Lower Bound for Function Computation in Distributed Networks

H. K. Dai[1(✉)] and M. Toulouse[2]

[1] Computer Science Department, Oklahoma State University, Stillwater, OK 74078, USA
dai@cs.okstate.edu
[2] Computer Science Department, Vietnamese-German University, Binh Duong New City, Vietnam
michel.toulouse@vgu.edu.vn

Abstract. Distributed computing network systems are modeled as graphs with which vertices represent compute elements and adjacency-edges capture their uni- or bi-directional communication. Distributed computation over a network system proceeds in a sequence of time-steps in which vertices update and/or exchange their values based on the underlying algorithm constrained by the time-(in)variant network topology. For finite convergence of distributed information dissemination and function computation in the model, we present a lower bound on the number of time-steps for vertices to receive (initial) vertex-values of all vertices regardless of underlying protocol or algorithmics in time-invariant networks via the notion of vertex-eccentricity.

Keywords: Distributed function computation
Linear iterative schemes · Information dissemination
Finite convergence · Vertex-eccentricity

1 Preliminaries

Distributed computation algorithms, decentralized data-fusion architectures, and multi-agent systems are modeled with a network of interconnected vertices that compute common value(s) based on initial values or observations at the vertices. Key computation and communication requirements for these network/system paradigms include that their vertices perform local/internal computations and regularly communicate with each other via an underlying protocol. Fundamental limitations and capabilities of these algorithms and systems are studied in the literature with viable applications in computer science, communication, and control and optimization (see, for examples, [1,3,4,7,8]). We give brief and informal descriptions of some example studies below:

1. Quantized consensus [5]: Consider an order-n network with an initial network-state in which each vertex assumes an initial (integer) value $x_i[0]$ for $i = 1, 2, \ldots, n$. The network achieves a quantized consensus when, at some later

T. K. Dang et al. (Eds.): FDSE 2018, LNCS 11251, pp. 371–384, 2018.
https://doi.org/10.1007/978-3-030-03192-3_28

time, all the n vertices simultaneously arrive with almost equal values y_i for $i = 1, 2, \ldots, n$ (that is, $|y_i - y_j| \leq 1$ for all $i, j \in \{1, 2, \ldots, n\}$) while preserving the sum of all initial values (that is, $\sum_{i=1}^{n} x_i[0] = \sum_{i=1}^{n} y_i$).
2. Collaborative distributed hypothesis testing [6]: Consider a network-system of n vertices (sensors/agents) that collaboratively determine the probability measure of a random variable based on a number of available observations/measurements. For the binary setting in deciding two hypotheses, each vertex collects measurement(s) and makes a preliminary (local) decision $d_i \in \{0, 1\}$ in favor of the two hypotheses for $i = 1, 2, \ldots, n$. The n vertices are allowed to communicate, and the network-system resolves with a final decision by, for example, the majority rule (that is, computes the indicator function of the event $\sum_{i=1}^{n} d_i > \frac{n}{2}$) in distributed fashion.
3. Solitude verification [3]: Consider an unlabeled network of n vertices (processes) in which each vertex is in one of a finite number of states: s_i for $i = 1, 2, \ldots, n$. Solitude verification on the network checks if a unique vertex with a given state s exists in the network, that is, computes the Boolean function for the equality $|\{i \in \{1, 2, \ldots, n\} \mid s_i = s\}| = 1$.

While there is a wide spectrum of algorithms in the literature that solve distributed computation problems such as the above, there are also studies that deal with algorithmic and complexity issues constrained by underlying time-(in)variant network topology, resource-limitations associated with vertices, time/space and communication tradeoffs, convergence criteria and requirements, etc. We present below a model of distributed computing systems and address the motivation of our study.

1.1 Model of Distributed Computing Systems

Most graph-theoretic definitions in this article are given in [2]. We will abbreviate "directed graph" and "directed path" to digraph and dipath, respectively.

We consider the topological model and algorithmics detailed in [7] for distributed function computation, and provide its abstraction components as follows:

1. Network topology: A distributed computing system is modeled as a digraph G with $V(G)$ and $E(G)$ denoting its sets of vertices and directed edges, respectively. Uni-directional communication on $V(G)$ is captured by the adjacency relation represented by $E(G)$: for all distinct vertices, $u, v \in V(G)$, $(u, v) \in E(G)$ if and only if vertex u can send information to vertex v (and v can receive information from u). Note that bi-directional communication between u and v is viewed as the co-existence of the two directed edge (u, v) and (v, u) in $E(G)$.

   Distributed computation over the network proceeds in a sequence of time-steps. At each time-step, all vertices update and/or exchange their values based on the underlying algorithm constrained by the network topology, which is assumed to be time-invariant.

