Message Passing Optimization of Harmonic Influence Centrality

Luca Vassio, Fabio Fagnani, Paolo Frasca, *Member*, *IEEE*, and Asuman Ozdaglar

***Abstract*—This paper proposes a new measure of node centrality in social networks, the *Harmonic Influence Centrality (HIC)*, which emerges naturally in the study of social influence over networks. Using an intuitive analogy between social and electrical networks, we introduce a distributed *message passing* algorithm to compute the HIC of each node. Although its design is based on theoretical results which assume the network to have no cycle, the algorithm can also be successfully applied on general graphs.**

***Index Terms*—Centrality measures, distributed algorithms, message passing, opinion dynamics, social networks.**

I. Introduction

A KEY issue in the study of networks is the identification of their most important nodes: the definition of prominence is based on a suitable function of the nodes, called *centrality measure*. The appropriate notion of centrality measure of a node depends on the nature of interactions among the agents situated in the network and the potential decision and control objectives. In this paper, we define a new measure of centrality, which we call *Harmonic Influence Centrality* (HIC) and which emerges naturally in the context of social influence. We explain why in addition to being descriptively useful, this measure answers questions related to the optimal placement of different agents or opinions in a network with the purpose of swaying average opinion. In large networks approximating real world social networks, computation of centrality measures of all nodes can be a challenging task. In this paper, we present a fully decentralized algorithm based on message passing for computing the HIC of all nodes, which converges to the correct values for trees (connected networks with no cycles), but can also be applied to general networks.

Our model of social influence builds on recent work [1], which characterizes opinion dynamics in a network consisting of stubborn agents who hold a fixed opinion equal to zero or one (i.e., type zero and type one stubborn agents) and regular agents who hold an opinion $x_i \in [0, 1]$ and update it as a weighted average of their opinion and those of their neighbors. We consider a special case of this model, where a fixed subset of the agents are type zero stubborn agents and the rest are regular agents. We define the HIC of node ℓ as the asymptotic value of the average opinion when node ℓ switches from a regular agent to a type one stubborn agent. This measure hence captures the long run influence of node ℓ on the average opinion of the social network.

The HIC measure, beside its natural descriptive value, is also the precise answer to the following network decision problem: suppose you would like to have the largest influence on long run average opinion in the network and you have the capacity to change one agent from regular to type one stubborn. Which agent should be picked for this conversion? This question has a precise answer in terms of HIC; the agent with the highest HIC should be picked.

Although the HIC measure is intuitive, its centralized computation in a large network would be challenging in terms of its informational requirements that involve the network topology and the location of type zero stubborn agents. We propose here a decentralized algorithm whereby each agent computes its own HIC based on local information. The construction of our algorithm uses a novel analogy between social and electrical networks by relating the Laplace equation resulting from social influence dynamics to the governing equations of electrical networks. Under this analogy, the asymptotic opinion of regular agent i can be interpreted as the voltage of node i, when type zero stubborn agents are kept at voltage zero and type one agents are kept at voltage one. This interpretation allows us to use properties of electrical circuits and provide a recursive characterization of HIC in trees. Using this characterization, we develop a so-called *message passing* algorithm (MPA) for its solution, which converges after at most a number of steps equal to the diameter of the tree. The algorithm we propose runs in a distributed and parallel way among the nodes, which do not need to know the topology of the whole network, and has a lower cost in terms of number of operations with respect to the centralized algorithm recently proposed in [15]. Although originally designed for trees, our algorithm can be employed in general networks. For regular networks with unitary resistances (corresponding to all agents placing equal weights on opinions of other agents), we prove that this algorithm converges (although not necessarily to the correct HIC values). Moreover, we show through simulations that the algorithm performs well also on general networks.

A. Related Works

In social science and network science there is a large literature on defining and computing centrality measures [6]. Among the most popular definitions, we mention degree centrality, node

Manuscript received September 20, 2013; revised January 02, 2014; accepted January 30, 2014. Date of publication February 5, 2014; date of current version April 9, 2014. Recommended by Associate Editor M. Chertkov.

L. Vassio is with the Dipartimento di Ingegneria Meccanica e Aerospaziale, Politecnico di Torino, 10129 Torino, Italy (e-mail: luca.vassio@polito.it).

F. Fagnani is with the Dipartimento di Scienze Matematiche, Politecnico di Torino, 10129 Torino, Italy (e-mail: fabio.fagnani@polito.it).

P. Frasca is with the Department of Applied Mathematics, University of Twente, 7522 NB Enschede, The Netherlands (e-mail: p.frasca@utwente.nl).

A. Ozaglar is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: asuman@mit.edu).

Digital Object Identifier 10.1109/TCNS.2014.2304870

betweenness, information centrality [11], and Bonacich centrality [3], which is related to the well-known Google PageRank algorithm. These notions have proven useful in a range of applications, but are not universally the appropriate concepts by any means. Our interest in opinion dynamics, thus motivates the choice to define a centrality measure for our purposes.

As opposed to centralized algorithms, the interest for distributed algorithms to compute centrality measures has risen more recently. In [7], a randomized algorithm is used to compute PageRank centrality. In [14], distributed algorithms are designed to compute node and edge betweenness on trees, but are not suitable for general graphs.

MPAs have been widely studied and are commonly used in artificial intelligence and information theory: they have demonstrated empirical success in numerous applications including low-density parity-check codes, turbo codes, free energy approximation, and satisfiability problems (see e.g., [10] [4] [13] [8]). In such algorithms, nodes are thought as objects with computational ability which can send and receive information from their neighbors.

Interesting connections between MPAs and electrical networks are discussed in [12]. Moreover, electrical networks have also been put in connection with social networks: two notable cases are the notion of resistance distance [5] and the characterization of random walk betweenness in [9].

B. Paper Outline

In Section II, we review our model of opinion dynamics with stubborn agents and we define our problem of interest, that is the optimal placement of a stubborn agent. In Section III, we review basic notions of electrical circuits and explain their connection with social networks. Section IV is devoted to apply this electrical analogy on tree graphs and to study the optimal placement problem in this important case. Next, in Section V, we describe the MPA to compute the optimal solution on trees, while in Section VI we consider its extension to general graphs: we provide both theoretical results, in Section VI-A, and numerical simulations, in Section VI-B. Final remarks are given in Section VII.

C. Notation

To avoid possible ambiguities, we here briefly recall some notation and a few basic notions of graph theory which are used throughout this paper. The cardinality of a (finite) set E is denoted by $|E|$ and when $E \subset F$, we define its complement as $E^c = \{f \in F | f \notin E\}$. A square matrix $\boldsymbol{P}$ is said to be non-negative when its entries P_{ij} are non-negative, substochastic when it is non-negative and $\sum_j P_{ij} \leq 1$ for every row i, and stochastic when it is non-negative and $\sum_j P_{ij} = 1$ for every i. We denote by $\mathbf{1}$ a vector of whose entries are all 1. An (undirected) graph G is a pair (I, E), where I is a finite set, whose elements are said to be the *nodes* of G and E is a set of unordered pairs of nodes called *edges*. The neighbor set of a node $i \in I$ is defined as $N_i := \{j \in I | \{i, j\} \in E\}$. The cardinality of the neighbor set $d_i := |N_i|$ is said to be the *degree* of node i. A graph in which all nodes have degree d is said to be d*-regular*. A path in G is a sequence of nodes $\gamma = (j_1, \dots j_s)$ such that $\{j_t, j_{t+1}\} \in E$, for every $t = 1, \dots, s-1$. The path γ is said to connect j_1 and j_s. The path γ is said to be simple if $j_h \neq j_k$ for $h \neq k$. A graph is connected if any pair of distinct nodes can be connected by a path (which can be chosen to be simple). The length of the shortest path between two nodes i and j is said to be the *distance* between them, and is denoted as $\mathrm{dst}(i, j)$. Consequently, the *diameter* of a connected graph is defined to be $\mathrm{diam}(G) := \max_{i,j \in I}\{\mathrm{dst}(i, j)\}$. A tree is a connected graph such that for any pair of distinct nodes there is just one simple path connecting them. Finally, given a graph $G = (I, E)$ and a subset $J \subseteq I$, the subgraph induced by J is defined as $G_{|J} = (J, E_{|J})$, where $E_{|J} = \{\{i, j\} \in E | i, j \in J\}$.

