Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 4192–4197

Optimal selection of the most informative nodes in Opinion Dynamics on Networks

Roberta Raineri *, **Giacomo Como** *, **Fabio Fagnani** *

* *Department of Mathematical Sciences "G.L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy (e-mail: {roberta.raineri, giacomo.como, fabio.fagnani}@polito.it)*

Abstract: Finding the optimal subset to observe in a network system is a fundamental problem in science and engineering, with a wide range of applications like monitoring spatial phenomena, control of epidemic spread, feature selection in machine learning, or active surveying in social studies. The goal of this paper is to address the subset selection problem on an Opinion Dynamics model where the variable of interest Y is the average opinion of the community. We consider the opinion vector X to be updated according to a Friedkin-Johnsen opinion dynamics model where every agent i is equipped with an original unknown belief u_i, which is assumed to be normally distributed, and a parameter λ_i describing its openness to interactions. The objective function of the optimization problem is the variance reduction from the observation of the steady-state opinions of a subset $\mathcal{K} \subseteq \mathcal{V}$ of agents. We show how this functional can be rewritten in terms of the Bonacich centrality and the cycle centrality of the agents in social network when the subset selection is of cardinality 1, providing particular graph-theoretic interpretations related to the network itself. In addition, first exploratory simulations highlight a behaviour which deviates from the one of known centrality measures depending on the choice of model parameters. Finally, we show that the submodularity of the functional is not guaranteed in our case and thus results taken from known literature are non-enforceable. This paves the way for further analysis.

Keywords: Opinion Dynamics, Centrality measures, Network centrality, Network systems, Subset selection, Probabilistic Graphical Models, Gaussian Random Fields.

1. INTRODUCTION

A crucial algorithmic question in many data-driven applications is that of selecting the most informative subset of observable variables, among a large set at disposal, in order to most accurately predict another quantity of interest. This is typically known as the subset selection problem. The setting is the following: a large number n of variables X_i can in principle be observed and we want to infer the value of another variable Y. What is known is the statistical description in terms of covariances among the X_i's and Y assumed to be prior knowledge to the problem. Based on this information, the goal is now to select a subset of $k << n$ variables to 'best' predict Y.

Applications are many (Krause et al. (2008b)), ranging from feature selection in machine learning, sparse approximation and compressed sensing in signal processing, sensor placement for environmental monitoring, risk assessing in medical studies, active surveying in social studies. Most recent applications deal with the smart testing problem aimed to find an efficient strategy to control epidemic spread(Batlle et al. (2022)).

In many of these scenarios, a natural way to formalize the optimization problem is to consider as target functional the variance reduction on Y, namely the difference between the variance of Y and the variance of Y conditioned on the observed variables, and let the optimal subset to select the one that maximizes this quantity.

This problem is known to be NP-hard and not submodular in general, see, e.g., Das and Kempe (2018) and Natarajan (1995). The work by Das and Kempe (2008) contains various results on the complexity of the problem when the graph associated to the covariance matrix has specific properties (trees, existence of large independent sets) and the analysis of a natural greedy algorithm (called forward regression) where variables are chosen in a recursive fashion, each time maximizing the variance reduction. They obtain bounds on the performance of the greedy algorithm for small covariances. They also discuss the existence of submodularity in the assumption of lack of the so called suppressor variables, a condition that is however not easy to be checked. In Das and Kempe (2018) authors prove an approximate submodularity capable of giving bounds on the greedy algorithm. In Ma et al. (2013), authors study the special case when the process X is a Gaussian graphical model with precision matrix given by the Laplacian of an undirected graph and show that in this case the variance reduction is submodular. The submodularity property is crucial because, as proven by Nemhauser et al. (1978), it guarantees a near-optimal solution adopting a greedy

⋆ Giacomo Como is also with the Department of Automatic Control, Lund University, Sweden. This research was carried on within the framework of the MIUR-funded *Progetto di Eccellenza* of the *Dipartimento di Scienze Matematiche G.L. Lagrange*, Politecnico di Torino, CUP: E11G18000350001.

10.1016/j.ifacol.2023.10.1767

approximation for maximizing the functional of interest subject to a cardinality constraint.

In this paper, we undertake a study of the subset selection problem in the context of a social network where the observable variables are the opinions of the various individuals and the quantity that we want to estimate is the average opinion Y of the community. Opinions of the individuals are assumed to be determined by their original beliefs, here assumed to be Gaussian random variables, and by the network interaction, in such a way to minimize what in social science is called the cognitive dissonance of the system.

