-meet message $(j, u, l_1, v, l_2)$ indicates that a length-$l_1$ walk sourced at node $u$ and a length-$l_2$ walk sourced at node $v$ meet. Also, assuming $u$ and $v$ are node IDs from 1 to $n$, the target machine that the message is sent to is the $\left\lfloor \frac{uM}{n} \right\rfloor$-th machine because this machine will compute $s(u, v)$. Each machine may then receive multiple copies of the walk-meet message $(j, u, l_1, v, l_2)$, and only one of them will be kept because each walk-meet event should be counted only once. For each unique message $(j, u, l_1, v, l_2)$ received, we add a value $\frac{|Z_{u,l_1}| \cdot |Z_{v,l_2}|}{|\hat{Z}_{u,l_1}| \cdot |\hat{Z}_{v,l_2}|} \cdot \frac{1}{z}$ to the SimRank value $\tilde{s}(u, v)$, where $z = \frac{\log 2n}{2(\epsilon - \frac{3}{n})^2}$ is the expected number of walk samples, $\hat{Z}_{u,l}$ denotes the set of actually generated length-$l$ random walks

sourced at $u$, and $Z_{u,l}$ is the expected set of length-$l$ random walks to be generated sourced at $u$.

We distinguish actual and expected samples because $Z_{u,l}$ may not be an integer. Particularly, we note that $|\hat{Z}_{u,l}| = \left\lceil z \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) \right\rceil$, and $|Z_{u,l}| = z \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c})$. When $z \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c})$ is not an integer, the expected walk set $Z_{u,l}$ is defined based on size-$\lceil |Z_{u,l}| \rceil$ set $Z_1$ and size-$\lfloor |Z_{u,l}| \rfloor$ set $Z_2$, such that there is probability $|Z_{u,l}| - \lfloor |Z_{u,l}| \rfloor$ that one walk of $Z_{u,l}$ is selected from set $Z_1$ and otherwise selected from set $Z_2$.

**Space Analysis for Walk-Meet Messages.** The walk-meet messages received by each machine are at most $\log_{1/\sqrt{c}} n$ times the number of the random walks started at the nodes handled by the machine. To see this, in the worst case, for each node of a walk $W$ generated from $v$ handled by the machine, there is a walk $W'$ from $u$ that meets $W$ at the node, creating a walk-meet message. Since each machine is responsible to compute SimRank values for $n/M$ nodes, the space per machine for receiving the message is $O\left(\frac{n}{M} \log_{\frac{1}{\sqrt{c}}} n \cdot \frac{\log 2n}{2(\epsilon - \frac{3}{n})^2}\right) = O\left(\frac{n \log^2 n}{M(\epsilon - \frac{3}{n})^2}\right)$. The total space cost for walk-meet messages is therefore $O\left(\frac{n \log^2 n}{M(\epsilon - \frac{3}{n})^2} \cdot M\right) = O\left(\frac{n \log^2 n}{(\epsilon - \frac{3}{n})^2}\right)$.

## 5 Summary of the Results

**Round Complexity**. The round complexity is dominated by the generation of parallel random walks, which is $O(\log^2 \log n)$ by Theorem 3.

**Space Complexity**. The total space cost for walk-meet messages is dominated by the random-walk generation space cost $O\left(m \log n + n^{1+o(1)} \log^{4.5} n + \frac{n \log^{3+o(1)} n}{2(\epsilon - \frac{3}{n})^2}\right)$ (see Theorem 3). We note that the total space cost is $\tilde{O}(m + n)$, indicating the existence of $M = \tilde{\Theta}(\frac{m+n}{S})$ for $S = n^\alpha$.

All these results give us Theorem 4 as well as Theorem 1.

**Theorem 4.** *Given a graph of $n$ nodes and $m$ edges, and a constant error $\epsilon > 0$, for sufficiently large $n$, Algorithm 1 can be run in MPC using $O\left(\log^2 \log n\right)$ communication rounds, with a total space per-round as $O\left(m \log n + n^{1+o(1)} \log^{4.5} n + \frac{n \log^{3+o(1)} n}{2(\epsilon - \frac{3}{n})^2}\right)$. The space per-machine needed is $S = n^\alpha$ such that $n^\alpha > \log^5 n / (\log \frac{1}{\sqrt{c}})$. Algorithm 1 is an imperfect sampler that fails with probability at most $O\left(n^{-1}\right)$.*

**Accuracy Analysis**. The accuracy guarantee of our random-walk sampling techniques resembles that of a simple Monte Carlo method, but we also incorporate the techniques of length-truncation and the rounding of the number of walks. We only include the main result as follows. The proof and the methodology for further reducing the threshold $\frac{3}{n}$ are detailed in the Appendix.

**Theorem 5.** *Algorithm 1 outputs SimRank values $\tilde{s}(u, v)$ with error at most $\epsilon$ ($\epsilon \geq \frac{3}{n}$) and with probability at least $1 - \frac{1}{n}$.*

## Acknowledgments

# References

[Antonellis *et al.*, 2008] Ioannis Antonellis, Hector Garcia-Molina, and Chi-Chao Chang. Simrank++: query rewriting through link analysis of the clickgraph (poster). In *WWW*, pages 1177–1178. ACM, 2008.

[Badola and Gupta, 2021] Kshitiz Badola and Mridul Gupta. Twitter spam detection using natural language processing by encoder decoder model. In *ICAIS*, pages 402–405. IEEE, 2021.

[Behnezhad *et al.*, 2019] Soheil Behnezhad, Laxman Dhulipala, Hossein Esfandiari, Jakub Lacki, and Vahab S. Mirrokni. Near-optimal massively parallel graph connectivity. In *FOCS*, pages 1615–1636. IEEE Computer Society, 2019.

[Cai *et al.*, 2008] Yuanzhe Cai, Pei Li, Hongyan Liu, Jun He, and Xiaoyong Du. S-simrank: Combining content and link information to cluster papers effectively and efficiently. In *ADMA*, volume 5139 of *Lecture Notes in Computer Science*, pages 317–329. Springer, 2008.

[Chen *et al.*, 2022] Rihan Chen, Bin Liu, Han Zhu, Yaoxuan Wang, Qi Li, Buting Ma, Qingbo Hua, Jun Jiang, Yunlong Xu, Hongbo Deng, et al. Approximate nearest neighbor search under neural similarity metric for large-scale recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 3013–3022, 2022.

[Chuanyan and Xiaoguang, 2021] Zhang Chuanyan and Hong Xiaoguang. Neural graph filtering for context-aware recommendation. In *Asian Conference on Machine Learning*, pages 969–984. PMLR, 2021.

[Czumaj *et al.*, 2021] Artur Czumaj, Peter Davies, and Merav Parter. Graph sparsification for derandomizing massively parallel computation with low space. *ACM Trans. Algorithms*, 17(2):16:1–16:27, 2021.