2. Resource capabilities in vertices: The digraph G of the network topology is vertex-labeled such that messages are identified with senders and receivers. The vertices of $V(G)$ are assumed to have sufficient computational capabilities and local storage. Generally we assume that: (1) all communications/transmissions between vertices are reliable and in correct sequence, and (2) each vertex may, in the current time-step, receive the prior-step transmission(s) from its in-neighbor(s), update, and send transmission(s) to its out-neighbor(s) in accordance to the underlying algorithm.
   The domain of all initial/input and observed/output values of the vertices of G is assumed to be an algebraic field $\mathbb{F}$.
3. Linear iterative scheme (for algorithmic lower- and upper-bound results): For a vertex $v \in V(G)$, denote by $x_v[k] \in \mathbb{F}$ the vertex-value of v at time-step $k = 0, 1, \ldots$. A function with domain $\mathbb{F}^{|V(G)|}$ and codomain $\mathbb{F}$ is computed in accordance to a linear iterative scheme. Given initial vertex-values $x_v[0] \in \mathbb{F}$ for all vertices $v \in V(G)$ as arguments to the function, at each time-step $k = 0, 1, \ldots$, each vertex $v \in V(G)$ updates (and transmits) its vertex-value via a weighted linear combination of the prior-step vertex-values constrained by neighbor-structures: for all $v \in V(G)$ and $k = 0, 1, \ldots$,

$$x_v[k+1] = \sum_{u \in V(G)} w_{vu} x_u[k],$$

   where the prescribed weights $w_{vu} \in \mathbb{F}$ for all $v, u \in V(G)$ that are subject to the adjacency-constraints $w_{vu} = 0$ (the zero-element of $\mathbb{F}$) if u is not adjacent to v (that is, $(u, v) \notin E(G)$); equivalently,

$$\text{transpose of } (x_v[k+1] \mid v \in V(G)) = W \cdot \text{ transpose of } (x_v[k] \mid v \in V(G))$$

   where the two vectors of vertex-values and W are indexed by a common discrete ordering of $V(G)$ with $W = [w_{vu}]_{(v,u) \in V(G) \times V(G)}$.

1.2 Motivation of Our Study

Based on the framework and its variants for distributed function computation, researches and studies are focused on mathematical interplays among:

- time-(in)variance of network-topology
- granularity of time-step: discrete versus continuous
- choice of base field: special (real or complexes) versus arbitrary (finite or infinite)
- characterization of calculable functions
- convergence criteria and rates (finite, asymptotic, and/or probabilistic)
- adoption and algebraic properties of weight-matrix for linear interactive schemes: random weight-matrix, spectrum of eigenvalues, base field, etc.
- resilience and robustness of computation algorithmics for network-topology in the presence/absence of malicious vertices

– lower and upper bounds on (linear) iteration required for the convergence of calculable functions.

Summarized results, research studies, and references are available in, for examples, [7–9,11].

Sundaram and Hadjicostis [7,8] present their research findings in the finite convergence of distributed information dissemination and function computation in the model with linear iterative algorithmics stated above, among other contributions in distributed function computation and data-stream transmission in the presences of noise and malicious vertices. More specifically, (1) they employ structural theories in observability and invertibility of linear systems over arbitrary finite fields to obtain lower and upper bounds on the number of linear iterations for achieving network consensus for finite convergence of arbitrary functions, and (2) the bounds are valid for all initial vertex-values of arbitrary finite fields as arguments to the functions in connected time-invariant topologies with almost all random weight-matrices.

For a time-invariant topology with underlying digraph G and a vertex $u \in V(G)$, denote by $\deg_{G,\mathrm{in}}(u)$ the in-degree of u in G, and by $\Gamma_{G,\mathrm{in}}(u)$ the in-neighbor of u in G; hence $\Gamma^*_{G,\mathrm{in}}(u)$ denotes the in-closure of u in G, that is,

$$\begin{aligned}\Gamma^*_{G,\mathrm{in}}(u) &= \cup_{\eta \geq 0} \Gamma^{\eta}_{G,\mathrm{in}}(u) \\ &= \{v \in V(G) \mid \text{ there exists a dipath in } G \text{ from } v \text{ to } u\}.\end{aligned}$$

Consider all possible families of directed trees that are: (1) a vertex-decomposition of $\Gamma^*_{G,\mathrm{in}}(u) - \{u\}$, and (2) rooted in (as subset of) $\Gamma_{G,\mathrm{in}}(u)$. Denote by:

$$\begin{aligned}\alpha_{G,u} = \min\{\max\{\mathrm{order}(T_i) \mid 1 \leq i \leq n\} \mid &\\ \{T_i\}_{i=1}^{n} \text{ is a family of directed trees that are: (1) a vertex-}&\\ \text{decomposition of } \Gamma^*_{G,\mathrm{in}}(u) - \{u\}, \text{ and (2) rooted in (as sub-}&\\ \text{set of) } \Gamma_{G,\mathrm{in}}(u)\}.&\end{aligned}$$

Their upper-bound result for a vertex $u \in V(G)$ is stated as follows: for every linear iterative scheme with random weight-matrix over a finite base field $\mathbb{F}$ of cardinality $|\mathbb{F}| \geq (\alpha_{G,u} - 1)(|\Gamma^*_{G,\mathrm{in}}(u)| - \deg_{G,\mathrm{in}}(u) - \frac{1}{2}\alpha_{G,u})$, then, with probability at least $1 - \frac{1}{|\mathbb{F}|}(\alpha_{G,u} - 1)(|\Gamma^*_{G,\mathrm{in}}(u)| - \deg_{G,\mathrm{in}}(u) - \frac{1}{2}\alpha_{G,u})$, the vertex u can calculate arbitrary functions of arbitrary initial vertex-values $x_v[0] \in \mathbb{F}$ for all $v \in \Gamma^*_{G,\mathrm{in}}(u)$ via the linear iterative scheme within a most $\alpha_{G,u}$ time-steps.

