Scaling Frustration Index and Corresponding Balanced State Discovery for Real Signed Graphs

Muhieddine Shebaro
m.shebaro@txstate.edu
Texas State University
San Marcos, Texas, USA

Jelena Tešić
jtesic@txstate.edu
Texas State University
San Marcos, Texas, USA

ABSTRACT

Structural balance modeling for signed graph networks presents how to model the sources of conflicts. The state-of-the-art has focused on computing the frustration index of a signed graph as a critical step toward solving problems in social and sensor networks and for scientific modeling. However, the proposed approaches do not scale to modern large sparse signed networks. Also, they do not address that there is more than one way in some networks to reach a consensus with the minimum number of edge-sign switches needed. We propose an efficient balanced state discovery algorithm and a network frustration computation that will discover the nearest balanced state for the *any* size of the graph network and compute the frustration of the network. The speedup of the proposed method is around 300 times faster than the state-of-the-art for signed graphs with hundreds of thousands of edges. The technique successfully scales to find the balanced states and frustration of the networks with millions of nodes and edges in real time where state-of-the-art fails.

PVLDB Reference Format:
Muhieddine Shebaro and Jelena Tešić. Frustration Index for Real Graphs. PVLDB, 17(1): XXX-XXX, 2023.
doi:XX.XX/XXX.XX

PVLDB Artifact Availability: The source code, data, and/or other artifacts are available at [6].

1 INTRODUCTION

Unstructured data requires a rich graph representation. The current state of the art for unsigned homogeneous graphs can handle trillions of edges and billions of nodes [42], while signed graph benchmarks are limited to thousands of nodes and edges. The signed networks model complex relationships through the interdependence between entities complementary to instance features. When opposing edges are included in a network, we can study social dynamics and stability concerning friendship and enmity in more depth [7, 37], or expand to new application domains, such as brain behavior [47]. However, signed graph benchmarks are currently too small and too similar in topology to real data, hindering progress in signed graph analysis [35, 38]. Singed graph extensions have been demonstrated for narrow-band tasks in finance [9], polypharmacy [40], bioinformatics [39] and sensor data analysis [15, 41]. To date, signed graph benchmark evaluations of the algorithms are small in size, do not have the same topology as signed graphs derived from accurate data, and the algorithms make assumptions that are not applicable in real signed networks [17, 51].

This paper focuses on scaling the discovery of balanced states and frustration index for the large signed networks. Balance theory represents a theory of changes in attitudes [3]: people's attitudes evolve in networks so that friends of a friend will likely become friends, and so will enemies of an enemy [3]. Heider established the foundation for social balance theory [28], and Harary established the mathematical foundation for signed graphs and introduced the k-way balance[14, 25]. Balance theory concepts have been used to predict edge sentiment, to recommend content and products, or to identify unusual trends [5, 19, 22, 32]. A network is considered balanced only if every fundamental cycle contains an even number of opposing edges. One measure of the network is its distance from its balanced state. A signed network's *frustration index* is the smallest number of edge-sign switches needed to reach a balanced state. The frustration index is one measure of network property in many scientific disciplines, that is, in chemistry [50], biology [31], brain studies [45], physical chemistry [53] and control [21]. The calculation of the frustration index can also be reduced to the maximum cut of the graph in a particular case of all opposing edges, which proved to be NP-hard [30]. State-of-the-art methods address the computation of the frustration index for signed graphs with up to 100,000 vertices [11], and the approach does not scale to modern large-scale sparse networks. The signed networks can have multiple nearest balanced states, e.g. multiple paths to a balanced state with the corresponding minimal or near-minimal number of edge switches to reach more than one nearest balanced state [44], and there is a way to discover the nearest balanced states in the

The paper contributions are:

- A novel algorithm based on discovering a minimal fundamental cycle basis to solve the discovery of the frustration index for any real-world signed network of any size or density.
- Extending the frustration cloud from a set to a *(key,value)* tuple collection $\mathcal{F}_\Sigma = \mathcal{B}{:}(\mathcal{C}, \mathcal{S})$ [44]. The nearest balanced states with their associated frequency and edge switches will be stored in the memory-bound frustrated cloud in the form of a tuple.
- Benchmark implementation and comparison of seven tree-sampling methods including a novel one to study the impact on frustration index computation.
- Novel scaling of frustration index computation to *any* size signed graphs: there is no adjacency matrix computation which allows us to evade the eigenvalue pollution.

Proceedings of the VLDB Endowment, Vol. 17, No. 1 ISSN 2150-8097.
doi:XX.XX/XXX.XX

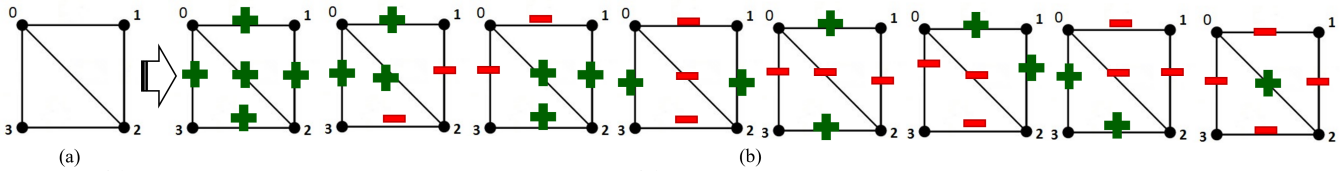

Figure 1: (a) Unsigned graph G with 4 vertices and 5 edges; (b) Eight possible edge sign combinations for G that are balanced.

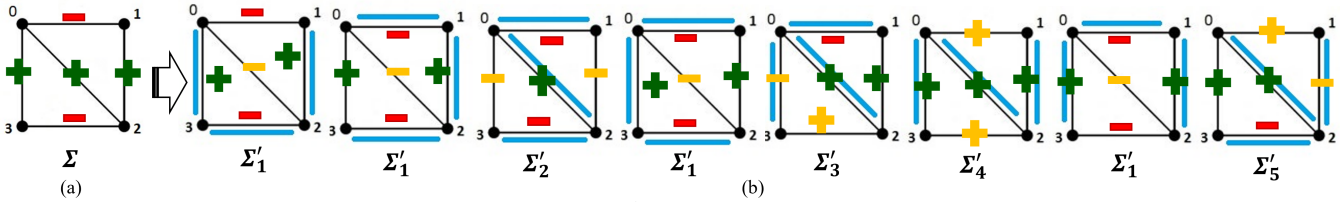

Figure 2: (a) Signed graph Σ; (b) Near-balanced states of Σ, $\Sigma'_i : i \in [1, 5]$ where blue lines illustrate the spanning tree and yellow signs note the edge sign change in Algorithm 1.

This paper introduces an algorithm to discover the nearest balanced state for signed graph of *any* size under real-time processing constraints. The outcome is the set of edges that need to change signs for the network to achieve a balanced state. The number of edges in the output set captures the frustration of the network. The paper is organized as follows: in Section 2, we summarize related work in the field; in Section 3, we present preliminary findings; in Section 4, we introduce a novel algorithm for computing the frustration index and its scalable version; in Section 5, we analyze different sampling approaches of the spanning tree and propose two new methods; in Section 7, we conduct experiments to evaluate the efficiency of the proposed algorithm on ten different signed graph benchmarks and apply and measure the findings on signed graphs derived from the 26 Amazon ratings and reviews datasets [26], and in Section 7.6 we summarize our findings.

2 RELATED WORK

Frustration Applications: In chemistry, the stability of fullerenes is related to the frustration index [50]. The frustration index has been used to measure how an incoherent system responds to perturbations in large-scale signed biological networks [31]. The frustration of the network has been determining the strength of agent commitment to make a decision and win the disorder in adversarial multi-agent networks [21]. The frustration in neuroplasticity assesses the development of brain networks, as studies have shown that a person's cognitive performance and the frustration of the brain network have a negative correlation [45]. In physical chemistry, the protein-protein interaction can be predicted using the frustration index of the protein-signed network [53]. Saberi et al. investigated the pattern for the formation of frustrating connections in different brain regions during multiple life stages[46].