[Fogaras and Rácz, 2005] Dániel Fogaras and Balázs Rácz. Scaling link-based similarity search. In *Proceedings of the 14th international conference on World Wide Web*, pages 641–650, 2005.

[Ghaffari *et al.*, 2019] Mohsen Ghaffari, Fabian Kuhn, and Jara Uitto. Conditional hardness results for massively parallel computation from distributed lower bounds. In *FOCS*, pages 1650–1663. IEEE Computer Society, 2019.

[Goodrich *et al.*, 2011] Michael T. Goodrich, Nodari Sitchinava, and Qin Zhang. Sorting, searching, and simulation in the mapreduce framework. In *ISAAC*, volume 7074 of *Lecture Notes in Computer Science*, pages 374–383. Springer, 2011.

[He *et al.*, 2010] Guoming He, Haijun Feng, Cuiping Li, and Hong Chen. Parallel simrank computation on large graphs with iterative aggregation. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010*, pages 543–552. ACM, 2010.

[Hoeffding, 1994] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. In *The collected works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

[Hours *et al.*, 2016] Hadrien Hours, Eric Fleury, and Márton Karsai. Link prediction in the twitter mention network: impacts of local structure and similarity of interest. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 454–461. IEEE, 2016.

[Jeh and Widom, 2002] Glen Jeh and Jennifer Widom. Simrank: a measure of structural-context similarity. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 538–543, 2002.

[Jiang *et al.*, 2017] Minhao Jiang, Ada Wai-Chee Fu, and Raymond Chi-Wing Wong. Reads: a random walk approach for efficient and accurate dynamic simrank. *Proceedings of the VLDB Endowment*, 10(9):937–948, 2017.

[Jin *et al.*, 2011] Ruoming Jin, Victor E Lee, and Hui Hong. Axiomatic ranking of network role similarity. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 922–930, 2011.

[Karloff *et al.*, 2010] Howard Karloff, Siddharth Suri, and Sergei Vassilvitskii. A model of computation for mapreduce. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 938–948. SIAM, 2010.

[Klauck *et al.*, 2014] Hartmut Klauck, Danupon Nanongkai, Gopal Pandurangan, and Peter Robinson. Distributed computation of large-scale graph problems. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, pages 391–410. SIAM, 2014.

[Kusumoto *et al.*, 2014] Mitsuru Kusumoto, Takanori Maehara, and Ken-ichi Kawarabayashi. Scalable similarity search for simrank. In *SIGMOD Conference*, pages 325–336. ACM, 2014.

[Lacki *et al.*, 2020] Jakub Lacki, Slobodan Mitrovic, Krzysztof Onak, and Piotr Sankowski. Walking randomly, massively, and efficiently. In *STOC*, pages 364–377. ACM, 2020.

[Lee *et al.*, 2012] Pei Lee, Laks V. S. Lakshmanan, and Jeffrey Xu Yu. On top-k structural similarity search. In *IEEE 28th International Conference on Data Engineering (ICDE 2012), Washington, DC, USA (Arlington, Virginia), 1-5 April, 2012*, pages 774–785. IEEE Computer Society, 2012.

[Li *et al.*, 2015] Zhenguo Li, Yixiang Fang, Qin Liu, Jiefeng Cheng, Reynold Cheng, and John CS Lui. Walking in the cloud: Parallel simrank at scale. *Proceedings of the VLDB Endowment*, 9(1):24–35, 2015.

[Liao *et al.*, 2022] Ningyi Liao, Dingheng Mo, Siqiang Luo, Xiang Li, and Pengcheng Yin. Scara: scalable graph neural networks with feature-oriented optimization. *Proceedings of the VLDB Endowment*, 15(11):3240–3248, 2022.

[Luo and Zhu, 2024] Siqiang Luo and Zulun Zhu. https://sites.google.com/view/massive-technical-report/, 2024.

[Luo *et al.*, 2012] Siqiang Luo, Yifeng Luo, Shuigeng Zhou, Gao Cong, and Jihong Guan. Disks: a system for distributed spatial group keyword search on road networks. *Proceedings of the VLDB Endowment*, 5(12):1966–1969, 2012.

[Luo *et al.*, 2014] Siqiang Luo, Yifeng Luo, Shuigeng Zhou, Gao Cong, and Jihong Guan. Distributed spatial keyword querying on road networks. In *Proceedings of the 17th International Conference on Extending Database Technology, EDBT 2014, Athens, Greece, March 24-28, 2014*, pages 235–246. OpenProceedings.org, 2014.

[Luo *et al.*, 2019] Siqiang Luo, Xiaokui Xiao, Wenqing Lin, and Ben Kao. Efficient batch one-hop personalized pageranks. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pages 1562–1565, 2019.

[Luo *et al.*, 2022] Siqiang Luo, Xiaowei Wu, and Ben Kao. Distributed pagerank computation with improved round complexities. *Information Sciences*, 607:109–125, 2022.

[Luo *et al.*, 2023] Siqiang Luo, Zichen Zhu, Xiaokui Xiao, Yin Yang, Chunbo Li, and Ben Kao. Multi-task processing in vertex-centric graph systems: Evaluations and insights. In *Proceedings 26th International Conference on Extending Database Technology, EDBT 2023, Ioannina, Greece, March 28-31, 2023*, pages 247–259. OpenProceedings.org, 2023.

[Luo, 2019] Siqiang Luo. Distributed pagerank computation: An improved theoretical study. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019*, pages 4496–4503. AAAI Press, 2019.

[Luo, 2020] Siqiang Luo. Improved communication cost in distributed pagerank computation–a theoretical study. In *International Conference on Machine Learning*, pages 6459–6467. PMLR, 2020.

[Maehara *et al.*, 2014] Takanori Maehara, Mitsuru Kusumoto, and Ken-ichi Kawarabayashi. Efficient simrank computation via linearizationpublication of this article pending inquiry. In *KDD*, pages 1426–1435. ACM, 2014.

[Mo and Luo, 2021] Dingheng Mo and Siqiang Luo. Agenda: Robust personalized pageranks in evolving graphs. In *CIKM*, pages 1315–1324. ACM, 2021.

[Nguyen *et al.*, 2015] Phuong Nguyen, Paolo Tomeo, Tommaso Di Noia, and Eugenio Di Sciascio. An evaluation of simrank and personalized pagerank to build a recommender system for the web of data. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1477–1482, 2015.

[Shao *et al.*, 2015] Yingxia Shao, Bin Cui, Lei Chen, Mingming Liu, and Xing Xie. An efficient similarity search framework for simrank over large dynamic graphs. *Proceedings of the VLDB Endowment*, 8(8):838–849, 2015.

[Shi *et al.*, 2020] Jieming Shi, Tianyuan Jin, Renchi Yang, Xiaokui Xiao, and Yin Yang. Realtime index-free single source simrank processing on web-scale graphs. *arXiv preprint arXiv:2002.08082*, 2020.

