Manipulating Black-Box Networks for Centrality Promotion

Wentao Li[♮], Min Gao[§], Fan Wu[§], Wenge Rong[‡], Junhao Wen[§], and Lu Qin[♮]

[♮]*AAII, FEIT, University of Technology Sydney, Australia*

[§]*Chongqing University, China*

[‡]*Beihang University, China*

[♮]wentao.li@student.uts.edu.au; [§]{gaomin, wufan, jhwen}@cqu.edu.cn;

[‡]w.rong@buaa.edu.cn; [♮]lu.qin@uts.edu.au

***Abstract*—Centrality measures are widely used to map each node to its importance in a network. For many practical applications, vital nodes bearing high centrality scores have superior positions over other nodes. To benefit from the positive impact of becoming a vital node, the problem of improving the centrality of the target node has attracted increasing attention. Many existing studies attack this problem by directly increasing the *centrality score* of the target node on the premise of knowing the network structure. However, these methods suffer from privacy issues due to their dependence on the network structure and may lose their effectiveness because other nodes can simultaneously increase the scores. Therefore, in this paper, we explore the following question: given a black-box network whose structure is unknown, is it possible to improve the *centrality ranking* (rather than the score) of a target node by implementing certain strategies? We provide an affirmative answer to this question. First, to avoid relying on the network structure for promotion, we propose strategies that freeze the original graph while appending nodes and edges just around the target node. Second, to guide strategies for effectively boosting centrality, we devise two principles that provide the target node with either the maximum gain or the minimum loss of centrality scores over other nodes. We prove that a strategy meeting the proposed principles is guaranteed to upgrade the target node's ranking. Extensive experiments were conducted to verify the effectiveness of the proposed strategies on black-box networks.**

I. Introduction

Networks have been extensively applied to model entities and their relations in the real-world [1]. As one of the essential aspects of network analysis, centrality reveals the basic properties of a network [2]: given a graph[1] $G(V, E)$, a **centrality measure** assigns a score to every node $v \in V$ to indicate the relative importance of this node [3]. Some commonly used centrality measures are degree centrality, closeness centrality, and betweenness centrality (see [4] and references therein).

As a tool for mapping each node to its corresponding significance, centrality measures are exploited to identify vital nodes on a network [4]. Vital node identification has many applications on networks [4–6], with examples ranging from advertisements on social networks [7] to outbreak control on epidemic networks [8]. In these applications, vital nodes with relatively high centrality scores have superior positions over other nodes on a network [9, 10].

Motivating Examples. This paper studies the issue of promoting a target node's centrality on a network to benefit from the positive impact of becoming a vital node (for fun and profit). The motivating examples for centrality promotion are listed below.

- *Closeness Promotion.* The closeness of a node is the reciprocal of the sum of the distances from this node to all nodes in the graph [5]. In co-authorship networks, authors with high closeness are likely to receive more citations [9, 11]. Moreover, research results published by authors with high closeness are prone to disseminate widely in the network [12]. Therefore, an author would be pleased to adopt promotion strategies to increase closeness, thereby having more research impact than colleagues with lower closeness.
- *Betweenness Promotion.* The betweenness of a node is the fraction of the shortest paths between node pairs that pass through this node [13]. In social networks, users with high betweenness are influential since their posted information (e.g., tweets) diffuses rapidly and widely [14]. Users with high betweenness can thus be requested to help spread information of others for wide dissemination. Hence, a user can employ promotion strategies to increase the betweenness to become more influential.
- *Coreness Promotion.* The coreness of a node is the largest integer k such that this node is contained in a subgraph in which each node has a degree not less than k [15]. In information networks, nodes with relatively high coreness act as blockers to prevent rumors from spreading throughout the system [16]. Consequently, a user can adopt promotion strategies to increase coreness for better control of rumor spreading than users with relatively lower coreness.
- *Eccentricity Promotion.* The eccentricity of a node is the reciprocal of the maximum distance from this node to all the nodes [6]. In sport team networks, compared with players with average eccentricity, players with high eccentricity can easily affect other teammates [17]. Accordingly, a player can adopt promotion strategies to increase eccentricity for a positive influence on other teammates' activities.

Existing Solutions. Current research normally inserts additional edges into the original graph to improve the *centrality score* of a target node, provided that the network topology is known. Examples can be found in [18] (for betweenness), [19] (for coreness), [9] (for closeness), and [20] (for eccentricity), etc. These studies are formed as follows: given a graph $G(V, E)$, a target node $t \in V$, and a budget b, select b edges from the non-existing edges $\widehat{E} = \{V^2 \setminus E\}$, thereby maximizing the centrality improvement of t. Due to the hardness of these problems, many researchers resort to greedy (approximation) algorithms to obtain suboptimal solutions for centrality promotion [18, 19, 21].

Motivations. Greedy algorithms are valuable to **network owners** [22] who have a complete view of the network structure to

Min Gao is the corresponding author.

[1]We use the terms "network" and "graph" interchangeably.

make a greedy decision for centrality improvement. In contrast, for privacy reasons [23, 24], a real-world network is more likely to be a black box for **network users** [22] — users have no access to the entire network structure. When the network structure is inaccessible, greedy algorithms are not candidates for users who want to increase centrality.

Furthermore, existing greedy algorithms normally improve a target node's centrality score[2] and are applicable when the score is critical [9]. For example, an increasing betweenness score for an airport always corresponds to an increasing volume of traffic and customers [25]. However, there are contexts where a high centrality ranking is desirable (see [4, 18] and our motivating examples). In this case, an increasing score *does not* necessarily mean an improved ranking [18] — other nodes can also enlarge their scores.

Challenges. To overcome the limitations of existing solutions, we manipulate **black-box networks** (networks with an unknown structure) to promote a target node's *centrality ranking*. Specifically, we have two goals: i) we do not rely on the knowledge of the network structure to make an improvement (thus it is feasible for network users); ii) we increase the centrality ranking (rather than the centrality score) of the target node.

These appealing goals are not easily achieved because we need to overcome the corresponding challenges: i) How to design a practical promotion strategy when the network structure is unknown? ii) How to ensure the promotion strategy is valid for a target node to increase its centrality ranking?

Our Solution. To address the first challenge, we propose promotion strategies that only append additional nodes/edges around the target node. These strategies negate the need to change (and refer to) the structure within the original graph, thus making promotion on black-box networks feasible. To solve the second challenge, we update the centrality scores of nodes to a different extent after promotion (the target node has the maximum gain or the minimum loss over all other nodes) to potentially improve the centrality ranking of the target node. Non-trivial theoretical analysis is undertaken to show that this simple yet elegant idea works well for centrality promotion.

Contributions. Our contributions are summarized as follows.

- *Formalization of the centrality promotion problem on black-box networks* (Section III). We discard dependency on the network structure for promotion. In addition, to ensure valid promotion, we aim to increase the ranking instead of the score of the target node.
- *Practically usable promotion strategies* (Section IV). We incorporate various promotion strategies into a general model, where only additional nodes and edges around the target node are inserted. Based on the structure between the inserted nodes, three strategies are proposed, i.e., multi-point, double-line, and single-clique strategies.
- *Theoretically effective promotion principles* (Section V). We propose the **maximum gain** (resp. the **minimum loss**) principle that provides the target node with the maximum score gain (resp. the minimum score loss) over other nodes. Then, for a specific centrality measure, the aforementioned principles *guide* the selection of an effective strategy: we verify that the strategy meeting a principle is theoretically guaranteed to boost the target node's ranking. Specifically, as given in Table I, the maximum gain principle guides the choice of the multi-point strategy and single-clique for betweenness and coreness, respectively; the minimum loss principle guides the choice of the multi-point strategy and double-line strategy for closeness and eccentricity, respectively.
- *Extensive empirical studies on real-world networks* (Section VII). We conduct experiments to validate the effectiveness of the proposed principle-guided strategies. The experiment results demonstrate that the proposed strategies are valid in improving the centrality ranking on black-box networks.

[2]The problem of ranking improvement on networks with a known structure is studied in [18], but only the hardness result is shown and no algorithm is given.

TABLE I: Principle-Guided Strategies

	Maximum Gain		Minimum Loss	
Centrality	Betweenness	Coreness	Closeness	Eccentrictiy
Strategy	Multi-Point	Single-Clique	Multi-Point	Double-Line

II. Preliminary

A. Notations

Given a graph $G(V_G, E_G)$ with the node set V_G and the edge set $E_G \subseteq V_G^2$, the node number is $n = |V_G|$ and the edge number is $m = |E_G|$. For $\forall v \in V_G$, the neighbors $N_G(v)$ of v are the nodes adjacent to v, i.e., $N_G(v) = \{u|(v,u) \in E_G\}$. Accordingly, the degree of v is defined as the number of nodes in $N_G(v)$ and is denoted as $deg_G(v) = |N_G(v)|$. Given a node set $S \subseteq V_G$, the subgraph induced by S is denoted as $G[S] = (S, E_S)$, where $(u,v) \in E_S$ if, and only if, $u, v \in S$ and $(u,v) \in E$. When inserting additional nodes Δ_V and edges Δ_E in the original graph G, we get an **updated graph** $G'(V', E') = G'(V \cup \Delta_V, E \cup \Delta_E)$.

Given a node pair (s,t) in G, the path $p_G(s,t)$ from s to t is a sequence of nodes, $\langle s = v_0, v_1, \ldots, v_k = t\rangle$, with $(v_i, v_{i+1}) \in E_G$, $i \in [0, k-1]$. The length of $p_G(s,t)$ is the number of edges on $p_G(s,t)$. The path from s to t with the minimum length is defined as the shortest path, whose length is defined as the **shortest distance** and is denoted as $dist_G(s,t)$.

For simplicity, we use V, E, $N(v)$, $deg(v)$, $p(s,t)$, $dist(s,t)$ to denote V_G, E_G, $N_G(v)$, $deg_G(v)$, $p_G(s,t)$, $dist_G(s,t)$ respectively. In this paper, we focus on unweighted and undirected graphs. Moreover, we assume that graphs are connected; otherwise, we work on the largest connected component.

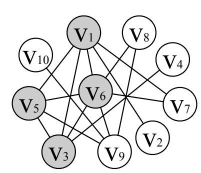


Fig. 1: Graph G

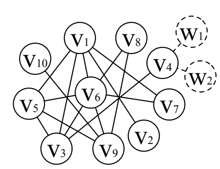


Fig. 2: Updated Graph G'

Example 2.1: *Fig. 1 shows an example graph G. For v_5, $N(v_5) = \{v_1, v_3, v_6, v_9\}$ and $deg(v_5) = |N(v_5)| = 4$. The subgraph $G[S]$ induced by $S = \{v_1, v_3, v_5, v_6\}$ is colored gray in Fig. 1. The neighbor of v_5 in $G[S]$ is $\{v_1, v_3, v_6\}$, and $deg_{G[S]}(v_5) = 3$. A path from v_5 and v_7 in G is $\langle v_5, v_1, v_7\rangle$ with length 2, which has the minimum length among all the paths from v_5 to v_7, and $dist(v_5, v_7) = 2$. Fig. 2 shows an updated graph G' with additional nodes $\Delta_V = \{w_1, w_2\}$ and edges $\Delta_E = \{(w_1, v_4), (w_2, v_4)\}$ attached to G.*

B. Centrality Measures

We concentrate on four types of centrality measures: closeness, eccentricity, betweenness, and coreness.

TABLE II: Description of Symbols

Symbol	Description
G, G'	Graph, and updated graph
$\mathbb{C}(v), \mathbb{C}'(v)$	Centrality score of a node v in G, G'
$\overline{\mathbb{C}}(v), \overline{\mathbb{C}}'(v)$	Reciprocal centrality score of a node v in G, G'
$\mathbb{R}(v), \mathbb{R}'(v)$	Centrality ranking of a node v in G, G'
$\Delta_{\mathbb{C}}(v)$	Score variation of v, i.e. $\mathbb{C}'(v) - \mathbb{C}(v)$
$\overline{\Delta}_{\mathbb{C}}(v)$	Reciprocal score variation of v, i.e. $\overline{\mathbb{C}}'(v) - \overline{\mathbb{C}}(v)$
$\Delta_{\mathbb{R}}(v)$	Ranking variation of v, i.e., $\mathbb{R}(v) - \mathbb{R}'(v)$
$\mathsf{BC}, \mathsf{RC}, \mathsf{CC}, \mathsf{EC}$	Assign $\mathbb{C}$ as betweenness, coreness, closeness, eccentricity
$\overline{\mathsf{BC}}, \overline{\mathsf{RC}}, \overline{\mathsf{CC}}, \overline{\mathsf{EC}}$	Assign $\overline{\mathbb{C}}$ as reciprocal score of $\mathsf{BC}, \mathsf{RC}, \mathsf{CC}, \mathsf{EC}$

Definition 2.1: (**Closeness** [5]) Given a graph $G(V, E)$, the closeness of a node $v \in V$ is $\mathsf{CC}(v) = \frac{1}{\sum_{u \in V} dist(v,u)}$.

Definition 2.2: (**Eccentricity** [2]) Given a graph $G(V, E)$, the eccentricity of a node $v \in V$ is $\mathsf{EC}(v) = \frac{1}{\max_{u \in V} dist(v,u)}$.

Definition 2.3: (**Betweenness** [13]) Given a graph $G(V, E)$, the betweenness of a node $v \in V$ is $\mathsf{BC}(v) = \sum_{(s,t) \in V^2, s \neq t \neq v} \frac{\sigma_v(s,t)}{\sigma(s,t)}$, where $\sigma(s,t)$ is the number of s-t shortest paths, and $\sigma_v(s,t)$ is the number of s-t shortest paths via v.