Computing the Frustration Index: Researchers have focused on calculating the exact frustration index, not finding the balanced state of the network. Calculating the frustration index is an NP-hard problem equivalent to calculating the ground state of the spin glass model on unstructured graphs [49]. Greedy algorithms have also been utilized to calculate the frustration index for networks as large as 1,000 and 10,000 nodes [**?**]. The frustration index for small fullerene graphs can be calculated in polynomial time[20], and the finding has been extended to estimate the genetic algorithm of the frustration index in [50]. Bansal et al. introduced the correlation clustering problem, which is a problem in computing the minimum number of frustrated edges for several subsets [12]. Aref et al. provided an exact algorithm to calculate the partial balance and frustration index with $O((2^b)|E|^2)$ complexity where b is a fixed parameter and $|E|$ is the number of edges [9]. Researchers have focused on approximating the frustration computation. Aref et. al introduced the approximate algorithm of complexity $O(\sqrt{log(|V|)})$, $|V|$ is the number of vertices that theoretically converged to the exact frustration index when tested on networks up to 2,000 edges [10]. Recent improvements in the algorithm include binary programming models and the use of multiple powerful mathematical solvers by Guribo [24], and the algorithm can handle up to $|E| = 100,000$ edges and compute the frustration index of the network in 10 hours [11]. For a million node Amazon ratings and reviews network, these approaches are still too complex.

3 PRELIMINARY DATA

This Section summarizes existing findings and approaches: the frustration cloud analysis of a signed graph in [44] and the efficient data structure and algorithm to efficiently compute fundamental cycles in [4]. *Graph* $G = (V, E)$ is a set of vertices V connected by the set of edges E. The number of vertices in the graph G is $|V|$, and the number of edges is $|E|$. *Path* is a sequence of distinct edges that connect a sequence of distinct vertices. *Cycle* is a path that begins and ends at the same vertex. *Connected graph* is a graph in which a path joins two vertices. *Subgraph* is a graph whose all edges and vertices are contained in a larger graph, for example, Path and Cycle. *Signed graph* Σ is a tuple of a graph $G = (V, E)$ and an edge signing function $\sigma : E \rightarrow \{+1, -1\}$. The edge can be positive + or negative $-$, $e \in [e^+, e^-]$, E^+, and E^- are sets of positive and negative edges of G. *Sign* of a subgraph is *product* of the edges signs. *Balanced signed graph* is a signed graph where every cycle is positive. *Harary bipartition* separates the vertices of the balanced graph into two sets such that the vertices of both sets internally agree with each other but disagree with the vertices of the other set[14]. *Near-balanced*

graph Σ' is a balanced graph that requires a minimum number of edge sign switches to produce a balanced graph from the signed graph Σ. *Frustration cloud* $\mathcal{F}_\Sigma$ is a set of all signed near-balanced graphs of Σ: $\mathcal{F}_\Sigma = \left\{\Sigma'_i | i \in [1, N] \wedge Fr(\Sigma'_i) = 0\right\}$.

Algorithm 1: Tree-Based Signed Graph Balancing

```
Data: Signed graph Σ
Data: Spanning tree T of Σ
Result: Balanced graph Σ'_T
1 forall edges e, e ∈ Σ \ T do
2     if fundamental cycle T ∪ e is negative then
3         switch edge sign in Σ'_T: e⁻ → e⁺; e⁺ → e⁻ ;
4 end
```

The unsigned graph G with four vertices and five edges is illustrated in Figure 1(a), and all possible combinations of edge signs that result in balanced signed graphs are illustrated in Figure 1(b). The frustration cloud $\mathcal{F}_\Sigma$ for a signed graph Σ is formed using the tree balance algorithm outlined in Alg. 1 and explained in detail in [44]. Balance only affects the signs of the edges as it negates some of them to reach a consensus state (line 3 in Alg.1). The minimum number of edge sig switches is the frustration cloud. Al illustrated in Figure 2(b), the minimal number of edge switches for Σ to reach the balanced state *Sigma'* is 1. Thus, the frustration index of the graph is 1. To discover each of these states, Alg. 1 should be run with *all* spanning trees T of Σ. Σ in Figure 2(a) has eight spanning trees, illustrated in Figure 2(b) with blue lines. The frustration cloud $\mathcal{F}_\Sigma$ encompasses all the individual closest balanced states that result from Alg. 1. Only five of these balanced graphs are nearest balanced states for a specifically signed graph Σ illustrated in Fig. 2(a), as illustrated by the frustration cloud $\mathcal{F}_\Sigma = \left\{\Sigma'_i | i \in [1, 5]\right\}$ in Fig. 2(b).

The number of spanning trees for Σ is eight, and the exact number for any graph G can be calculated in polynomial time as the determinant of a matrix derived from the chart, using Kirchhoff's matrix-tree theorem [52]. For a small social network such as the Highland Tribes [43] with 16 vertices and 29 positive and 29 opposing edges, the number of spanning trees is 402, 506, 278, 163 [44]. The *graphB* algorithm was introduced in [44] as a scaled adjustment of the Alg. 1 **for** loop in line 1: instead of all trees, the algorithm loops over k spanning trees discovered using the method M (Alg. 2 line 1). Next, the *graphB+* algorithm scaled the computation of fundamental cycles for the spanning tree T in Alg. 1. If T is a spanning tree of Σ and e is an edge of Σ that does not belong to T, then *fundamental Cycle* C_e, defined by e, is the cycle consisting of e together with the straightforward Path in T connecting the endpoints of e. If $|V|, v \in V$ denotes the number of edges and $|E|, e \in E$ the number of vertices in Σ, there are precise $|E| - (|V| - 1)$ fundamental cycles, one for each edge that does not belong to T. Each C_e is linearly independent of the remaining cycles because it includes an edge e that is not present in any other fundamental process. The *graphB+* algorithm was introduced in [4] as an efficient way to balance the signed graph given the spanning tree T. The main benefits of *graphB+* are that labels can be computed with linear time complexity, only require a linear amount of memory, and that the running time for balancing a cycle is linear in the length of the cycle times the vertex degrees but *independent* of the size of the graph. The algorithm speeds up the balancing to more than 14 million fundamental cycles identified, traversed, and balanced. This paper proposes the *graphB++* algorithm to scale the frustration cloud computation to compute the frustration index for large real signed graphs.

4 SCALABLE COMPUTATION OF THE FRUSTRATION INDEX

In this Section, we propose *graphB++*, an efficient algorithm that computes balanced states and frustration index. The *graphB++* algorithm builds on the *graphB* and *graphB+* baseline. First, the algorithm integrates different tree-sampling approaches, as outlined in Alg. 2) for frustration index computation. Next, the *graphB++* algorithm scales the calculation of the frustration index and associated optimal balanced state by iteratively keeping in memory only the subset of nearest balanced states with the smallest number of edge negations, as outlined in Alg. 3. The result of *graphB++* is that the frustration index and the closest balanced associated state can be computed in minutes for the *any* large signed graph with millions of vertices and edges.

Algorithm 2: Graph Balancing and Frustration Index

```
Data: Signed graph Σ and spanning trees sampling
      method M
Result: frustration index Fr(Σ)=min_i(S) and frustration
        cloud F_Σ = B:(C, S).
1  Generate set T_{M^k} of k spanning trees of Σ using M;
2  Empty F_Σ; ;
3  foreach spanning trees T, T ∈ T_k do
4      Find nearest balanced state Σ'_T using Alg. 1;
5      s = edge signs difference count from Σ to Σ'_T;
6      Transform Σ'_T balanced state to string B ;
7      if B ∉ B then
8          add key B to B;
9          S(B) = s;
10         C(B) = 1 ;
11     else
12         C(B)++ ;
13 end
```

First, we extend the definition of frustration cloud $\mathcal{F}_\Sigma$ (Sect. 3) from a set to a *(key,value)* tuple collection $\mathcal{F}_\Sigma = \mathcal{B}:(\mathcal{C}, \mathcal{S})$. The key is the unique balanced state $\mathcal{B}(i)$, and the value is the count of balanced states occurring in iteration $\mathcal{C}(i)$, and the edge count switches to the balanced state $\mathcal{S}(i)$. In each balancing iteration, we examine the resulting balance state (Alg. 2). Σ'_T in relation to $\mathcal{B}$. To make the process more efficient, we represent the balanced state Σ'_T as a string B. The balanced state Σ'_T is represented as the 3 edges vectors (src,tgt,sign). If an edge i is defined by two vertices (u, v) and a sign s, the algorithm *graphB+* balances the graph by storing the edges as src(i)=u, tgt(i)=v, sign(i)=s [4].