Sundaram conjectures in [7] that $\alpha_{G,u}$ may also serve as a lower bound on the number of time-steps for a vertex $u \in V(G)$ to receive the initial vertex-values of all $v \in \Gamma^*_{G,\mathrm{in}}(u)$ regardless of underlying protocol or algorithmics. Hence, linear iterative schemes are time-optimal in disseminating information over arbitrary time-invariant connected networks.

Toulouse and Minh [10] refute the conjecture via the notion of rank-step sequences for linear iterative schemes over connected network with an explicit counter-example in Fig. 1.

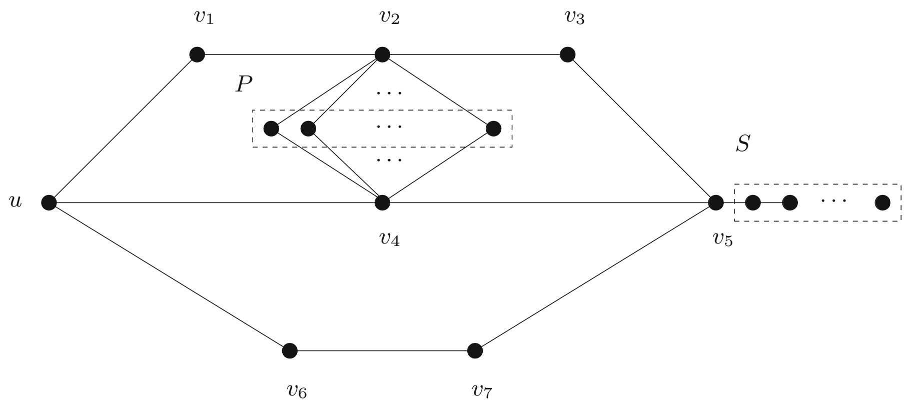

Fig. 1. A counter-example graph, in which the embedded parallel component P and serial component S satisfying $\mathrm{order}(S) = \lfloor \frac{\mathrm{order}(P)}{2} \rfloor + 1$, to the lower-bound conjecture in terms of $\alpha_{G,u}$ in [7].

In order to complement the explicitly constructed counter-example to the lower-bound conjecture on the number of time-steps for distributed function computation and information dissemination with respect to a given vertex, we present in this article a lower bound on the number of time-steps for a vertex $u \in V(G)$ to receive the initial vertex-values of all $v \in \Gamma^*_{G,\mathrm{in}}(u)$ regardless of underlying protocol or algorithmics in a time-invariant network via the notion of vertex-eccentricity.

2 Revised Lower Bound for Distributed Function Computation and Information Dissemination

Consider an arbitrary vertex $u \in V(G)$, and assume a non-trivial $\Gamma^*_{G,\mathrm{in}}(u)$ ($|\Gamma^*_{G,\mathrm{in}}(u)| > 1$) hereinafter. We develop a lower bound on the number of time-steps required for the vertex u to receive the (initial) vertex-values of all vertices of $\Gamma^*_{G,\mathrm{in}}(u)$ (regardless of underlying protocol, including linear iterative schemes). See Fig. 2 for an example of $\Gamma^*_{G,\mathrm{in}}(u)$.

For two vertices u and v of G, $\overrightarrow{d}_G(u, v)$ denotes the directed distance from u to v in G, that is,

$$\overrightarrow{d}_G(u, v) = \begin{cases} \text{length of a shortest dipath from } u \text{ to } v \text{ in } G & \text{if exists,} \\ \infty & \text{otherwise.} \end{cases}$$

For a vertex u of G, $e_{G,\mathrm{in}}(u)$ denotes the in-eccentricity of u in G, which is the maximum directed distance from a vertex to u in G, that is,

$$e_{G,\mathrm{in}}(u) = \max\{\underbrace{\overrightarrow{d}_G(v, u)}_{\text{minimum length of a dipath from } v \text{ to } u \text{ in } G} \mid v \in V(G)\}.$$

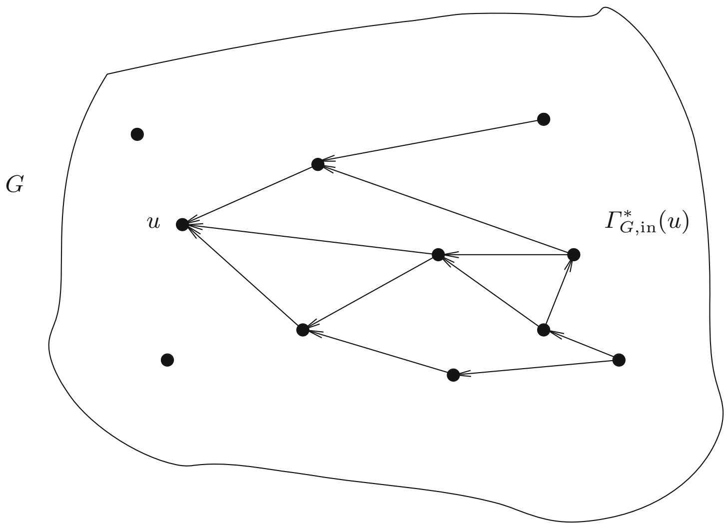

Fig. 2. For a vertex u in a digraph G: an example organization of the in-closure $\Gamma^*_{G,\text{in}}(u)$ of u in G.