Definition 2.4: (**Coreness** [15]) Given a graph $G(V, E)$, the coreness of a node $v \in V$, i.e., $\mathsf{RC}(v)$, is the largest k, such that there is a subgraph $G[S]$ includes v, and each node $u \in S$ has a degree not less than k in $G[S]$, that is $deg_{G[S]}(u) \geq k$.

Example 2.2: *For the graph G in Fig. 1, the shortest distance from v_1 to $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_9, v_{10}\}$ are $\{0, 1, 1, 2, 1, 1, 1, 2, 2, 3\}$. Then, the closeness of v_1 is $\mathsf{CC}(v_1) = \frac{1}{0+1+1+2+1+1+1+2+2+3} = \frac{1}{14}$. The eccentricity of v_1 is $\mathsf{EC}(v_1) = \frac{1}{3}$ since 3 is the largest distance from v_1 to other nodes. To obtain the betweenness of v_1, we enumerate all the node pairs in the graph. For example, for pair (v_3, v_7), $\delta(v_3, v_7) = 2$ since there are two shortest paths between them; $\delta_{v_1}(v_3, v_7) = 1$ since there is one v_3-v_7 shortest path via v_1. Then, $\frac{\delta_{v_1}(v_3,v_7)}{\delta(v_3,v_7)} = \frac{1}{2}$ for pair (v_3, v_7). We summarize over all the pairs and obtain $\mathsf{BC}(v_1) = 9.5$. The coreness of v_1 is 3 because there is a subgraph $G[S]$ with nodes $S = \{v_1, v_3, v_5, v_6\}$ that includes v_1. The degree of each node in $G[S]$ is not less than 3, and we cannot find another subgraph $G[S']$ containing v_1, and each node in $G[S']$ has a degree greater than 3.*

III. Problem Formulation

In this section, we first introduce some concepts when a graph G is updated to G', and then the centrality promotion problem is formally defined. For ease of understanding, some commonly used symbols are summarized in Table II.

Score Variation. Given a graph $G(V, E)$, we define the centrality measure $\mathbb{C}$ as a function that maps a node $v \in V$ to the real value $\mathbb{C}(v)$. For example, centrality measures CC, EC, BC, and RC defined in Section II are centrality functions. For $\forall v \in V$, $\mathbb{C}(v)$ is the **centrality score** of v. We define the **reciprocal centrality score** of v as the reciprocal of $\mathbb{C}(v)$ and denote it as $\overline{\mathbb{C}}(v) = \frac{1}{\mathbb{C}(v)}$. When Δ_V and Δ_E are inserted to transform G to an updated graph $G'(V', E') = G'(V \cup \Delta_V, E \cup \Delta_E)$, we denote the centrality score of v in G' as $\mathbb{C}'(v)$ and the reciprocal centrality score as $\overline{\mathbb{C}}'(v) = \frac{1}{\mathbb{C}'(v)}$, for $\forall v \in V'$.

For a node $v \in V$ with $\mathbb{C}(v)$ in G and $\mathbb{C}'(v)$ in G', we define the **score variation** of v as the difference between the centrality score of v in G' and G and denote it as $\Delta_{\mathbb{C}}(v)$, i.e., $\Delta_{\mathbb{C}}(v) = \mathbb{C}'(v) - \mathbb{C}(v)$. We also define the **reciprocal score variation** of v as the difference between the reciprocal centrality score of v in G' and G and denote it as $\overline{\Delta}_{\mathbb{C}}(v)$, i.e., $\overline{\Delta}_{\mathbb{C}}(v) = \overline{\mathbb{C}}'(v) - \overline{\mathbb{C}}(v)$. For a node $w \in \{V' \setminus V\}$, or $w \in \Delta_V$, w does not appear in G, and we denote $\mathbb{C}(w) = 0$ and $\overline{\mathbb{C}}(w) = \emptyset$[3]. Then we have $\Delta_{\mathbb{C}}(w) = \mathbb{C}'(w)$ and $\overline{\Delta}_{\mathbb{C}}(w) = \overline{\mathbb{C}}'(w)$.

TABLE III: Score Variations and Ranking Variations (CC)

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	w_1	w_2
$\mathsf{CC}(v)$	$\frac{1}{14}$	$\frac{1}{22}$	$\frac{1}{15}$	$\frac{1}{23}$	$\frac{1}{14}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{16}$	$\frac{1}{24}$	0	0
$\mathsf{CC}'(v)$	$\frac{1}{20}$	$\frac{1}{30}$	$\frac{1}{19}$	$\frac{1}{25}$	$\frac{1}{20}$	$\frac{1}{18}$	$\frac{1}{26}$	$\frac{1}{26}$	$\frac{1}{24}$	$\frac{1}{34}$	$\frac{1}{35}$	$\frac{1}{35}$
$\mathbb{R}(v)$	2	8	4	9	2	1	6	6	5	10	11	11
$\mathbb{R}'(v)$	3	9	2	6	3	1	7	7	5	10	11	11

Example 3.1: *Table III shows the closeness of nodes in Fig. 1 (denoted as G) and in Fig. 2 (denoted as G'). Here, $\mathbb{C}$ is set to CC. For node v_4 in V, $\mathsf{CC}(v_4) = \frac{1}{23}$ and $\overline{\mathsf{CC}}(v_4) = 23$; $\mathsf{CC}'(v_4) = \frac{1}{25}$ and $\overline{\mathsf{CC}}'(v_4) = 25$. Therefore, $\Delta_{\mathbb{C}}(v_4) = \mathsf{CC}'(v_4) - \mathsf{CC}(v_4) = \frac{1}{25} - \frac{1}{23} = -\frac{2}{575}$, $\overline{\Delta}_{\mathbb{C}}(v_4) = \overline{\mathsf{CC}}'(v_4) - \overline{\mathsf{CC}}(v_4) = 25 - 23 = 2$. For node $w_1 \in \Delta_V$, $\mathsf{CC}(w_1) = 0$, $\Delta_{\mathbb{C}}(w_1) = \mathsf{CC}'(w_1) = \frac{1}{35}$; $\overline{\mathsf{CC}}(w_1) = \emptyset$ and $\overline{\Delta}_{\mathbb{C}}(w_1) = \overline{\mathsf{CC}}'(w_1) = 35$.*

Ranking Variation. Given a graph $G(V, E)$, we define the **centrality ranking** $\mathbb{R}(v)$ of a node $v \in V$ as the position of $\mathbb{C}(v)$ in the ordered centrality scores of nodes in V (sorted in non-increasing order), i.e., $\mathbb{R}(v) = |\{u | u \in V, \mathbb{C}(u) > \mathbb{C}(v)\}| + 1$, where $\mathbb{R}(v) = 1$ means v has the highest score while $\mathbb{R}(v) = |V|$ implies v has the smallest score. When converting $G(V, E)$ to $G'(V', E')$, we denote the centrality ranking of $v \in V'$ as $\mathbb{R}'(v)$, i.e., $\mathbb{R}'(v) = |\{u | u \in V', \mathbb{C}'(u) > \mathbb{C}'(v)\}| + 1$.

For a node $v \in V$ with $\mathbb{R}(v)$ in G and $\mathbb{R}'(v)$ in G', we define the **ranking variation** of v as the difference between the ranking of v in G and G' and denote it as $\Delta_{\mathbb{R}}(v)$, that is, $\Delta_{\mathbb{R}}(v) = \mathbb{R}(v) - \mathbb{R}'(v)$. For a node $w \in \Delta_V$, w does not appear in G, and we derive $\mathbb{R}(w)$ by assuming $\mathbb{C}(v) = 0$ in G, thereby obtaining the ranking variation of w.

Example 3.2: *In Table III, for node $v_4 \in V$, $\mathbb{R}(v_4) = 9$ and $\mathbb{R}'(v_4) = 6$. Thus, $\Delta_{\mathbb{R}}(v_4) = \mathbb{R}(v_4) - \mathbb{R}'(v_4) = 9 - 6 = 3$. For node $w_1 \in \Delta_V$, $\mathbb{R}(w_1) = 11$ since we assume $\mathbb{C}(w_1) = 0$ in G; $\mathbb{R}'(w_1) = 11$ in G'. Hence, $\Delta_{\mathbb{R}}(w_1) = \mathbb{R}(w_1) - \mathbb{R}'(w_1) = 0$.*

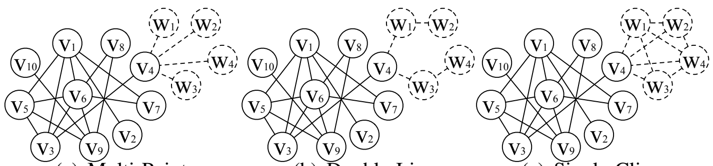
Problem Definition. This paper focuses on exploring whether it is possible to improve the centrality ranking of a target node by adopting a promotion strategy. A **promotion strategy** is an operation to transform $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ by inserting a certain structure (composed of Δ_E and (or) Δ_V) into G. For a target node $t \in V$, a strategy is **effective** if the centrality ranking of t is upgraded, that is $\Delta_{\mathbb{R}}(t) = \mathbb{R}(t) - \mathbb{R}'(t) > 0$. Formally, the problem studied in this paper is the following.

> Given a black-box network $G(V, E)$, a target node $t \in V$, and a centrality measure $\mathbb{C}$, we aim to investigate whether there exists a promotion strategy to convert $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, thereby improving the centrality ranking of t to make $\Delta_{\mathbb{R}}(t) > 0$.

IV. Practically Usable Promotion Strategies

This section provides practically usable promotion strategies. The promotion strategy is defined as an operation that modifies G to an updated graph G'. Note that we assume *the network structure is unknown*. To meet this requirement, we attach a structure (i.e., Δ_V, Δ_E) around the target node t to reshape G

[3]For $\forall w \in \Delta_V$, setting $\overline{\mathbb{C}}(w) = \emptyset$ may bring ambiguity. But this equation is only for convention's sake, and we will not use the value of $\overline{\mathbb{C}}(w)$ in the sequel.



(a) Multi-Point (b) Double-Line (c) Single-Clique

Fig. 3: Strategies for v_4

while avoiding changing the original topology of nodes in G. To model promotion strategies with this requirement, we engage a triple tuple [$\mathsf{target\ node}, \mathsf{promotion\ size}, \mathsf{type}$] $= [t, p, T]$.

- $\mathsf{target\ node}$ is the node t to be upgraded.
- $\mathsf{promotion\ size}$ is the number p of inserted nodes Δ_V, that is, $p = |\Delta_V|$.
- type specifies the structure T of nodes within Δ_V, and from T, we derive Δ_E.

Different assignments of type lead to different types of promotion strategies. We will introduce three strategies where type is assigned as multiple points, double lines, and a single clique, respectively.

Algorithm 1: Multi-Point Strategy.

Input: Graph $G(V, E)$, $[t, p, \mathsf{multiple\ points}]$
Output: Graph $G'(V \cup \Delta_V, E \cup \Delta_E)$
1 $\Delta_V \leftarrow p$ points with no inter-connection;
2 **for** *each node* $w \in \Delta_V$ **do**
3 $\quad \Delta_E \leftarrow (\Delta_E \cup (t, w))$;
4 **return** $G'(V \cup \Delta_V, E \cup \Delta_E)$

<u>*Multi-Point Strategy.*</u> If type is assigned as "$\mathsf{multiple\ points}$", we get a multi-point strategy, whose detail is given in Algorithm 1. First, p nodes without interconnection are introduced as Δ_V (Line 1). Then, for each node $w \in \Delta_V$, we insert an edge (w, t) between w to the target node t to form Δ_E (Line 2-3).

Example 4.1: *For graph G in Fig. 1, the updated graph G' using the multiple-point strategy* $[v_4, 4, \mathsf{multiple\ points}]$ *is shown in Fig. 3(a). In G', four points $\{w_1, w_2, w_3, w_4\}$ that are not interconnected are attached to the target node v_4.*

Algorithm 2: Double-Line Strategy.

Input: Graph $G(V, E)$, $[t, p, \mathsf{double\ lines}]$
Output: Graph $G'(V \cup \Delta_V, E \cup \Delta_E)$
1 $\Delta_V \leftarrow p$ points with no inter-connection;
2 Divide Δ_V into two disjoint subsets $S1$ and $S2$ with equal sizes;
3 **for** $w_i \in S_1$, *where* $i = 1 : |S1| - 1$ **do**
4 $\quad$ **if** $i = 1$ **then**
5 $\quad\quad \Delta_E \leftarrow (\Delta_E \cup (t, w_i))$;
6 $\quad \Delta_E \leftarrow (\Delta_E \cup (w_i, w_{i+1}))$;
7 **for** $w_i \in S_2$, *where* $i = 1 : |S_2| - 1$ **do**
8 $\quad$ **if** $i = 1$ **then**
9 $\quad\quad \Delta_E \leftarrow (\Delta_E \cup (t, w_i))$;
10 $\quad \Delta_E \leftarrow (\Delta_E \cup (w_i, w_{i+1}))$;
11 **return** $G'(V \cup \Delta_V, E \cup \Delta_E)$

<u>*Double-Line Strategy.*</u> If type is assigned as "$\mathsf{double\ lines}$", we get a double-line strategy, whose detail is given in Algorithm 2. First, p nodes constitute Δ_V (Line 1). Then, we divide the inserted nodes Δ_V into two equal-size[4] subsets S_1 and S_2 such that $|S_1| = |S_2|$, $S_1 \cup S_2 = \Delta_V$, $S_1 \cap S_2 = \emptyset$ (Line 2). For S_1 (Line 3), we connect the first node $w_1 \in S_1$ to t (Line 4-5), and for other nodes but the last node $w_{|S_1|}$, we connect w_i to w_{i+1} thus to form a line (Line 6); the same procedure happens for nodes in S_2 (Line 7-10).