Our contribution is twofold. We show that our process X is a Gaussian random field on an augmented graph obtained triangulating the graph that describes the social network, with a precision matrix that may possess positive off diagonal values. For the special case when the subset selection is of cardinality 1, we give an interpretation of the variance reduction in terms of the classical Bonacich centrality of a graph and the cycle centrality, first introduced in Talamas and Tamuz (2017), to study optimal intervention problems in quadratic games. Finally, we observe that our case does not satisfy the assumptions of the problem described in Ma et al. (2013), as in our case the precision matrix is not a graph Laplacian, thus the analysis of submodularity is more complicate. In particular, we show by some examples that the functional is not submodular in the general case. Neverthless, the introduced centrality measure is relevant pointing out a behaviour that strays from the one of common known ones.

2. MODEL

Consider a network of agents $\mathcal{V}$ whose interactions are encoded by row-stochastic matrix P in $\mathbb{R}_+^{\mathcal{V}\times\mathcal{V}}$.[1] Let the sparsity pattern of P be represented by the directed graph $\mathcal{G}_P = (\mathcal{V}, \mathcal{E})$ such that $(i, j) \in \mathcal{E}$ if and only if $P_{ij} > 0$.

Every agent i is equipped with an original belief u_i and a parameter λ_i in $(0, 1)$ describing its openness to social interaction. The equilibrium opinion of the social system is a vector x in $\mathbb{R}^{\mathcal{V}}$ such that

$$x_i = \lambda_i \sum_{j\in\mathcal{V}} P_{ij}x_j + (1 - \lambda_i)u_i\,.$$

If we indicate with x and u the corresponding vectors of opinions and original beliefs, with A the matrix such that $A_{ij} = \lambda_i P_{ij}$, and with B the diagonal matrix such that $B_{ii} = 1 - \lambda_i$, this can be written as

$$x = Ax + Bu\,.$$

As the spectral radius of A is strictly below 1, such equation has just one solution that can be represented as

$$x = (I - A)^{-1}Bu\,. \tag{1}$$

The value x determined by (1) has various sociological interpretations. It can be shown to be the unique Nash equilibrium of a quadratic game where individuals have as utility function their cognitive dissonance

$$u_i(x) = \lambda_i \sum_{j\in\mathcal{V}} P_{ij}(x_i - x_j)^2 + (1 - \lambda_i)(x_i - u_i)^2\,.$$

[1] Recall that a row-stochastic matrix P is a nonnegative square matrix such that $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ is the all-1 vector.

It also coincides with the asymptotic outcome of the Friedkin-Johnsen's opinion dynamics model

$$x(t + 1) = Ax(t) + Bu\,.$$

In this paper, we assume the belief vector to be random vector U having multivariate Gaussian distribution $N(\mu, \Sigma)$ with mean vector μ and covariance matrix Σ. The corresponding opinions vector

$$X = (I - A)^{-1}BU \tag{2}$$

then has itself multivariate Gaussian distribution $N(\nu, C)$ with mean vector $\nu = (I - A)^{-1}B\mu$ and covariance matrix

$$C = (I - A)^{-1}B\Sigma B(I - A')^{-1}\,. \tag{3}$$

Observe that, since

$$M = (I - A)^{-1}B = \sum_{k=0}^{+\infty} A^k B$$

is a nonnegative matrix, then the covariance matrix

$$C = M\Sigma M'$$

of the opinions vector is nonnegative whenever the covariance matrix Σ of the original beliefs is nonnegative.

On the other hand, the precision matrix $W = C^{-1}$ of the opinions vector satisfies

$$W = (M')^{-1}\Sigma^{-1}M^{-1} = D - (A'D + DA) + A'DA\,, \tag{4}$$

where $D = B^{-1}\Sigma^{-1}B^{-1}$. Notice that W is not an L-matrix in general, i.e., its extra-diagonal elements are not necessarily all nonpositive. E.g., in the special case when the original beliefs are independent and identically distributed with variance $\sigma_i^2 = 1$, and $\lambda_i = \alpha$ in $(0, 1)$ for every i in $\mathcal{V}$, then

$$W_{ij} = -\frac{\alpha}{(1-\alpha)^2}(P_{ij} + P_{ji}) + \frac{\alpha^2}{(1-\alpha)^2}\sum_{k\in\mathcal{V}} P_{ki}P_{kj} > 0\,,$$

for every two nodes $i \neq j$ that are not directly linked to each other in $\mathcal{G}_P$ (so that $P_{ij} = P_{ji} = 0$), but have a common in-neighbor k in $\mathcal{V} \setminus \{i, j\}$ (so that $P_{ki}P_{kj} > 0$).