II. Opinion Dynamics and Stubborn Agent Placement

Consider a connected graph $G = (I, E)$. Nodes in I will be thought as agents who can exchange information through the available edges $\{i, j\} \in E$. Each agent $i \in I$ has an opinion $x_i(t) \in \mathbb{R}$ possibly changing in time $t \in \mathbb{N}$. We assume a splitting $I = S \cup R$ with the understanding that agents in S are *stubborn* agents not changing their opinions, whereas those in R are *regular* agents whose opinions modify in time according to the consensus dynamics

$$x_i(t+1) = \sum_{j \in I} Q_{ij} x_j(t) \quad \forall i \in R$$

where $Q_{ij} \geq 0$ for all $i \in R$ and for all $j \in I$ and $\sum_j Q_{ij} = 1$ for all $i \in R$. The scalar Q_{ij} represents the relative weight that agent i places on agent j's opinion. We will assume that Q only uses the available edges in G, more precisely, our standing assumption will be that

$$Q_{ij} = 0 \Leftrightarrow \{i, j\} \notin E. \tag{1}$$

A basic example is obtained by choosing for each regular agent uniform weights along the edges incident to it, i.e., $Q_{ij} = d_i^{-1}$ for all $i \in R$ and $\{i, j\} \in E$. Assembling opinions of regular and stubborn agents in vectors, denoted by $\boldsymbol{x}^R(t)$ and $\boldsymbol{x}^S(t)$, we can rewrite the dynamics in a more compact form as

$$\begin{aligned} \boldsymbol{x}^R(t+1) &= \boldsymbol{Q}^{11}\boldsymbol{x}^R(t) + \boldsymbol{Q}^{12}\boldsymbol{x}^S(t) \\ \boldsymbol{x}^S(t+1) &= \boldsymbol{x}^S(t) \end{aligned}$$

where the matrices $\boldsymbol{Q}^{11}$ and $\boldsymbol{Q}^{12}$ are non-negative matrices of appropriate dimensions.

Using the adaptivity assumption (1), it is standard to show that $\boldsymbol{Q}^{11}$ is a substochastic asymptotically stable matrix (e.g., spectral radius < 1). Henceforth, $\boldsymbol{x}^R(t) \to \boldsymbol{x}^R(\infty)$ for $t \to +\infty$ with the limit opinions satisfying the relation

$$\boldsymbol{x}^R(\infty) = \boldsymbol{Q}^{11}\boldsymbol{x}^R(\infty) + \boldsymbol{Q}^{12}\boldsymbol{x}^S(0) \tag{2}$$

which is equivalent to

$$\boldsymbol{x}^R(\infty) = (\boldsymbol{I} - \boldsymbol{Q}^{11})^{-1}\boldsymbol{Q}^{12}\boldsymbol{x}^S(0) \tag{3}$$

where $\boldsymbol{I}$ is the identity matrix of appropriate dimension. Notice that $[(\boldsymbol{I} - \boldsymbol{Q}^{11})^{-1}\boldsymbol{Q}^{12}]_{hk} = [\sum_n (\boldsymbol{Q}^{11})^n \boldsymbol{Q}^{12}]_{hk}$ is always

non-negative and is nonzero, if and only if there exists a path in G connecting the regular agent h to the stubborn agent k and not touching other stubborn agents. Moreover, the fact that $\boldsymbol{P}$ is stochastic easily implies that $\sum_k [(\boldsymbol{I}-\boldsymbol{Q}^{11})^{-1}\boldsymbol{Q}^{12}]_{hk}=1$ for all $h\in R$: asymptotic opinions of regular agents are thus convex combinations of the opinions of stubborn agents.

In this paper, we will focus on the situation when $S=S^0\cup\{\ell\}$ and $R=I\backslash S$ assuming that $x_i(0)=0$ for all $i\in S^0$, while $x_\ell(0)=1$, i.e., there are two types of stubborn agents: one type consisting of those in set S^0 that have opinion 0 and the other type consisting of the single agent ℓ that has opinion 1. We investigate how to choose ℓ in $I\backslash S^0$ in such a way to maximize the influence of opinion 1 on the limit opinions. More precisely, let us denote as $x_i^{R,\ell}(\infty)$ the asymptotic opinion of the regular agent $i\in R$ under the above stubborn configuration, and define the objective function

$$H(\ell):=\sum_{i\in R} x_i^{R,\ell}(\infty). \tag{4}$$

Notice that the subset R itself is actually a function of ℓ, however we have preferred not to indicate such dependence to avoid too heavy notations. The *Optimal Stubborn Agent Placement* (OSAP) is then formally defined as

$$\max_{\ell\in I\backslash S^0} H(\ell). \tag{5}$$

In this optimization problem, for any different choice of ℓ, the block matrices $\boldsymbol{Q}^{11}$ and $\boldsymbol{Q}^{12}$ change and a new matrix inversion $(\boldsymbol{I}-\boldsymbol{Q}^{11})^{-1}$ needs to be performed. Such matrix inversions require global information about the network and are not feasible for a distributed implementation. In this paper, we propose a fully decentralized algorithm for the computation of asymptotic opinions and for solving the optimization problem which is based on exploiting a classical analogy with electrical circuits. Under such analogy, we can interpret $x_i^{R,\ell}(\infty)$ as the voltage at node i when nodes in S^0 are kept at voltage 0, while ℓ is kept at voltage 1. This interpretation results to be quite useful as it allows to use typical "tricks" of electrical circuits (e.g., parallel and series reduction, and glueing).

In order to use the electrical circuit analogy, we will need to make an extra "reciprocity" assumption on the weights Q_{ij} assuming that they can be represented through a symmetric matrix $\boldsymbol{C}\in\mathbb{R}^{I\times I}$ (called the conductance matrix) with non-negative elements and $C_{ij}>0$ iff $\{i,j\}\in E$ by imposing

$$Q_{ij}=\frac{C_{ij}}{\sum_j C_{ij}},\quad i\in I,\quad j\in I. \tag{6}$$

The value $C_{ij}=C_{ji}$ can be interpreted as a measure of the "strength" of the relation between i and j. For two regular nodes connected by an edge, the interpretation is a sort of reciprocity in the way the nodes trust each other. Notice that C_{ij} when $i\in S$ is not used in defining the weights, but is anyhow completely determined by the symmetry assumption. Finally, the terms Q_{ij} when $i,j\in S$ do not play any role and for simplicity, we can assume they are all equal to 0. By the definition (6) and from matrix $\boldsymbol{C}$, we are actually defining a square matrix $\boldsymbol{Q}\in\mathbb{R}^{I\times I}$. Compactly, if we consider the diagonal matrix $\boldsymbol{D}_{C1}\in\mathbb{R}^{I\times I}$ defined by $(D_{C1})_{ii}=(C1)_i$, where $\mathbf{1}$ is all ones vector with appropriate dimension, the extension is obtained by putting $\boldsymbol{Q}=\boldsymbol{D}_{C1}^{-1}\boldsymbol{C}$. The matrix $\boldsymbol{Q}$ is said to be a *time-reversible* stochastic matrix in the probability jargon. The special case of uniform weights considered before fits in this framework, by simply choosing $\boldsymbol{C}=\boldsymbol{A}_G$, where $\boldsymbol{A}_G$ is the adjacency matrix of the graph. In this case, all edges have equal strengths and the resulting time-reversible stochastic matrix $\boldsymbol{Q}$ is known as the simple random walk (SRW) on G.

III. The Electrical Network Analogy

In this section, we briefly recall the basic notions of electrical circuits and we illustrate their relation with our problem. A connected graph $G=(I,E)$ together with a conductance matrix $\boldsymbol{C}\in\mathbb{R}^{I\times I}$ can be interpreted as an electrical circuit, where an edge $\{i,j\}$ has electrical conductance $C_{ij}=C_{ji}$ (and thus resistance $R_{ij}=C_{ij}^{-1}$). The pair $(G,\boldsymbol{C})$ will be called an *electrical network* from now on.

An *incidence matrix* on $\boldsymbol{G}$ is any matrix $\boldsymbol{B}\in\{0,+1,-1\}^{E\times I}$ such that $\boldsymbol{B1}=0$ and $B_{ei}\neq 0$ iff $i\in e$. It is immediate to see that given $e=\{i,j\}$, the eth row of $\boldsymbol{B}$ has all zeroes except B_{ei} and B_{ej}: necessarily one of them will be $+1$ and the other one -1 and this will be interpreted as choosing a direction in e from the node corresponding to $+1$ to the one corresponding to -1. Define $\boldsymbol{D}_C\in\mathbb{R}^{E\times E}$ to be the diagonal matrix, such that $(D_C)_{ee}=C_{ij}=C_{ji}$ if $e=\{i,j\}$. A standard computation shows that $\boldsymbol{B}^*\boldsymbol{D}_C\boldsymbol{B}=\boldsymbol{D}_{C1}-\boldsymbol{C}$.