[4]when p is odd, we make $|S_1| - |S_2| = 1$. We omit this situation for simplicity.

Example 4.2: *For graph G in Fig. 1, the updated graph G' using the double-line strategy* $[v_4, 4, \mathsf{double\ lines}]$ *is shown in Fig. 3(b). In G', two lines $\{w_1, w_2\}$ and $\{w_3, w_4\}$ are attached to the target node v_4.*

Algorithm 3: Single-Clique Strategy.

Input: Graph $G(V, E)$, $[t, p, \mathsf{single\ clique}]$
Output: Graph $G'(V \cup \Delta_V, E \cup \Delta_E)$
1 $\Delta_V \leftarrow p$ points with no inter-connection;
2 **for** $w_i \in \Delta_V$, *where* $i = 1 : |\Delta_V|$ **do**
3 $\quad \Delta_E \leftarrow (\Delta_E \cup (w_i, t))$;
4 $\quad$ **for** $w_j \in \Delta_V$, *where* $j = 1 : |\Delta_V|$ **do**
5 $\quad\quad$ **if** $i < j$ **then**
6 $\quad\quad\quad \Delta_E \leftarrow (\Delta_E \cup (w_i, w_j))$;
7 **return** $G'(V \cup \Delta_V, E \cup \Delta_E)$

<u>*Single-Clique Strategy.*</u> If type is assigned as "$\mathsf{single\ clique}$", we get a single-clique strategy, as shown in Algorithm 3. First, p nodes establish Δ_V (Line 1). Then, for $\forall w_i \in \Delta_V$ (Line 2), we connect w_i to t (Line 3) and all other nodes in Δ_V.

Example 4.3: *For graph G in Fig. 1, the updated graph G' using the single-clique strategy* $[v_4, 4, \mathsf{single\ clique}]$ *is shown in Fig. 3(c). In G', a clique (i.e., complete graph) with $\{w_1, w_2, w_3, w_4\}$ and the target node v_4 is formed.*

Remark 1: This paper investigates whether a black-box network can be manipulated for centrality promotion. To this end, *we keep the promotion strategies simple enough* to emphasize that the black-box network is easy to control even with straightforward strategies. There are other equally important topics, such as the detectability of strategies and the maximal promotion effect under certain budgets. However, all these topics are underpinned by our research — only when it is possible to raise the centrality ranking on black-box networks are these topics of research significance. We leave these topics as future work.

V. THEORETICALLY EFFECTIVE PROMOTION PRINCIPLES

Section IV introduces several practical promotion strategies. But for a specific centrality measure, how to elect *an* effective promotion strategy to guarantee that the target node's ranking will be improved? We respond to this question by giving two principles in Section V-A as a guide for choosing the strategy. Then, we provide examples of how these principles are exploited in some commonly used centrality measures. For simplicity, the proofs in this section are moved to the Appendix.

A. Promotion Principles

<u>*Maximum Gain Principle.*</u> The mechanism of the first principle is straightforward yet non-trivial: when inserting additional nodes into G using a strategy, we try to increase the centrality scores of all nodes in G to a different extent. When the growth of the centrality scores on the target node t is more significant than those of all other nodes, the final scores of the target node t will potentially exceed other high centrality nodes — thus, resulting in an effective ranking upgrade. This scheme is called the maximum gain principle, which is formally defined below.

Definition 5.1: (**Maximum Gain Principle**) Given a graph $G(V, E)$, a promotion strategy $[t, p, T]$ that converts $G(V, E)$ to an updated graph $G'(V \cup \Delta_V, E \cup \Delta_E)$ fulfills the maximum gain principle for a centrality measure $\mathbb{C}$ if

- (Maximum Property) for an arbitrary size p, $\Delta_{\mathbb{C}}(t) \geq \Delta_{\mathbb{C}}(v) \geq 0$, for $\forall v \in V$;
- (Dominance Property) $\mathbb{C}'(t) \geq \mathbb{C}'(w)$, for $\forall w \in \Delta_V$;
- (Boost Property) there exists a size p' such that when $p > p'$, $\mathbb{C}'(t) > \mathbb{C}'(v)$, for some node $v \in V$ with $\mathbb{C}(v) > \mathbb{C}(t)$.

We now interpret this principle. For nodes v with scores no larger than that of t in G, the *maximum property* ensures the scores of nodes v will not overtake t in the updated graph G'; the *boost property* guarantees that in G', the score of t will exceed at least one node whose score is larger than t in G. These two properties ensure the ranking of t will be advanced by at least one among the nodes in V. The dominance property forces the ranking of nodes w in Δ_V not to exceed t in G'. As a result, the ranking of t in G' will definitely be upgraded compared to that in G. The effectiveness is formally presented in Theorem 5.1.

Theorem 5.1: *If a strategy $[t, p, T]$ fulfills the maximum gain principle for a centrality measure $\mathbb{C}$, then it converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ to make $\Delta_{\mathbb{R}}(t) > 0$.*

Minimum Loss Principle. The mechanism of the second principle is similar to the first one but with a reversed logic: when inserting nodes into G, we try to decrease the centrality scores of all nodes in G to different degrees. When the decline of the centrality scores on the target node t is less significant than all other nodes, the final scores of the target node t will potentially exceed the other high centrality nodes in G, thus, producing an effective ranking upgrade. This scheme is denoted as the minimum loss principle and is formally defined below.

Definition 5.2: (**Minimum Loss Principle**) Given a graph $G(V, E)$, a promotion strategy $[t, p, T]$ that converts $G(V, E)$ to an updated graph $G'(V \cup \Delta_V, E \cup \Delta_E)$ fulfills the minimum loss principle for a centrality measure $\mathbb{C}$ if

- (Minimum Property) for an arbitrary size p, $\overline{\Delta}_{\mathbb{C}}(v) \geq \overline{\Delta}_{\mathbb{C}}(t) \geq 0$[5], for $\forall v \in V$;
- (Dominance Property) $\mathbb{C}'(t) \geq \mathbb{C}'(w)$, for $\forall w \in \Delta_V$;
- (Boost Property) there exists a size p' such that when $p > p'$, $\mathbb{C}'(t) > \mathbb{C}'(v)$, for some node $v \in V$ with $\mathbb{C}(v) > \mathbb{C}(t)$.

The interpretation of the minimum loss principle resembles that of the maximum gain principle. The *minimum property* guarantees that the ranking of nodes with scores no larger than t are not ranked higher than t in G', while the *boost property* ensures that the ranking of t will be advanced by at least one. The *dominance property* ensures the ranking of t is not lower than the nodes in Δ_V. Together, the ranking of t will be improved definitely in G', which is given in Theorem 5.2.

Theorem 5.2: *If a strategy $[t, p, T]$ fulfills the minimum loss principle for a centrality measure $\mathbb{C}$, then it converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ to make $\Delta_{\mathbb{R}}(t) > 0$.*

Remark 2: Given a strategy $[t, p, T]$, for both principles, the first two properties are fulfilled for every promotion size p. However, the boost property only requires the existence of a certain size p'. We can theoretically compute the value of p' (see Lemma 5.3, Lemma 5.6, Lemma 5.9, and Lemma 5.12), with which the ranking of the target node will definitely increase. Empirically, in Section VII, we observe that a small size (e.g., 16) is sufficient to ensure an effective improvement using our proposed strategies.

[5]When G is modified to G', a positive $\overline{\Delta}_{\mathbb{C}}(v)$ means a decrease in the score of v, for $\forall v \in V$.

TABLE IV: Example of Maximum Gain Principle (BC)

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	w_1	w_2	w_3	w_4
$\mathsf{BC}(v)$	9.5	0	8	0	4	13	0	0	8.5	0	0	0	0	0
$\mathbb{R}(v)$	2	6	4	6	5	1	6	6	3	6	6	6	6	6
$\mathsf{BC}'(v)$	15.5	0	40	42	8	23	0	0	12.5	0	0	0	0	0
$\mathbb{R}'(v)$	4	7	2	1	6	3	7	7	5	7	7	7	7	7

Usage of Promotion Principles. We discuss how to select from the above two principles for a specific centrality measure. We apply the maximum gain principle when the insertion of new nodes does not reduce the centrality score of any node in G. For example, centrality measures such as betweenness and coreness possess this feature (see [26] and references therein). Their promotions will be introduced in Section V-B.

We adopt the minimum loss principle when the insertion of new nodes does not increase the centrality score of any node in G. Centrality measures such as closeness and eccentricity admit this characteristic (see [26] and references therein). Their upgrades will be presented in Section V-C.

B. Maximum Gain Principle-Guided Strategies

For centrality measures (e.g., betweenness and coreness) where the maximum gain principle takes effect, Theorem 5.1 explains that the strategy meeting the maximum gain principle will provide an effective promotion. Specifically, the maximum gain principle guides the selection of the multi-point strategy for betweenness and the single-clique strategy for coreness.

Multi-Point Strategy for Betweenness Centrality. We justify that a multi-point strategy is effective for betweenness promotion by verifying that it satisfies the three properties given in the maximum gain principle (Definition 5.1). Specifically, Lemma 5.1 verifies the maximum property, Lemma 5.2 validates the dominance property, and Lemma 5.3 justifies the boost property.

Lemma 5.1: $[t, p, \textsf{multiple points}]$ *that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ fulfills the maximum property for* BC*, that is, for an arbitrary size p, $\Delta_{\mathbb{C}}(t) \geq \Delta_{\mathbb{C}}(v) \geq 0$, for $\forall v \in V$.*

Lemma 5.2: $[t, p, \textsf{multiple points}]$ *that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ fulfills the dominance property for* BC*, that is, for an arbitrary size p, $\mathsf{BC}'(t) \geq \mathsf{BC}'(w)$, for $\forall w \in \Delta_V$.*

Lemma 5.3: $[t, p, \textsf{multiple points}]$ *that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$ fulfills the boost property for* BC*, that is, there exists a size $p' = \sqrt{\mathsf{BC}(v) - \mathsf{BC}(t)} + 1$ such that when $p > p'$, $\mathsf{BC}'(t) > \mathsf{BC}'(v)$, for some node $v \in V$ with $\mathsf{BC}(v) > \mathsf{BC}(t)$.*

Based on Lemmas 5.1-5.3 and Theorem 5.1, the effectiveness of the multi-point strategy for BC is derived in Theorem 5.3.

Theorem 5.3: *The multi-point strategy fulfills the maximum gain principle for betweenness centrality* (BC)*. Hence, given a graph $G(V, E)$, $[t, p, \textsf{multiple points}]$ converts $G(V, E)$ to an updated graph $G'(V \cup \Delta_V, E \cup \Delta_E)$ to make $\Delta_{\mathbb{R}}(t) > 0$.*

Example 5.1: *For betweenness promotion, Table IV shows the effect of the multi-point strategy $[v_4, 4, \textsf{multiple points}]$ that transforms G in Fig. 1 to G' in Fig. 3(a). (i) The maximum property holds since $\Delta_{\mathbb{C}}(v_4) = 42$ is the maximum for nodes in V. (ii) The dominance property holds since $\mathsf{BC}'(w_1) = \mathsf{BC}'(w_2) = \mathsf{BC}'(w_3) = \mathsf{BC}'(w_4) = 0 < \mathsf{BC}'(v_4)$. (iii) The boost property holds since for v_5 with $\mathsf{BC}(v_5) > \mathsf{BC}(v_4)$, we have $\mathsf{BC}'(v_5) < \mathsf{BC}'(v_4)$ (when the size $p = 4 > p' = \sqrt{\mathsf{BC}(v_5) - \mathsf{BC}(v_4)} + 1 = 3$). Moreover, v_4 now becomes the node with the highest betweenness in the updated graph and its ranking variation $(6 - 1 = 5)$ is larger than zero.*

TABLE V: Example of Minimum Loss Principle (CC)

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	w_1	w_2	w_3	w_4
$\overline{\mathsf{CC}}(v)$	14	22	15	23	14	12	18	18	16	24	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
$\mathbb{R}(v)$	2	8	4	9	2	1	6	6	5	10	11	11	11	11
$\overline{\mathsf{CC}}'(v)$	26	38	23	27	26	24	34	34	32	44	39	39	39	39
$\mathbb{R}'(v)$	3	9	1	5	3	2	7	7	6	14	10	10	10	10

Single-Clique Strategy for Coreness Centrality. We verify that a single-clique strategy satisfies the maximum gain principle for coreness promotion by Lemma 5.4 (maximum property), Lemma 5.5 (dominance property), and Lemma 5.6 (boost property).

Lemma 5.4: $[t, p, \mathsf{single\ clique}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the maximum property for* RC*, that is, for an arbitrary size* p, $\Delta_{\mathbb{C}}(t) \geq \Delta_{\mathbb{C}}(v) \geq 0$, *for* $\forall v \in V$.

Lemma 5.5: $[t, p, \mathsf{single\ clique}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the dominance property for* RC*, that is, for an arbitrary size* p, $\mathsf{RC}'(t) \geq \mathsf{RC}'(w)$, *for* $\forall w \in \Delta_V$.

Lemma 5.6: $[t, p, \mathsf{single\ clique}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the boost property for* RC*, that is, there exists a size* $p' = \mathsf{RC}(v) + 1$ *such that when* $p > p'$, $\mathsf{RC}'(t) > \mathsf{RC}'(v)$, *for some node* $v \in V$ *with* $\mathsf{RC}(v) > \mathsf{RC}(t)$.