For *graphB++* implementation, we propose an efficient transform (O(|E|)) of the balanced state output Σ' to the string hash key B for comparison with other balanced states (Alg. 2 line 5). First, the

triple edge vector (src(iThis),sign(i)) is inserted into a set of tuple data structures to organize the edges and prepare for string conversion automatically. Then, it is transformed to a string format "src(i)->tgt(i): sign(i)," and then all edge strings are concatenated in order, separated by the delimiter "|" and stored as the B key in $\mathcal{B}$. If B is in $\mathcal{B}$, we increase the corresponding $C(B)$ value count, where B is the existing balanced state Σ'_T. If Σ'_T is not in $\mathcal{B}$, we add (Σ'_T,(1,number of switched edge signs) pair to the collection. If the state was previously unseen, we add the new balanced state inefficient matrix format to the hashmap as a string key as illustrated in Alg. 2. Then, we add 1 to the end of the count stack C and add the number of edge switches in the graph for this balanced state to the frustration cloud frequency stack $\mathcal{S}$. These two values (count stack and frequency stack) are stored as a pair, and the value of the hashmap of the balanced state string as a key is that pair. If the balanced state exists in $\mathcal{B}$, we increase the count at the same string key in C only (the first element of the pair is modified), as illustrated in the Algorithm 2. Now, Alg. 2 line 5 takes $O(E)$ for comparison. And the minimum number of edge switches in all balanced states defines the frustration index, and we compute the frustration index as Fr(Σ)=min($\mathcal{S}$). As the number of iterations increases, more elements will be added to $\mathcal{B}$, and the space complexity is linear space complexity. This can become an issue for graphs with millions of nodes and vertices as the frustration cloud is too big for the main memory.

Algorithm 3: Graph Balancing and Frustration Index

Data: Σ signed graph; M tree sampling method
Result: frustration index $Fr_\Sigma = min_i(\mathcal{S}(i))$ and frustration cloud $\mathcal{F}_\Sigma$ =($\mathcal{B}(i), C(i), \mathcal{S}(i)$), $i \leq \mathcal{F}_{max}$

1 ;
2 Generate set $\mathcal{T}_k$ of k spanning trees of Σ ;
3 Determine $\mathcal{F}_{max}$, $|\mathcal{F}_\Sigma| < CAP$;
4 Empty matrix $\mathcal{B}$, empty count vector C, and empty edge switch count vector $\mathcal{S}$, $i = 0$, frInd=0. ;
5 **foreach** *T spanning tree, $T \in \mathcal{T}_k$* **do**
6 Find nearest balanced state Σ'_T using Alg. 1 ;
7 frInd = # Σ'_T switch edge signs (Alg. 1, line 3) ;
8 **if** $frInd < max_i(\mathcal{S}(i))$ **then**
9 **if** $\Sigma_T \notin \mathcal{B}$ **then**
10 **if** $i < \mathcal{F}_{max}$ **then**
11 i++ ;
12 **else**
13 i takes the index of current $max_i(\mathcal{S}(i))$;
14 $\mathcal{S}(i)$ = frInd ;
15 $\mathcal{B}(i) = \Sigma_T$;
16 $C(i) = 1$;
17 **else**
18 C_i ++ ;
19 **end**

Thus, for large graphs, we adapt Alg. 2 as outlined in Alg. 3 to scale the computation. First, we compute the number of balanced states that we can keep in the memory as $\mathcal{F}_{max}$ (Alg. 3, line 2). Next, we keep only the $\mathcal{F}_{max}$ closest balanced states in $\mathcal{B}$, their count through all k iterations in C, and the number of switched edges for each of them in $\mathcal{S}(i)$. Note that there can be different balanced states of Σ with the same number of switched edges. Finally, we compute the frustration index as a minimum of $\mathcal{S}$). The proposed approach scales well with the size of the signed graph, as we limit the number of the nearest balanced states that we keep in the memory based on their frustration index and capacity of the frustration cloud hashmap in-memory storage, as illustrated in Algorithm 3. The comparison of balanced states now has up to k iterations times $\mathcal{F}_{max}$ closest balanced states. We determine $\mathcal{F}_{max}$ so that the size of the frustration cloud in memory is smaller than CAP. In experiments, we define CAP as 75% of the total RAM size under the assumption that some vital system processes are running in the background. The algorithmic complexity remains $O(|E|log(|V|)d_a)$ where d_a is the average degree of a vertex, $|V|$ is the number of vertices, and $|E|$ is the number of edges [4].

5 SAMPLING SPANNING TREES

In thisSection, we propose to utilize the randomization and hybridization of the standard tree sampling approaches to maximize the chances of discovering the optimal nearest balanced state in Alg. 2. The following techniques have been dominantly used in maze generation. **Depth first search (DFS)** algorithm [16] begins the traverse at the root node (0, in Fig. 3(1) and proceeds through the nodes as far as possible until it reaches the node with all the nearby nodes visited. In Fig. 3(3), node two is set as a root. The same spanning trees are generated, and the reiteration process fails to capture all the unique spanning trees from Figure 2. This effect seems to apply to DFS and those observed in iterations 2 and 6, with the exact edges of 0-1, 0-3, and 1-2. **Breadth first search (BFS)** algorithm [16] is a graph traversal approach in which the algorithm first passes through all nodes on the same level before moving on to the next level. BFS algorithm starts at a source node 0 in Fig. 3(1), constructs a spanning tree having 0-1,0-2 and 0-3 as edges. Re selecting node 0 as the root in Fig. 3(5) repeats the same process. This nullifies the point of reiterating through the nodes, since it does not diversify the unique spanning trees. The main drawback of efficient implementation of the efficient vertex search is the *static* order in accessing vertices. As a result, the efficient spanning tree algorithm repeats the exact tree sampling and misses several unique trees in Fig. 3.

We propose to use the randomized algorithms as follows: in each iteration, we shuffle and randomize a node's neighborhood using a uniformly distributed random seed number before applying a static algorithm. Figure 4 illustrates the idea that a node establishes a link to the first unvisited node based on the randomized order of the adjacency list in the network. The randomized breadth first search example in Figure 4 produces the *exact* result as the breadth first search. In the first iteration, regardless of the order of neighboring nodes in the adjacency lists, the edges 0-1, 0-2, and 0-3 will always be selected. Re-shuffling neighboring unvisited nodes on the adjacency lists of the RBFS sampler does not matter since all nodes will be traversed regardless, as illustrated in Figure 4. Randomized DFS ensures that even when the same node is selected as a root, there is a high chance that it will produce a completely different spanning

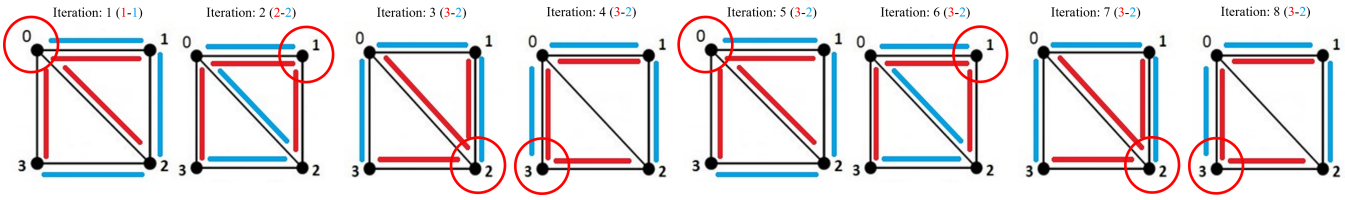

Figure 3: Breadth First Search (red lines) and Depth First Search (blue lines) spanning tree construction approaches result in a different number of the unique nearest balanced states retrieved for iteration. In this example, BFS produces 3 unique balanced states in 8 iterations, while DFS produces 2 unique balanced trees.

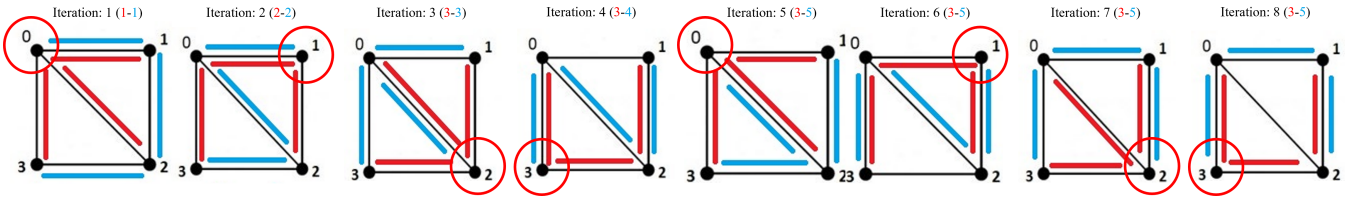

Figure 4: Randomized Breadth First Search (red lines) and Randomized Depth First Search (blue lines) spanning tree construction approaches result in a different number of the unique nearest balanced states retrieved for iteration. In this example, RBFS produces the same balanced states as BFS in Figure 1 while RDFS captures all five unique and balanced states.

tree due to random edge selection, e.g., edges 0-2,0-3, and 0-1 in iteration 3 and edges 0-1, 0-3, and 1-2 in iteration 7 in Figure 4.