On the electrical network $(G,\boldsymbol{C})$, we now introduce current flows. Consider a vector $\boldsymbol{\eta}\in\mathbb{R}^I$, such that $\boldsymbol{\eta}^*\mathbf{1}=0$: we interpret η_i as the input current injected at node i (if negative being an outgoing current). Given $\boldsymbol{C}$ and $\boldsymbol{\eta}$, we can define the voltage $\boldsymbol{W}\in\mathbb{R}^I$ and the current flow $\boldsymbol{\Phi}\in\mathbb{R}^E$ in such a way that the usual Kirchoff and Ohm's law are satisfied on the network. Compactly, they can be expressed as

$$\begin{cases}\boldsymbol{B}^*\boldsymbol{\Phi}=\boldsymbol{\eta}\\ \boldsymbol{D}_C\boldsymbol{B}\boldsymbol{W}=\boldsymbol{\Phi}\end{cases}.$$

Notice that Φ_e is the current flowing on edge e and sign is positive iff flow is along the conventional direction individuated by $\boldsymbol{B}$ on edge e. Coupling the two equations we obtain $(\boldsymbol{D}_{C1}-\boldsymbol{C})\boldsymbol{W}=\boldsymbol{\eta}$ which can be rewritten as

$$\boldsymbol{L}(\boldsymbol{Q})\boldsymbol{W}=\boldsymbol{D}_{C1}^{-1}\boldsymbol{\eta} \tag{7}$$

where $\boldsymbol{L}(\boldsymbol{Q}):=\boldsymbol{I}-\boldsymbol{Q}$ is the so-called Laplacian of $\boldsymbol{Q}$. Since the graph is connected, $\boldsymbol{L}(\boldsymbol{Q})$ has rank $|I|-1$ and $\boldsymbol{L}(\boldsymbol{Q})\mathbf{1}=0$. This shows that (7) determines $\boldsymbol{W}$ up to translations. Notice that $(L(Q)W)_i=0$, for every $i\in I$ such that $\eta_i=0$. For this reason, in analogy with the Laplacian equation in continuous spaces, $\boldsymbol{W}$ is said to be *harmonic* on $\{i\in I|\eta_i=0\}$. Clearly, given a subset $S\subseteq I$ and a $\boldsymbol{W}\in\mathbb{R}^I$ which is harmonic on S^c, we can always interpret $\boldsymbol{W}$ as a voltage with input currents given by $\boldsymbol{\eta}=\boldsymbol{D}_{C1}\boldsymbol{L}(\boldsymbol{Q})\boldsymbol{W}$ which will necessarily be supported on S. Actually, $\boldsymbol{W}$ is the only voltage harmonic on S^c and with assigned values on S.

It is often possible to replace an electrical network by a simplified one without changing certain quantities of interest. An useful operation is *gluing:* if we merge vertices having the same voltage into a single one, while keeping all existing edges, voltages, and currents are unchanged, because current never flows between vertices with the same voltage. Another useful operation is replacing a portion of the electrical network connecting two nodes h, k by an *equivalent resistance*, a single resistance denoted as R^{eff}_{hk} which keeps the difference of voltage $W(h) - W(k)$ unchanged. Following our interpretation of (6), we see $(R^{\text{eff}}_{hk})^{-1}$ as measuring of the strength of the relation between two nodes h, k. Two basic cases of this operation consist in removing nodes of degree two by adding resistances (series law) and replacing multiple edges between two nodes with a single one having conductance equal to the sum of the various conductances (parallel law). These techniques will be extensively used in deriving our algorithm.

A. Social Networks as Electrical Networks

We are now ready to state the relationship between social and electrical networks. Consider a connected graph $G = (I, E)$, a subset of stubborn agents $S \subseteq I$, and a stochastic time-reversible matrix $\boldsymbol{Q}$ having conductance matrix $\boldsymbol{C}$. Notice that relation (2) can also be written as

$$\boldsymbol{L}(\boldsymbol{Q}) \begin{pmatrix} \boldsymbol{x}^{R,\ell}(\infty) \\ \boldsymbol{x}^{S}(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\theta} \end{pmatrix} \tag{8}$$

for some suitable vector $\boldsymbol{\theta} \in \mathbb{R}^S$ (where $\boldsymbol{\theta}$ represents the initial opinions of the stubborn agents). Comparing with (7), it follows that $\boldsymbol{x}^{R,\ell}(\infty)$ can be interpreted as the voltage at the regular agents when stubborn agents have fixed voltage $\boldsymbol{x}^S(0)$ or, equivalently, when input currents $\boldsymbol{D}_{C\mathbf{1}}\boldsymbol{\theta}$ are applied to the stubborn agents. Because of (8), the vector $\boldsymbol{x}^{R,\ell}(\infty)$ is said to be the *harmonic* extension of $x_S(0)$ and the function H defined in (4) the *HIC*.

Thanks to the electrical analogy as we can compute the asymptotic opinion of the agents by computing the voltages in the graph seen as an electrical network. From now on, we will stick to this equivalence and we will exclusively talk about an electrical network $(G, \boldsymbol{C})$ with a subset $S^0 \subseteq I$ of nodes at voltage 0. For any $\ell \in I \backslash S_0$, $W^{(\ell)}$ denotes the voltage on I such that $W^{(\ell)}(i) = 0$ for every $i \in S^0$ and $W^{(\ell)}(\ell) = 1$. Using this notation and the association between limiting opinions and electric voltages provided in (7) and (8), we can express the HIC of node ℓ as

$$H(\ell) = \sum_{i \in I, i \neq \ell} W^{(\ell)}(i).$$

Consider the following simple example.

Example 1 (Line Graph): Consider the electrical network (L, A_L), where L is the line graph $L = (\{0, \ldots, n\}, E_L)$. Moreover, assume $S^0 = \{0\}$ and $S^1 = \{n\}$. Then, the induced voltage W satisfies $W(i) = \frac{i}{n}$, and thus $H(n) = \frac{n+1}{2}$.

The following result, which is an immediate consequence of (8), provides a formula that is useful in computing the voltages.

Lemma 1 (Voltage Scaling): Consider the electrical network $(G, \boldsymbol{C})$ and a subset $S^0 \subseteq I$. Let $W^{(\ell)}$ be the voltage that is 0 on S^0 and 1 on ℓ. Let W be another voltage, such that $W(s) = w_0$ if $s \in S^0$ and $W(\ell) = w_1$. Then, for every node $i \in I$ it holds

$$W(i) = w_0 + (w_1 - w_0)W^{(\ell)}(i). \tag{9}$$

IV. The Electrical Analogy on Trees

The case when the graph G is a *tree* is very important to us, since we are able to prove a number of useful results on the solution of the OSAP problem and it will play a pivotal role in the introduction of our MPA for general graphs. To begin with, we establish some useful notation and basic results.

In the foregoing, we assume to have fixed a tree $T = (I, E)$, a conductance matrix $\boldsymbol{C}$ and the subset $S^0 \subseteq I$ of 0 voltage nodes. We next define subsets of nodes of the given tree, which enable us to isolate the effects of the upstream and downstream neighbors of a node in computing its harmonic centrality and its voltage. Given a pair of distinct nodes $i, j \in I$, we let $I^{<ij}$ denote the subset of nodes that form the subtree rooted at node i that does not contain node j. Formally

$$I^{<ij} := \{h \in I | \text{the simple path from } h \text{ to } j \text{ goes through } i\}.$$

We also define $I^{ij>} := I^{<ji}$, $I^{ij<} := (I^{ij>})^c \cup \{j\}$, $I^{>ij} := I^{ji<}$, and $I^{i<j} := I^{ij<} \cap I^{>ij}$. Fig. 1 illustrates these definitions.

The induced subtrees is denoted using the same apex $T^{<ij}$ and so on; similarly the HIC of the nodes on each of the trees above is denoted as $H^{<ij}(\cdot)$ and so on. Finally, we use the notation $R^{\text{eff}}_{<ij}$ to denote the effective resistance inside $T^{<ij}$ between $S^0 \cap I^{<ij}$ (considered as a unique collapsed node) and node i. We will conventionally interpret this resistance as infinite in case $S^0 \cap I^{<ij} = \emptyset$. Analogously, we define $R^{\text{eff}}_{ij>} := R^{\text{eff}}_{<ji}$.

Given a pair of distinct nodes $i, j \in I$, consider the two voltages $W^{(i)}$ and $W^{(j)}$. If we restrict them to $T^{<ij}$, we may interpret them as two voltages on $T^{<ij}$ which are 0 on S^0 and take values in node i, respectively, $W^{(i)}(i) = 1$ and $W^{(j)}(i)$. It follows by applying Lemma 1 that

$$W^{(j)}(\ell) = W^{(j)}(i)W^{(i)}(\ell) \quad \forall \ell \in I^{<ij}. \tag{10}$$

Moreover, $W^{(j)}(i)$ can be computed through effective resistances replacing the circuit determined by the subtree $T^{<ij}$ by an equivalent circuit represented by a line graph with three nodes S^0, i, and j as in Fig. 2. We recall that collapsing all nodes of S^0 in a single node is possible as they all have the same voltage. Moreover, by definition, the edge $\{S^0, i\}$ has resistance $R^{\text{eff}}_{<ij}$, while $\{i, j\}$ has resistance R^{eff}_{ij}. Therefore, as the current flowing along the two edges is the same, Ohm's law yields

$$\frac{W^{(j)}(j)}{R^{\text{eff}}_{<ij} + R^{\text{eff}}_{ij}} = \frac{W^{(j)}(i)}{R^{\text{eff}}_{<ij}}$$

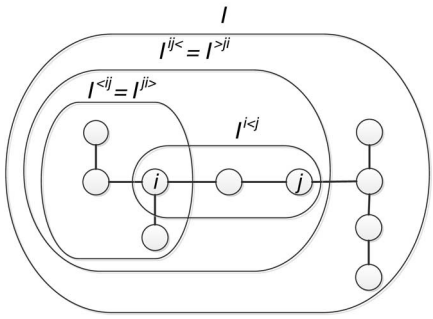


Fig. 1. Example of tree presenting the notation of the subsets of I.