Following the above-stated distributed-computation framework as in [8] and for their conjecture, we develop a lower-bound result based on the notion of eccentricity (instead of "order" or "size" as in the conjecture):

1. For every (linear or non-linear) iteration scheme, in which a vertex's value or information is transmitted to its out-neighbors via their incidence directed edges in unit time-step, requires at least $e_{G,\text{in}}(u)$ time-steps for vertex u to access values/information of all the vertices in $\Gamma^*_{G,\text{in}}(u)$. Thus, $e_{G,\text{in}}(u)$ serves as a lower bound on the number of time-steps required for function-computation by vertex u via such iteration scheme.
2. In accordance with the distributed framework for our function-computation, we show below that:

$$e_{G,\text{in}}(u) = 1 + \min\{\max\{\underbrace{e_{T_i,\text{in}}(\text{root}(T_i))}_{=\,\text{depth}(T_i)} \mid 1 \leq i \leq n\} \mid \{T_i\}_{i=1}^{n} \text{ is a family of directed trees that are: (1) a vertex-decomposition of } \Gamma^*_{G,\text{in}}(u) - \{u\}\text{, and (2) rooted in (as subset of) } \Gamma_{G,\text{in}}(u)\}.$$

We illustrate an example organization of $\Gamma^*_{G,\text{in}}(u) - \{u\}$ in a family of vertex-disjoint directed trees in Fig. 3.

To show the above equality for $e_{G,\text{in}}(u)$, we prove the two embedded inequalities in the following sections.

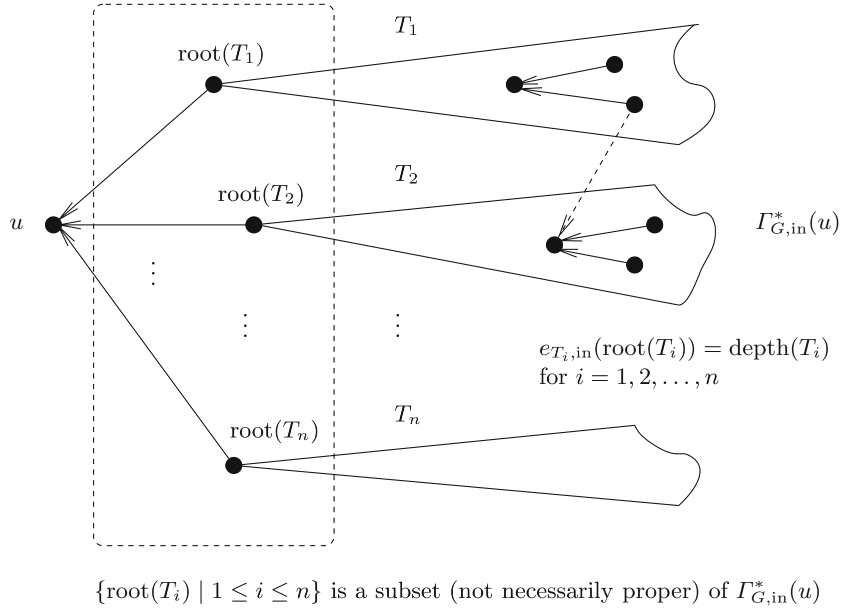

Fig. 3. For a vertex u in a digraph G: an example organization of $\Gamma^*_{G,\text{in}}(u) - \{u\}$ in a family $\{T_i\}_{i=1}^n$ of directed trees that are: (1) a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$, and (2) rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$.

2.1 Upper Bound for Vertex-Eccentricity

We first prove that:

$$e_{G,\text{in}}(u) \leq 1 + \min\{\max\{\underbrace{e_{T_i,\text{in}}(\text{root}(T_i))}_{=\ \text{depth}(T_i)} \mid 1 \leq i \leq n\} \mid \{T_i\}_{i=1}^n \text{ is a family of directed trees that are: (1) a vertex-decomposition of } \Gamma^*_{G,\text{in}}(u) - \{u\}\text{, and (2) rooted in (as subset of) } \Gamma_{G,\text{in}}(u)\};$$

equivalently,

$$e_{G,\text{in}}(u) \leq 1 + \max\{\underbrace{e_{T_i,\text{in}}(\text{root}(T_i))}_{=\ \text{depth}(T_i)} \mid 1 \leq i \leq n\}$$

for arbitrary family of directed trees, $\{T_i\}_{i=1}^n$, which are a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$.

Consider an arbitrary family of directed trees, $\{T_i\}_{i=1}^n$, which are a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$. The in-eccentricity $e_{G,\text{in}}(u)$ of u in G is realized by a dipath P from a vertex $v \in \Gamma^*_{G,\text{in}}(u) - \{u\}$ to u in G. Since $\{T_i\}_{i=1}^n$ is a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$, we have $v \in V(T_i)$ for some $i \in \{1, 2, \ldots, n\}$. We depict the scenario in Fig. 4.