Randomized Depth First Search (RDFS) algorithm transforms DFS into a non-deterministic algorithm by eliminating the static ordering of the adjacency lists. The time complexity of the DFS is known to be O($|V|+|E|$), where $|V|$ is the number of vertices and $|E|$ is the number of edges in the signed network. The algorithm also runs in linear time O(n), where n is the number of nodes adjacent to a specific node in the network, so the total time complexity is O($|V| + |E|$). **Aldous-Broder algorithm** produces a random uniform spanning tree by performing a random walk on a finite graph with any initial vertex and stops after all vertices have been visited [29]. The partial rejection tree sampling algorithm randomly selects neighboring nodes in the graph. If the cycle is present in the selection, all edges of the loops are removed, and neighbors are selected again [33]. RDFS shuffles the static order of the neighboring nodes in the adjacency list and then visits the first unvisited node. For the popular **Kruskal's algorithm** [2] [1], we intend to generate random spanning trees by assigning random weights to every edge in each iteration before running the algorithm. It is used to find the minimum spanning tree of a connected and weighted graph. The goal to find a subset of edges that connect all the vertices of the graph while minimizing the total weight or cost of the tree. Similar to Kruskal's algorithm, randomizing the weights of **Prim's algorithm** [23] [1] can also generate random spanning trees. It grows the minimum spanning trees from a single arbitrary node by incrementally adding the node that is closest to the existing MST (whichever edge connecting the current node of the MST to its neighborhood has the least weight is going to be added) .

RDFS-BFS sampler combines the best of both worlds. Breadth-first search (BFS) finds the balanced state with a minimum number of edge switches, and randomized depth-first search (RDFS) diversifies the spanning trees and recovers more unique balanced states. We propose a new algorithm, the RDFS-BFS sampler, to minimize the frustration index and maximize the number of unique stable states to increase algorithmic chances of finding the optimal state among all the nearest balanced states. The proposed algorithm in Alg. 4 and introduces the randomization at each iteration on what algorithmic step to execute next.

Algorithm 4: Hybridized RDFS-BFS Sampling

Data: Signed graph Σ and a root Node n
Result: Spanning tree T of Σ
Get uniformly distributed random number 0 or 1, z ;
if z *is* 0 **then**
  Run BFS algorithm [13]
else if z *is* 1 **then**
  Run RDFS algorithm

6 COMPLEXITY ANALYSIS

graphB++ that is based on the original implementation of graphB+ (BFS) has a complexity of O(|E| * log(|V| * d)) time, where |E| is the number of edges, |V| is the number of vertices, and d is the average spanning-tree degree of the vertices on each cycle. The code for scaling the processing and saving of balanced states in the memory-bound frustration cloud as well as approximating the frustration index which builds upon graphB+ adds O(|E|). O(|E| * log(|V| * d)) is still the dominant term. On the other hand, for graphB++ that is adapted for other tree-sampling techniques, the reimplemented vertex relabeling step takes O(|V| + |E|) instead because DFS is used to perform the pre-order traversal on a random spanning tree generated by a custom sampler of certain complexity. For the edge relabeling, specifically for assigning the beg and end ranges on each edge to any spanning tree, it has been reimplemented which compromised the efficiency resulting in a complexity of approximately O(|V||E|*α) where α is the average depth from a certain vertex of an

SNAP	vertices	edges			vertex degrees			attributes	
[38]	$\|V\|$	$\|E\|$	cycles	% positive	average	median	max	density	bal_3
Σ Fig. 2(a)	4	5	2	60	2.5	2.5	3	0.833	0.0
test10 [4]	10	13	4	53.85	2.6	2.5	4	0.288	0.5
highland [43]	16	58	43	50	7.25	7.5	10	0.483	0.868
sampson18 [48]	18	112	95	54.4	12.44	12.50	16	0.732	0.6
rainFall [17]	306	93,636	93,331	68.78	305.00	305	305	1.0	0.717
S&P1500 [17]	1,193	711,028	709,836	75.13	1,192	1,192	1,192	0.833	0.718
wikiElec [38]	7,539	112,058	104,520	73.33	28.16	15	1,079	0.004	0.798
wikiRfa [38]	7,634	175,787	168,154	77.91	43.99	13	1,233	0.005	0.717
epinions [38]	119,130	704,267	585138	83.23	11.82	2	3,558	< 0.001	0.890
slashdot [38]	82,140	500,481	418342	77.03	12.19	2	2,548	< 0.001	0.856

Table 1: SNAP Signed graph Largest connected component (LCC) attributes. $|V|$ is the number of vertices and $|E|$ is the number of edges in the largest connected component LCC; The label *% positive* is the number of positive edges divided by e; Vertex degree statistics are calculated in terms of the average, mean and maximum node degree; graph density $d = 2 * |E|/(|V| * |V-1|)$ and bal_3 is the percentage of triangles in the graph that are balanced.

edge to the deepest relabeled vertex where the assignment of the end range of the edge takes place. Both reimplementations have been connected to the efficient fundamental cycle balancing method [4] with complexity O(|E| * log(|V| * d)). Likewise, the same code aforementioned was employed for index computation which takes O(|E|). Hence, the total time complexity for the adapted version of graphB++ is O(|V||E|$^*\alpha$) unless the complexity of the selected custom sampler is high enough to exceed this complexity. Table 2 shows the time complexity for each tree-sampling technique.

7 EXPERIMENTS

7.1 Setup and Implementation

Setup The operating system used for all experimental evaluations is Linux Ubuntu 20.04.3 running on the 11th Gen Intel(R) Core(TM) i9-11900K @ 3.50GHz with 16 physical cores. It has one socket, two threads per core, and eight cores per socket. The architecture is X86_x64. GPU is Nvidia GeForce having 8GB memory. Its driver version is 495.29.05, and the CUDA version is 11.5. The cache configuration is L1d : 384 KiB, L1i : 256 KiB, L2 : 4 MiB, L3 : 16 MiB. The CPU op is 32-bit and 64-bit.

Implementation The baseline implementation is based on the published Binary Linear Programming (BLP) code [8]. The binary linear model runs on a Jupyter notebook in Python [8] and is based on a Guribo mathematical solver and covers several parameters [24]. The dependencies of the binary terms in the objective function of the AND and XOR models are considered using AND constraints and two standard XOR constraints per edge, respectively. Two replacements in the ABS model's objective function linearized two absolute value terms [11]. The code [8] was run with the following modifications: (1) the method parameter is -1, automatic; (2) the lazy parameter is 1 with enabled speedup; (3) *multiprocessing.cpu_count()* is added as the thread parameter, and (4) the time limit for the model run is set to up to 40 hours. The code provided [8] generates random graphs based on the specified number of nodes, edges and probability of negative edges. Our improvements to the code allow for the code to (1) accept the same input format as *graphB++*; and to (2) detect and eliminates duplicates, inconsistencies, self-loops, and invalid signs in the input graph. The **graphB++** implementation extends the open source implementation [4] to include and test proposed tree sampling strategies while keeping the original speed-up optimization for finding fundamental cycles intact. The implementation of tree sampling strategies and the adaptation of **graphB++** to work with them are highly optimized in C++. This is achieved by employing the least number of loops possible, incorporating OpenMP directives for parallelization, and freeing dynamically allocated objects when unused. The *graphB++* also implements frustration cloud computation based on the memory size in Alg. 3. This is done by retrieving the total RAM size from the Linux system during runtime and the current resident set size in which the current execution is consuming. Then, these two values are compared in each iteration to determine the limit of the number of balanced states to save. The code is released on the anonymous gitHub [6], and the data is publicly available [26, 34, 38].