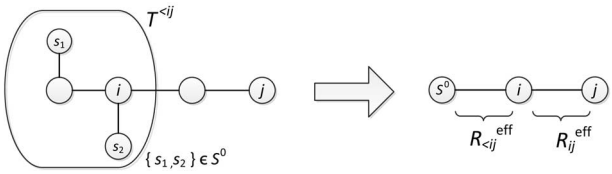


Fig. 2. Subtree equivalently represented as a line graph.

yielding

$$W^{(j)}(i) = \frac{R^{\text{eff}}_{<ij}}{R^{\text{eff}}_{<ij} + R^{\text{eff}}_{ij}} \tag{11}$$

(equal to 1 in case $S^0 \cap I^{<ij} = \emptyset$). From relations (10) and (11), later on we will derive iterative equations for the computation of voltages on a tree. In the rest of this section, we prove some properties of the OSAP on a tree, which provide significant simplifications at the computational level.

A. Nodes in S^0 can be Assumed to be Leaves

A first general remark to be made is that nodes in S^0 break the computation of the HIC into separate noninteracting components. Indeed, the induced subgraph $T_{|I\setminus S_0}$ is a forest composed of subtrees $\{T_h = (J_h, E_h)\}_{h\in\{1,\dots,n\}}$. For every $h \in \{1,\dots,n\}$, define the set S^0_h as the set of type 0 stubborn nodes that are adjacent to nodes in J_h in the graph G, that is

$$S^0_h := \{i \in S^0 | \exists j \in J_h : \{i,j\} \in E\}.$$

Then, define the tree $\widehat{T}_h = G|_{S^0_h \cup I_h}$, which is therefore the tree T_h augmented with its type 0 stubborn neighbors in the original graph G. An example of this procedure is shown in Fig. 3. The following result shows that it is sufficient to compute the HIC on the subtrees $\widehat{T}_h$; in its statement we denote by $\widehat{H}_h$ the HIC function on $\widehat{T}_h$.

Proposition 2 (Tree Decomposition): For every $h = 1,\dots,n$, and for very $\ell \in I_h$

$$\widehat{H}_h(\ell) = H(\ell) \quad \forall \ell \in I_h.$$

Proof: Given $\ell \in I_h$, the voltage $W^{(\ell)}(j)$ is zero for every j such that the (unique) path from j to ℓ goes through a type 0 stubborn node. This implies that $W^{(\ell)}(j) = 0$, for every $j \in I\setminus I_h$. This observation proves the result. □

Thanks to this property as it is sufficient to study the HIC on the subtrees $\widehat{T}_h$ and then compute

$$\max_{\ell\in I\setminus S^0} H(\ell) = \max_{h\in\{1,\dots,n\}} \max_{\ell\in I_h} \widehat{H}(\ell).$$

Consequently, we will assume from now on that *stubborn agents are located in the leaves*, without any loss of generality.

B. Further Properties in Case of Unitary Resistances

In this paragraph, we present a number of results characterizing the OSAP in the special, but important, case when all resistances are equal (without further loss of generality, we set their value to be one). We recall that the social interpretation of this case is when all edges expressing relationship among regular nodes have equal strength.

Before proceeding, let us first consider the trivial situation in which there is *only one type 0 stubborn* (which is a leaf of T, without loss of generality). In this case, the maximizer of H in I is the node *adjacent* to the node in S^0. Indeed, let s be the node in S^0 and i its unique neighbor: then $W^{(i)}(\ell) = 1$ for all $\ell \neq s$, that is $H(i) = |I| - 1$, which is obviously the largest achievable value. In order to deal with more than one stubborn, we need the following preliminary result.

Lemma 3: Let $T = (I, E)$ be a tree with unitary resistances and $a, b \in I$. Let W be the voltage such that $W(a) = 0$ and $W(b) = 1$. Then, the effective resistance between a and b satisfies

$$R^{\text{eff}}_{ab} \leq 2\sum_{i\in I} W(i) - 1.$$

Proof: Consider the simple path $a, j_1, \dots, j_{t-1}, b$ connecting a to b in T. By Example 1, we clearly have that $W(j_s) = s/t$. Therefore

$$2\sum_{i\in I} W(i) - 1 \geq 2\sum_{s=1}^{t-1} W(j_s) + 1 = t = R^{\text{eff}}_{ab}.$$

□

We are now ready to state and prove the main result of this section, which provides a set of necessary conditions for a node to be a solution to OSAP: provided there are at least two stubborn, the solution lies on a path connecting two of them and is a node of degree at least three, unless all paths connecting two stubborn are made of nodes of degree two.

Theorem 4 (Necessary Conditions for OSAP): Consider the OSAP problem (5) over a tree $T = (I, E)$ with all unitary resistances. Assume that $S^0 \subseteq I$ is a subset of leaves and that $|S^0| \geq 2$. Define

$$K = \{i \in I | \exists s', s'' \in S^0 \text{such that } i \text{ belongs to a simple path between } s' \text{ and } s''\}.$$

Then,

$$\arg\max_{\ell\in I} H(\ell) = \arg\max_{\ell\in K} H(\ell).$$

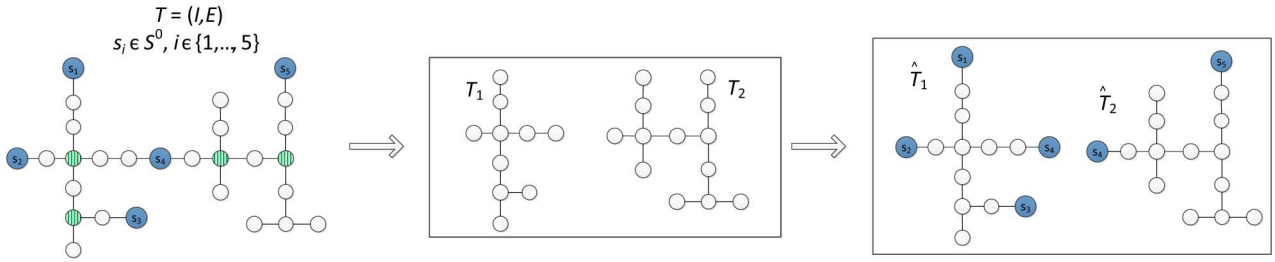


Fig. 3. Tree with type 0 stubborn nodes in blue, together with its decomposition according to Proposition 2.

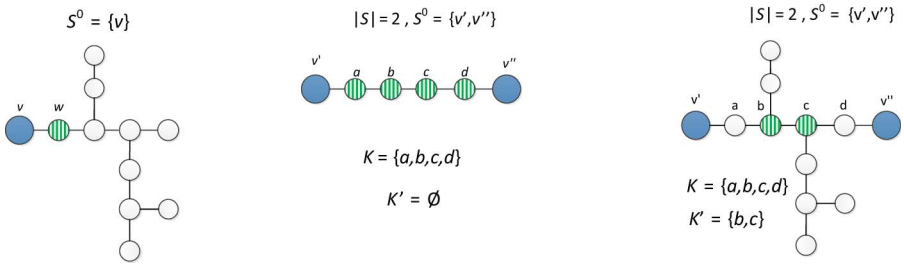


Fig. 4. Three different trees: blue nodes are stubborn and the ones with green stripes are the candidate solutions.

Moreover, if $K' := \{i \in K | d_i \geq 3\} \neq \emptyset$, then

$$\arg\max_{\ell \in I} H(\ell) = \arg\max_{\ell \in K'} H(\ell).$$

Proof: Let $j \notin K$. Then, there exists $i \in K$ such that every path from j to S^0 contains i. We claim that $H(i) > H(j)$, proving that the optimum must belong to K. Harmonic influences can be computed as follows

$$H(j) = \sum_{y \in I^{<ij} \setminus \{i\}} W^{(j)}(y) + \sum_{x \in I^{<j}} W^{(j)}(x) + |I^{ij>}| - 1$$
$$H(i) = \sum_{y \in I^{<ij} \setminus \{i\}} W^{(i)}(y) + |I^{<ij}| + |I^{ij>}| - 1.$$

We can then observe that, as a consequence of the scaling formula (10)

$$\sum_{y \in I^{<ij} \setminus \{i\}} W^{(j)}(y) < \sum_{y \in I^{<ij} \setminus \{i\}} W^{(i)}(y)$$

while, clearly, $\sum_{x \in I^{<ij}} W^{(j)}(x) \leq |I^{<ij}|$. These inequalities prove the claim.