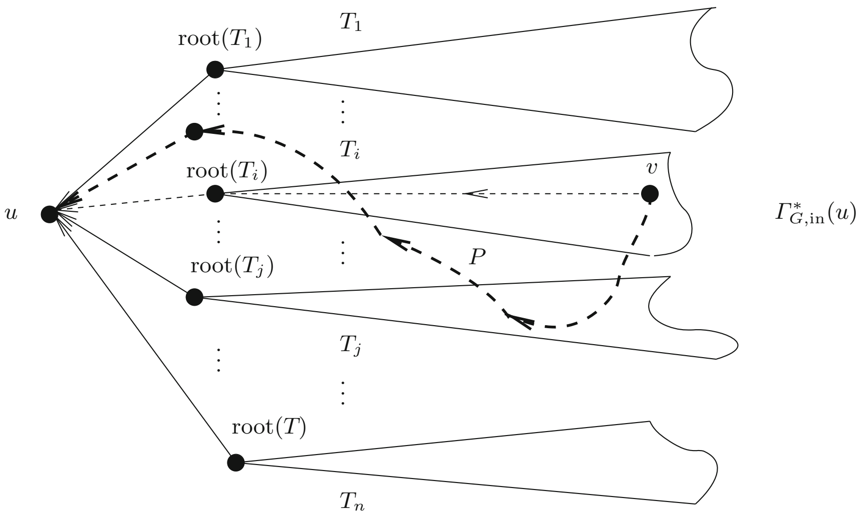

Fig. 4. For a vertex u in a digraph G: the in-eccentricity $e_{G,\text{in}}(u)$ of u in G is realized by a dipath P from a vertex $v \in \Gamma^*_{G,\text{in}}(u) - \{u\}$ to u in G.

Now,

$$\begin{aligned}\underbrace{e_{G,\text{in}}(u)}_{\text{in } G} = \underbrace{\text{length}(P)}_{\text{in } G} = \underbrace{\overrightarrow{d}_G(v,u)}_{\text{in } G} &\le \text{length}((\text{unique}) \text{ dipath from } v \text{ to root}(T_i) \text{ in } T_i \text{ concatenated} \\ &\quad\ \text{with directed edge } (\text{root}(T_i), u)) \text{ since } T_i \text{ is a sub-digraph of} \\ &\quad\ \text{the digraph vertex-spanned by } \Gamma^*_{G,\text{in}}(u) \\ &\le \text{depth}(T_i) + 1 \\ &\le 1 + \max\{\text{depth}(T_i) \mid 1 \le i \le n\}\end{aligned}$$

as desired.

2.2 Lower Bound for Vertex-Eccentricity

To show the reverse inequality:

$$e_{G,\text{in}}(u) \ge 1 + \min\{\max\{\underbrace{e_{T_i,\text{in}}(\text{root}(T_i))}_{=\,\text{depth}(T_i)} \mid 1 \le i \le n\} \mid \begin{array}{l}\{T_i\}_{i=1}^n \text{ is a family of directed trees that are:} \\ (1) \text{ a vertex-decomposition of } \Gamma^*_{G,\text{in}}(u) - \{u\}, \text{ and} \\ (2) \text{ rooted in (as subset of) } \Gamma_{G,\text{in}}(u)\},\end{array}$$

it suffices to construct a family $\{T_i\}_{i=1}^n$ of directed trees that are a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$, and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$, such that:

$$e_{G,\text{in}}(u) \geq 1 + \max\{\text{depth}(T_i) \mid 1 \leq i \leq n\}.$$

We proceed with an inductive construction of a sequence $(P_1, P_2, \ldots)$ of dipaths with common end-vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \ldots)$ is organized as a family $\{T_1, T_2, \ldots, T_i\}$, where $i \geq 1$, of directed trees such that:

1. The family $\{T_1, T_2, \ldots, T_i\}$ consists of mutually vertex-disjoint directed trees with their roots in $\Gamma_{G,\text{in}}(u)$,
2. Each directed tree in the family provides a shortest dipath (in G) for each of its vertices to u, that is, for every vertex $v \in V(T_j)$ where $j \in \{1, 2, \ldots, i\}$, the (unique) dipath from v to root(T_j) in T_j yields $\overrightarrow{d}_G(v, u)$:

   length((unique) dipath from v to root(T_j) concatenated with directed edge $(\text{root}(T_j), u)) = \overrightarrow{d}_G(v, u)$,

   and
3. The in-eccentricity of u in G is bounded below as:

$$e_{G,\text{in}}(u) \geq 1 + \max\{\text{depth}(T_1), \text{depth}(T_2), \ldots, \text{depth}(T_i)\}.$$

See an example configuration in Fig. 5.

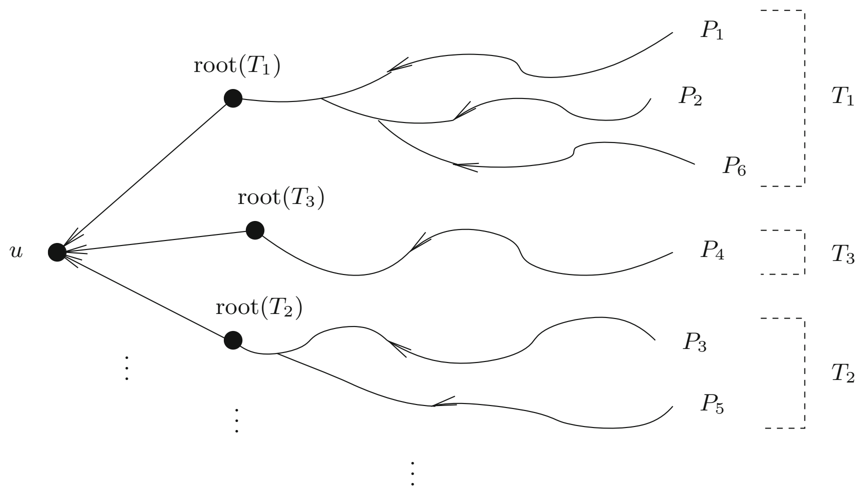

Fig. 5. For a vertex u in a digraph G: an inductive construction of a sequence $(P_1, P_2, \ldots)$ of dipaths with common end-vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \ldots)$ is organized as a family $\{T_1, T_2, \ldots, T_i\}$, where $i \geq 1$, of directed trees that satisfies the stated conditions in items 1, 2, and 3.