7.2 Benchmark Datasets

Signed Graphs benchmarks used in the experiments are **SNAP** [38], **Konect** [34], and **Amazon** ratings [26]

SNAP signed graphs [38] characteristics are described in Table 1. The graph constructed from the Highland Tribes data, *highland*, has a high number of balanced triangles (86.8% from Tab. 1), is small in size (3 communities, 16 vertices), and exhibits high clusterability. The *sampson* signed graph captures 18 monks and their pairwise attitudes with 63 positive and 49 negative edges [48]. The *rainfall* signed graph is a fully connected dense graph of pairwise interrelationships of Australian rainfall between 306 locations. A matrix contains positive and negative correlations between stations [27]. The *S&P1500* graph models the correlations of stocks over 2 years [27]. The **wikiElec**, the signed graph of the Wikipedia administrator votes has a low density (0.04), and 79.8% of the triangles in the network are balanced [38]. The **wikiRfa**, the Wikipedia manager election [38] has 7,634 nodes and 105,160 edges with an average degree of 43.99 [38]. The **slashdot** signed graph has 1 connected component, 82, 140 vertices, 385, 515 positive, and 114, 966 negative edges [38]. Half of the nodes have two or fewer edges, and the graph density is 0.1 %. The **epinions** signed graph we are analyzing has 131, 828 vertices, 592, 236 positive and 118, 974 negative edges, and low density (under 0.1%) [38]. Both *slashdot* and *epinions* qualify as

Sampler	BFS	DFS	RDFS	Hybrid	Kruskal	Prim	AB
Complexity	O($\|V\|+\|E\|$)	O($\|V\|+\|E\|$)	O($\|V\|+\|E\|$)	O($\|V\|+\|E\|$)	O($\|E\|log\|V\|$) or O($\|E\|log\|E\|$)	O($\|V\|^2$)	O($\|V\|$)

Table 2: The time complexities of SEVEN tree-sampling techniques incorporated in graphB++

	Largest Connected Component Graph								
Konect [34]	vertices $\|V\|$	edges $\|E\|$	cycles $\|E\|-\|V\|+1$	# % positive	average degree	median degree	# max degree	# density	# bal_3
Sampson	18	126	145	51.32	14.0	14	17	0.82	0.66
ProLeague	16	120	105	49.79	15.0	15	15	1.0	0.53
DutchCollege	32	422	391	31.51	26.375	28	31	0.85	0.90
Congress	219	521	303	80.44	4.75	3	33	0.021	0.97
BitcoinAlpha	3,775	14,120	10,346	93.64	7.48	2	511	0.001	0.88
BitcoinOTC	5,875	21,489	15,615	89.98	7.315	2	795	0.001	0.89
Chess	7,115	55,779	48,665	32.53	15.67	7	181	0.002	0.57
TwitterReferendum	10,864	251,396	240,533	94.91	46.28	12	2784	0.004	0.91
SlashdotZoo	79,116	467,731	388,616	76.092	11.82	2	2534	0.0001	0.85
Epinions	119,130	704,267	585,138	85.29	11.82	2	3558	<0.0001	0.88
WikiElec	7,066	100,667	93,602	78.77	28.49	4	10657	0.004	0.74
WikiConflict	113,123	2,025,910	1,912,788	43.31	35.81	4	20153	0.0003	0.52
WikiPolitics	137,740	715,334	577,595	87.88	10.38	2	10715	<0.001	0.90

Table 3: Konect Largest Connected Component (LCC) graph attributes [34]. $|E|$ is the number of vertices, and $|V|$ is the number of edges in the LCC of the graph. The *% positive* label marks the percentage of positive edges in the LCC. Vertex degree statistics are calculated regarding the average, mean, and maximum node degree in the LCC. The graph density d is $2 * |E|$ by $|V| * |V-1|$. The bal_3 is the percentage of triangles in the graph balanced in LCC.

a large and sparse graph with a power-law degree distribution [18]. The median degree in sparse networks is usually much lower than the average degree, and all four [38] graphs are considered sparse.

Konect graph [34] characteristics are described in Table 3. emph-Sampson is a directed network that contains ratings between monks related to a crisis in a cloister (or monastery) in New England (USA) which led to the departure of several of the monks [34]. *ProLeague* are results of football games in Belgium from the Pro League in 2016/2017 in the form of a directed signed graph. Nodes are teams; each directed edge from A to B denotes that team A played at home against team B. The edge weights are the goal difference, and thus positive if the home team wins, negative if the away team wins, and zero for a draw [34]. *DutchCollege* is a directed network that contains friendship ratings between 32 university freshmen who mostly did not know each other before starting university. Each student was asked to rate the other at seven different time points. A node represents a student, and an edge between two students shows that the left rated the right. The edge weights show how good their friendship is in the eye of the left node. The weight ranges from -1 for the risk of conflict to +3 for best friend [34]. *Congress* is a signed network where nodes are politicians speaking in the United States Congress, and a directed edge denotes that a speaker mentions another speaker [34]. In the *Chess* network, each node is a chess player and a directed edge represents a game with the white player having an outgoing edge and the black player having an ingoing edge. The weight of the edge represents the outcome [34]. *BitcoinAlpha* is a user-user trust/distrust network from the Bitcoin Alpha platform on which Bitcoins are traded [34].*BitcoinOTC* is a user-user trust/distrust network, from the Bitcoin OTC platform, on which Bitcoins are traded [34]. *TwitterReferendum* captures data from Twitter concerning the 2016 Italian Referendum. Different stances between users signify a negative tie, while the same stances indicate a positive link [36]. *WikiElec* is the network of users from the English Wikipedia that voted for and against each other in admin elections [34]. *Slashdot* is the reply network of the technology website Slashdot. Nodes are users, and edges are replies [34]. The edges of *WikiConflict* represent positive and negative conflicts between users of the English Wikipedia [34]. *WikiPolitics* is an undirected signed network that contains interactions between the users of the English Wikipedia that have edited pages about politics. Each interaction, such as text editing and votes, is given a positive or negative value [34]. *Epinions* is the trust and distrust network of Epinions, an online product rating site. It incorporates individual users connected by directed trust and distrust links [34]. The characteristics of the signed graphs are listed in Table **??**. Note the overlap with SNAP data and a subtle difference in hwo graphs were derived.

Amazon ratings and reviews data [26] provides rating information between 0 (low) and 5 (high) of the Amazon user on different products. We have transformed the graphs to 18 signed bipartite graphs, where ratings 5 and 4 imply a positive edge, rating 3 and 2 give no edge, and ratings 0 and 1 give a negative edge. The graph characteristics are described in Table 8 and result in tens of millions of vertices and edges in a signed graph.

7.3 Experiment: Selecting the Spanning Tree Method

In this experiment, we compare the timing and frustration computation of *graphB++* implementation of Alg. 2 of **SEVEN** different tree sampling methods described in Section 5 and look for the most effective and efficient samplign method for the frustration index computation. The seven methods are: Prim, Krushkal, Breadth First Search (BFS), Depth First Search (DFS), Randomized DFS (RDFS), Hybrid RDFS-BFS and Aldous-Broder sampling. Note that *graphB++* runs are non-deterministic, and we run the methods multiple times. The frustration computed and completion time are always the same for smaller graphs and within 0.1% for larger graphs. We compare the findings to the BLP baseline implementation for SNAP datasets. The resulting frustration index for different methods are outlined in Table 1.

	rainFall	**S&P 1500**	**wikiElec**	**wikiRfa**	**epinions**	**slashdot**
BFS	**10,217**	**134,515**	24,827	43,971	**100,450**	117,587
RDFS	20,047	326,957	51,197	78,215	276,584	205,236
DFS	22,879	351,135	50,617	78,151	275,211	205,089
Hybrid	**10,217**	134,863	**23,970**	**42,573**	118,588	**115,932**
Kruskal	18,576	318,733	38,255	71,990	200,264	188,495
AB	19,403	323,657	47,252	74,438	246941	198785
Prim	21,312	355,819	51,732	79,482	N/A	N/A
BLP	10,150	176,965	29,257	26,778	N/A	77,283

Table 4: SNAP frustration for 1000 iterations of graphB++ with SEVEN different tree samplers introduced in Section 5, and the baseline. AB stands for Aldous-Broder. BFS and Hybrid consistently perform the best.

The results are summarized in Table 4 in terms of the actual frustration index value and in Figure 7 in terms of the computed frustration index as a percentage of the total number of edges in the graph. BFS-spanning trees produce balanced states of minimum edge switches, and DFS-spanning trees make trees with maximum edge switches, as evident from the frustration computed in Figure 7. RDFS/Kruskal/Aldous-Broder frustration scores are slightly better due to the randomization step. BFS discovers the optimal trees for the frustration computation, but they are repetitive.

The timing is reported on the log 10 scale in seconds in Figure 8 as BLP takes 40 hours for larger datasets (far right navy bar in Fig. 8). RDFS-BFS hybrid approach is competitive with BFS in terms of frustration index (green and blue bars in Figure 7) with the small timing overhead for large graphs (Fig. 8): BFS produces 117,587 frustrations while BFS-RDFS produces 115,932 frustrations in 1000 iterations for the slashdot dataset. The Prim approach is too slow for large datasets, and the baseline BLP takes too long or it does not complete. Thus, we opted to compare only the top 6 methods for the Konect benchmarks and we tabulated the results in Table 5, where the frustration index produced by different tree-sampling techniques has been compared. Since the time complexity of Prim is $O(V^2)$, the number of iterations is set to 1 and it was very inefficient and slow for large graphs such as WikiConflict. Moreover, Aldous-Broder did not terminate for DutchCollege because, in uncommon scenarios, this sampler would get stuck looping when performing a random walk. After all, the current vertex's neighbors are already visited.