Next, we need to prove that the optimal node has degree larger than two, that is, it belongs to K' (provided this set is not empty). Consider a path between two nodes in S^0 and a maximal string of consecutive nodes of degree two in this path, denoted as $j_1, j_2, \ldots, j_n$. Let $a, b \in S^0 \cup K'$ be the two nodes such that $\{a, j_1\}, \{j_n, b\} \in E$. For any $x = 1, \ldots, n$, consider $W^{(j_x)}$ and notice that, because of formula (10), there hold

$$W^{(j_x)}(\ell) = W^{(j_x)}(a) W^{(a)}(\ell) \quad \forall \ell \in I^{<ab}$$
$$W^{(j_x)}(\ell) = W^{(j_x)}(b) W^{(b)}(\ell) \quad \forall \ell \in I^{ab>}.$$

Combining these equations with Example 1, we can compute

$$\begin{aligned} H(j_x) &= \sum_{\ell \in I^{<ab}} W^{(j_x)}(\ell) + \sum_{\ell \in I^{ab>}} W^{(j_x)}(\ell) + \sum_{\ell \in I^{j_1 <j_n}} W^{(j_x)}(\ell) \\ &= W^{(j_x)}(a) H^{<ab}(a) + W^{(j_x)}(b) H^{ab>}(b) \\ &\quad + \frac{1+W_a}{2}(x-1) + \frac{1+W_b}{2}(n-x) + 1. \end{aligned}$$

Since, by (11), $W(a) = \frac{R^{\text{eff}}_{<ab}}{R^{\text{eff}}_{<ab}+x}$ and $W(b) = \frac{R^{\text{eff}}_{ab>}}{R^{\text{eff}}_{ab>}+n-x+1}$, we deduce that

$$\begin{aligned} H(j_x) = &\frac{R^{<ab}}{R^{<ab}+x}\left(H^{<ab}(a) + \frac{x-1}{2}\right) + \frac{x-1}{2} \\ &+ \frac{R^{ab>}}{R^{ab>}+n-x+1}\left(H^{ab>}(b) + \frac{n-x}{2}\right) + \frac{n-x}{2} + 1. \end{aligned}$$

Notice that the above expression naturally determines an extension for $x \in [1, n] \subseteq \mathbb{R}$ and it is straightforward to compute that

$$\begin{aligned} \frac{\partial^2}{\partial x^2} H(j_x) = &\frac{2R^{<ab}(H^{<ab}(a) - \frac{R^{<ab}+1}{2})}{(R^{<ab}+x)^2} \\ &+ \frac{2R^{ab>}(H^{ab>}(b) - \frac{R^{ab>}+1}{2})}{(R^{ab>}+x)^2}. \end{aligned}$$

Since Lemma 3 implies that $2H^{<ab}(a) \geq R^{<ab}+1$ and $2H^{ab>}(b) \geq R^{ab>}+1$, this second derivative is non-negative and then $H(j_x)$ is convex in x. We conclude that the value in a or in b of the HIC is greater or equal than the value on the nodes having degree two. By this statement, the proof is complete. □

As a consequence of the necessary conditions in Theorem 4, in order to solve the OSAP it is sufficient to compute the HIC only on a subset of nodes, which can be much smaller than the whole node set of the social network. Example of the effectiveness of

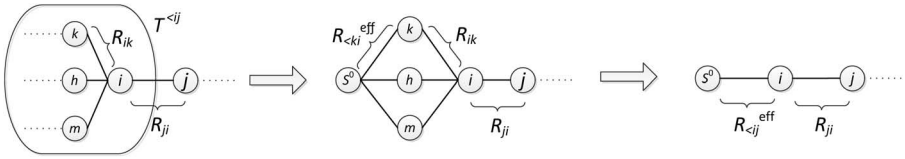


Fig. 5. Equivalent representation of parallel paths between i and S^0.

this result are provided in Figs. 3 and 4, where the candidate nodes to solve problem (5) are pictured with green stripes: in the former example, the candidates are reduced from 22 to 4.

V. Message Passing on Trees

In this section, we design an MPA, which computes the HIC of every node of a tree in a distributed way. We begin by outlining the structure of a general MPA on a tree. Preliminarily, define any node i in the graph as the *root*. In the first phase, messages are passed inwards: starting at the leaves, each node passes a message along the unique edge toward the root node. The tree structure guarantees that it is possible to obtain messages from all other neighbor nodes before passing the message on. This process continues until the root i has obtained messages from all its neighbors. The second phase involves passing the messages back out: starting at the root, messages are passed in the reverse direction. The algorithm is completed when all leaves have received their messages.

Next, we show how this approach can be effective in our problem. Take a generic root node $i \in I \backslash S^0$ and, for every $j \in N_i$, notice the following iterative structure of the subtree rooted in i and not containing j:

$$I^{<ij} = \bigcup_{k \in N_i \backslash \{j\}} I^{<ki} \cup \{i\}$$

This relation, together with (10), yields

$$\begin{aligned} H^{<ij}(i) &= \sum_{k \in N_i \backslash \{j\}} \sum_{\ell \in T^{<ki}} W^{(i)}(\ell) + 1 \\ &= \sum_{k \in N_i \backslash \{j\}} W^{(i)}(k) H^{<ki}(k) + 1. \end{aligned} \tag{12}$$

Forthermore, (11) yields

$$W^{(i)}(k) = \frac{R^{\text{eff}}_{<ki}}{R^{\text{eff}}_{<ki} + R_{ik}} \tag{13}$$

where, we conventionally assume that $R^{\text{eff}}_{<ki} = \infty$ and $W^{(i)}(k) = 1$ if $S^0 \cap I^{<ki} = \emptyset$. On the other hand, also effective resistances $R^{\text{eff}}_{<ki}$ admit an iterative representation. Indeed, replace $T^{<ij}$ with the equivalent circuit consisting of nodes: S^0, i, j, and all the nodes $k \in N_i \backslash \{j\}$. Between S^0 and i, there are $|N_i| - 1$ parallel (length 2) paths each passing through a different $k \in N_i \backslash \{j\}$ and having resistance $R^{\text{eff}}_{<ki} + R_{ik}$ (see Fig. 5). Therefore, using the parallel law for resistances we obtain

$$R^{\text{eff}}_{<ij} = \left(\sum_{k \in N_i \backslash \{j\}} \frac{1}{R^{\text{eff}}_{<ki} + R_{ik}} \right)^{-1}. \tag{14}$$

The three relations (12), (13), and (14) determine an iterative algorithm to compute $H(i)$ at every node i starting from leaves and propagating to the rest of the graph. More precisely, define $H^{i \to j} := H^{<ij}(i)$, and $W^{i \to j} := W^{(i)}(j)$ to be thought as messages sent by node i to node j along the edge $\{i, j\}$. From (12), (13), and (14), we easily obtain the following iterative relations

$$\begin{aligned} H^{i \to j} &= \sum_{k \in N_i \backslash \{j\}} W^{k \to i} H^{k \to i} + 1 \\ W^{i \to j} &= \left(1 + R_{ij} \sum_{k \in N_i \backslash \{j\}} \frac{1 - W^{k \to i}}{R_{ik}} \right)^{-1}. \end{aligned} \tag{15}$$

Notice that a node i can only send messages to a neighbor j after it has received messages $H^{k \to i}$ and $W^{k \to i}$ from all its neighbors k but j. Iteration can start at leaves (having just one neighbor) with the following initialization step

$$\begin{aligned} H^{i \to j} &= \mathbf{1}_{i \notin S^0} \\ W^{i \to j} &= \mathbf{1}_{i \notin S^0} \end{aligned} \tag{16}$$

where we denote by $\mathbf{1}_{i \notin S^0}$ a vector indexed in I which has entry 1 if $i \notin S^0$ and entry 0 elsewhere. Notice that each regular agent i can finally compute $H(i)$ by the formula

$$H(i) = \sum_{k \in N_i} W^{k \to i} H^{k \to i} + 1.$$

Clearly this algorithm terminates in finite time, because as soon as a node i has received all messages from its neighbors, it can compute the correct value of the HIC. It is then interesting to estimate the convergence time of the algorithm, in terms of number of steps, as well as the number of messages to be sent by the nodes and the number of operations to be performed.

Proposition 5 (Complexity of MPA): Consider the MPA defined in (15) and (16) on tree T. Then

1) the algorithm converges after at most $\operatorname{diam}(T)$ steps;
2) the number of messages (triples) sent by node i is d_i;
3) the number of messages (triples) sent in the whole network is not larger than $2|I|$;
4) the number of operations performed by each node is $O(d_i^2)$;
5) the number of operations in the whole network is $O(\sum_{i \in I \backslash S^0} d_i^2)$.

Proof: The time before $H(i)$ converges is clearly equal to the distance between i and the furthest leave; then, the first claim follows. It is immediate to observe that, if we count a triple as a message, the number of messages sent by each node is d_i. Then, across the whole network, we have $\sum_{i\in I} d_i = 2|E| = 2|I| - 2$ messages, because T is a tree. The number of operations for each node i to compute a message for a neighbor j is proportional to $d_i - 1$. As node i has to compute d_i different messages, the computational cost for a node i is of the order $O(d_i^2)$. The last claim follows by summing over all the network. □

A centralized algorithm to compute the HIC in any connected network was proposed in [15]. Being centralized, this algorithm requires full knowledge of the topology of the graph. The computational complexity of the algorithm [15] is $O((|I| - |S^0|)^3)$. As $\sum_{i\in I\setminus S^0} d_i^2 \leq |I\setminus S^0| d_{\max}^2$, provided $d_{\max}$ denotes the largest degree, on graphs with bounded degrees the MPA has a much smaller complexity $O(|I| - |S^0|)$. The main drawback of MPA is its limitation to trees. In the Section VI, we will work toward removing this restrictive assumption.