Based on Lemmas 5.4-5.6 and Theorem 5.1, the effectiveness of the single-clique strategy for RC is derived in Theorem 5.4.

Theorem 5.4: *The single-clique strategy fulfills the maximum gain principle for coreness centrality* (RC). *Hence, given a graph* $G(V, E)$, $[t, p, \mathsf{single\ clique}]$ *converts* $G(V, E)$ *to an updated graph* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *to make* $\Delta_{\mathbb{R}}(t) > 0$.

C. Minimum Loss Principle-Guided Strategies

For centrality measures (e.g., closeness and eccentricity) where the minimum loss principle takes effect, Theorem 5.2 explains that the strategy meeting the minimum loss principle will provide an effective promotion. Specifically, the minimum loss principle guides the selection of the multi-point strategy for closeness and the double-line strategy for eccentricity.

Multi-Point Strategy for Closeness Centrality. We demonstrate that a multi-point strategy is effective for closeness improvement by verifying it meets the three properties given in the minimum loss principle (Definition 5.2). Specifically, Lemma 5.7 justifies the minimum property, Lemma 5.8 verifies the dominance property, and Lemma 5.9 validates the boost property.

Lemma 5.7: $[t, p, \mathsf{multiple\ points}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the minimum property for* CC*, that is, for an arbitrary size* p, $\overline{\Delta}_{\mathbb{C}}(v) \geq \overline{\Delta}_{\mathbb{C}}(t) \geq 0$, *for* $\forall v \in V$.

Lemma 5.8: $[t, p, \mathsf{multiple\ points}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the dominance property for* CC*, that is, for an arbitrary size* p, $\mathsf{CC}'(t) \geq \mathsf{CC}'(w)$, *for* $\forall w \in \Delta_V$.

Lemma 5.9: $[t, p, \mathsf{multiple\ points}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the boost property for* CC*, that is, there exists a size* $p' = \frac{\overline{\mathsf{CC}}(t) - \overline{\mathsf{CC}}(v)}{dist_G(v,t)}$ *such that when* $p > p'$, $\mathsf{CC}'(t) > \mathsf{CC}'(v)$, *for some node* $v \in V$ *with* $\mathsf{CC}(v) > \mathsf{CC}(t)$.

Based on Lemmas 5.7-5.9 and Theorem 5.2, the effectiveness of the multi-point strategy for CC is derived in Theorem 5.5.

Theorem 5.5: *The multi-point strategy fulfills the minimum loss principle for closeness centrality* (CC). *Hence, given a graph* $G(V, E)$, $[t, p, \mathsf{multiple\ points}]$ *converts* $G(V, E)$ *to an updated graph* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *to make* $\Delta_{\mathbb{R}}(t) > 0$.

Example 5.2: *For closeness promotion, Table V shows the effect of the multi-point strategy* $[v_4, 4, \mathsf{multiple\ points}]$ *that transforms* G *in Fig. 1 to* G' *in Fig. 3(a). (i) The minimum property holds since* $\overline{\Delta}_{\mathbb{C}}(v_4) = 4$ *is the minimum for nodes in* V. *(ii) The dominance property holds since* $\mathsf{CC}'(w_1) = \mathsf{CC}'(w_2) = \mathsf{CC}'(w_3) = \mathsf{CC}'(w_4) = \frac{1}{39} < \mathsf{CC}'(v_4)$. *(iii) The boost property holds since for* v_2 *with* $\mathsf{CC}(v_2) > \mathsf{CC}(v_4)$, *we have* $\mathsf{CC}'(v_2) < \mathsf{CC}'(v_4)$ *(when the size* $p = 4 > p' = \frac{\overline{\mathsf{CC}}(v_4) - \overline{\mathsf{CC}}(v_2)}{dist_G(v_4,t_2)} = \frac{1}{3}$). *Moreover, the positive ranking variation* $(9 - 5 = 4)$ *of* v_4 *confirms the effectiveness.*

Double-Line Strategy for Eccentricity Centrality. We reveal that a double-line strategy satisfies the minimum loss principle for eccentricity promotion by Lemma 5.10 (minimum property), Lemma 5.11 (dominance property), and Lemma 5.12 (boost property).

Lemma 5.10: $[t, p, \mathsf{double\ lines}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the minimum property for* EC*, that is, for an arbitrary size* p, $\overline{\Delta}_{\mathbb{C}}(v) \geq \overline{\Delta}_{\mathbb{C}}(t) \geq 0$, *for* $\forall v \in V$.

Lemma 5.11: $[t, p, \mathsf{double\ lines}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the dominance property for* EC*, that is, for an arbitrary size* p, $\mathsf{EC}'(t) \geq \mathsf{EC}'(w)$, *for* $\forall w \in \Delta_V$.

Lemma 5.12: $[t, p, \mathsf{double\ lines}]$ *that converts* $G(V, E)$ *to* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *fulfills the boost property for* EC*, that is, there exists a size* $p' = 2 \times \mathsf{EC}(t)$ *such that when* $p > p'$, $\mathsf{EC}'(t) > \mathsf{EC}'(v)$, *for some node* $v \in V$ *with* $\mathsf{EC}(v) > \mathsf{EC}(t)$.

Based on Lemmas 5.10-5.12 and Theorem 5.2, the effectiveness of the double-line strategy for EC is derived in Theorem 5.6.

Theorem 5.6: *The double-line strategy fulfills the minimum loss principle for eccentricity centrality* (EC). *Hence, given a graph* $G(V, E)$, $[t, p, \mathsf{double\ lines}]$ *converts* $G(V, E)$ *to an updated graph* $G'(V \cup \Delta_V, E \cup \Delta_E)$ *to make* $\Delta_{\mathbb{R}}(t) > 0$.

VI. Related Work

A. Centrality Promotion

Promotion with Known Network Structures. Existing studies typically modify the original graph (by adding edges) when network structures are known. Bergamini et al. [18] presented the hardness result of maximizing the betweenness score and devised a greedy algorithm. Given a coreness score, the issue of maximizing the number of nodes whose coreness is no smaller than that given score has been studied in [19]. In addition, Crescenzi et al. [9] provided hardness and algorithmic results to improve a target node's closeness score. Furthermore, approximation algorithms have been designed in [20] to maximize a target node's eccentricity score.

While existing solutions apply to network owners who have a complete view of network structures, we attack this problem from the perspective of network users who do not have access to the entire network structure, that is, users plan to modify the graph without referring to the network structure, thereby influencing the centrality calculated by the network owners. To the best of our knowledge, this is the first work to explore centrality promotion on a black-box network.

Greedy Algorithm for Betweenness Promotion. As an example of how greedy algorithms are exploited in boosting centrality, we present the algorithm developed in [18], which is used to increase the betweenness score and is denoted as Greedy. Given a graph $G(V, E)$, a target node t, and a budget b, Greedy works in b rounds. In each round, it selects an edge from $\{V^2 \setminus E\}$. The output of Greedy is the selected b edges (denoted as B). The specific procedure of Greedy is detailed as follows:

1) Initialize $G'(V, E)$ as $G(V, E)$ and compute $\mathsf{BC}(t)$ on G.
2) For each node $v \in \{V' \setminus N_{G'}(t)\}$,
   a) Add (v, t) into G' temporarily and compute $\mathsf{BC}'(t)$ on G'.
   b) Compute $\Delta_{\mathbb{C}}(t|v) = \mathsf{BC}'(t) - \mathsf{BC}(t)$ and then remove (v, t) from G'.
3) Select the node v with the largest $\Delta_{\mathbb{C}}(t|v)$ and insert (v, t) into both G and B.
4) Stop after b rounds; otherwise go to step 1).

In step 1), we initialize the graph G' as G and calculate $\mathsf{BC}(t)$ of t on G. In Step 2), for each node v that is not directly connected to t in G', i.e., $v \in \{V' \setminus N_{G'}(t)\}$, we temporarily add the edge (v, t) into G'. We then calculate $\mathsf{BC}'(t)$ and obtain the betweenness improvement $\Delta_{\mathbb{C}}(t|v) = \mathsf{BC}'(t) - \mathsf{BC}(t)$ of t on G', followed by removing this edge. Then, node v with the maximum $\Delta_{\mathbb{C}}(t|v)$ is selected and we insert the edge (v, t) into both G and the answer set B. This process stops after b rounds; if not, we go to step 1) for a new iteration.

Note that in step 2), Greedy needs to compute $\Delta_{\mathbb{C}}(t|v)$ (for $\forall v \in \{V' \setminus N_{G'}(t)\}$), which necessitates the knowledge of the network structure to calculate $\mathsf{BC}'(t)$ and $\mathsf{BC}(t)$. Instead, our method does not involve any computation on the network and thus is feasible when the network structure is unknown. We compare our method with Greedy in Section VII.

B. Centrality Measures

There are other centrality measures in the literature [4]. The harmonic centrality of a node is the sum of the reciprocal shortest distance from this node to all other nodes [27]. Katz centrality determines the node's importance by evaluating the number of nodes that can reach it through a path, with a penalization on the path length [28]. Current-flow betweenness centrality evaluates node importance by applying the electrical current model for information spreading [13]. Our principles are applicable to guide the selection of suitable strategies for these centrality measures (see Section V-A for guidance) — provided the strategies conform to the proposed principles, they are guaranteed to be effective.

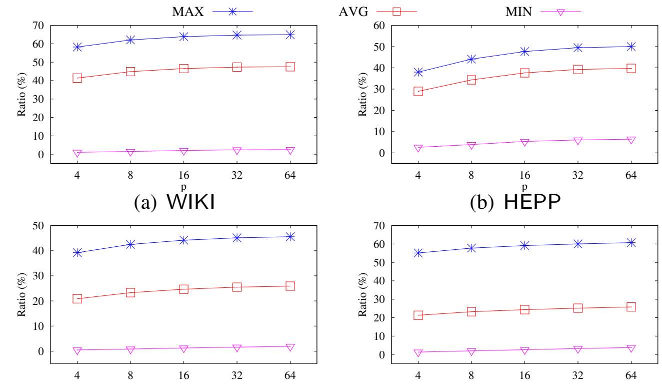
VII. Experiments

This section evaluates the performance of principle-guided strategies on real-world graphs. We first present some experimental settings, and then test the maximum gain principle and minimum loss principle in Section VII-A and Section VII-B, respectively, followed by a comparison with the greedy algorithm in Section VII-C.

TABLE VI: Description of Datasets

Name	Dataset	n	m	Diameter	Degeneracy
WIKI	Wiki-vote	7,066	100,736	7	53
HEPP	CA-HepPh	11,204	117,619	13	238
EPIN	Epinions	75,877	405,739	15	67
SLAS	Slashdot	77,360	469,180	12	54

<u>*Dataset.*</u> We conducted experiments on four real-world networks[6]. The details of these datasets are given in Table VI. We assume the graphs are undirected. If not, the edge directions are ignored. For a disconnected graph, we performed experiments on the largest connected component. Table VI presents the node number (n) and edge number (m) of the largest connected component of a graph G. The diameter of G is the largest reciprocal eccentricity score of all nodes, i.e., $\max_{v \in V}(\overline{\mathsf{EC}}(v))$ [29]. The degeneracy of G is the largest coreness score of all nodes, i.e., $\max_{v \in V}(\mathsf{RC}(v))$ [15]. The algorithms used in this paper were implemented using NetworkX [30] and teexGraph[7] [29].



(a) WIKI (b) HEPP (c) EPIN (d) SLAS

Fig. 4: Relative Ranking Variations (BC)

<u>*Metric.*</u> We propose strategies to promote a centrality measure $\mathbb{C}$. To assess the effectiveness of the proposed strategies in upgrading centrality ranking, a simple metric is the ranking variation $\Delta_{\mathbb{R}}(t)$ for the target node t. However, $\Delta_{\mathbb{R}}(t)$ is an absolute value that does not consider the graph size; a desirable alternative would be the relative ranking variation *Ratio*, which is the ratio of $\Delta_{\mathbb{R}}(t)$ to the node number n, that is,

$$Ratio = \frac{\Delta_{\mathbb{R}}(t)}{n} \times 100\%.$$

A. Testing Maximum Gain Principle

We examine strategies that satisfy the maximum gain principle: the multi-point strategy for betweenness and the single-clique strategy for coreness. For all experiments, each time we randomly selected one target node from the graph. We repeated the process ten times to report the results. Specifically, the same experiment was conducted ten times, each time for a selected target node. For each target node, we set the promotion size p from 4, 8, 16, 32, to 64 to insert Δ_V (and Δ_E).

<u>*Exp 1: Betweenness Centrality Promotion.*</u> This set of experiments confirms that the multi-point strategy satisfies the three properties of the maximum gain principle to promote betweenness (BC) on real graphs.

<u>*Exp 1-1: Maximum Property.*</u> Recall that the maximum property indicates that target node t has a score variation no smaller than that of the other nodes in V. To this end, we compare the score variations between target nodes t and node v, whose score variation is maximal among the nodes in $\{V \setminus t\}$. Due to space limitations, we only present the results for five target nodes (by assigning new IDs from 1 to 5) on WIKI and HEPP in Table VII — each row of Table VII is for one target node, and each target node represents an independent experiment.