Basis step: For P_1, we may employ a dipath from a vertex, say v, in $\Gamma^*_{G,\text{in}}(u) - \{u\}$ to u in G that realizes $\overrightarrow{d}_G(v,u)$. Then, designate P_1 as such a path, and $T_1 = \{P_1 - \{u\}\}$.

For the family $\{T_1\}$, we can verify the above-stated three items 1, 2 (via "shortest dipath in G" enjoys "optimal substructure property in G" by typical cut-and-paste argument), and 3.

Induction step: Assume that we have constructed a sequence $(P_1, P_2, \ldots, P_j)$ of dipaths with common end-vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \ldots, P_j - \{u\})$ is organized as a family $\{T_1, T_2, \ldots, T_i\}$, where $j \geq i \geq 1$, of directed trees that satisfies the above-stated items 1, 2, and 3.

If the family $\{T_1, T_2, \ldots, T_i\}$ yields a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$, then the inductive construction is complete. Thus, we may assume that there exists a vertex $v \in (\Gamma^*_{G,\text{in}}(u) - \{u\}) - \cup_{\eta=1}^{i} V(T_\eta)$. We construct a desired dipath P_{j+1} from v to u in G as follows.

First, consider a dipath P from v to u in G that realizes $\overrightarrow{d}_G(v,u)$ (that is, $\text{length}(P) = \overrightarrow{d}_G(v,u)$). Observe that,

$$\text{length}(P) = \overrightarrow{d}_G(v,u) \leq \underbrace{\max\{\overrightarrow{d}_G(v,u) \mid v \in V(G)\}}_{e_{G,\text{in}}(u)}.$$

Consider the two cases of P based on its possible intersection with the constructed directed forest/family $\{T_1, T_2, \ldots, T_i\}$—which are shown in Fig. 6.

Case 1: $V(P) \cap \cup_{\eta=1}^{i} V(T_\eta) = \emptyset$. From the above observation that $\text{length}(P) \leq e_{G,\text{in}}(u)$, hence for P_{j+1}, we may employ P by designating $P_{j+1} = P$ and $T_{i+1} = \{P_{j+1} - \{u\}\}$ as in the basis step. We can verify the above-stated items 1, 2, and 3 for the augmented family $\{T_1, T_2, \ldots, T_{i+1}\}$.

Case 2: $V(P) \cap \cup_{\eta=1}^{i} V(T_\eta) \neq \emptyset$. Denote the first entrance of the dipath P into $\cup_{\eta=1}^{i} V(T_\eta)$ by w, say $w \in V(P) \cap V(T_k)$ for some $k \in \{1, 2, \ldots, i\}$.

With the denotations/labelings in Fig. 7, we have two possible dipaths from w to u: (1) the dipath:

$$Q = \underbrace{\text{(unique) dipath from } w \text{ to root}(T_k)}_{\text{contained in } T_k} \quad \text{concatenated with the directed edge } (\text{root}(T_k), u),$$

and (2) the dipath P_2 such that:

$$P = \underbrace{\text{dipath from } v \text{ to } w}_{\text{via vertices in } (\Gamma^*_{G,\text{in}}(u) - \{u\}) - \cup_{\eta=1}^{i} V(T_\eta)} \text{concatenated with} \underbrace{\text{dipath } P_2}_{\text{from } w \text{ to } u \text{ in } G}.$$

What can we say about length(Q) versus length(P_2)? They must be equal—via a proof by contradiction as follows:

1. Suppose that $\text{length}(Q) < \text{length}(P_2)$: The dipath from v, via w, to u formed by the concatenation of P_1 (v to w) and Q (w, via root(T_k), to u)

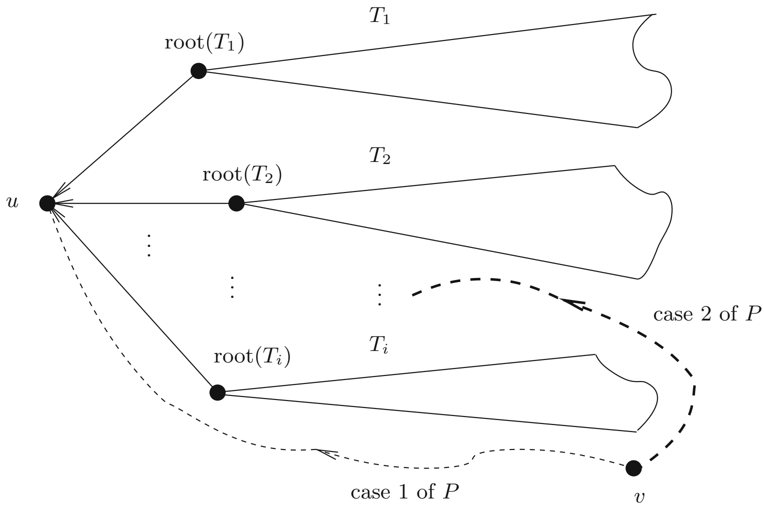

Fig. 6. For a vertex u in a digraph G: assume the inductive construction of a sequence $(P_1, P_2, \ldots, P_j)$ of dipaths that results in a family $\{T_1, T_2, \ldots, T_i\}$, where $j \geq i \geq 1$, of mutually vertex-disjoint directed trees with their roots in $\Gamma^*_{G,\text{in}}(u) - \{u\}$ that satisfies the stated conditions in items 1, 2, and 3, then, for a vertex $v \in (\Gamma^*_{G,\text{in}}(u) - \{u\}) - \cup_{\eta=1}^{i} V(T_\eta)$, construct a desired dipath P_{j+1} from v to u in G by considering a dipath P from v to u in G with length$(P) = \overrightarrow{d}_G(v, u)$ in two cases.