[Song *et al.*, 2018] Junshuai Song, Xiongcai Luo, Jun Gao, Chang Zhou, Hu Wei, and Jeffrey Xu Yu. Uniwalk: Unidirectional random walk based scalable simrank computation over large graph. *IEEE Trans. Knowl. Data Eng.*, 30(5):992–1006, 2018.

[Symeonidis *et al.*, 2021] Panagiotis Symeonidis, Lidija Kirjackaja, and Markus Zanker. Session-based news recommendations using simrank on multi-modal graphs. *Expert Systems with Applications*, 180:115028, 2021.

[Tian and Xiao, 2016] Boyu Tian and Xiaokui Xiao. Sling: A near-optimal index structure for simrank. In *Proceedings of the 2016 International Conference on Management of Data*, pages 1859–1874, 2016.

[Wang *et al.*, 2020a] Hanzhi Wang, Zhewei Wei, Ye Yuan, Xiaoyong Du, and Ji-Rong Wen. Exact single-source simrank computation on large graphs. In *SIGMOD Conference*, pages 653–663. ACM, 2020.

[Wang *et al.*, 2020b] Yue Wang, Ruiqi Xu, Zonghao Feng, Yulin Che, Lei Chen, Qiong Luo, and Rui Mao. DISK: A distributed framework for single-source simrank with accuracy guarantee. *Proc. VLDB Endow.*, 14(3):351–363, 2020.

[Wang *et al.*, 2021] Hanzhi Wang, Zhewei Wei, Yu Liu, Ye Yuan, Xiaoyong Du, and Ji-Rong Wen. Exactsim: benchmarking single-source simrank algorithms with high-precision ground truths. *The VLDB Journal*, 30(6):989–1015, 2021.

[Xu *et al.*, 2017] Linchuan Xu, Xiaokai Wei, Jiannong Cao, and Philip S. Yu. On learning mixed community-specific similarity metrics for cold-start link prediction. In *WWW (Companion Volume)*, pages 861–862. ACM, 2017.

[Zhang *et al.*, 2017] Zhipeng Zhang, Yingxia Shao, Bin Cui, and Ce Zhang. An experimental evaluation of simrank-based similarity search algorithms. *Proceedings of the VLDB Endowment*, 10(5):601–612, 2017.

[Zheng *et al.*, 2013] Weiguo Zheng, Lei Zou, Yansong Feng, Lei Chen, and Dongyan Zhao. Efficient simrank-based similarity join over large graphs. *Proceedings of the VLDB Endowment*, 6(7):493–504, 2013.

[Zhou *et al.*, 2019] Kai Zhou, Tomasz P. Michalak, and Yevgeniy Vorobeychik. Adversarial robustness of similarity-based link prediction. In *ICDM*, pages 926–935. IEEE, 2019.

[Zhu *et al.*, 2022] Zulun Zhu, Jiaying Peng, Jintang Li, Liang Chen, Qi Yu, and Siqiang Luo. Spiking graph convolutional networks. In *IJCAI*, pages 2434–2440. ijcai.org, 2022.

[Zhu *et al.*, 2024] Zulun Zhu, Sibo Wang, Siqiang Luo, Dingheng Mo, Wenqing Lin, and Chunbo Li. Personalized pageranks over dynamic graphs–the case for optimizing quality of service. In *Proceedings of the 2024 IEEE 40th International Conference on Data Engineering*, 2024.

## A  Limitation and Societal Impact

**Limitation**. In this paper, we provide a detailed theoretical analysis based on a computational model and offer no experimental part. However, the structure of this paper is to follow the convention of MPC-related study, which only discusses the theoretical computational complexity. Furthermore, this kind of analysis might be the most attractive point in the MPC model, which acts as a general model and can be adapted in many cases.

**Societal Impact**. Our work offers a theoretical method to reduce communication rounds with limited machine space and no significant social negative impact can be caused to the best of our knowledge.

## B  Frequently used notations

| Notation | Description |
|----------|-------------|
| $G(V,E)$ | Directed graph $G$ with node set $V$ and edge set $E$ |
| $n, m$ | $n = |V|, m = |E|$ |
| $s(u,v)$ | SimRank score between node $u$ and $v$ |
| $\tilde{s}(u,v)$ | Estimated SimRank score between node $u$ and $v$ |
| $l$ | Length or steps of random walks |
| $\epsilon$ | Error threshold for SimRank scores |
| $c$ | Decay factor in the definition of SimRank scores |
| $M$ | Number of all machines |
| $S$ | Available space per machine |
| $\alpha$ | Constant factor to measure the space |
| $N_l$ | Actual number of Length-$l$ random walks |

Table 2: Frequently used notations.

## C  Missing Complexity Analysis

In Table 1, We explore the existing distributed or parallel SimRank algorithms and analyze them in terms of communication rounds and space per-machine. The analysis is based on synchronized round communication. The total space cost per round is the sum of the space occupied by each machine. The space per-machine refers to the worst case space cost in a machine. Due to unbalanced workloads, the worst-case per-machine space complexity can be the same as the total space complexity. We note that all these three algorithms aim to sample random walks in parallel for calculating the endpoint distribution (CloudWalker) or the meeting probability of two paths (UniWalk and DISK). Following existing analysis in UniWalk [Song *et al.*, 2018], we assume one random walk step in a distributed system costs one communication round.

### C.1  CloudWalker

CloudWalker [Li *et al.*, 2015] uses the following definition to compute SimRank:

$$\boldsymbol{S} = c\boldsymbol{P}^\top \boldsymbol{S}\boldsymbol{P} + \boldsymbol{D} = \sum_{t=0}^{\infty} c^t \boldsymbol{P}^{\top t} \boldsymbol{D}\boldsymbol{P}^t, \qquad (2)$$

where $\boldsymbol{S}$ is the SimRank matrix that $\boldsymbol{S}(i,j)$ is SimRank score between the $i$-th node and $j$-th node, $\boldsymbol{P}$ is the transition matrix of $G^T$ which is a transpose of the input graph $G$, and $\boldsymbol{D}$ is the diagonal correction matrix that CloudWalker aims to

compute first. The method of CloudWalker mainly includes the *calculation of $\boldsymbol{D}$* and *calculation of single-source Sim-Rank*.

**Calculation of $\boldsymbol{D}$.** CloudWalker relies on a method that combines the Jacobi method and Monte Carlo method to calculate the diagonal correction matrix $\boldsymbol{D}$ whose each row corresponds to the correction values of one node to the other nodes. Under the assumption that the input graph cannot be fit in a single machine ($S = n^\alpha$, $0 < \alpha < 1$), CloudWalker cannot process all nodes at a time and adopts the distributed structure to store the data for each node. Particularly, each machine will process $b$ nodes at a time and $b$ is assigned a node number to be processed. As a result, CloudWalker needs to perform its algorithms for $\frac{n}{b}$ times to process all nodes. We note that in the process of estimating $\boldsymbol{D}$, CloudWalker needs to sample $N = (\log n)/\epsilon_p^2$ length-$l$ random walks for each node, where $\epsilon_p$ is a parameter that controls the accuracy. When sampling $N$ random walks, as we assume machine space $S < n$, each machine may need communication with other machines to access the adjacent nodes which might be stored in other machines. For example, assuming that $u \to v$ is an edge in $E$ and $u$, $v$ are stored in two different machines, a walk-step $u \to v$ requires one communication round. Therefore sampling a length-$l$ random walk needs $O(l)$ communication rounds. After running $I$ iterations for $n/b$ times, the overall communication rounds of CloudWalker is $O(nIl/b)$.