It can be found that on both datasets and at various sizes p, the score variation of target node t is larger than v. For example, on WIKI and when the size p is 4, the score variation is 56532 for t (with ID 1), while the score variation of v is 3512.2: the

[6]downloaded from http://snap.stanford.edu/data/index.html

[7]https://github.com/franktakes/teexgraph

TABLE VII: Score Variations of V (BC)

Dataset	ID	4		8		16		32		64	
		t	v	t	v	t	v	t	v	t	v
WIKI	1	56,532	3,512.2	113,096	7,046.4	226,320	14,092.8	453,152	28,185.5	908,352	56,371.1
	2	56,532	46,979.1	113,096	93,958.2	226,320	187,916.5	453,152	375,833	908,352	751,666
	3	56,532	20,037	113,096	40,073.9	226,320	80,147.9	453,152	160,295.8	908,352	320,591.6
	4	56,532	11,867.3	113,096	23,734.6	226,320	47,469.3	453,152	94,938.5	908,352	189,877
	5	56,532	9,493.3	113,096	18,986.5	226,320	37,973	453,152	75,946	908,352	151,892
HEPP	1	89,636	7,737.7	179,304	15,475.3	358,736	30,950.7	717,984	61,901.4	1,438,016	123,802.8
	2	89,636	8,241.5	179,304	16,483.1	358,736	32,966.1	717,984	65,932.2	1,438,016	131,864.5
	3	89,636	89,616	179,304	179,232	358,736	358,464	717,984	713,628	1,438,016	1,433,856
	4	89,636	50,027.8	179,304	100,055.5	358,736	200,111.1	717,984	400,222.2	1,438,016	800,444.4
	5	89,636	75,419.4	179,304	150,838.8	358,736	301,677.5	717,984	603,355.1	1,438,016	1,206,710.2

TABLE VIII: Scores of Target Nodes and Δ_V (BC)

Dataset	ID	4		8		16		32		64	
		t	w	t	w	t	w	t	w	t	w
WIKI	1	179,818.3	0	236,382.3	0	349,606.3	0	576,438.3	0	1,031,638.3	0
	2	56,532.7	0	113,096.7	0	226,320.7	0	453,152.7	0	908,352.7	0
	3	56,593.2	0	113,157.2	0	226,381.2	0	453,213.2	0	908,413.2	0
	4	57,139.1	0	113,703.1	0	226,927.1	0	453,759.1	0	908,959.1	0
	5	58,305.5	0	114,869.5	0	228,093.5	0	454,925.5	0	910,125.5	0
HEPP	1	89,636	0	179,304	0	358,736	0	747,984	0	1,438,016	0
	2	257,043.9	0	346,711.9	0	526,143.9	0	885,391.9	0	1,605,423.9	0
	3	89,636	0	179,304	0	358,736	0	717,984	0	1,438,016	0
	4	103,552.4	0	193,220.4	0	372,652.4	0	731,900.4	0	1451,932.4	0
	5	89,636	0	179,304	0	358,736	0	717,984	0	1,438,016	0

TABLE IX: Score Variations of V (RC)

Dataset	ID	4		8		16		32		64	
		t	v	t	v	t	v	t	v	t	v
WIKI	1	1	0	5	0	13	0	29	0	61	1
	2	0	0	1	0	9	0	25	0	57	1
	3	0	0	2	0	10	0	26	0	58	1
	4	2	0	6	0	14	0	30	1	62	1
	5	2	0	6	0	14	0	30	1	62	1
HEPP	1	2	1	6	1	14	1	30	1	62	1
	2	0	0	2	1	10	1	26	1	58	1
	3	3	0	7	0	15	0	31	0	63	0
	4	0	0	4	1	12	1	28	1	60	1
	5	1	1	5	1	13	1	29	1	61	1

variation of t is more than ten times that of v. This means all nodes in $\{V \setminus t\}$ have *score variations* no larger than that of t.

Exp 1-2: Dominance Property. The dominance property reveals that after inserting Δ_V into G to form G', target node t has a betweenness score no smaller than that of nodes in Δ_V. Thus, we compare the betweenness score between target node t and the nodes in Δ_V in G', and the results are in Table VIII.

Each row of Table VIII illustrates the comparison of betweenness score between target node t and the node w whose score is the maximum in Δ_V. It can be found that $\mathsf{BC}'(t)$ is much larger than $\mathsf{BC}'(w)$. Moreover, the score of w (which is maximum in Δ_V) is zero in Table VIII. This confirms that the target node has a *score* no smaller than those of nodes in Δ_V.

Exp 1-3: Boost Property. The boost property guarantees that, in the updated graph G' and when size p is sufficient, $\mathsf{BC}(t)'$ will exceed some node v in V whose $\mathsf{BC}(v)$ is larger than $\mathsf{BC}(t)$ in G. Combining the former two properties, the boost property's effect is reflected in the positive ranking upgrade of t. Thus, we show the relative ranking variations (*Ratio*) for ten target nodes t on the four tested graphs in Fig. 4, where the maximum, average, and minimum *Ratio* of these target nodes are reported.

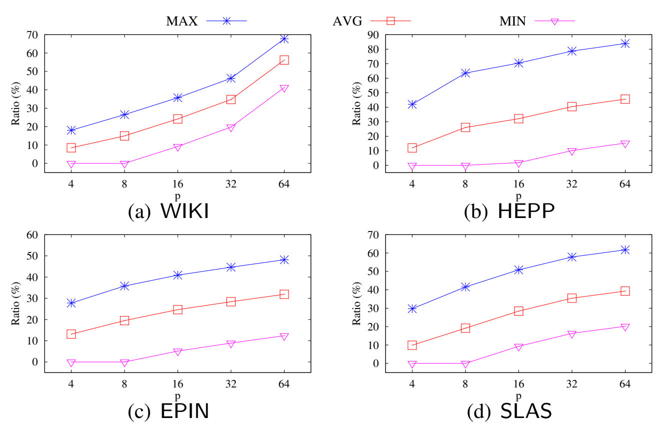
Fig. 4 shows that the value of *Ratio* increases with the increasing size p on all datasets. Furthermore, the *Ratio* of all target nodes is larger than zero at various sizes p on the graphs we used. For example, on HEPP, with only 8 inserted nodes, the maximum *Ratio* of the target node exceeds 44.1%. This means that, in the updated graph G', the target node's ranking advances by more than 4940 ($44.1\% \times 11204 > 4940$).

Exp 2: Coreness Centrality Promotion. The second set of experiments verifies that the single-clique strategy meets three properties of the maximum gain principle to promote coreness (RC) on real graphs.

Exp 2-1: Maximum Property. We compute the score variations of nodes in V and report the results for five target nodes (IDs from 1 to 5) on WIKI and HEPP in Table IX. Each row of Table IX represents a comparison between a certain target node

TABLE X: Scores of Target Nodes and Δ_V (RC)

Dataset	ID	4		8		16		32		64	
		t	w	t	w	t	w	t	w	t	w
WIKI	1	4	4	8	8	16	16	32	32	64	64
	2	7	4	8	8	16	16	32	32	64	64
	3	6	4	8	8	16	16	32	32	64	64
	4	4	4	8	8	16	16	32	32	64	64
	5	4	4	8	8	16	16	32	32	64	64
HEPP	1	4	4	8	8	16	16	32	32	64	64
	2	6	4	8	8	16	16	32	32	64	64
	3	4	4	8	8	16	16	32	32	64	64
	4	4	4	8	8	16	16	32	32	64	64
	5	4	4	8	8	16	16	32	32	64	64



(a) WIKI (b) HEPP (c) EPIN (d) SLAS

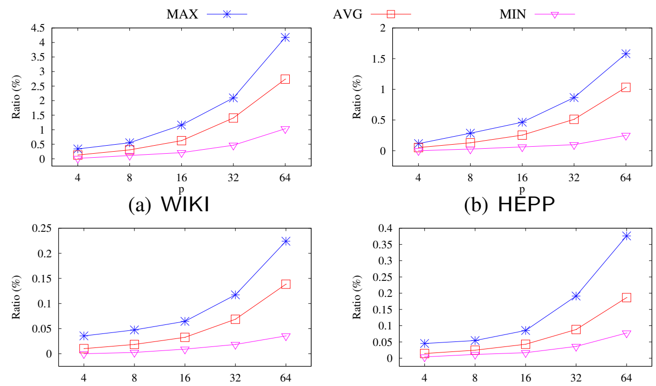
Fig. 5: Relative Ranking Variations (RC)

t and node v, whose score variation is the maximum among nodes in $\{V \setminus t\}$. It can be observed that the *score variation* of target node t is no smaller than those of nodes in $\{V \setminus t\}$ under various sizes p, which is consistent with the maximum property.

Exp 2-2: Dominance Property. Each row of Table X gives a comparison of the coreness score between target node t and node w whose score is maximal in Δ_V. We can observe that, in the updated graph G', the score of t is not less than that of w (and thus all nodes in Δ_V) at various p, which conforms to

TABLE XI: Reciprocal Score Variations of V (CC)

Dataset	ID	4		8		16		32		64	
		t	v	t	v	t	v	t	v	t	v
WIKI	1	4	8	8	16	16	32	32	64	64	128
	2	4	8	8	16	16	32	32	64	64	128
	3	4	8	8	16	16	32	32	64	64	128
	4	4	8	8	16	16	32	32	64	64	128
	5	4	8	8	16	16	32	32	64	64	128
HEPP	1	4	8	8	16	16	32	32	64	64	128
	2	4	8	8	16	16	32	32	64	64	128
	3	4	8	8	16	16	32	32	64	64	128
	4	4	8	8	16	16	32	32	64	64	128
	5	4	8	8	16	16	32	32	64	64	128



(a) WIKI (b) HEPP (c) EPIN (d) SLAS

Fig. 6: Relative Ranking Variations (CC)

the dominance property.

Exp 2-3: Boost Property. Fig. 5 shows the relative ranking variations ($Ratio$) for ten target nodes on all four datasets. We report the maximum, minimum, and average $Ratio$ of these target nodes. From Fig. 5, we find that all the target nodes successfully upgrade their centrality rankings given a suitable size p — the $Ratio$ value is larger than zero at that size. For example, on WIKI, the minimum $Ratio$ among the ten target nodes is larger than 9.2% when p is only 16. Combining the former two properties, the non-negative ranking improvement implies that the single-clique strategy fulfills the boost property.

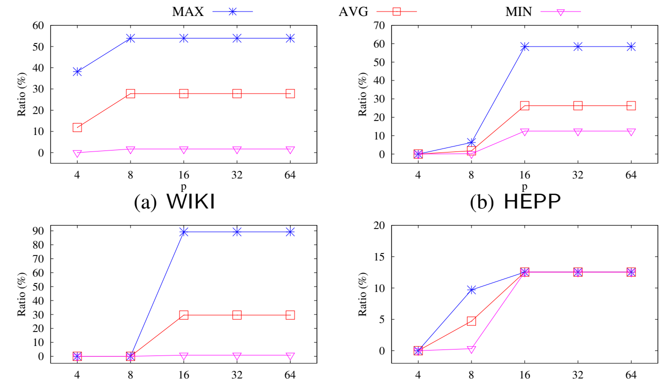
B. Testing Minimum Loss Principle

These experiments study the strategies that fulfill the minimum loss principle: the multi-point strategy for closeness and the double-line strategy for eccentricity. The experimental settings are similar to the test of the maximum gain principle.

Exp 3: Closeness Centrality Promotion. The third set of experiments demonstrates that the multi-point strategy satisfies the three properties of the minimum loss principle to promote closeness (CC) on real graphs.

Exp 3-1: Minimum Property. Note that the minimum property indicates that target node t has a reciprocal score variation no larger than that of other nodes. For this purpose, we compare the reciprocal score variation between target node t and node v, which has the smallest reciprocal score variation in $\{V \setminus t\}$. Due to space constraints, Table XI shows the results for five target nodes (by assigning new IDs from 1 to 5) on WIKI and HEPP.

Each row of Table XI represents a separate experiment in which a target node t is selected to be promoted. Table XI shows that the *reciprocal score variation* of target node t is smaller than the variation of node v, whose variation is the minimum in $\{V \setminus t\}$. For example, for target node t with ID 1 and at the size 8 on WIKI, the reciprocal score variation of t is 8, while the variation of v is 16. This implies that the multi-point strategy fulfills the minimum property.



(a) WIKI (b) HEPP (c) EPIN (d) SLAS

Fig. 7: Relative Ranking Variations (EC)

Exp 3-2: Dominance Property. Each row of Table XII compares the closeness score between target node t and node w (with the largest score in Δ_V). It can be seen that the *score* of each target node t is no smaller than that of w at various p in the updated graph G'. For example, on WIKI, the closeness of t with ID 1 is $\frac{1}{23450}$, while the closeness of node w with the largest score in Δ_V is $\frac{1}{30518}$. These results indicate that the multi-point strategy satisfies the minimum property.

Exp 3-3: Boost Property. Fig. 6 shows the relative ranking variations ($Ratio$) for ten target nodes on four datasets, where the maximum, minimum, and average $Ratio$ of these target nodes are reported. This figure shows that the $Ratio$ is larger than zero for most target nodes at various p. For example, when $p = 16$, all target nodes have $Ratio$ values larger than zero. Combining the first two properties, the positive $Ratio$ suggests that the multi-point strategy meets the boost property.

Exp 4: Eccentricity Centrality Promotion. The fourth set of experiments demonstrates that the double-line strategy conforms to the three properties of the minimum loss principle to upgrade eccentricity (EC) on real graphs.

Exp 4-1: Minimum Property. Table XIII depicts the reciprocal score variations $\Delta_{\mathbb{C}}$ for t and v (whose variations are the smallest in $\{V \setminus t\}$). We find that the *reciprocal score variation* of target node t is not larger than those of nodes in $\{V \setminus t\}$ on WIKI and HEPP. These results reveal that target node t has the minimum loss of all the nodes in V.