VI. Message Passing on General Graphs

MPAs are commonly designed on trees, but also implemented with some modification over general graphs. In many cases, the application is just empirical, without a proof of convergence. We will see in this section how to apply the MPA to every graph, with suitable modifications in order to manage the new issues. Namely, we design an "iterative" version of the MPA of Section V, which can run on every network, regardless of the presence of cycles. We show that for regular graphs with unitary resistances, this algorithm converges (but not necessarily to the correct HIC values as we demonstrate next). We also present simulation results that show the algorithm effectiveness in computing the HIC on families of graphs with cycles.

We let the nodes send their messages at every time step, so that we denote them as

$$W^{i\to j}(t), \quad H^{i\to j}(t), \quad \text{for all } t \geq 0.$$

The dynamics of messages are

$$H^{i\to j}(t+1) = \sum_{k\in N_i\setminus\{j\}} W^{k\to i}(t) H^{k\to i}(t) + 1 \tag{17a}$$

$$W^{i\to j}(t+1) = \left(1 + R_{ij} \sum_{k\in N_i\setminus\{j\}} \frac{1 - W^{k\to i}(t)}{R_{ik}}\right)^{-1} \tag{17b}$$

(where $R_{ij} = C_{ij}^{-1}$ are the edge resistances) if $i\notin S^0$ and

$$H^{i\to j}(t+1) = 0 \tag{18a}$$

$$W^{i\to j}(t+1) = 0 \tag{18b}$$

otherwise. The initialization is

$$H^{i\to j}(0) = \mathbf{1}_{i\notin S^0}$$
$$W^{i\to j}(0) = \mathbf{1}_{i\notin S^0}. \tag{19}$$

By these definitions of messages, we have defined the MPA for general graphs. Additionally, we should define a termination criterion: for instance, the algorithm may stop after a number of steps which is chosen *a priori*. At every time t, each agent i can compute an approximate $H(i)^{(t)}$ by the formula

$$H(i)^{(t)} = \sum_{k\in N_i} W^{k\to i}(t) H^{k\to i}(t) + 1.$$

This new algorithm clearly converges to the HIC if the graph is a tree, and the convergence time is not larger than the diameter of the graph. Otherwise, the algorithm is not guaranteed to converge: furthermore, if the algorithm happens to converge, then the convergence value may be different from the HIC.

In order to illustrate the issues caused by the presence of cycles, we can use the so-called *computation trees* [13], which are constructed in the following way. Given a graph G, we focus on a "root" node, and for all $t \in \mathbb{N}$, we let the nodes at distance t from the root (the level t of the tree) be the nodes whose messages reach the root after t iterations of the MPA. Note that if the graph G is a tree, the computation tree is just equal to G; otherwise, it has a number of nodes which diverges when t goes to infinity. As an example, Fig. 6 shows the first four levels of a sample computation tree. In our MPA, each node i is computing its own HIC *in the computation tree instead than on the original graph*. As the number of levels of the computation tree diverges, the computation procedure may not converge, and—if converging–may not converge to the harmonic influence in the original graph.

A. MPA on Regular Graphs

Section VI-A is devoted to prove the following convergence result.

Theorem 6 (Convergence of MPA): Consider a connected graph $G = (I, E)$ with unitary resistances and $S^0 \subseteq I$. Assume, moreover, that $d_i = d$ for all nodes $i \in I\setminus S^0$. Then, the MPA algorithm described by (17), (18), and (19) converges.

We start analyzing the behavior of the W variables, which is independent from the H variables. Notice that under the assumptions of Theorem 6, the relations (17b)–(18b) simplifies to

$$W^{i\to j}(t+1) = \begin{cases} \left(d - \sum_{k\in N_i\setminus\{j\}} W^{k\to i}(t)\right)^{-1}, & \text{if } i\notin S^0 \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

Lemma 7: Under the assumptions of Theorem 6, there exist numbers $W^{i\to j} \in [0, 1[$ satisfying, for all $\{i, j\} \in E$, the fixed point relations

$$W^{i\to j} = \begin{cases} \left(d - \sum_{k\in N_i\setminus\{j\}} W^{k\to i}\right)^{-1}, & \text{if } i\notin S^0 \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

and such that $W^{i\to j}(t) \to W^{i\to j}$ as $t \to +\infty$ for all $\{i, j\} \in E$.

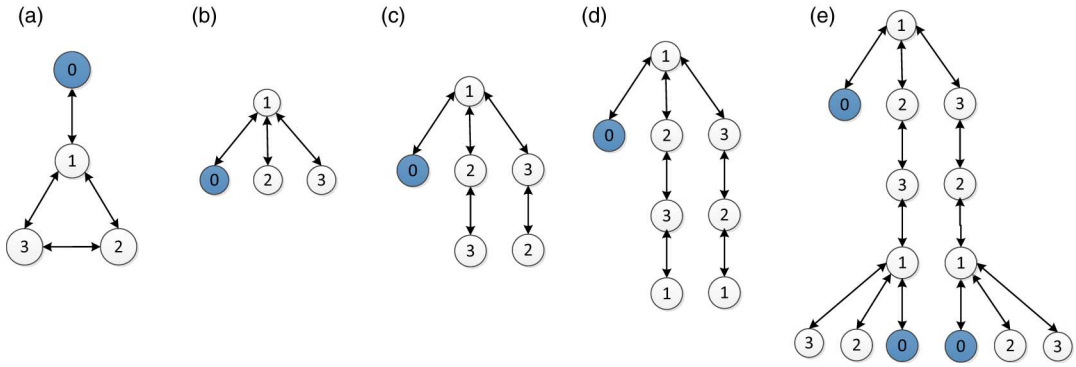


Fig. 6. Graph and computation trees of an MPA from root 1: (a) original graph, (b) computation tree (first iteration), (c) computation tree (second iteration), (d) computation tree (third iteration), and (e) computation tree (fourth iteration).

Proof: We already know that the sequence $W^{i\to j}(t)$ is bounded, $0 \le W^{i\to j}(t) \le 1$ for all $\{i,j\} \in E$ and for all $t \ge 0$. Notice now that $W^{i\to j}(1) \le W^{i\to j}(0)$ for all $\{i,j\} \in E$. On the other hand, it is immediate to check, from expression (20), that the following inductive step holds true:

$$W^{i\to j}(t) \le W^{i\to j}(t-1) \Rightarrow W^{i\to j}(t+1) \le W^{i\to j}(t).$$

This implies that $W^{i\to j}(t)$ is a decreasing sequence for all $\{i,j\} \in E$: we thus get convergence to a limit satisfying (21). Finally, to show that $W^{i\to j} < 1$ for all $\{i,j\}$, we notice that

$$W^{i\to j} = 1 \Rightarrow W^{k\to i} = 1 \quad \forall k \in N_i \backslash \{j\}.$$

Iterating this argument and being the graph connected, we obtain the absurd statement that $W^{h\to \ell} = 1$ for $h \in S^0$ and some $\ell \in N_h$.

We can actually say more on the limit numbers $W^{i\to j}$ s.

Lemma 8: Under the assumptions of Theorem 6, the numbers $W^{i\to j}$s defined in Lemma 7 satisfy the bounds

$$\sum_{k\in N_i\backslash\{j\}} W^{k\to i} < 1 \quad \forall i \in I \quad \forall j \in N_i.$$

Proof: For every $\{i,j\} \in E$, define $\tilde{W}^{i\to j} = \sum_{k\in N_i\backslash\{j\}} W^{k\to i}$. From (21), we can easily obtain the iterative relation

$$\tilde{W}^{i\to j} = \sum_{k\in (N_i\backslash\{j\})\cap (S^0)^c} \left(d - \sum_{k\in N_i\backslash\{j\}} \tilde{W}^{k\to i} \right)^{-1}.$$

Suppose that

$$\alpha = \tilde{W}^{i\to j} = \max\{\tilde{W}^{h\to k} | \{h,k\} \in E, h \notin S^0\}.$$

Clearly, by Lemma 7, we have $\alpha \in [0, d-1[$ and we easily obtain the relation

$$\alpha \le \frac{|(N_i\backslash\{j\}) \cap (S^0)^c|}{d-\alpha} \le \frac{d-1}{d-\alpha} \tag{22}$$

which yields $\alpha \in [0,1]$. Suppose by contradiction that $\alpha = 1$ and notice that inequalities in (22) would yield

$$\tilde{W}^{i\to j} = 1 \Rightarrow (N_i\backslash\{j\}) \cap S^0 = \emptyset$$

and

$$\tilde{W}^{k\to i} = 1 \quad \forall k \in N_i\backslash\{j\}.$$

Iterating this argument, we easily obtain the result that there is no path from nodes in S^0 to i, contradicting the fact that G is connected. □

Finally, we analyze the behavior of the sequences $H^{i\to j}(t)$.