Konect [34]	**Frustration Index for Spanning Tree Method**					
Sampling	**BFS**	**RDFS**	**DFS**	**Hybrid**	**Kruskal**	**AB**
Sampson	**35**	37	39	37	38	37
ProLeague	13	13	20	13	13	13
DutchCollege	2	2	2	2	2	N/A
Congress	21	18	53	19	bf 18	19
BitcoinAlpha	1,105	3,487	4,371	**1,022**	1,716	1,585
BitcoinOTC	**1,827**	6,364	6,422	1,999	2,506	3,367
Chess	20,991	23,800	23,842	**20,685**	23,212	23,753
TwitterRef.	**16,183**	104,122	93,014	27,938	40,723	37,670
Slashdot Zoo	**109,930**	190,720	191,936	111,270	172,902	188,072
Epinions	**100,646**	276,883	280,758	117,829	198,989	244,482
WikiElec	**22,289**	45,094	45,024	23,284	36,024	39,916
Wiki Conflict	252,400	933,814	907,660	**217,385**	631,438	780,447
Wiki Politics	86,833	268,518	266,613	**80,389**	172,859	224,145

Table 5: Frustration index approximated for Konect datasets [34] using graphB++ for *SIX* sampling techniques (BFS, RDFS, DFS, Hybrid, Kruskal, Aldous-Broder) in 1000 iterations in Aglorithm 3.

Konect [34]	**Computation Time for Spanning Tree Method**					
Sampling	**BFS**	**RDFS**	**DFS**	**Hybrid**	**Kruskal**	**AB**
Sampson	**0.0003s**	0.00084s	0.00053s	0.00106s	0.00055s	0.00057s
ProLeague	**0.00027s**	0.0008s	0.0035s	0.00088s	0.0005s	0.0005s
DutchCollege	**0.0008s**	0.00216s	0.00179s	0.00288s	0.00158s	N/A
Congress	**0.00102s**	0.00505s	0.00301s	0.00629s	0.00317s	0.00340s
BitcoinAlpha	**0.024s**	0.126s	0.103s	0.143s	0.098s	0.102s
BitcoinOTC	**0.041s**	0.265s	0.229s	0.268s	0.214s	0.231s
Chess	0.085s	0.388s	0.283s	0.375s	0.250s	0.282s
TwitterRef.	0.457s	2.826s	2.277s	2.566s	2.111s	2.155s
SlashdotZoo	0.838s	19.131s	13.803s	15.102s	9.849s	11.880s
Epinions	1.368s	20.794s	12.795s	16.715s	11.902s	13.373s
WikiElec	0.16859s	0.63977s	0.487s	0.559s	0.462s	0.502s
WikiConflict	6.503s	94.102s	65.282s	77.045s	40.720s	41.293s
WikiPolitics	1.582s	21.585s	17.374s	19.613s	15.925s	17.010s

Table 6: Average Frustration Index computation time per iteration using spanning tree sampling methods for 1000 iterations for Konect data in Table 5. BFS sampling method is the fastest. The algorithm is highly parallelizable.

7.4 Iteration Timing

Next, we evaluate the efficiency of the proposed algorithm by comparing *graphB++* frustration and timing if the number of iterations increases. Figure 5 shows the change in performance for the two best tree sampling methods when the number of iterations grows. As discussed in Section 5, more iterations will not impact BFS sampling in smaller graphs. The frustration shows a slight improvement for the larger graphs for both methods when the number of iterations increases in Figure 5. Note that the *graphB++* running time linearly increases with the number of iterations, as the experiment summarizes in Figure 6. The frustration improvements are insignificant on the sampled signed graph to justify the increased iteration count.

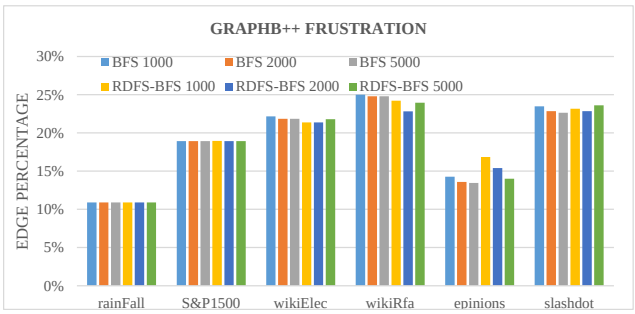

Figure 5: *graphB++* frustration for six benchmark datasets, two spanning three sampling approaches (BFS and RDFS-BFS), and three different iteration counts (1000, 2000, and 5000).

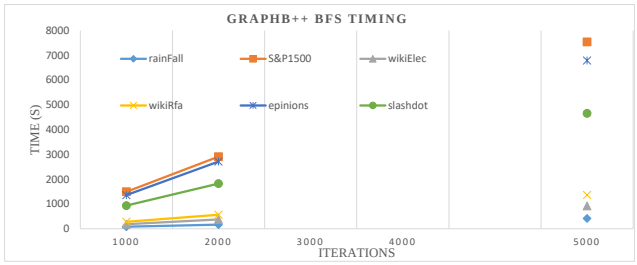

Figure 6: *graphB++* time complexity for BFS sampling method for three different iteration counts, 1000, 2000, and 5000.

SNAP [38]	Cycles $\lvert E\rvert - \lvert V\rvert + 1$	BLP index	BLP time	graphB++ index	graphB++ time
Σ Fig. 2(a)	2	1	0.03s	1	0.04s
test10 [4]	4	2	0.06s	2	0.08s
highland [43]	43	7	0.035s	7	0.13s
sampson18 [48]	95	43	0.06s	39	0.27s
rainFall [17]	93,331	10,150	7.5hrs	10,271	83.4s
S&P1500 [17]	709,836	176,965*	14.15*hrs	134,515	1478s
wikiElec [38]	104,520	29,257*	40hrs	24,827	184s
wikiRfa [38]	168,154	26,778*	40 hrs	43,971	281s
epinions [38]	585,138	N/A	N/A	100,450	1360s
slashdot [38]	418,342	77,283*	40 hrs	117,587	937s

Table 7: SNAP Signed graph baseline performance as a function of the number of fundamental cycles $\lvert E\rvert - \lvert V\rvert + 1$ for the baseline Binary Linear Programming (BLP) [11] and proposed *graphB++* algorithm with BFS tree sampling for frustration index computation. The BLP baseline never completed for epinions and never converged beyond Guribo heuristic for S&P1500 and wikiElec and wikiRfa datasets (*).

7.5 graphB++ vs. Baseline

In this experiment, we compare the baseline BLP [11] with graphB++ implementation with breadth first search (BFS) spanning tree sampling in 1000 iterations in terms of the frustration index and the time it takes to compute the frustration index for 10 benchmark graphs in Table 3. The space complexity of BLP is $O(|V|^2)$, where

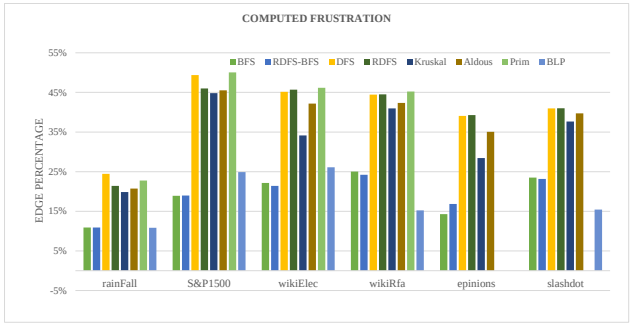

Figure 7: Frustration index computed using Binary Linear Programming (BLP) [11] and graphB++ 1000 iterations for different tree sampling methods over different real large signed graphs except Prim (1 iteration). BLP never finished computing the frustration index for epinions and sp1500 within the 40 hours allocated.

$|V|$ is the number of vertices on the graph. The researchers also stated that real signed graphs with up to 100,000 edges could be solved in 10 hours [11]. Since the frustration cloud computation can fit in the memory, we ran the *graphB++* implementation of Algorithm 2 with the breadth-first search (BFS) spanning tree. All external processes are closed to prevent interference with time measurements. The measurement for both methods includes the time it takes to input the file, process it, and output the results.

The results are outlined in the last four columns of Table 3. BLP and graphB++ computation for small graphs was fast and close. Both methods retrieve correct frustration indices for the three datasets. For the sampson18 dataset, BLP heuristic gets stuck in the local minima and results in 43 instead of 39 for the frustration index. During the optimization process, BLP model's best bound increases from 9 to 40.55088 while the incumbent value is at 43. After, the best bound also increases to reach an integer solution of 43 resulting in a gap 0%, and larger than 39 frustration *graphB++* produces. The baseline code fails for two graphs with over 700,000 edges, as described in Table 3. The BLP code fails for the sparse epinions (over 700,000 vertices) and produces no results, where graphB++ finds 1000 nearest balanced states of the graph, the most optimal one with frustration 100,450 in under 23 minutes. BLP code on the fully connected S&P1500 signed graph produces a heuristic frustration estimate of 176,965 after 14.15 hours of processing and fails during the model-solving phase. On the other hand, *graphB++* finds 1000 near-balanced states of the network and associated frustration 134,515 in 24.6 minutes (Table 3). The most extensive signed graph in which we were able to run the linear binary solver was the fully connected rainFall network [17] due to $|E| < 100,000$ [11] algorithm limitation. The linear binary solver finished with a frustration index of 10,150 and a time of 26,054.25s (7.5 hours) while the *graphB++*'s computed frustration index as 10,217 in 83.58s (0.02 hours), more than 300 times faster.