Exp 4-2: Dominance Property. Table XIV compares the eccentricity score between target node t and node w (whose score is the maximum in Δ_V) in the updated graph G'. It can be observed that the *score* of t is not less than that of w (and thus all nodes in Δ_V).

Exp 4-3: Boost Property. Fig. 7 shows the ranking variations ($Ratio$) of ten target nodes on four datasets. Fig. 7 shows that most target nodes' $Ratio$ values are positive at various sizes of p, although some $Ratio$ values are not significant for some small p. For example, when p is 16, on EPIN, the maximum $Ratio$ is 89.2%, while the minimum $Ratio$ is 0.8%. The non-negative $Ratio$ confirms that the ranking of target node t exceeds some node in $\{V \setminus t\}$ whose score is larger than t in the original graph, thereby resulting in an effective improvement.

TABLE XII: Scores of Target Nodes and Δ_V (CC)

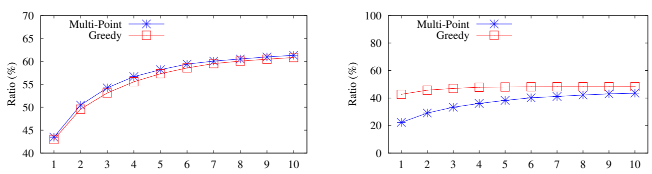
Dataset	ID	4		8		16		32		64	
		t	w	t	w	t	w	t	w	t	w
WIKI	1	$\frac{1}{23,450}$	$\frac{1}{30,518}$	$\frac{1}{23,454}$	$\frac{1}{30,526}$	$\frac{1}{23,462}$	$\frac{1}{30,542}$	$\frac{1}{23,478}$	$\frac{1}{30,574}$	$\frac{1}{23,510}$	$\frac{1}{30,638}$
	2	$\frac{1}{25,883}$	$\frac{1}{32,951}$	$\frac{1}{25,887}$	$\frac{1}{32,959}$	$\frac{1}{25,895}$	$\frac{1}{32,975}$	$\frac{1}{25,911}$	$\frac{1}{33,007}$	$\frac{1}{25,943}$	$\frac{1}{33,071}$
	3	$\frac{1}{25,788}$	$\frac{1}{32,856}$	$\frac{1}{25,792}$	$\frac{1}{32,864}$	$\frac{1}{25,800}$	$\frac{1}{32,880}$	$\frac{1}{25,816}$	$\frac{1}{32,912}$	$\frac{1}{25,848}$	$\frac{1}{32,976}$
	4	$\frac{1}{25,782}$	$\frac{1}{32,850}$	$\frac{1}{25,786}$	$\frac{1}{32,858}$	$\frac{1}{25,794}$	$\frac{1}{32,874}$	$\frac{1}{25,810}$	$\frac{1}{32,906}$	$\frac{1}{25,842}$	$\frac{1}{32,970}$
	5	$\frac{1}{27,254}$	$\frac{1}{34,322}$	$\frac{1}{27,258}$	$\frac{1}{34,330}$	$\frac{1}{27,266}$	$\frac{1}{34,346}$	$\frac{1}{27,282}$	$\frac{1}{34,378}$	$\frac{1}{27,314}$	$\frac{1}{34,442}$
HEPP	1	$\frac{1}{45,845}$	$\frac{1}{57,051}$	$\frac{1}{45,849}$	$\frac{1}{57,059}$	$\frac{1}{45,857}$	$\frac{1}{57,075}$	$\frac{1}{45,873}$	$\frac{1}{57,107}$	$\frac{1}{45,905}$	$\frac{1}{57,171}$
	2	$\frac{1}{49,887}$	$\frac{1}{61,093}$	$\frac{1}{49,891}$	$\frac{1}{61,101}$	$\frac{1}{49,899}$	$\frac{1}{61,117}$	$\frac{1}{49,915}$	$\frac{1}{61,149}$	$\frac{1}{49,947}$	$\frac{1}{61,213}$
	3	$\frac{1}{56,192}$	$\frac{1}{67,398}$	$\frac{1}{56,196}$	$\frac{1}{67,406}$	$\frac{1}{56,204}$	$\frac{1}{67,422}$	$\frac{1}{56,220}$	$\frac{1}{67,454}$	$\frac{1}{56,252}$	$\frac{1}{67,518}$
	4	$\frac{1}{53,701}$	$\frac{1}{64,907}$	$\frac{1}{53,705}$	$\frac{1}{64,915}$	$\frac{1}{53,713}$	$\frac{1}{64,931}$	$\frac{1}{53,729}$	$\frac{1}{64,963}$	$\frac{1}{53,761}$	$\frac{1}{65,027}$
	5	$\frac{1}{56,837}$	$\frac{1}{68,043}$	$\frac{1}{56,841}$	$\frac{1}{68,051}$	$\frac{1}{56,849}$	$\frac{1}{68,067}$	$\frac{1}{56,865}$	$\frac{1}{68,099}$	$\frac{1}{56,897}$	$\frac{1}{68,163}$

TABLE XIII: Reciprocal Score Variations of V (EC)

Dataset	ID	4		8		16		32		64	
		t	v	t	v	t	v	t	v	t	v
WIKI	1	0	0	0	0	2	4	10	12	26	28
	2	0	0	0	0	2	4	10	12	26	28
	3	0	0	0	0	2	4	10	12	26	28
	4	0	0	0	0	2	4	10	12	26	28
	5	0	0	0	0	2	4	10	12	26	28
HEPP	1	0	0	0	0	0	0	7	8	23	24
	2	0	0	0	0	0	0	7	8	23	24
	3	0	0	0	0	0	0	7	8	23	24
	4	0	0	0	0	0	0	7	8	23	24
	5	0	0	0	0	0	1	7	9	23	25

TABLE XIV: Scores of Target Nodes and Δ_V (EC)

Dataset	ID	4		8		16		32		64	
		t	w	t	w	t	w	t	w	t	w
WIKI	1	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	2	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	3	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	4	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	5	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
HEPP	1	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	2	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	3	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	4	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$
	5	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{16}$	$\frac{1}{17}$	$\frac{1}{32}$	$\frac{1}{33}$



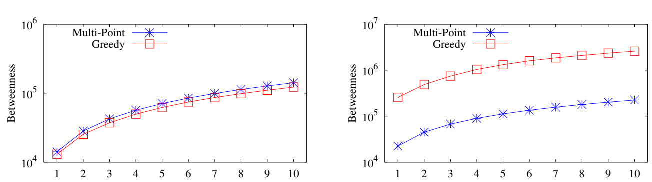
(a) WIKI (b) HEPP

Fig. 8: Comparison of Relative Ranking Variations

C. Comparison with a Greedy Algorithm

To further examine the effectiveness of the proposed strategies, we compare our multi-point strategy (denoted as Multi-Point) with the greedy algorithm (Greedy) for betweenness promotion. The comparisons for other centrality measures are omitted due to space limitations. Details of Greedy are presented in Section VI, and the description of Multi-Point is in Algorithm 1.

Although Greedy inserts edges into the graph while our method Multi-Point inserts nodes into the graph, both methods increase the betweenness score when the graph is modified. In this case, we test the effect of both methods in terms of **promotion size** p (i.e., the edge number for Greedy and the node number for Multi-Point). As suggested in [18], we select five target nodes t (with initially low betweenness scores) for boosting and set promotion size p ranging from 1 to 10. The promotion effect is evaluated by score variation and relative ranking variation (*Ratio*) before and after the insertion. The results are reported as the average value over five target nodes.



(a) WIKI (b) HEPP

Fig. 9: Comparison of Score Variations

Exp-5: Comparison of Relative Ranking Variation. We show the average relative ranking variation (*Ratio*) of Greedy and Multi-Point on WIKI and HEPP in Fig. 8. On WIKI, Multi-Point's *Ratio* outperforms Greedy's at all p. For example, when 10 nodes/edges are inserted, Multi-Point's *Ratio* is 61.3%, while Greedy's *Ratio* is 60.8%. On HEPP, although Greedy's *Ratio* is always better than Multi-Point's, the gap narrows as p increases. For instance, when $p = 1$, Greedy's *Ratio* is 1.9 times better than that of Multi-Point, while Greedy's *Ratio* is 1.1 times better than Multi-Point's when $p = 10$. This means our proposed Multi-Point is comparable with Greedy regarding the ranking promotion.

Exp-6: Comparison of Score Variation. We present the average score variation in Fig. 9. On WIKI, the average score variation of Multi-Point is slightly better than that of Greedy at various p. Nevertheless, the difference is relatively small: the variation of Multi-Point is only 1.15 times larger than that of Greedy when $p = 10$. On the other hand, on HEPP, Greedy outperforms Multi-Point significantly: when $p = 10$, the Greedy' score variation is over 11.5 times better than that of Multi-Point. This is reasonable because Multi-Point lacks the network structure for promotion, and its aim is for ranking promotion. In this case, Greedy is applicable when score promotion is the primary goal.

VIII. Conclusion

This paper provides an affirmative answer to whether it is possible to improve a target node's centrality ranking on a black-box network. By designing feasible strategies, we eliminate the dependence on the network structure for promotion. The effectiveness of promotion strategies is supported by the maximum gain principle and minimum loss principle. Extensive experimental studies on real-world networks confirm that the principle-guided strategies effectively improve the target nodes' centrality ranking for various centrality measures. We hope our research can pave the way for more attention on manipulating black-box networks for fun and profit.

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.
[2] P. Hage and F. Harary, "Eccentricity and centrality in networks," *Social Networks*, vol. 17, no. 1, pp. 57–63, 1995.
[3] V. Nicosia, R. Criado, M. Romance, G. Russo, and V. Latora, "Controlling centrality in complex networks," *Scientific Reports*, vol. 2, p. 218, 2012.
[4] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, "Vital nodes identification in complex networks," *Physics Reports*, vol. 650, pp. 1–63, 2016.
[5] P. W. Olsen, A. G. Labouseur, and J.-H. Hwang, "Efficient top-k closeness centrality search," in *2014 IEEE 30th International Conference on Data Engineering*. IEEE, 2014, pp. 196–207.
[6] W. Li, M. Qiao, L. Qin, Y. Zhang, L. Chang, and X. Lin, "Exacting eccentricity for small-world networks," in *2018 IEEE 34th International Conference on Data Engineering (ICDE)*. IEEE, 2018, pp. 785–796.
[7] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB)*, vol. 1, no. 1, p. 5, 2007.
[8] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Physical Review E*, vol. 65, no. 3, p. 036104, 2002.
[9] P. Crescenzi, G. D'angelo, L. Severini, and Y. Velaj, "Greedily improving our own closeness centrality in a network," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 11, no. 1, p. 9, 2016.
[10] P. D. Meo, K. Musial-Gabrys, D. Rosaci, G. M. Sarne, and L. Aroyo, "Using centrality measures to predict helpfulness-based reputation in trust networks," *ACM Transactions on Internet Technology (TOIT)*, vol. 17, no. 1, p. 8, 2017.
[11] E. Yan and Y. Ding, "Applying centrality measures to impact analysis: A coauthorship network analysis," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 10, pp. 2107–2118, 2009.
[12] G. Maia, P. O. V. de Melo, D. L. Guidoni, F. S. Souza, T. H. Silva, J. M. Almeida, and A. A. Loureiro, "On the analysis of the collaboration network of the brazilian symposium on computer networks and distributed systems," *Journal of the Brazilian Computer Society*, vol. 19, no. 3, p. 361, 2013.
[13] M. E. Newman, "A measure of betweenness centrality based on random walks," *Social Networks*, vol. 27, no. 1, pp. 39–54, 2005.
[14] S. Wu, L. Gong, W. Rand, and L. Raschid, "Making recommendations in a microblog to improve the impact of a focal user," in *Proceedings of the sixth ACM conference on Recommender systems*, 2012, pp. 265–268.
[15] D. Wen, L. Qin, Y. Zhang, X. Lin, and J. X. Yu, "I/o efficient core graph decomposition: application to degeneracy ordering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 1, pp. 75–90, 2018.
[16] J. Borge-Holthoefer and Y. Moreno, "Absence of influential spreaders in rumor dynamics," *Physical Review E*, vol. 85, no. 2, p. 026116, 2012.
[17] F. M. Clemente, F. M. L. Martins, R. S. Mendes *et al.*, *Social network analysis applied to team sports analysis*. Springer, 2016.
[18] E. Bergamini, P. Crescenzi, G. D'angelo, H. Meyerhenke, L. Severini, and Y. Velaj, "Improving the betweenness centrality of a node by adding links," *Journal of Experimental Algorithmics (JEA)*, vol. 23, pp. 1–5, 2018.
[19] R. Chitnis and N. Talmon, "Can we create large k-cores by adding few edges?" in *International Computer Science Symposium in Russia*. Springer, 2018, pp. 78–89.
[20] S. Perumal, P. Basu, and Z. Guan, "Minimizing eccentricity in composite networks via constrained edge additions," in *MILCOM 2013-2013 IEEE Military Communications Conference*. IEEE, 2013, pp. 1894–1899.
[21] G. D'Angelo, M. Olsen, and L. Severini, "Coverage centrality maximization in undirected networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 501–508.
[22] A. Korolova, R. Motwani, S. U. Nabar, and Y. Xu, "Link privacy in social networks," in *Proceedings of the 17th ACM conference on Information and knowledge management*, 2008, pp. 289–298.
[23] C. Zhang, J. Sun, X. Zhu, and Y. Fang, "Privacy and security for online social networks: challenges and opportunities," *IEEE Network*, vol. 24, no. 4, pp. 13–18, 2010.
[24] L. Fang and K. LeFevre, "Privacy wizards for social networking sites," in *Proceedings of the 19th International Conference on World Wide Web*. ACM, 2010, pp. 351–360.
[25] G. Malighetti, G. Martini, S. Paleari, and R. Redondi, "The impacts of airport centrality in the eu network and inter-airport competition on airport efficiency," 2009.
[26] F. Bonchi, G. De Francisci Morales, and M. Riondato, "Centrality measures on big graphs: Exact, approximated, and distributed algorithms," in *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016, pp. 1017–1020.
[27] M. Marchiori and V. Latora, "Harmony in the small-world," *Physica A: Statistical Mechanics and its Applications*, vol. 285, no. 3-4, pp. 539–546, 2000.
[28] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.
[29] F. W. Takes and W. A. Kosters, "Determining the diameter of small world networks," in *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 1191–1196.
[30] A. Hagberg, P. Swart, and D. S Chult, "Exploring network structure, dynamics, and function using networkx," Los Alamos National Lab.(LANL), Los Alamos, NM (United States), Tech. Rep., 2008.
[31] U. Brandes, "A faster algorithm for betweenness centrality," *Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001.
[32] A. E. Saríyüce, B. Gedik, G. Jacques-Silva, K.-L. Wu, and Ü. V. Çatalyürek, "Streaming algorithms for k-core decomposition," *Proceedings of the VLDB Endowment*, vol. 6, no. 6, pp. 433–444, 2013.
[33] D. Ouyang, L. Qin, L. Chang, X. Lin, Y. Zhang, and Q. Zhu, "When hierarchy meets 2-hop-labeling: efficient shortest distance queries on road networks," in *Proceedings of the 2018 International Conference on Management of Data*. ACM, 2018, pp. 709–724.