Then in the Jacobi method, CloudWalker needs $O(blN)$ space to store the walk state matrix $A$ in every single machine. In addition, they need to store all the nodes and edges, incurring $O(m + n)$ costs. Putting together, the space cost per round is $O\big((bl\log n)/\epsilon_p^2 + n + m\big)$. Furthermore, each machine takes $O\big((bl\log n)/\epsilon_p^2\big)$ space to store $A$. Assuming the graph is equally distributed across the machines, then the per-machine space is $O\big((bl\log n)/\epsilon_p^2 + (m+n)/M\big)$.

**Calculation of single-source SimRank.** In Algorithm 4 of [Li *et al.*, 2015], CloudWalker conducts the computation of Single-source SimRank based on the obtained diagonal correction matrix $\boldsymbol{D}$. In Lines 4-13, CloudWalker again utilizes the random walks to compute the vectors corresponding to SimRank results, which needs to consume $O(l^2)$ communication rounds. The rationale of the communication round consumption is also that walk sampling requires to access other nodes which may be resided on multiple machines. Similar to the computation of $\boldsymbol{D}$, $b$ nodes will be processed at a time and the number of overall communication rounds is $O(\frac{nl^2}{b})$. For the reason that $O(bl^2N)$ space needs to be used for storing the computation vectors related to random walks, the total space complexity is $O\big((bl^2\log n)/\epsilon_p^2 + n + m\big)$.

To conclude, CloudWalker needs $O(nl(I + l)/b)$ communication rounds and $O\big((bl^2\log n)/\epsilon_p^2 + n + m\big)$ total space complexity, where each machine requires $O\big((bl^2\log n)/\epsilon_p^2 + (m+n)/M\big)$ space. Furthermore, CloudWalker needs to make a trade-off between the per-machine-space memory cost and the number of communication rounds. A special case is when $b$ equals $n$, CloudWalker can finish the SimRank within $O(Il + l^2)$ rounds where each machine requires $O(n\log n)$ space. Finally, CloudWalker only provides the error bound when estimating

$D$ and truncating the length $l$; they did not show any accuracy guarantees for computing SimRank scores [Wang *et al.*, 2020b].

## C.2 UniWalk

The application scope of UniWalk [Song *et al.*, 2018] is *undirected graph*. To compute the SimRank score $s(u,v)$, Uniwalk adopts the Monte Carlo method by sampling a certain number of Length-$L$ random walks and calculating the expected meeting distance of two bidirectional Length-$L$ random walks from two nodes $u$ and $v$. The main contribution of UniWalk is to convert sampling $O(nN)$ bidirectional Length-$L$ walks to only $O(N)$ Length-$2L$ unidirectional walks. A rectified factor is introduced to make the two kinds of walks equivalent. $L$ is empirically set in UniWalk, and UniWalk takes $l = 2L$ rounds to generate length-$2L$ unidirectional walks. By sampling $N = \frac{c^2(1-c^L)^2 d_{\max}^2 \log 2n}{2\epsilon^2(1-c)^2 d_{\min}^2}$ random walks, where $d_{\max}$ and $d_{\min}$ are the maximum and minimum degree in $G$ respectively. UniWalk can guarantee with high probability that the computed SimRank values have at most $\epsilon$ error. In the distributed version of UniWalk (Algorithm 3 in [Song *et al.*, 2018]), to guarantee $O(l)$ communication rounds, we need to apply Algorithm 3 $N$ times in parallel. Then in the worst case, Line 21 generates $N$ iterations and Line 23 needs to send a message of size $O(l)$. We note that $O(Nl)$ is already the worst-case total message space cost generated during the UniWalk process. Hence, both the per-machine space cost and total space cost are $O(Nl) = O(\frac{n^2 l \log n}{\epsilon^2})$ because $d_{max} = O(n)$ and $d_{min} \geq 1$.

## C.3 DISK

Similar to CloudWalker, DISK [Wang *et al.*, 2020b] also utilizes random walks to estimate the diagonal correction matrix $D$ and then performs the single-source SimRank computation. We also demonstrate the processes from the perspective of the **calculation of $D$** and **calculation of single-source SimRank**.

**Calculation of $D$.** The following analysis is based on the results in Section 4 of [Wang *et al.*, 2020b]. DISK explores the physical meaning of $D$ and conducts a more efficient algorithm to estimate the diagonal correction matrix $D$. Particularly, it transfers the calculation of $D$ into the meeting probability of two $\sqrt{c}$-walks starting from each node $v \in V$. The main advantage of this method is that the estimation accuracy of $D$ can be guaranteed. Then DISK build a tree-based method to sample $N = O(\frac{\log n}{\epsilon_d^2})$ trees when we set the failure probability $\delta_d = \frac{1}{n}$ for each node $v \in V$. The trees can also be interpreted as the $\sqrt{c}$-walks. Since the height of the trees is expected to be $O(\log n)$ and each level of one tree will require one communication round, the overall communication rounds after building $N$ trees can be $O(\frac{\log^2 n}{\epsilon_d^2})$. For the space cost, DISK needs to save the entire tree (random walk), which is $O(\log n)$, and the space per machine can be $O(\frac{m \log n}{M})$.

**Calculation of single-source SimRank.** The final single-source computation is based on the truncated formulation of
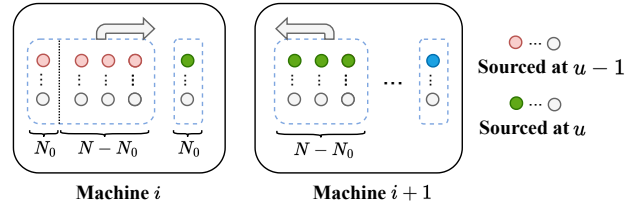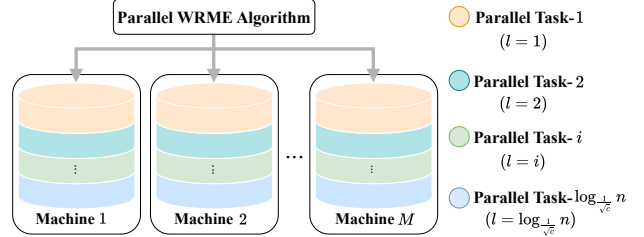


Figure 2: Swapping $N - N_0$ random walks.



Figure 3: Each machine divides its space to handle multiple tasks in parallel.