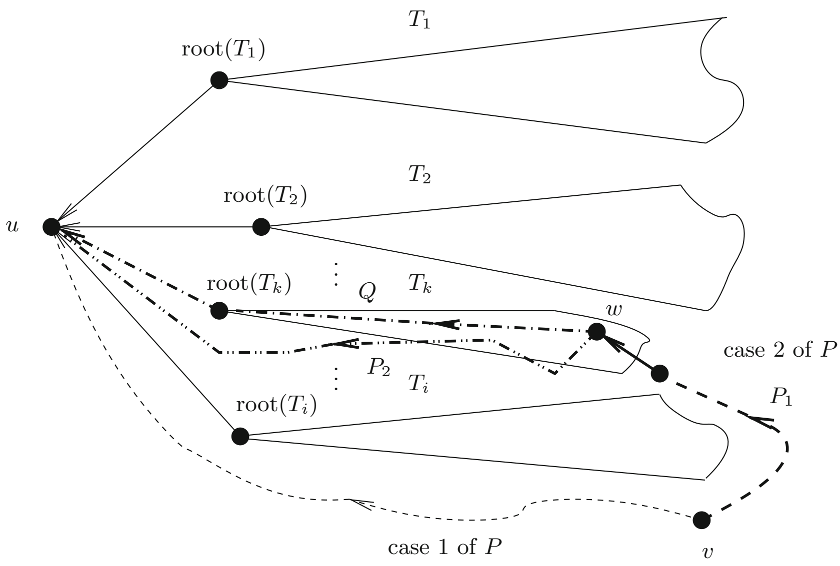

Fig. 7. For a vertex u in a digraph G: case 2 of P with $V(P) \cap \cup_{\eta=1}^{i} V(T_\eta) \neq \emptyset$ is considered.

is a shorter dipath than P—which contradicts to the assumption that $\text{length}(P) = \overrightarrow{d}_G(v, u)$.

2. Suppose that $\text{length}(Q) > \text{length}(P_2)$: The existence of such dipath P_2 from v to u in G contradicts to the above item 2 that $\text{length}(Q) = \overrightarrow{d}_G(v, u)$.

Now, we let:

$$P_{j+1} = \underbrace{\text{dipath } P_1 \text{ from } v \text{ to } u \text{ in } G}_{\text{via vertices in } (\Gamma^*_{G,\text{in}}(u) - \{u\}) - \cup_{\eta=1}^{i} V(T_\eta)} \text{concatenated with} \underbrace{\text{dipath } Q \text{ from } w \text{ to } u}_{\text{via vertices in } T_k},$$

and include the dipath $P_{j+1} - \{u\}$ into the directed tree T_k.
We can check/verify the above-stated items 1, 2, and 3:

1. The statement is obvious,
2. The condition follows from that "shortest dipath in G" enjoys "optimal substructure property in G", and
3. By noting that:
$$e_{G,\text{in}}(u) \geq \overrightarrow{d}_G(v, u) = \text{length}(P).$$

This completes the inductive construction, and we have shown the reverse inequality.

3 Concluding Remarks

We can derive a lower bound on $e_{G,\text{in}}(u)$ from the knowledge of the maximum in-degree of G (vertex-spanned by $\Gamma^*_{G,\text{in}}(u)$), which yields a (possibly weaker) lower bound on the number of time-steps for vertex u to access values/information of all the vertices in $\Gamma^*_{G,\text{in}}(u)$.

Denote by $\Delta_{G,\text{in}}(u)$ (≥ 1) the maximum in-degree of the subdigraph of G vertex-spanned by $\Gamma^*_{G,\text{in}}(u)$.

Organize the in-closure of u in G as the sequence of successive in-neighbors as illustrated in Fig. 8, we have the following inequality:

$$|\Gamma^*_{G,\text{in}}(u)| \leq 1 + \Delta_{G,\text{in}}(u) + \Delta_{G,\text{in}}(u)^2 + \cdots + \Delta_{G,\text{in}}(u)^{e_{G,\text{in}}(u)-1}.$$

Consider the two cases for the $\Delta_{G,\text{in}}(u)$-value. When $\Delta_{G,\text{in}}(u) = 1$:

$$|\Gamma^*_{G,\text{in}}(u)| \leq e_{G,\text{in}}(u).$$

When $\Delta_{G,\text{in}}(u) \geq 2$:

$$|\Gamma^*_{G,\text{in}}(u)| \leq \frac{\Delta_{G,\text{in}}(u)^{e_{G,\text{in}}(u)} - 1}{\Delta_{G,\text{in}}(u) - 1},$$

which gives a lower bound on $e_{G,\text{in}}(u)$:

$$\log_{\Delta_{G,\text{in}}(u)}((\Delta_{G,\text{in}}(u) - 1)|\Gamma^*_{G,\text{in}}(u)| + 1) \leq e_{G,\text{in}}(u).$$