APPENDIX

A. Proofs for Section V-A.

Proofs of Theorem 5.1. We prove it via the properties in Definition 5.1.

- Maximum Property implies that $\mathbb{C}'(v) \leq \mathbb{C}'(t)$ (or equivalently, $\mathbb{R}'(t) \leq \mathbb{R}'(v)$), for $\forall v \in V$ with $\mathbb{C}(v) \leq \mathbb{C}(t)$, since otherwise $\Delta_{\mathbb{C}}(v) = \mathbb{C}'(v) - \mathbb{C}(v)$ is larger than $\Delta_{\mathbb{C}}(t) = \mathbb{C}'(t) - \mathbb{C}(t)$, contradiction.
- Dominance Property indicates that $\mathbb{R}'(t) \leq \mathbb{R}'(w)$, for $\forall w \in \Delta_V$.
- Boost Property shows that when $p > p'$, $\mathbb{R}'(t) > \mathbb{R}'(v)$, for at least a node $v \in V$ with $\mathbb{C}(v) > \mathbb{C}(t)$.

Suppose there are num nodes with scores larger than t in G, i.e., $num = |\{v \in V | \mathbb{C}(v) > \mathbb{C}(t)\}|$, then there are at most $num - 1$ nodes in V with scores larger than t in G' (by Boost Property). Hence, $\mathbb{R}(t) = num + 1$ and $\mathbb{R}'(t) \leq num + 1 - 1$. Therefore, $\Delta_{\mathbb{R}}(t) = \mathbb{R}(t) - \mathbb{R}'(t) \geq 1 > 0$.

Proofs of Theorem 5.2. The proof is similar to that of Theorem 5.1.

B. Proofs for Strategy for Betweenness Centrality

We first give some supporting lemmas and then provide proofs for Lemmas 5.1-5.3, and Theorem 5.3. When $G(V, E)$ is converted to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma(a, b)$ and $\sigma'(a, b)$ denote the number of a-b shortest paths in G and G', respectively; $\sigma_v(a, b)$ and $\sigma'_v(a, b)$ denote the number of a-b shortest paths via v in G and G', respectively.

Lemma S.1: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, any a-b shortest path does not pass through a node in Δ_V, for $\forall (a, b) \in V^2$.

Proof: If $w \in \Delta_V$ is on an a-b shortest path, then the a-b path can be shortened by connecting a to t then to b (to bypass w), contradiction.

Lemma S.2: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma'(a, b) = \sigma(a, b)$, for $\forall (a, b) \in V^2$.

Proof: Since we only append nodes around t and do not change edges within V, $\sigma'(a, b) \geq \sigma(a, b)$; $\sigma'(a, b) \leq \sigma(a, b)$ since otherwise a new a-b path must contain $w \in \Delta_V$, which contradicts with Lemma S.1.

Lemma S.3: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma'_v(a, b) = \sigma_v(a, b)$, for $\forall v \in V$, $\forall (a, b) \in V^2$, $v \neq a \neq b$.

Proof: Lemma S.2 indicates there is no increase in shortest path number for node pairs in V. (i) If v is not on any a-b shortest path, $\sigma'_v(a, b) = \sigma_v(a, b) = 0$; (ii) If v is on some a-b shortest paths, by [31], $\sigma_v(a, b) = \sigma(v, a) \times \sigma(v, b)$ equals $\sigma'_v(a, b) = \sigma'(v, a) \times \sigma'(v, b)$.

Lemma S.4: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma'_t(a, b) = \sigma'(a, b)$, for $\forall (a, b) \in \{\Delta_V \times \{V \cup \Delta_V\}\} \cup \{\{V \cup \Delta_V\} \times \Delta_V\}$, $a \neq b \neq t$.

Proof: We only prove the case for $(a, b) \in \{\Delta_V \times \{V \cup \Delta_V\}\}$ since the other case is symmetric. If $\sigma'_t(a, b) \neq \sigma'(a, b)$, then node a connects to some node $w \neq t$ to reach b, which means $deg_{G'}(a) \geq 2$. This contradicts with the fact that $deg_{G'}(a) = 1$ (a only connects to t).

Lemma S.5: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma'_v(a, b) = 0$, for $\forall (a, b) \in \Delta_V^2$, $\forall v \in \{V \setminus t\}$, $a \neq b \neq v$.

Proof: Suppose $\sigma'_v(a, b) \neq 0$, then the a-b shortest path via v is shortened by connecting a to t then to b, contradiction.

Lemma S.6: Given $[t, p, \textsf{multiple points}]$ that converts $G(V, E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\sigma'_w(a, b) = 0$, for $\forall w \in \Delta_V$, $\forall (a, b) \in \{V \cup \Delta_V\}^2$, $a \neq b \neq w$.

Proof: If $\sigma'_w(a, b) \neq 0$, $deg_{G'}(w) \geq 2$ because w is on some a-b path, which contradicts with the fact that $deg_{G'}(w) = 1$.

Proof of Lemma 5.1. Denote $\frac{\sigma_v(a,b)}{\sigma(a,b)}$ as $\mathsf{PBC}_v(a, b)$ and $\frac{\sigma'_v(a,b)}{\sigma'(a,b)}$ as $\mathsf{PBC}'_v(a, b)$. Given $u \in V$, by Lemma S.2 and Lemma S.3, $\mathsf{PBC}_u(a, b) = \mathsf{PBC}'_u(a, b)$, for $\forall (a, b) \in V^2$, $a \neq b \neq u$; since Δ_V do not exist in G, $\mathsf{PBC}_u(a, b) = 0$, for $\forall (a, b) \in \{\Delta_V \times \{V \cup \Delta_V\}\} \cup \{\{V \cup \Delta_V\} \times \Delta_V\}$, $a \neq b \neq u$. We divide V into $\{t\}$ and $\{V \setminus t\}$. For t, by lemma S.4, $\mathsf{PBC}'_t(a, b) = 1$, for $\forall (a, b) \in \{\Delta_V \times \{V \cup \Delta_V\}\} \cup \{\{V \cup \Delta_V\} \times \Delta_V\}$. Thus, $\Delta_{\mathbb{C}}(t) = \mathsf{BC}'(t) - \mathsf{BC}(t) = \Sigma_{(a,b) \in V^2 \cup \{V \times \Delta_V\} \cup \{\Delta_V \times V\} \cup \Delta_V^2} (\mathsf{PBC}'_t(a, b) - \mathsf{PBC}_t(a, b)) = 0 + 2(|V| - 1) \times |\Delta_V| + |\Delta_V| \times (|\Delta_V| - 1)$. For $v \in \{V \setminus t\}$, (i) for $\forall (a, b) \in \Delta_V^2$, $\mathsf{PBC}'_v(a, b) = 0$ by Lemma S.5; (ii) for $\forall (a, b) \in \{\Delta_V \times V\} \cup \{V \times \Delta_V\}$, $\mathsf{PBC}'_v(a, b) \leq 1$ [31]. Thus, $\Delta_{\mathbb{C}}(v) = \mathsf{BC}'(v) - \mathsf{BC}(v) = \Sigma_{(a,b) \in V^2 \cup \{V \times \Delta_V\} \cup \{\Delta_V \times V\} \cup \Delta_V^2} (\mathsf{PBC}'_v(a, b) -$

$\mathsf{PBC}_v(a,b)) \leq 0 + 2(|V|-1) \times |\Delta_V| + 0$. Consequently, $\Delta_{\mathbb{C}}(t) - \Delta_{\mathbb{C}}(v) \geq |\Delta_V| \times (|\Delta_V|-1) \geq 0$.

Proof of Lemma 5.2. For $w \in \Delta_V$, $\mathsf{BC}'(w) = 0$ by Lemma S.6. Then, $\mathsf{BC}'(t) \geq \mathsf{BC}'(w)$ since betweenness is non-negative [31].

Proof of Lemma 5.3. From proof of Lemma 5.1, $\Delta_{\mathbb{C}}(t) - \Delta_{\mathbb{C}}(v) \geq |\Delta_V| \times (|\Delta_V|-1) \geq (|\Delta_V|-1)^2$. Given $v \in \{V \setminus t\}$ with $\mathsf{BC}(v) > \mathsf{BC}(t)$, $\Delta_{\mathbb{C}}(t) - \Delta_{\mathbb{C}}(v) > \mathsf{BC}(v) - \mathsf{BC}(t)$ derives $\mathsf{BC}'(v) < \mathsf{BC}'(t)$. Thus, $\mathsf{BC}'(t) > \mathsf{BC}'(v)$, if $|\Delta_V| = p > p' = \sqrt{\mathsf{BC}(v) - \mathsf{BC}(t)} + 1$.

Proof of Theorem 5.3. By Lemmas 5.1-5.3 and Theorem 5.1.

C. Proofs for Strategy for Coreness Centrality

We first present supporting lemmas and then give proofs for Lemmas 5.4-5.6, and Theorem 5.4

Lemma S.7: Given a graph $G(V,E)$, if $v \in V$ is contained in a clique S with size $|S| = k$, then $\mathsf{RC}(v) \geq k-1$.

Proof: For $\forall v \in S$, $deg_S(v) \geq k-1$, and then $\mathsf{RC}(v) \geq k-1$ in G.

Lemma S.8: Given $[t, p, \mathsf{single\ clique}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\mathsf{RC}'(w) = |\Delta_V|$, for $\forall w \in \Delta_V$.

Proof: $\mathsf{RC}'(w) \geq |\Delta_V|$ by Lemma S.7 since t and $w \in \Delta_V$ form a clique; $\mathsf{RC}'(w) \leq |\Delta_V|$ as w's coreness is bounded by its degree [32].

Lemma S.9: Given $[t, p, \mathsf{single\ clique}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, if $\Delta_{\mathbb{C}}(t) = 0$, then $\Delta_{\mathbb{C}}(v) = 0$ for $\forall v \in \{V \setminus t\}$.

Proof: For $\forall v \in \{V \setminus t\}$, let S be the maximal subgraph in G, s.t., $v \in S$ and $deg_S = \mathsf{RC}(v)$; let S' be the maximal subgraph in G', s.t., $v \in S'$ and $deg_{S'} = \mathsf{RC}'(v)$. When $\Delta_{\mathbb{C}}(v) > 0$, $\{S' \setminus S\} \subseteq \{t \cup \Delta_V\}$. The fact v connects Δ_V only by t means $t \in \{S' \setminus S\}$, thus $\Delta_{\mathbb{C}}(t) \neq 0$.

Lemma S.10: Given $[t, p, \mathsf{single\ clique}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $\Delta_{\mathbb{C}}(v) = \mathsf{RC}'(v) - \mathsf{RC}(v) \leq 1$, for $\forall v \in \{V \setminus t\}$.

Proof: We reuse the symbols S and S' as in the proof of Lemma S.9. The facts $\{S' \setminus S\} \subseteq \{t \cup \Delta_V\}$ and v has no direct connection with Δ_V indicate the degree of nodes $v \in S$ can be reduced by at most one by deleting $\{S' \setminus S\}$ from S', hence, $\mathsf{RC}(v) \leq \mathsf{RC}'(v) - 1$.

Proof of Lemma 5.4. By Lemma S.9, if $\Delta_{\mathbb{C}}(t) = 0$, $\Delta_{\mathbb{C}}(v) = 0$, $\forall v \in \{V \setminus t\}$; By Lemma S.10, if $\Delta_{\mathbb{C}}(t) \neq 0$, then $\Delta_{\mathbb{C}}(v) \leq 1 \leq \Delta_{\mathbb{C}}(t)$.

Proof of Lemma 5.5. By Lemma S.7, $\mathsf{RC}'(t) \geq |\Delta_V|$ (t and Δ_V form a clique); By Lemma S.8, $\mathsf{RC}'(w) = |\Delta_V|$, for $\forall w \in \Delta_V$.