Equation 2 as:

$$\mathbf{S} \approx \sum_{t=0}^{K} c^t \boldsymbol{P}^{\top t} \boldsymbol{D} \boldsymbol{P}^t, \qquad (3)$$

where $K$ notes the truncated terms for approximation. In order to obtain the final SimRank score, $O(K)$ communication rounds will be consumed to calculate the result of Equation 3. For the reason that each node also needs to maintain $O(K)$ space for the above equation, the space for each machine will be $O(\frac{Kn}{M})$.

Hence we summarize that the number of overall communication rounds for DISK is $O\left(\frac{\log^2 n}{\epsilon_d^2} + K\right)$. Then in the space analysis, DISK needs total space $O(m \log n + Kn)$ and space $O(\frac{m \log n + Kn}{M})$ per machine to store the trees. The estimation error of DISK is controlled as smaller than $\frac{c(1-c^K)\epsilon_d}{1-c} + c^{K+1}$ with high probability, which corresponds to our statistic in Table 1.

## D Missing Proofs

### D.1 Proof of Theorem 3

*Proof.* For any instance $n$, let $\beta$ satisfy that $n^\beta = n^\alpha / \log_{\frac{1}{\sqrt{c}}} n$. If $n^\beta > (\log^3 n \cdot \log_{\frac{1}{\sqrt{c}}} n)$ (equivalently, $n^\alpha > \log^4 n \log_{\frac{1}{\sqrt{c}}} n = \frac{\log^5 n}{\log \frac{1}{\sqrt{c}}}$), then $\log_{\frac{1}{\sqrt{c}}} n = o(S)/\log^3 n$, and hence Theorem 2 is applicable for $S = n^\beta$. Applying Theorem 2 with $S = n^\beta$ still gives us $O(\log^2 \log n + \log^2 l)$ round complexity and $O(m + n^{1+o(1)} l^{3.5} + Nnl^{2+o(1)})$ space because the two complexities hides the $\alpha$ (or $\beta$) factor. Also, when applying Theorem 2 with $S = n^\beta$, the number of machines $M$ satisfies that $M \cdot n^\beta = \tilde{\Theta}(m + n)$ by the requirement of the MPC model. Then, $n^\alpha \cdot M = \tilde{\Theta}(m + n)$ also holds, which means the number of machines used is as desired for MPC when $= n^\alpha$. Now, for each of the $M$ machines each having space $n^\alpha$, we divide the machine

space into $\log_{\frac{1}{\sqrt{c}}} n$ subspaces, and for each subspace we apply WRME algorithms (Theorem 2) in parallel, respectively for $l = 1, 2, \ldots, \log_{\frac{1}{\sqrt{c}}} n$.

This indicates that each WRME algorithm runs with $S = n^\beta$. The worst-case round complexity corresponds to the WRME algorithm running with the longest length, which is $\log_{\frac{1}{\sqrt{c}}} n$. By Theorem 2, we have the round complexity

$$
\begin{aligned}
O\left(\lceil \log^2 \log n + \log^2 l \rceil\right) &= O\left(\left\lceil \log^2 \log n + \log^2 \log_{\frac{1}{\sqrt{c}}} n \right\rceil\right) \\
&= O\left(\log^2 \log n\right) \quad (4)
\end{aligned}
$$

The last step holds because $c$ is a constant. For each WRME instance, a machine space $n^\beta$ is sufficient. Considering all $\log_{\frac{1}{\sqrt{c}}} n$ WRME instances run in parallel, the space per machine is $n^\beta \cdot \log_{\frac{1}{\sqrt{c}}} n = n^\alpha$.

For the total space cost, we need to sum up all the space costs for each call of the WRME algorithm. Therefore, the space bound is calculated by applying Theorem 2 for $1 \leq l \leq \log_{\frac{1}{\sqrt{c}}} n$, and for each $l$, the space cost is expanded $\frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c})$ times (the number of random walks):

$$
\sum_{l=1}^{\log_{\frac{1}{\sqrt{c}}} n} O\left(m + n^{1+o(1)} l^{3.5} + N_l n l^{2+o(1)}\right)
$$

$$
= \sum_{l=1}^{\log_{\frac{1}{\sqrt{c}}} n} O\left(m + n^{1+o(1)} l^{3.5} + \frac{\log 2n}{2(\epsilon - \frac{3}{n})^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) n l^{2+o(1)}\right)
$$

$$
= O\left(m \log n + n^{1+o(1)} \log^{4.5} n + \frac{n \log^{3+o(1)} n}{2(\epsilon - \frac{3}{n})^2}\right) \quad (5)
$$

The last step holds because $\sum_{l=1}^{\log_{\frac{1}{\sqrt{c}}} n} (\sqrt{c})^l (1 - \sqrt{c}) < 1$. The space cost per machine is expanded at most $\log_{\frac{1}{\sqrt{c}}} n$ compared with the per-machine cost needed for Theorem 2, and thus still being strongly sub-linear to $n$. $\square$

### D.2 Proof of Lemma 1

*Proof.* Consider that there exists a distributed algorithm $\mathcal{A}$ that works on $M = \tilde{O}(\frac{m+n}{n^\alpha})$ machines with per-machine space $S = \Theta(n^\alpha)$ for $0 < \alpha < 1$ using $O(g(\alpha) \cdot f(n))$ communication rounds, where $g(\alpha)$ is a constant function that is computed based on $\alpha$. By the definition of big-$O$ notation, there must exist constants $C$, $C'$ and $n_0$, such that for any $n' > n_0$, the per-machine space $S$ used in $\mathcal{A}$ is at most $C \cdot n'^\alpha$, and the number of rounds used in $\mathcal{A}$ is at most $C' \cdot g(\alpha) \cdot f(n')$.

Let $\beta = \log_{n'}(C \cdot n'^\alpha)$, and we have $n'^\beta = C \cdot n'^\alpha$. Hence, Algorithm $\mathcal{A}$ is conducted on machines with space $n'^\beta$ using at most $C' \cdot g(\alpha) \cdot f(n')$ rounds. The number of machines $M = \tilde{O}(\frac{m+n'}{n'^\alpha}) = \tilde{O}(\frac{m+n'}{C \cdot n'^\alpha}) = \tilde{O}(\frac{m+n'}{n'^\beta})$. Hence, Algorithm $\mathcal{A}$ is a valid MPC algorithm for $S = n'^\beta$ when the number of graph nodes is $n'$, and the number of rounds used is a constant factor multiplied by $f(n')$.

We fix $\beta$ and let $M^*$ be the maximum number of machines used considering all $n' > n_0$, and it is easy to see $M = \tilde{O}(\frac{m+n'}{n'^\beta})$.

Meanwhile, note that $\beta = \alpha + \log_{n'}(C)$. When $n'$ is sufficiently large, $\beta$ can be arbitrarily close to $\alpha$. Therefore, for any $\beta \in (0, 1)$ we can always find $\alpha = \beta - \log_{n'}(C') > 0$ for sufficiently large $n'$, and hence, we can construct the corresponding MPC algorithm $\mathcal{A}$.