Proposition 9: Under the assumptions of Theorem 6, there exist numbers $H^{i\to j} \in [0,1[$ satisfying, for all $\{i,j\} \in E$, the fixed point relations

$$H^{i\to j} = \begin{cases} \sum\limits_{k\in N_i\backslash\{j\}} W^{k\to i} H^{k\to i} + 1, & \text{if } i \notin S^0 \\ 0, & \text{otherwise} \end{cases}$$

and such that $H^{i\to j}(t) \to H^{i\to j}$ as $t \to +\infty$ for all $\{i,j\} \in E$.

Proof: It is convenient to gather the sequences $H^{i\to j}(t)$ into a vector sequence $\boldsymbol{H}(t) \in \mathbb{R}^E$ and rewrite the iterative relation in (17a) as

$$\boldsymbol{H}(t+1) = \boldsymbol{W}(t)\boldsymbol{H}(t) + \mathbf{1}_{(S^0)^c}$$

where $\boldsymbol{W}(t) \in \mathbb{R}^{E\times E}$ is given by

$$W(t)_{i\to j, h\to k} := \begin{cases} W^{h\to k}(t), & \text{if } k = i \notin S^0, h \ne j \\ 0, & \text{otherwise.} \end{cases}$$

We know from Lemmas 7 and 8 that $W(t)$ converges to a matrix $\boldsymbol{W} \in \mathbb{R}^{E\times E}$ with non-negative elements and satisfying the row relations

$$\sum_{\{h,k\}\in E} W_{i\to j, h\to k} = \begin{cases} \sum\limits_{h\in N_i\backslash\{j\}} W^{h\to i}, & \text{if } \notin S^0 \\ 0, & \text{otherwise} \end{cases}$$

Notice that $\boldsymbol{W}$ is an asymptotically stable substochastic matrix such that

$$\|\boldsymbol{W}\|_\infty = \max\left\{\sum_{\{h,k\}\in E} W_{i\to j,h\to k} | \{i,j\}\in E\right\} < 1.$$

Straightforward calculus considerations then yield convergence of $\boldsymbol{H}(t)$. □

B. Simulations

We have performed extensive simulations of our algorithm on well-known families of random graphs such as Erdős–Rényi and Watts–Strogatz, obtaining very encouraging results. First, the algorithm is convergent in every test. Second, in many cases the computed values of HIC are very close to the correct values, which we can obtain by the benchmark algorithm in [15]. In order to make this claim more precise, we define two notions of error. We denote by $H(i)$ the correct HIC of node i and by $\widehat{H}^{(t)}(i)$ the output of the algorithm in node i after t steps. We define the *mean deviation error* at time step t as

$$e_{\text{dev}}(t) = \frac{\sum_{i\in I} |H(i) - \widehat{H}^{(t)}(i)|}{|I|}.$$

Additionally, as we are interested in the OSAP problem, we are specially concerned about obtaining the right ranking of the nodes, in terms of HIC. We thus define the *mean rank error* at time step t as

$$e_{\text{rank}}(t) = \frac{\sum_{i\in I} |\text{rank}_H(i) - \text{rank}_{\widehat{H}^{(t)}}(i)|}{|I|}$$

where for a function $f : I \to \mathbb{R}$, we denote by $\text{rank}_f(i)$ the position of i in the vector of nodes, sorted according to the values of f.

We now move on to describe some simulation results in more detail. As the stopping criterion, we ask that the mean difference between the output of two consecutive steps is below a threshold, chosen as 10^{-5}. As the topology of the graphs, we choose random graphs generated according to the well-known Erdős–Rényi model: $\mathcal{G}(n,p)$ is a graph with n nodes such that every pair of nodes is independently included in the edge set with probability p (for further details see [2]). For simplicity, all resistances are set equal to one.

First, we consider an example of Erdős–Rényi random graph with $n = 15$ and $p = 0.2$; nodes 1, 2, and 3 are stubborn in S^0. In spite of the presence of several cycles in the sampled graph, the algorithm finds the maximum of the HIC correctly; see Fig. 7. Fig. 8 plots the mean deviation error and the mean ranking error as functions of time steps in the same experiment: after just four steps the ranking error reaches a minimum value. Note that although the obtained ranking is not entirely correct, the three nodes with highest HIC are identified and the HIC profile is well approximated.

The true HIC is smaller than the approximated one and the deviation error is localized on some node. These facts, which can be widely observed in our simulations, can be explained by thinking of the computation trees as in Section VI. Indeed, for each node i the MP algorithm actually computes the HIC on the computation tree rooted in i. The computation tree is closer to the actual graph when the node i is farther from cycles or it belongs to fewer of them; then also the computed HIC will be closer to the true one. Moreover, when the number of iterations grows the computation tree has an increasing number of nodes: even if they are far from the node considered, they contribute to overestimate its centrality.

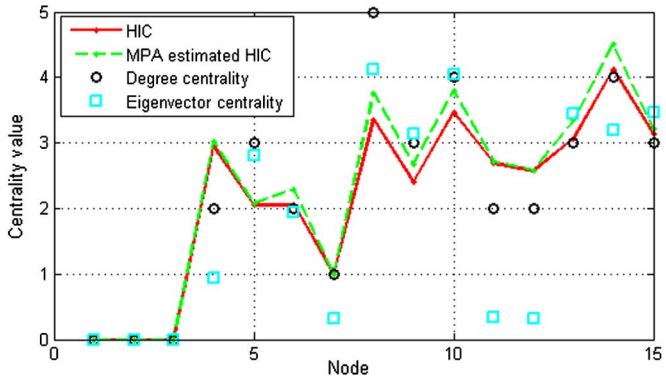


Fig. 7. Comparison between the actual values of H in the nodes of an Erdős–Rényi graph with 15 nodes and the values estimated by the MPA. Degree and eigenvector centralities are also shown.

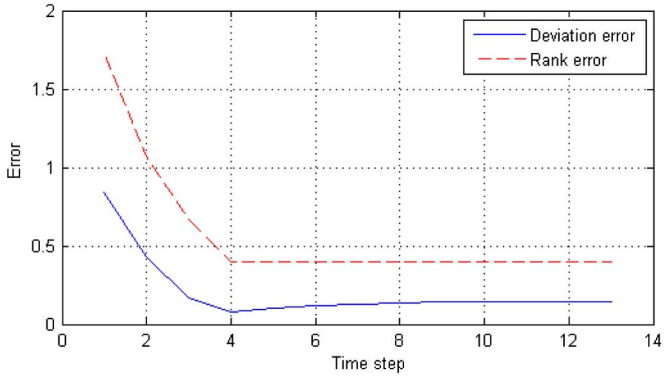


Fig. 8. Mean deviation error and mean ranking error of the MPA as functions of time steps on the same graph as in Fig. 7. The stopping condition is reached after 13 steps.

Next, we also present simulations on larger Erdős–Rényi random graphs: we let $n = 500$ and consider (A) $p = \ln(n)/n \approx 0.012$ and (B) $p = 0.1$. It is known [2] that in the regime of case (A), the graph is guaranteed to be connected for large n, while in case (B) the graph, besides being connected, has many more cycles and its diameter is smaller. We plot the time evolution of the error for cases (A) and (B) in Figs. 9 and 10, respectively. In both cases the node with the higher HIC has been correctly identified by the MPA, and the mean ranking error is below 3.

In all these examples, the deviation error first decreases and then slightly increases as a function of time. This observation corresponds to the evolution of the computation trees: first they grow to represent relatively well the neighbourhood of the neighbors, but after a certain number of steps too many "ghost" nodes are included, thus worsening the computed approximation.

In order to stress the accuracy of our results, we can compare the HIC computed through the MPA to other centralities, which can be computed in a distributed way and may be considered as

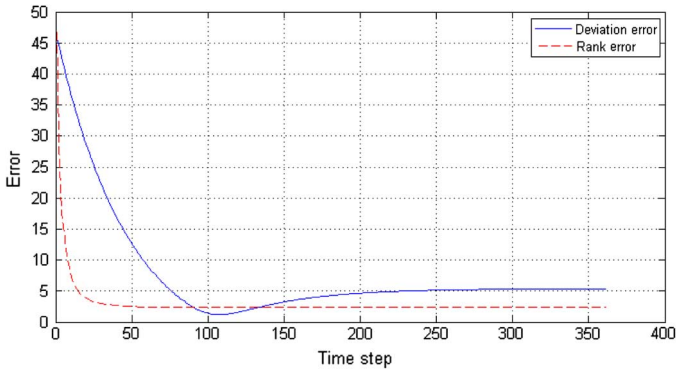


Fig. 9. Mean deviation error and mean ranking error of the MPA as functions of time steps on a large Erdős–Rényi graph with low connectivity.