Proof of Lemma 5.6. By Lemma S.7, $\mathsf{RC}'(t) \geq |\Delta_V|$; By Lemma S.10, $\Delta_{\mathbb{C}}(v) \leq 1$, for $\forall v \in \{V \setminus t\}$. Given $\mathsf{RC}(v) > \mathsf{RC}(t)$, $\mathsf{RC}'(t) \geq |\Delta_V| > \mathsf{RC}(v) + 1 \geq \mathsf{RC}'(v)$ when $|\Delta_V| = p > p' = \mathsf{RC}(v) + 1$.

Proof of Theorem 5.4. By Lemmas 5.4-5.6 and Theorem 5.1.

D. Proofs for Strategy for Closeness Centrality

We first provide supporting lemmas and then present proofs for Lemmas 5.7-5.9, and Theorem 5.5.

Lemma S.11: Given $[t, p, \mathsf{multiple\ points}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $dist_{G'}(a,t) + dist_{G'}(t,b) = dist_{G'}(a,b)$, for $\forall (a,b) \in \{\{V \cup \Delta_V\} \times \Delta_V\} \cup \{\Delta_V \times \{V \cup \Delta_V\}\}$, $a \neq b$.

Proof: We verify the case for $(a,b) \in \{\{V \cup \Delta_V\} \times \Delta_V\}$ since the other case is similar. If $dist_{G'}(a,t) + dist_{G'}(t,b) \neq dist_{G'}(a,b)$, then any a-b shortest path bypasses t in G', which contradicts with the fact that b only connects to t, for $\forall b \in \Delta_V$, $b \neq a$.

Lemma S.12: Given $[t, p, \mathsf{multiple\ points}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $dist_{G'}(a,b) = dist_G(a,b)$, for $\forall (a,b) \in V^2$.

Proof: $dist_{G'}(a,b) \leq dist_G(a,b)$ since edges within V do not change; $dist_{G'}(a,b) \geq dist_G(a,b)$ since otherwise a new a-b shortest path via $w \in \Delta_V$ can be shortened by connecting a to t then to b, contradiction.

Proof of Lemma 5.7. For $\forall u \in V$, (i) $dist_{G'}(u,a) = dist_G(u,a)$, for $\forall a \in V$; (ii) $dist_G(u,a) = 0$ since Δ_V is not in G, for $\forall a \in \Delta_V$. Thus, $\overline{\Delta}_{\mathbb{C}}(u) = \Sigma_{a \in \{V \cup \Delta_V\}} dist_{G'}(u,a) - dist_G(u,a) = \Sigma_{a \in \Delta_V} dist_{G'}(u,a)$. Consequently, for $\forall v \in \{V \setminus t\}$, $\overline{\Delta}_{\mathbb{C}}(v) - \overline{\Delta}_{\mathbb{C}}(t) = \Sigma_{a \in \Delta_V}(dist_{G'}(v,a) - dist_{G'}(t,a)) = \Sigma_{a \in \Delta_V} dist_{G'}(v,t) > 0$.

Proof of Lemma 5.8. For $\forall w \in \Delta_V$, by Lemma S.11, $\mathsf{CC}'(w) = \frac{1}{\Sigma_{a \in \{V \cup \Delta_V\}} dist_{G'}(w,a)} = \frac{1}{\Sigma_{a \in \{V \cup \Delta_V\}} \{dist_{G'}(w,t) + dist_{G'}(t,a)\}} < \frac{1}{\Sigma_{a \in \{V \cup \Delta_V\}} dist_{G'}(t,a)} = \mathsf{CC}'(t)$.

Proof of Lemma 5.9. From proof of Lemma 5.7, $\overline{\Delta}_{\mathbb{C}}(v) - \overline{\Delta}_{\mathbb{C}}(t) = \Sigma_{a \in \Delta_V} dist_{G'}(v,t) = |\Delta_V| dist_G(v,t)$ (by Lemma S.12), for $\forall v \in \{V \setminus t\}$. Given $\mathsf{CC}(v) > \mathsf{CC}(t)$, $\mathsf{CC}'(v) < \mathsf{CC}'(t)$ (or equivalently, $\overline{\mathsf{CC}}'(v) > \overline{\mathsf{CC}}'(t)$) when $\overline{\Delta}_{\mathbb{C}}(v) - \overline{\Delta}_{\mathbb{C}}(t) = |\Delta_V| dist_G(v,t) > \overline{\mathsf{CC}}(t) - \overline{\mathsf{CC}}(v)$, or $|\Delta_V| = p > p' = \frac{\overline{\mathsf{CC}}(t) - \overline{\mathsf{CC}}(v)}{dist_G(v,t)}$.

Proof of Theorem 5.5. By Lemmas 5.7-5.9 and Theorem 5.2.

E. Proofs for Strategy for Eccentricity Centrality

We first present some supporting lemmas and then show proofs for Lemmas 5.10-5.12, and Theorem 5.6. We define the **aggregate distance** from v to a set of nodes $S \subseteq V$ as $dist_G(v,S) = \max_{u \in S} dist_G(v,u)$.

Lemma S.13: Given a graph $G(V,E)$, $\overline{\mathsf{EC}}(v) = \max(dist_G(v,S), dist_G(v, \{V \setminus S\}))$, for $\forall v \in V$, $\forall S \subseteq V$.

Proof: $\overline{\mathsf{EC}}(v) = \max_{u \in \{S \cup \{V \setminus S\}\}} dist_G(u,v) = \max(dist_G(v,S), dist_G(v, \{V \setminus S\}))$.

Lemma S.14: Given a graph $G(V,E)$, $\overline{\mathsf{EC}}(a) \leq dist(a,b) + \overline{\mathsf{EC}}(b)$, for $\forall (a,b) \in V^2$.

Proof: Let c be the node, s.t., $dist_G(a,c) = \overline{\mathsf{EC}}(a)$, then $\overline{\mathsf{EC}}(a) = dist_G(a,c) \leq dist_G(a,b) + dist_G(b,c) \leq dist_G(a,b) + \overline{\mathsf{EC}}(b)$.

Lemma S.15: Given $[t, p, \mathsf{double\ lines}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, i) $dist_{G'}(a,b) = dist_{G'}(a,t) + dist_{G'}(t,b)$, for $\forall (a,b) \in \{\Delta_V \times \{V \cup \Delta_V\} \cup \{V \cup \Delta_V\} \times \Delta_V\}$, $dist_{G'}(a,b) > dist_{G'}(a,t)$; ii) $dist_{G'}(b, \Delta_V) = dist_{G'}(b,t) + dist_{G'}(t, \Delta_V)$, $b \in V$.

Proof: Δ_V consists of two disjoint sets S_1 and S_2 (on each of the double lines). Moreover, t is the cut node among S_1, S_2, and $\{V \setminus t\}$ in G' — deleting t separates these parts. Therefore, $dist_{G'}(a,t) + dist_{G'}(t,b) = dist_{G'}(a,b)$ by the cut property [33], for $\forall (a,b) \in \{\Delta_V \times \{V \cup \Delta_V\} \cup \{V \cup \Delta_V\} \times \Delta_V\}$, $dist_{G'}(a,b) > dist_{G'}(a,t)$ (to avoid a, b being both in S_1 or S_2); In addition, for $\forall b \in V$, $dist_{G'}(b, \Delta_V) = \max_{w \in \Delta_V} dist_{G'}(b,w) = \max_{w \in \Delta_V} dist_{G'}(b,t) + dist_{G'}(t,w) = dist_{G'}(b,t) + dist_{G'}(t, \Delta_V)$.

Lemma S.16: Given $[t, p, \mathsf{double\ lines}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, $dist_{G'}(a,b) = dist_G(a,b)$, for $\forall (a,b) \in V^2$; $dist_{G'}(a,S) = dist_G(a,S)$, for $\forall a \in V, \forall S \subseteq V$.

Proof: $dist_{G'}(a,b) \leq dist_G(a,b)$ since edges within V do not change; $dist_{G'}(a,b) \geq dist_G(a,b)$ since otherwise a new a-b shortest path via $w \in \Delta_V$ can be shortened by connecting a to t then to b. For $S \subseteq V$, by aggregating over $b \in S$, $dist_{G'}(a,S) = dist_G(a,S)$.

Lemma S.17: Given $[t, p, \mathsf{double\ lines}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, if $\mathsf{EC}'(t) < \mathsf{EC}(t)$, then $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V)$; otherwise, $\overline{\mathsf{EC}}'(t) = dist_{G'}(t,V) = dist_G(t,V)$.

Proof: By Lemma S.16, $dist_{G'}(t,V) = dist_G(t,V)$. $\overline{\mathsf{EC}}'(t) = \max(dist_{G'}(t,V), dist_{G'}(t, \Delta_V)) = \max(\overline{\mathsf{EC}}(t), dist_{G'}(t, \Delta_V))$. Thus, $\mathsf{EC}'(t) < \mathsf{EC}(t)$ (resp. $\overline{\mathsf{EC}}'(t) > \overline{\mathsf{EC}}(t)$) implies $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V)$, and $\mathsf{EC}'(t) = \mathsf{EC}(t)$ indicates $\overline{\mathsf{EC}}'(t) = \overline{\mathsf{EC}}(t) = dist_{G'}(t,V) = dist_G(t,V)$.

Lemma S.18: Given $[t, p, \mathsf{double\ lines}]$ that converts $G(V,E)$ to $G'(V \cup \Delta_V, E \cup \Delta_E)$, if $\mathsf{EC}'(t) < \mathsf{EC}(t)$, then $\mathsf{EC}'(v) < \mathsf{EC}(v)$, for $\forall v \in V$.

Proof: If $\mathsf{EC}'(t) < \mathsf{EC}(t)$, then $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V)$ by Lemma S.17. This means $dist_{G'}(t, \Delta_V) > dist_{G'}(t,V) = dist_G(t,V)$. For $\forall v \in \{V \setminus t\}$, $dis_{G'}(v, \Delta_V) = dist_{G'}(v,t) + dist_{G'}(t, \Delta_V) > dist_G(v,t) + dist_G(t,V) \geq \overline{\mathsf{EC}}(v)$ (by Lemma S.14). Then, $\overline{\mathsf{EC}}'(v) \geq dis_{G'}(v, \Delta_V) > \overline{\mathsf{EC}}(v)$, and thus $\mathsf{EC}'(v) < \mathsf{EC}(v)$.

Proof of Lemma 5.10. We discuss two cases. (i) If $\overline{\Delta}_{\mathbb{C}}(t) = 0$, for $\forall v \in \{V \setminus t\}$, $\overline{\mathsf{EC}}'(v) = \max(dist_{G'}(v,V), dist_{G'}(v, \Delta_V)) \geq dist_{G'}(v,V) = dist_G(v,V) = \overline{\mathsf{EC}}(v)$. Then, $\overline{\Delta}_{\mathbb{C}}(v) \geq \overline{\Delta}_{\mathbb{C}}(t) = 0$. (ii) If $\overline{\Delta}_{\mathbb{C}}(t) > 0$, by Lemma S.17, $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V)$; by Lemma S.18, $\overline{\mathsf{EC}}'(v) = dist_{G'}(v, \Delta_V)$ because $\mathsf{EC}'(v) < \mathsf{EC}(v)$, for $\forall v \in \{V \setminus t\}$. Thus, $\overline{\Delta}_{\mathbb{C}}(v) - \overline{\Delta}_{\mathbb{C}}(t) = (dist_{G'}(v, \Delta_V) - \overline{\mathsf{EC}}(v)) - (dist_{G'}(t, \Delta_V) - \overline{\mathsf{EC}}(t)) = dist_{G'}(v,t) - \overline{\mathsf{EC}}(v) + \overline{\mathsf{EC}}(t) = dist_G(v,t) - \overline{\mathsf{EC}}(v) + \overline{\mathsf{EC}}(t) \geq 0$ (according to Lemma S.14).

Proof of Lemma 5.11. There are two cases. (i) If $\mathsf{EC}'(t) = \mathsf{EC}(t)$, then by Lemma S.17, $\mathsf{EC}'(t) = dist_{G'}(t,V)$. For $\forall w \in \Delta_V$, $\overline{\mathsf{EC}}'(w) \geq dist_{G'}(w,V) = dist_{G'}(w,t) + dist_{G'}(t,V) \geq \overline{\mathsf{EC}}'(t)$, and thus $\mathsf{EC}'(w) \geq \mathsf{EC}'(t)$. (ii) If $\mathsf{EC}'(t) \geq \mathsf{EC}(t)$, then by Lemma S.17, $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V) = \frac{|\Delta_V|}{2}$ (maximum distance to nodes on each line). For $\forall w \in \Delta_V$, $\overline{\mathsf{EC}}'(w) \geq dist_{G'}(w, \Delta_V) \geq \frac{|\Delta_V|}{2}$ (distance to nodes on the other line), and it follows $\mathsf{EC}'(w) \leq \mathsf{EC}'(t)$.

Proof of Lemma 5.12. If $\overline{\mathsf{EC}}'(t) = dist_{G'}(t, \Delta_V) = \frac{|\Delta_V|}{2} > dist_{G'}(t,V)$, by Lemma S.18, $\overline{\mathsf{EC}}'(v) = dist_{G'}(v, \Delta_V) > \overline{\mathsf{EC}}(v)$, and by Lemma S.15, $\overline{\mathsf{EC}}'(v) = dist_{G'}(v,t) + dist_{G'}(t, \Delta_V) > \overline{\mathsf{EC}}'(t)$. Then, when $p = |\Delta_V| > p' = 2 \times \mathsf{EC}(t)$, $\mathsf{EC}'(t) > \mathsf{EC}'(v)$.

Proof of Theorem 5.6. By Lemmas 5.10-5.12 and Theorem 5.2.