We can then safely replace symbol $\beta$ with $\alpha$ and the lemma holds. $\square$

### D.3 Proof of Lemma 3

*Proof.* Suppose Walk $W$ and Walk $W'$ meet at their $i$-th step at node $v_i$, then there is a decomposed tuple $T = (j, i, v_i, l_1, u)$ from $W$, and a decomposed tuple $T' = (j, i, v_i, l_2, v)$ from $W'$. They share the first three elements. Furthermore, we prove by contraction that there is a decomposed tuple $T^*$ originated from $W$ but $T^*$ is closer to $T'$ in the sorting order. Since the first three elements of $T$ and $T'$ are the same, then $T^*$ must also have the same first three elements. However, both $T$ and $T^*$ are decomposed from the walk $W$, but any two tuples decomposed from the same walk must have different elements in the first three positions of the tuple. Contradiction ensues. $\square$

### D.4 Proof of Theorem 4

*Proof.* The proof is immediate from Theorem 3 and all the analysis results from Section 4. $\square$

### D.5 Proof of Theorem 5

We aim to show that our algorithm can obtain the SimRank values with at most $\epsilon$ absolute error compared with the true value. Recall that we sample $N_l = \left\lceil \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) \right\rceil$ length-$l$ ($1 \leq l \leq \log_{\frac{1}{\sqrt{c}}}$) walks for each node. We will utilize two subsections to demonstrate how we bound the estimation error.

**Controlling Truncated Error**

Based on the Monte Carlo method, we will show that checking the meeting probability by sampling truncated $\sqrt{c}$-decay walks from $u$ and $v$ gives an $\epsilon$ ($\epsilon \geq \frac{3}{n}$) absolute error. The derivation is based on the Hoeffding Inequality as follows:

**Lemma 4.** *Let $X_1, \ldots, X_z$ be independent bounded random variables with $X_i \in [0, 1]$ for all $i$. Then*

$$
\mathbf{Pr}\left[\left|\frac{1}{z} \sum_{i=1}^{z} (X_i - E[X_i])\right| \geq \lambda\right] \leq 2e^{(-2z\lambda^2)}. \quad (6)
$$

We note that the probability that a pair of $\sqrt{c}$ walks from node $u$ and node $v$ meet is the SimRank value $s(u, v)$, and we let $X_i$ be the event that the $i$-th pair $\sqrt{c}$-walks from $u$ and $v$ meet. Hence, $\mathbb{E}[X_i] = s(u, v)$. By simple application of Hoeffding Inequality and the estimation $\tilde{s}(u, v) = \frac{\sum_{i=1}^{z} X_i}{z}$, we have

$$
\begin{aligned}
&\mathbf{Pr}\left[|\tilde{s}(u, v) - s(u, v)| \geq \lambda\right] \\
=&\mathbf{Pr}\left[\left|\frac{\sum_{i=1}^{z} X_i}{z} - \frac{\sum_{i=1}^{z} E(X_i)}{z}\right| \geq \lambda\right] \leq 2e^{-2z\lambda^2} \quad (7)
\end{aligned}
$$

Let $z = \frac{\log 2n}{2\lambda^2}$, we have $e^{-2z\lambda^2} = \frac{1}{n}$. However, our algorithm does not involve the $\sqrt{c}$-walks that are longer than $\log_{\frac{1}{\sqrt{c}}} n$. To quantify the effect of truncation, we note that sampling a $\sqrt{c}$-decay walk is equivalent to sampling a length-$l$-walk with probability $(\sqrt{c})^l (1 - \sqrt{c})$ for $l \geq 0$. By calculating the probability where a $\sqrt{c}$-walk is less than $\log_{\frac{1}{\sqrt{c}}} n$, we have:

$$\mathbf{Pr}\left[l \leq \log_{\frac{1}{\sqrt{c}}} n\right] = \sum_{l=0}^{\log_{\frac{1}{\sqrt{c}}} n} (\sqrt{c})^l \cdot (1 - \sqrt{c}) > 1 - \frac{1}{n} \quad (8)$$

The above equation demonstrates that the probability of a random walk longer than $\log_{\frac{1}{\sqrt{c}}} n$ is smaller than $\frac{1}{n}$. Then we set $N_l = \left\lceil \frac{\log 2n}{2\lambda^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) \right\rceil$ for length-$l$ ($0 \leq l \leq \log_{\frac{1}{\sqrt{c}}} n$) based on this probability result, and truncate the length at $\log_{\frac{1}{\sqrt{c}}} n$. Next we show that this simplification will impact less than $\frac{3}{n}$ absolute value of the SimRank value.

We denote $C_{u,l_1,v,l_2}$ as the number of meets between $u$'s length-$l_1$ walks and $v$'s length-$l_2$ walks, if we directly perform the random walks based on the expected numbers, i.e., $|Z_{u,l}|$. Let $\phi = \log_{1/\sqrt{c}} n$. We can transform Equation 7 as follows:

$$\mathbf{Pr}\left[|\tilde{s}(u,v) - s(u,v)| < \lambda\right]$$
$$= \mathbf{Pr}\left[-\lambda < \frac{\sum_{l_1=0}^{+\infty} \sum_{l_2=0}^{+\infty} C_{u,l_1,v,l_2}}{z} - s(u,v) < \lambda\right]$$
$$= \mathbf{Pr}\left[-z\lambda < \sum_{l_1=0}^{\phi} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2} + \sum_{l_1=0}^{\phi} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2} + \right.$$
$$\left. \sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2} + \sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2} - s(u,v) \cdot z < z\lambda\right]$$
$$\geq 1 - 2e^{-2z\lambda^2} \quad (9)$$

By Equation 8, there are fewer than $\frac{z}{n}$ random walks sourced at nodes $u$ (resp. $v$) whose lengths are longer than $\phi = \log_{1/\sqrt{c}} n$. Hence, there are at most $\frac{z}{n}$ meets of paired-walks when one walk is longer than $\phi = \log_{1/\sqrt{c}} n$. Formally,

$$\sum_{l_1=0}^{\phi} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2} \leq \frac{z}{n}$$

$$\sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2} \leq \frac{z}{n}$$

$$\sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2} \leq \frac{z}{n}$$

Then we have

$$0 < \frac{\sum_{l_1=1}^{\phi} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2}}{z} + \frac{\sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=1}^{\phi} C_{u,l_1,v,l_2}}{z}$$
$$+ \frac{\sum_{l_1=\phi+1}^{+\infty} \sum_{l_2=\phi+1}^{+\infty} C_{u,l_1,v,l_2}}{z} < \frac{3}{n}. \quad (10)$$

If we apply Equation 10 in Equation 9, we have:

$$\mathbf{Pr}\left[|\tilde{s}(u,v) - s(u,v)| < \lambda\right]$$
$$\leq \mathbf{Pr}\left[-\lambda - \frac{3}{n} < \frac{\sum_{l_1=0}^{\phi} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2}}{z} - s(u,v) < \lambda\right]$$
$$\leq \mathbf{Pr}\left[-\lambda - \frac{3}{n} < \frac{\sum_{l_1=0}^{\phi} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2}}{z} - s(u,v) < \lambda + \frac{3}{n}\right]$$
$$\quad (11)$$