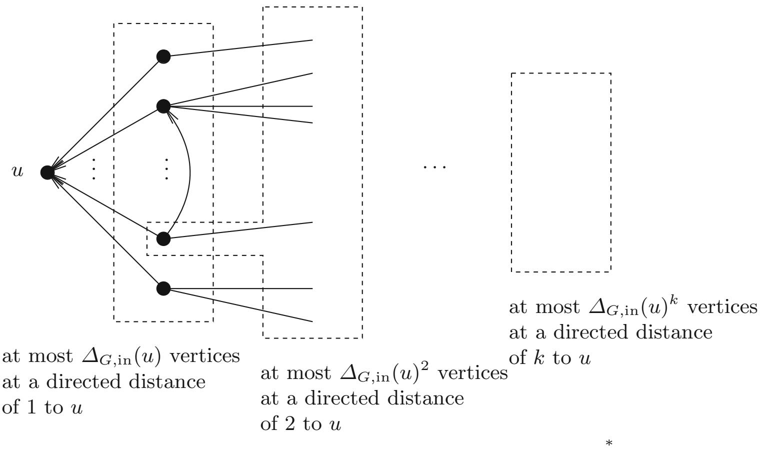

Fig. 8. For a vertex u in a digraph G: organize vertices of the in-closure $\Gamma^*_{G,\text{in}}(u)$ of u in G according to their directed distances to u.

We can obtain desired lower bounds in analogous fashion with similar graph-parameters such a regularity in-degree, and maximum and regularity degrees.

The lower bound of $e_{G,\text{in}}(u)$ time-steps for vertex u to collect values/information from all the vertices in $\Gamma^*_{G,\text{in}}(u)$ is stated in accordance with the min-max distributed framework as:

$$e_{G,\text{in}}(u) = 1 + \min\{\max\{\underbrace{e_{T_i,\text{in}}(\text{root}(T_i))}_{=\,\text{depth}(T_i)} \mid 1 \le i \le n\} \mid$$

$\{T_i\}_{i=1}^{n}$ is a family of directed trees that are: (1) a vertex-decomposition of $\Gamma^*_{G,\text{in}}(u) - \{u\}$, and (2) rooted in (as subset of) $\Gamma_{G,\text{in}}(u)\}$.

However, it is not necessary to compute $e_{G,\text{in}}(u)$, directly or indirectly, through an underlying optimal directed forest whose maximum depth yielding $e_{G,\text{in}}(u)$. Instead, with a given/fixed topology of directed forest (described in above fashion) underlying a distributed iteration scheme, we can obtain a stronger lower bound with a distributed computation of the maximum depth $D_{\max}$ among all the directed trees in the forest, and note that:

> number of time-steps for vertex u to access values/information from all vertices in the given directed forest $\ge 1 + D_{\max} \ge e_{G,\text{in}}(u)$.

In addition to the probabilistic upper-bound result on the number of time-steps for (general) distributed function computation via linear iterative schemes with random weight-matrix, Sundaram and Hadjicostis [7,8] employ observability theory of linear systems to study the linear-functional case for distributed computation (of linear functions), and achieve an upper bound via the minimal polynomial of the underlying weight-matrix.

Toulouse and Minh [10] study the linear functional case with prescribed time-invariant network-topology over random weight-matrices, and obtain various empirical upper-bound results.

References

1. Ayaso, O., Shah, D., Dahleh, M.A.: Information theoretic bounds for distributed computation over networks of point-to-point channels. IEEE Trans. Inf. Theor. **56**(12), 6020–6039 (2010)
2. Bondy, J.A., Murty, U.S.R.: Graph Theory, Volume 244 of Graduate Texts in Mathematics. Springer, London (2008)
3. Fich, F.E., Ruppert, E.: Hundreds of impossibility results for distributed computing. Distrib. Comput. **16**(2–3), 121–163 (2003)
4. Hendrickx, J.M., Olshevsky, A., Tsitsiklis, J.N.: Distributed anonymous discrete function computation. IEEE Trans. Autom. Control **56**(10), 2276–2289 (2011)
5. Kashyap, A., Basar, T., Srikant, R.: Quantized consensus. Automatica **43**(7), 1192–1203 (2007)
6. Katz, G., Piantanida, P., Debbah, M.: Collaborative distributed hypothesis testing. Computing Research Repository, abs/1604.01292 (2016)
7. Sundaram, S.: Linear iterative strategies for information dissemination and processing in distributed systems. Ph.D. thesis, University of Illinois at Urbana-Champaign (2009)
8. Sundaram, S., Hadjicostis, C.N.: Distributed function calculation and consensus using linear iterative strategies. IEEE J. Sel. Areas Commun. **26**(4), 650–660 (2008)
9. Sundaram, S., Hadjicostis, C.N.: Distributed function calculation via linear iterative strategies in the presence of malicious agents. IEEE Trans. Autom. Control **56**(7), 1495–1508 (2011)
10. Toulouse, M., Minh, B.Q.: Private communication (2018)
11. Toulouse, M., Minh, B.Q.: Applicability and resilience of a linear encoding scheme for computing consensus. In: Muñoz, V.M., Wills, G., Walters, R.J., Firouzi, F., Chang, V. (eds.) Proceedings of the Third International Conference on Internet of Things, Big Data and Security, IoTBDS 2018, Funchal, Madeira, Portugal, 19–21 March 2018, pp. 173–184. SciTePress (2018)