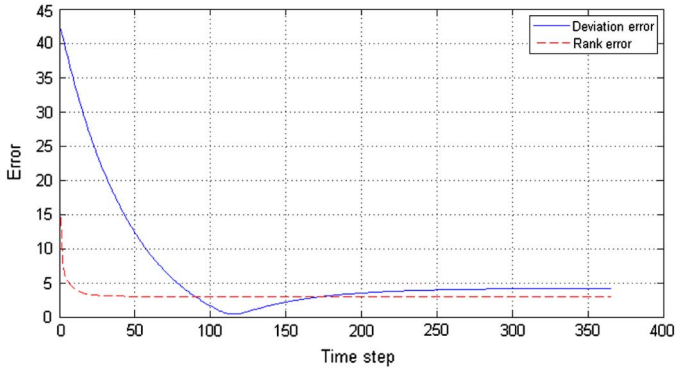


Fig. 10. Mean deviation error and mean ranking error of the MPA as functions of time steps on a large Erdős–Rényi graph with high connectivity.

reasonable approximations of the HIC. The first naive option is the degree centrality, that is, the number of neighbors. This is exactly what the MPA computes after one time step and the results in terms of deviation and rank error can be read on Figs. 8, 9, and 10. The experiments indicate that degrees are in general insufficient to describe the HIC and can at best be used as rough approximations. A second option is the eigenvector centrality: our experiments show it to be an unreliable approximation of the HIC, as it gives fairly large mean rank errors of 1.9, 18.7, and 9.1, respectively, for the three experiments shown before. Similar observations on degree and eigenvector centralities can be drawn from Fig. 7, which includes the node-by-node values of degree and eigenvector centralities (the latter is re-scaled by the maximum of the true HIC).

These measures are inadequate to our problem for the following main reason: both the degree centrality and the eigenvector centrality evaluate the influence of a node within a network, but they do not consider the different role of stubborn nodes, treating them as normal nodes.

VII. Conclusion

In this paper, we have proposed a centrality measure on graphs related to the consensus dynamics in the presence of stubborn agents, the HIC: this definition of centrality quantifies the influence of a node on the opinion of the global network. Although our setting assumes all stubborn except one to have the same value, the approach can be extended to more complex configurations, as discussed in [15]. Thanks to an intuitive analogy with electrical networks, which holds true for time-reversible dynamics, we have obtained several properties of HIC on trees. As an application of these results, we have proposed an MPA to compute the node which maximizes centrality. We have proved the algorithm to be exact on trees and to converge on any regular graph. Furthermore, numerical simulations show a good performance of the algorithm beyond the theoretical results. Further research should be devoted to extend the analysis of the algorithm beyond the scope of the current assumptions to include general networks with cycles, varied degree distributions, and directed edges.

References

[1] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Math. Oper. Res.*, vol. 38, no. 1, pp. 1–27, 2013.

[2] R. Albert and A. Barabasi, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, 2002.

[3] P. Bonacich, "Power and centrality: A family of measures," *Amer. J. Sociol.*, vol. 92, no. 5, pp. 1170–1182, 1987.

[4] A. Braunstein, R. Mezard, and R. Zecchina, "Survey propagation: An algorithm for satisfiability," *Random Struct. Algorithms*, vol. 27, no. 2, pp. 201–226, 2005.

[5] E. Estrada and N. Hatano, "Resistance distance, information centrality, node vulnerability, and vibrations in complex networks," in *Network Science*, E. Estrada, M. Fox, D. J. Higham, and G.-L. Oppo, Eds. London: Springer-Verlag, 2010, pp. 13–29.

[6] N. E. Friedkin, "Theoretical foundations for centrality measures," *Amer. J. Sociol.*, vol. 96, no. 6, pp. 1478–1504, 1991.

[7] H. Ishii and R. Tempo, "Distributed randomized algorithms for the Page Rank computation," *IEEE Trans. Automat. Control*, vol. 55, no. 9, pp. 1987–2002, Sep. 2010.

[8] M. Mézard and A. Montanari, *Information, Physics, and Computation.* USA: Oxford University Press, 2009.

[9] M. E. J. Newman, "A measure of betweenness centrality based on random walks," *Social Netw.*, vol. 27, no. 1, pp. 39–54, 2005.

[10] J. Pearl, in *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd ed. Massach usetts, USA: Morgan Kaufmann, 1988.

[11] K. Stephenson and M. Zelen, "Rethinking centrality: Methods and examples," *Social Netw.*, vol. 11, no. 1, pp. 1–37, 1989.

[12] P. O. Vontobel and H.-A. Loeliger, "On factor graphs, and electrical networks," in *Mathematical Systems Theory in Biology, Communication, Computation, and Finance*, J. Rosenthal and D. S. Gilliam, Eds. London: Springer-Verlag, 2003, pp. 469–492.

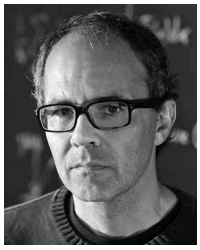
[13] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inform. Theory*, pp. 1–28, 1962.

[14] W. Wang and C. Y. Tang, "Distributed computation of node, and edge betweenness on tree graphs," in *Proc. IEEE Conf. Decision Control.* Florence, Italy, Dec. 2013, pp. 43–48.

[15] E. Yildiz, D. Acemoglu, A. Ozdaglar, A. Saberi, and A. Scaglione, "Binary opinion dynamics with stubborn agents," *ACM Tran. Econ. Comput.*, vol. 1, no. 4, 2013.

Luca Vassio received the M.Sc. degree in mathematical modelling for engineering from the Politecnico di Torino, Torino, Italy, in 2012. During the studies, he was hosted by the Massachusetts Institute of Technology, Cambridge, MA, USA, and the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

In 2009 and 2010, he collaborated in the design and construction of a hydrogen fuel cell low-consumption vehicle (Team H2polito). Currently, he is employed by the Politecnico di Torino for a research project jointly with the company General Electric (GE) Avio focused on optimization of aeronautical low-pressure turbines. His mainly research interests are optimization algorithms, metaheuristics, operations research, graph theory, dynamics over networks, and applications to aeronautical and automotive engineering.



Fabio Fagnani received the Laurea degree in mathematics from the University of Pisa, Pisa, Italy, and from Scuola Normale Superiore of Pisa, Pisa, Italy, in 1986. He got the Ph.D. degree in mathematics from the University of Groningen, Groningen, the Netherlands, in 1991.

From 1991 to 1998, he was an Assistant Professor of mathematical analysis with Scuola Normale Superiore. In 1997, he was a Visiting Professor with Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. Since 1998, he has been with the Politecnico of Torino, where he is currently (since 2002) a Full Professor of mathematical analysis. From 2006 to 2012, he has acted as a Coordinator of the Ph.D. program Mathematics for Engineering Sciences of Politecnico di Torino, Torino, Italy. From June 2012, he is the Head of the Department of Mathematical Sciences, Politecnico di Torino. His current research topics are on cooperative algorithms and dynamical systems over graphs, inferential distributed algorithms, and opinion dynamics. He has published over 50 refereed papers on international journals, he has delivered invited conferences in many international workshops and conferences and in many universities. He regularly serves as a Referee in most of the top journal of his research area: IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *SIAM Journal on Control and Optimization*, and *Automatica*. He is an Associate Editor of IEEE TRANSACTIONS ON NETWORK SYSTEMS. He is a Member of the International Program Committee for the events NECSYS.

Paolo Frasca (M'13) received the Ph.D. degree from Politecnico di Torino, Torino, Italy, in 2009. Between 2008 and 2013, he has held Research and Visiting Positions with the University of California, Santa Barbara, CA, USA, at the Istituto per le Applicazioni del Calcolo, Consiglio Nazionale delle Ricerche (IAC-CNR), Rome, and at the Politecnico di Torino.

Since June 2013, he has been an Assistant Professor with the Department of Applied Mathematics, University of Twente, Enschede, the Netherlands. His research interests are mainly in the theory of network systems and cyber-physical systems, with applications in robotic, sensor, infrastructural, and social networks. He has (co)authored more than 30 journal and conference papers and has given several invited talks at international institutes and events, both in Europe and in the USA, on these topics.

Dr. Frasca is serving as a Subject Editor for the International Journal of Robust and Nonlinear Control and in the Conference Editorial Boards of the IEEE Control Systems Society and of the European Control Association. He is a recipient of the 2013 SIAG/CST Best SICON Paper Prize.

Asuman Ozdaglar received the B.S. degree in electrical engineering from the Middle East Technical University, Ankara, Turkey, in 1996, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1998 and 2003, respectively.

Currently, she is a Professor in the Electrical Engineering and Computer Science Department, Massachusetts Institute of Technology. She is also a Member of the Laboratory for Information and Decision Systems and the Operations Research Center. Her research expertise includes optimization theory, with emphasis on nonlinear programming and convex analysis, game theory, with applications in communication, social, and economic networks, distributed optimization and control, and network analysis with special emphasis on contagious processes, systemic risk, and dynamic control.

Prof. Ozdaglar served as a Member of the Board of Governors of the Control System Society and as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. Currently, she is the area Co-Editor for a new area for the Journal Operations Research, entitled "Games, Information and Networks" and the Chair of the Control System Society Technical Committee "Networks and Communication Systems." She is the Co-Author of the book entitled "Convex Analysis and Optimization" (Athena Scientific, 2003). She is the recipient of a Microsoft fellowship, the MIT Graduate Student Council Teaching award, the NSF Career award, the 2008 Donald P. Eckman award of the American Automatic Control Council, the Class of 1943 Career Development Chair, a 2011 Kavli Fellowship of the National Academy of Sciences, and the inaugural Steven and Renee Finn Innovation Fellowship.