For the reason that we only measure the meetings happening at walk-lengths less than $\phi = \log_{1/\sqrt{c}} n$, we set $\tilde{s}(u,v) = \frac{\sum_{l_1=0}^{\phi} \sum_{l_2=0}^{\phi} C_{u,l_1,v,l_2}}{z}$ as our estimation result. Lastly we have the following accuracy guarantee:

$$\mathbf{Pr}\left[|\tilde{s}(u,v) - s(u,v)| < \lambda + \frac{3}{n}\right] \geq 1 - 2e^{-2z\lambda^2} \quad (12)$$

We simplify this result by setting $\lambda = \epsilon - \frac{3}{n}$, and we have:

$$\mathbf{Pr}\left[|\tilde{s}(u,v) - s(u,v)| \geq \epsilon\right] \leq 2e^{-2z\left(\epsilon - \frac{3}{n}\right)^2}, \quad (13)$$

where $z = \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2}$ and $\epsilon \geq \frac{3}{n}$. Based on this result, we guarantee that by sampling $N_l = \left\lceil \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) \right\rceil$ length-$l$ ($0 \leq l \leq \log_{\frac{1}{\sqrt{c}}}$) walks for each node, we can obtain the SimRank values $\tilde{s}(u,v)$ with at most $\epsilon$ error ($\epsilon \geq \frac{3}{n}$).

**Rounding Up Error**
we consider that the actual number of random walks for each length $l$ is $\left\lceil \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c}) \right\rceil$ because $\frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^l \cdot (1 - \sqrt{c})$ may not be an integer. The rounding up of the samples for length-$l$ makes our SimRank estimation biased, which is denoted by a followed example. We let $\hat{Z}_{u,l_1}$ denote the set of actually generated length-$l_1$ random walks sourced at $u$, where $|\hat{Z}_{u,l_1}| = \left\lceil \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^{l_1} \cdot (1 - \sqrt{c}) \right\rceil$. And we denote the expected set of length-$l_1$ random walks as $Z_{u,l_1}$ with $|Z_{u,l_1}| = \frac{\log 2n}{2\left(\epsilon - \frac{3}{n}\right)^2} \cdot (\sqrt{c})^{l_1} \cdot (1 - \sqrt{c})$. When two random walks selected from $Z_{u,l_1}$ and $Z_{v,l_2}$ meet, rounding up $|Z_{u,l_1}|$ to $|\hat{Z}_{u,l_1}|$ and $|Z_{v,l_2}|$ to $|\hat{Z}_{v,l_2}|$ will change the distribution of walk length. As a result, the meeting probability of $u$ and $v$ will be affected and degrade our accuracy. We give a result as follows:

**Lemma 5.** *Each time when two random walks selected from $\hat{Z}_{u,l_1}$ and $\hat{Z}_{v,l_2}$ meet, this meeting will only contribute $\frac{|Z_{u,l_1}| \cdot |Z_{v,l_2}|}{|\hat{Z}_{u,l_1}| \cdot |\hat{Z}_{v,l_2}|} \cdot \frac{1}{z}$ score to the SimRank value of $\tilde{s}(u,v)$.*

*Proof.* When we estimate the SimRank $\tilde{s}(u,v)$ values based on the changed sample numbers $|\hat{Z}_{u,l_1}|$ and $|\hat{Z}_{v,l_2}|$, we still need to guarantee the meeting probability is the same as using the expected sample numbers $|Z_{u,l_1}|$ and $|Z_{v,l_2}|$:

$$\frac{\hat{C}_{u,l_1,v,l_2}}{|\hat{Z}_{u,l_1}| \cdot |\hat{Z}_{v,l_2}|} = \frac{C_{u,l_1,v,l_2}}{|Z_{u,l_1}| \cdot |Z_{v,l_2}|}. \quad (14)$$

Here $\hat{C}_{u,l_1,v,l_2}$ means the actual meeting numbers where $u$' length-$l_1$ walks meet with $v$' length-$l_2$ walks when we run the actual random walks. The difference between $\hat{C}_{u,l_1,v,l_2}$ and $C_{u,l_1,v,l_2}$ is that $C_{u,l_1,v,l_2}$ means the meeting number when running the theoretical random walks. Then we derive the SimRank value $\tilde{s}(u,v)$ according to:

$$
\begin{aligned}
\tilde{s}(u,v) &= \frac{\sum_{l_1=1}^{\log_{\frac{1}{\sqrt{c}}} n} \sum_{l_2=1}^{\log_{\frac{1}{\sqrt{c}}} n} C_{u,l_1,v,l_2}}{z} \\
&= \frac{\sum_{l_1=1}^{\log_{\frac{1}{\sqrt{c}}} n} \sum_{l_2=1}^{\log_{\frac{1}{\sqrt{c}}} n} \hat{C}_{u,l_1,v,l_2} \cdot \frac{|Z_{u,l_1}| \cdot |Z_{v,l_2}|}{|\hat{Z}_{u,l_1}| \cdot |\hat{Z}_{v,l_2}|}}{z}
\end{aligned}
$$
(15)

Here $\hat{C}_{u,l_1,v,l_2}$ is derived by counting the times where walks from the actual sample sets $\hat{Z}_{u,l_1}$ and $\hat{Z}_{v,l_2}$ meet. This result proves the claim.

$\square$

### D.6 Reducing Threshold $\frac{3}{n}$ and Failure Probability $\frac{1}{n}$

In Theorem 5, we note that we can further reduce the error threshold (e.g., $\frac{3}{n}$) and the failure probability (e.g. $\frac{1}{n}$) by increasing the length of walks and enlarging the number of samples, respectively.

**Reduce the error threshold.** Assuming we extend the maximum length to $p \cdot \log_{\frac{1}{\sqrt{c}}} n$, where $p$ is a constant and $p \geq 1$, then we can transform Equation 8 as following:

$$
\begin{aligned}
\mathbf{Pr}\left[l \leq p \cdot \log_{\frac{1}{\sqrt{c}}} n\right] &= \sum_{l=0}^{p \cdot \log_{\frac{1}{\sqrt{c}}} n} (\sqrt{c})^l \cdot (1 - \sqrt{c}) \\
&= 1 - \sqrt{c}^{p \cdot \log_{\frac{1}{\sqrt{c}}} n + 1} > 1 - \frac{1}{n^p}
\end{aligned}
$$
(16)

As a result, the probability of a random walk longer than $p \cdot \log_{\frac{1}{\sqrt{c}}} n$ is smaller than $\frac{1}{n^p}$. Moreover, the upper bound in Equation 10 can be $\frac{3}{n^p}$ and our error in Theorem 5 becomes $\epsilon$ ($\epsilon \geq \frac{3}{n^p}$, $p \geq 1$).