The linear binary solver fails to complete the computations for wikiElec within 40 hours: a proximity value of the frustration index was estimated to be between 21,299 and 29,257, where *graphB++* calculated the frustration 24,827 and provides an associated balanced state in 185s (3min+). Note that the BLP code produces a heuristics

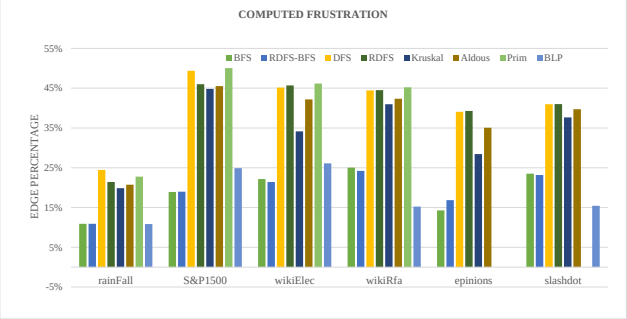

Figure 8: Binary Linear Programming (BLP) [11] and graphB++ 1000 iterations timing comparison in calculating frustration index for different tree sampling methods over different real large signed graphs except Prim (1 iteration). BLP crashed at input processing for epinions and did not complete the heuristic computation in 40 hours allocated for wikiElec, wikiRfa, and slashdot: the index* returned is a heuristic estimate on the upper range.

frustration estimate in 40 hours for wikiRfa and slashdot (it fails for epinion) and a gap with *no associated balanced state* as a guide on how actually to balance the graph. For wikiRfa and slashdot, the heuristic upper bound computed in 40 hours is much lower than discovered balanced states. On the other hand, *graphB++* finds 1000 unique near-balanced states for wikiRfa and slashdot in 5min and 15min, respectively, and offers frustration as a measure of the nearest balanced state it discovered in the process, see Table 3 for details.

How does the discovered *graphB++* compare with the BLP baseline? The *graphB++* outputs the corresponding balanced state in the same time frame, while the BLP code does not. The *graphB++* computes the nearest balanced states in minutes for large graphs compared to hours for BLP if the computation does not fail. For eight out of ten graphs tested, *graphB++* frustration is exact (4), close to actual (rainFall), or better than the heuristic (wikiElec, S&P1500 and epinions). Next, we investigate how the tree sampling methods influence the discovered state frustration. In this experiment, we have shown that the proposed *graphB++* balanced state discovery is equal to or superior to state of the art for small networks and efficient for more extensive networks, and scales for large networks both in terms of processing time and producing outcome where BLP either fails or makes heuristics without the associated balanced state.

7.6 Scaling Balanced State Discovery

In this experiment, we implement Alg. 3 and set the *CAP* to 75% of the total RAM size. We apply it to Amazon data in Table 8. The BLP model only worked and converged for the smallest 3 Amazon signed networks, the Core5 reviews in Tab. 8. All Amazon rating signed graphs have a number of vertices |V| higher than 300,000, and BLP outputs a memory error before initializing the model. The algorithm attempts to construct an adjacency matrix that does not fit into memory for any graph with more than 100,000 vertices. For smaller signed graphs BLP algorithm converges within 40 hours and

Amazon	|V|	|E|	BLP		graphB++ scale	
Ratings			index	time	index	time
Books	9,973,735	22,268,630	N/A	N/A	3,146,316	19hrs
Electronics	4,523,296	7,734,582	N/A	N/A	1,025,401	7.8 hrs
Jewelry	3,796,967	5,484,633	N/A	N/A	613,129	6hrs
TV	2,236,744	4,573,784	N/A	N/A	636,568	5.2hrs
Vinyl	1,959,693	3,684,143	N/A	N/A	412,859	4.4hrs
Outdoors	2,147,848	3,075,419	N/A	N/A	264,497	4hrs
Android App	1,373,018	2,631,009	N/A	N/A	386,947	3.6hrs
Games	1,489,764	2,142,593	N/A	N/A	173,063	2.2hrs
Automotive	950,831	1,239,450	N/A	N/A	85,859	50min
Garden	735,815	939,679	N/A	N/A	70,690	32.2min
Baby	559,040	892,231	N/A	N/A	106,092	30.1min
Digital Music	525,522	702,584	N/A	N/A	34,019	23min
Instant Video	433,702	572,834	N/A	N/A	32,001	20.7min
Musical Inst.	355,507	457,140	N/A	N/A	24,959	14.7min
Amazon	|V|	|E|	BLP		graphB++	
Reviews			index	time	index	time
Digital Music	9,109	64,706	10,311	38.5 hrs	19,926	101s
Instant Video	6,815	37,126	6,001	40hrs	10,833	101s
Musical Instr	2,329	10,261	1,162	136.4s	2,311	17.3s

Table 8: Amazon ratings and reviews graph characteristics [26]: The number of vertices, edges, and cycles reflect the number in the largest connected component of each dataset.

finds the optimal frustration index. The *graphB++* scale algorithm recovers the balanced state and associated frustration insonds for small graphs and in minutes for under 2 million edges; see the last column of Table 8. The serialized process takes about an extra hour for each 1 million edges, and the processing, for a fixed *CAP*, is linear in time with the number of edges and vertices, see Figure 9. The most extensive graph we have processed is Amazon books with close to 10 million vertices and over 22 million edges, and it took 19 hours to find the nearest balanced state with frustration 3,146,316. The *graphB++* processing times of the six signed benchmark graphs (Figure **??**) and the 17 Amazon graphs (Tab. 8) exhibit the same linear trend with the size of the graph. In conclusion, we have demonstrated that graphB++ can scale and find the nearest balanced state for *any* size of signed graph.

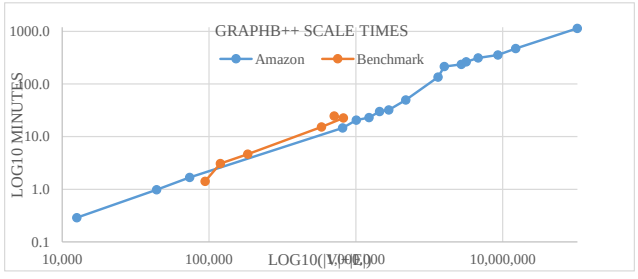

Figure 9: *graphB++* time complexity is linear with graph size |V|+|E| for a fixed number of iterations (1000) for the benchmark (Tab. 3) and Amazon (Tab 8) signed graphs.

8 SUMMARY AND FUTURE WORK

There is more than one way to achieve balance in the network. The frustration index characterizes the optimal nearest balanced state where the minimum edge switches are required to achieve balance in the network. It has been shown that the tree-spanning approach to graph balancing produces the nearest balanced states, e.g., there can be no other balanced state nearest balanced state derived from [44]. In this paper, we extend this finding and propose a novel algorithm for discovering the nearest balanced states for any graph size in a fraction of the time. Our approach converges to the global optimum for the small graphs that the state-of-the-art binary linear programming (BLP) model computes. BLP does not work for graphs larger than 100,000 vertices while *graphB++* seamlessly scales with the graph size to discover one or more nearest balanced states for the network. The state might not be optimal for a minimal number of edge switches, but it is close to optimal, and the algorithm produces a list of edges to switch to achieve the balanced state. In addition, the iterations of the algorithm can be parallelizable; while we report the result on one computer when we input 1000 iterations, the timing we should consider is 1 iteration if we get to spawn the jobs in parallel for large signed graphs.

REFERENCES

[1] geeksforgeeks. https://www.geeksforgeeks.org/. Accessed: 2023-06-01.

[2] *17 On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem (1956).* 2021.

[3] Robert P. Abelson and Milton J. Rosenberg. Symbolic psycho-logic: A model of attitudinal cognition. *Behavioral Science*, 3(1):1–13, 1958.

[4] Ghadeer Alabandi, Jelena Tešić, Lucas Rusnak, and Martin Burtscher. Discovering and balancing fundamental cycles in large signed graphs. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '21, New York, NY, USA, 2021. Association for Computing Machinery.

[5] Victor Amelkin and Ambuj K Singh. Fighting opinion control in social networks via link recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 677–685, 2019.