**Reduce the failure probability.** The failure probability is derived from Equation 13. We can simply reduce the failure probability by sampling more random walks. In particular, with the error threshold $\frac{3}{n^p}$, by setting $z = (\log 2n^q)/2\left(\epsilon - \frac{3}{n^p}\right)^2$, where $q$ is a constant and $q \geq 1$. We can transform Equation 13 as follows:

$$
\mathbf{Pr}\left[|\tilde{s}(u,v) - s(u,v)| \geq \epsilon\right] \leq 2e^{-2z(\epsilon - \frac{3}{n^p})^2} = \frac{1}{n^q} \quad (17)
$$

In summary, by setting the maximum length to $p \cdot \log_{\frac{1}{\sqrt{c}}} n$ and sampling $z = (\log 2n^q)/2\left(\epsilon - \frac{3}{n^p}\right)^2$, we guarantee $\epsilon$ ($\epsilon \geq \frac{3}{n^p}$) error with $1 - \frac{1}{n^q}$ probability, where $p, q \geq 1$. Clearly, this transformation only increase the number of random walks and walk length by a constant factor, which does not impact the round complexity and space complexity. For example, increasing the number of samples to

$(\log 2n^2)/2\left(\epsilon - \frac{3}{n^2}\right)^2$ and extending the maximum length to $2\log_{\frac{1}{\sqrt{c}}} n$, we can guarantee $\epsilon$ ($\epsilon \geq \frac{3}{n^2}$) error with $1 - \frac{1}{n^2}$ probability.

## E   Warm-Up: A $O(\log n)$-Round Algorithm with Pseudo-linear Space

In this section, we present an algorithm that directly adapts the $\sqrt{c}$-walk-based computation for single-source SimRank approximation in the MPC model. We show that by properly setting the parameters, the algorithm can finish the computation in $O(\log n)$ rounds with high probability (i.e., with a probability at least $1 - \frac{1}{n}$) while it requires pseudo-linear space-per-machine $S = \tilde{O}(n)$.

The main idea of the algorithm is to adapt the $\sqrt{c}$-walk based approach and apply the Monte Carlo method by setting the sample size $N = \frac{\log 2n^2}{2\epsilon^2} = O(\frac{\log n}{\epsilon^2})$, i.e., sampling $\sqrt{c}$-walks from the given source node $u$ and any other node $v$ over the reverse graph $\bar{G}$. By the Monte Carlo method, we need to create $N$ pairs of random walks from $u$ and $v$ and determine how many paired-up walks meet. The probability of the walk-meet, $\frac{H}{N}$, estimates $s(u,v)$. Consider a random variable $X_i = 1$ if the two walks in the $i$-th pair meet, otherwise $X_i = 0$. By Hoeffding's Inequality [Hoeffding, 1994] we have the following inequality:

$$
\begin{aligned}
\mathbf{Pr}\left[\left|s(u,v) - \frac{H}{N}\right| \geq \epsilon\right] &= \mathbf{Pr}\left[\left|\sum_{1 \leq i \leq N} \mathbb{E}(X_i) - \sum_{1 \leq i \leq N} X_i\right| \geq \epsilon\right] \\
&\leq 2e^{-2N\epsilon^2} \\
&= 1/n^2
\end{aligned}
$$
(18)

Since there are $n$ SimRank estimations, by union bound, all the estimated SimRank has an error at most $\epsilon$ with a probability of at least $1 - \frac{1}{n^2} \cdot n = 1 - \frac{1}{n}$.

Next, we show that sampling $N$ random walks respectively from each node can be performed in parallel, and concurrently walking one step for all walks can be finished in $O(1)$ rounds with machine space $S = \tilde{O}(n)$. Particularly, we assign each edge $(u,v)$ with an ID $u \cdot n + v$. In MPC we can use $O(1)$ communication rounds to sort [Goodrich *et al.*, 2011] the edges based on their IDs and store them from Machine 1 to Machine $M$ in order. To separate the neighbor set of consecutive nodes, we insert a special edge placeholder $(v,0)$ with ID $v \cdot n$ for every node $v$ and sort them together with the actual edges. After sorting, the placeholder $(v,0)$ will be placed right before the first edge starting with $v$. By sorting, each edge also knows its rank in the ordering. Then node $v$ has $r_{v+1} - r_v - 1$ neighbors, where $r_v$ (resp. $r_{v+1}$) denotes the rank of $(v,0)$ (resp. $v + 1$). Evaluating $\{r_{v+1} - r_v - 1 | v \in V\}$ can be done in 2 rounds: suppose each machine has $S = \tilde{O}(n)$, then all $\{r_v | v \in V\}$ can be sent to the same machine to compute $\{r_{v+1} - r_v - 1 | v \in V\}$ and send $r_{v+1} - r_v - 1$ back to the machine holding edge $(v,0)$. This will not break the $\tilde{O}(n)$ space as the size of the set $\{r_{v+1} - r_v - 1 | v \in V\}$ is $O(n)$. Furthermore, to store the random walks, when the walk path $P$ ending at node $u$ extends

from node $u$ to $v$ along an edge $(u, v)$, the path $P \cup (u, v)$ will be sent to the machine holding $(v, 0)$.

Since the random walks are $\sqrt{c}$-decay, there is at most probability $\epsilon^3/n^3$ that a random walk has a length at least $3 \log_{1/\sqrt{c}} \frac{n}{\epsilon}$. Hence all the $Nn$ random walks have length $O(\log \frac{n}{\epsilon})$ with high probability. Noting that it takes $O(l)$ rounds to compute length-$l$ random walks, the round complexity of computing all the random walks is $O(\log \frac{n}{\epsilon})$ with high probability. Furthermore, as there are $Nn = O(\frac{n \log n}{\epsilon^2})$ random walks and each walk has a length $O(\log \frac{n}{\epsilon})$ with high probability, the messages sent among the machines in a round is at most $O(\log n \cdot Nn) = O(\frac{n \log n \log \frac{n}{\epsilon}}{\epsilon^2}) = \tilde{O}(n)$ with high probability (given $\epsilon$ is a constant).

Finally, to compute $\{s(u, v) | v \in V\}$ it remains to check in parallel how many pairs of walk meet, where each pair contains one walk from $u$ and the other walk from $v$. As we will explain shortly in Section 4.1, this step can also be finished in $O(1)$ rounds.

In summary, the algorithm we presented can finish approximating SimRank computation in $O(\log \frac{n}{\epsilon})$ rounds with high probability over MPC with $S = O(\frac{n \log n \log \frac{n}{\epsilon}}{\epsilon^2})$. One can see that directly adapting the existing random-walk-based algorithms is hard to break the $O(\log n)$-round barrier because in such designs each random walk step costs $O(1)$ rounds and a random walk has $O(\log n)$ steps with high probability. Meanwhile, it is also challenging to further reduce the space-per-machine because there can be hub nodes that are passed through by many random walks sourced at different nodes.