[6] Anonimoys. Codebase for scaling frustration index and corresponding balanced state discovery for real signed graphs paper. https://anonymous.4open.science/r/graphBplusplus-547A/README.md, 2023.

[7] T. Antal, P. L. Krapivsky, and S. Redner. Social balance on networks: The dynamics of friendship and enmity. *Physica D: Nonlinear Phenomena*, 224(1):130–136, 2006.

[8] Samin Aref. frustration-index-xor. https://github.com/saref/frustration-index-XOR, 2021.

[9] Samin Aref, Andrew J. Mason, and Mark C. Wilson. An exact method for computing the frustration index in signed networks using binary programming. *CoRR*, abs/1611.09030, 2016.

[10] Samin Aref, Andrew J. Mason, and Mark C. Wilson. A modeling and computational study of the frustration index in signed networks. *Networks*, 75(1):95–110, 2020.

[11] Samin Aref and Zachary P Neal. Identifying hidden coalitions in the us house of representatives by optimally partitioning signed networks based on generalized balance. *Scientific reports*, 11(1):1–9, 2021.

[12] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. *Machine Learning*, 56(1-3):89 – 113, 2004.

[13] Martin Burtscher. graphB+: Balancing algorithm for large graphs. https://userweb.cs.txstate.edu/~burtscher/research/graphB/, 2021.

[14] Dorwin Cartwright and Frank Harary. Structural balance: a generalization of Heider's theory. *Psychological Rev.*, 63:277–293, 1956.

[15] Sergio Casas, Cole Gulino, Renjie Liao, and Raquel Urtasun. Spagnn: Spatially-aware graph neural networks for relational behavior forecasting from sensor data. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9491–9497. IEEE, 2020.

[16] Thomas H. Cormen. *Introduction to algorithms.* MIT Press, 2009.

[17] Mihai Cucuringu, Apoorv Vikram Singh, Déborah Sulem, and Hemant Tyagi. Regularized spectral methods for clustering signed networks. *Journal of Machine Learning Research*, 22(264):1–79, 2021.

[18] Lorenzo Dall'Amico, Romain Couillet, and Nicolas Tremblay. A unified framework for spectral clustering in sparse graphs. *Journal of Machine Learning Research*, 22:1–56, 2021.

[19] Tyler Derr, Zhiwei Wang, Jamell Dacon, and Jiliang Tang. Link and interaction polarity predictions in signed networks. *Social Network Analysis and Mining*, 10(1):1–14, 2020.

[20] Tomislav Došlić and Damir Vukičević. Computing the bipartite edge frustration of fullerene graphs. *Discrete Applied Mathematics*, 155(10):1294–1301, 2007.

[21] Angela Fontan and Claudio Altafini. Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength. *2018 IEEE Conference on Decision and Control (CDC), Decision and Control (CDC), 2018 IEEE Conference on*, pages 109 – 114, 2018.

[22] Kiran Garimella, Tim Smith, Rebecca Weiss, and Robert West. Political polarization in online news consumption. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 152–162, 2021.

[23] Saul I. Gass and Michael C. Fu, editors. *Prim's Algorithm*, pages 1160–1160. Springer US, Boston, MA, 2013.

[24] GUROBI. Gurobi optimization. https://www.gurobi.com/.

[25] Frank Harary and Dorwin Cartwright. On the coloring of signed graphs. *Elemente der Mathematik*, 23:85–89, 1968.

[26] Ruining He and Julian McAuley. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on WWW*, pages 507–517, 2016.

[27] Yixuan He, Gesine Reinert, Songchao Wang, and Mihai Cucuringu. SSSNET: semi-supervised signed network clustering. *CoRR*, abs/2110.06623, 2021.

[28] Fritz Heider. Attitudes and cognitive organization. *J. Psychology*, 21:107–112, 1946.

[29] Yiping Hu, Russell Lyons, and Pengfei Tang. A reverse aldous-broder algorithm. *ANNALES DE L INSTITUT HENRI POINCARE-PROBABILITES ET STATISTIQUES*, 57(2):890 – 900, 2021.

[30] Falk Hüffner, Nadja Betzler, and Rolf Niedermeier. Separator-based data reduction for signed graph balancing. *Journal of combinatorial optimization*, 20(4):335–360, 2010.

[31] Giovanni Iacono and Claudio Altafini. Average frustration and phase transition in large-scale biological networks: a statistical physics approach. *IFAC Proceedings Volumes*, 43(14):320–325, 2010. 8th IFAC Symposium on Nonlinear Control Systems.

[32] Ruben Interian, Ruslan G Marzo, Isela Mendoza, and Celso C Ribeiro. Network polarization, filter bubbles, and echo chambers: An annotated review of measures, models, and case studies. *arXiv preprint arXiv:2207.13799*, 2022.

[33] Mark Jerrum. Fundamentals of partial rejection sampling, 2021.

[34] Jérôme Kunegis. KONECT – The Koblenz Network Collection. In *Proc. Int. Conf. on World Wide Web Companion*, pages 1343–1350, 2013.

[35] Jérôme Kunegis. Konect: the koblenz network collection. In *Proceedings of the 22nd international conference on world wide web*, pages 1343–1350, 2013.

[36] Mirko Lai, Viviana Patti, Giancarlo Ruffo, and Paolo Rosso. Stance evolution and twitter interactions in an italian political debate. In Max Silberztein, Faten Atigui, Elena Kornyshova, Elisabeth Métais, and Farid Meziane, editors, *Natural Language Processing and Information Systems*, pages 15–27, Cham, 2018. Springer International Publishing.

[37] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 641–650. ACM, 2010.

[38] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.

[39] Rui Li, Xin Yuan, Mohsen Radfar, Peter Marendy, Wei Ni, Terence J. O'Brien, and Pablo M. Casillas-Espinosa. Graph signal processing, graph neural network and graph learning on biological data: A systematic review. *IEEE Reviews in Biomedical Engineering*, pages 1–1, 2021.

[40] Taoran Liu, Jiancong Cui, Hui Zhuang, and Hong Wang. Modeling polypharmacy effects with heterogeneous signed graph convolutional networks. *Applied Intelligence*, 51:8316–8333, 2021.

[41] Tianya Liu, Andong Sheng, Guoqing Qi, and Yinya Li. Admissible bipartite consensus in networks of singular agents over signed graphs. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(8):2880–2884, 2021.

[42] John Palowitch, Anton Tsitsulin, Brandon Mayer, and Bryan Perozzi. Graphworld: Fake graphs bring real insights for gnns. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '22, New York, NY, USA, 2022. Association for Computing Machinery.

[43] Kenneth Read. Cultures of the central highlands, new guinea. *Southwestern Journal of Anthropology*, 10(1):1–43, 1954.

[44] Lucas Rusnak and Jelena Tešić. Characterizing attitudinal network graphs through frustration cloud. *Data Mining and Knowledge Discovery*, 6, November 2021.

[45] Majid Saberi, Reza Khosrowabadi, Ali Khatibi, Bratislav Misic, and Gholamreza Jafari. Requirement to change of functional brain network across the lifespan. *PLoS ONE*, 16(11):1 – 19, 2021.

[46] Majid Saberi, Reza Khosrowabadi, Ali Khatibi, Bratislav Misic, and Gholamreza Jafari. Pattern of frustration formation in the functional brain network. *NETWORK NEUROSCIENCE*, 6(4):1334 – 1356, 2022.

[47] Majid Saberi, Reza Khosrowabadi, Ali Khatibi, Bratislav Mišić, and Gholamreza Jafari. Topological impact of negative links on the stability of resting-state brain network. *Scientific Reports*, 11(1):2176, 2021.

[48] S. Sampson. A novitiate in a period of change: An experimental and case study of relationships. Ph.D. thesis, Cornell University, 1968.

[49] Michael T Schaub, Neave O'Clery, Yazan N Billeh, Jean-Charles Delvenne, Renaud Lambiotte, and Mauricio Barahona. Graph partitions and cluster synchronization in networks of oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(9):094821, 2016.

[50] Z Seif and MB Ahmadi. Computing frustration index using genetic algorithm. *Communications in Mathematical and in Computer Chemistry*, 71:437–443, 2014.

[51] Maria Tomasso, Lucas Rusnak, and Jelena Tešić. Advances in scaling community discovery methods for signed graph networks. *Journal of Complex Networks*, 10(3), 06 2022. cnac013.

[52] William T. Tutte. *Graph theory*, volume 21 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1984.

[53] Xiaozhou Zhou, Haoyu Song, and Jingyuan Li. Residue frustration based prediction of protein-protein interactions using machine learning. *Journal of physical chemistry*, 126(8):1719–727, 2022.