Remark 1. Notice that, in the terminology of probabilistic graphical models (see, e.g., Wainwright and Jordan (2008), Lauritzen (1996)), the multivariate normal variable X is viewed as a Gaussian random field defined on an underlying undirected concentration graph $\mathcal{G}^* = (\mathcal{V}, \mathcal{E}^*)$ with the same sparsity pattern as the precision matrix W, i.e., $W_{ij} \neq 0$ for some $i \neq j$ if and only if $\{i, j\}$ in $\mathcal{E}^*$ is an undirected link in $\mathcal{G}^*$.

We highlight a particular relation between the original graph $\mathcal{G}$ and $\mathcal{G}^*$. Firstly, define the converse graph $\mathcal{G}'$ of the directed graph $\mathcal{G}$, that is retriven from $\mathcal{G}$ keeping the same set of vertices $\mathcal{V}$ and reversing the orientation of all the edges in $\mathcal{E}$. Then, $\mathcal{G}^*$ is obtained from $\mathcal{G}'$ by adding links between all nodes that share an out-neighbor and then making all links undirected. In the literature, $\mathcal{G}^*$ is defined as the moral graph of $\mathcal{G}'$.

3. PROBLEM

We focus on the problem of estimating the arithmetic average of the opinions vector

$$Y = \frac{1}{n}\sum_{i\in\mathcal{V}} X_i$$

from the observation of the opinions of a subset $\mathcal{K} \subseteq \mathcal{V}$ of the agents, with $|\mathcal{K}| < n$.

Formally, let $X_{\mathcal{K}} = (X_i)_{i\in\mathcal{K}}$ be the restriction of the opinions vector X to the observed set $\mathcal{K}$. Consider the map $F : 2^{\mathcal{V}} \to \mathbb{R}^+$, defined by

$$F(\mathcal{K}) = Var(Y) - \mathbb{E}[Var(Y|X_{\mathcal{K}})]\,,$$

returning the variance reduction in the variable Y from the observation of the opinions $X_{\mathcal{K}}$ as a function of the set of observed agents $\mathcal{K} \subseteq \mathcal{V}$. Then, for a given positive integer s, we aim at studying the optimization problem

$$\max_{|\mathcal{K}| \leq s} \; F(\mathcal{K})\,. \tag{5}$$

Remark 2. It is a well known fact that the best estimator (in mean square error sense) of Y that is measurable w.r. to $X_{\mathcal{K}}$ is given by $\hat{Y} = E[Y|X_{\mathcal{K}}]$. Moreover,

$$\mathbb{E}[Y - \hat{Y}]^2 = \mathbb{E}[Var(Y|X_{\mathcal{K}})]$$

This implies that problem (5) is equivalent to finding the subset $\mathcal{K}$ such that $|\mathcal{K}| \leq s$ that allows for an estimation of Y with minimal mean square error.

Optimization problem (5) is in fact an instance of more general subset selection problems, that are known to be NP-hard (Das and Kempe (2018)). While some of the literature on subset selection problems has in fact studied variance reduction problems for Gaussian random vectors, such as, e.g., Ma et al. (2013), such works usually assume that the precision matrix W is an L-matrix. In this case, it is known that the functional $F(\mathcal{K})$ is submodular, i.e.,

$$F(\mathcal{A} \cup \{k\}) - F(\mathcal{A}) \geq F(\mathcal{B} \cup \{k\}) - F(\mathcal{B}) \tag{6}$$

for every $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$ and $k \in \mathcal{V}\backslash\mathcal{B}$.

However, as already explained in Section 2, the precision matrix W of the opinions vector X fails to be an L-matrix, so that the aforementioned results do not apply to our model. In fact, as the following examples show, the functional $F(\mathcal{K})$ is not in general submodular in our case.

Example 1. Let us take as example the undirected line graph in Fig. 1.

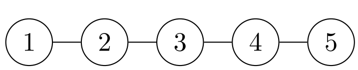


Fig. 1. Undirected line graph with 5 nodes.

Consider the subsets of nodes $\mathcal{A} = \{3\}, \mathcal{B} = \{2, 3\}, \{k\} = \{4\}$ and define the function

$$\Phi = F(\mathcal{A} \cup \{k\}) - F(\mathcal{A}) - F(\mathcal{B} \cup \{k\}) - F(\mathcal{B})\,.$$

Analysing the sign of function Φ we can deduce if the inequality in (6) is satisfied (i.e, F is submodular iff $\Phi \geq 0$). As shown in Fig. 2, for any choice of the model parameter λ, which represents the openness to interactions of agents, the function $\Phi(\lambda) < 0$ and so the variance reduction is not submodular.

We highlight a more complex behaviour if we take as example the cycle graph with 5 nodes. In that case we observe that the submodularity of the variance reduction depends on the choice of the model parameter λ as shown in Fig. 3.

4. MAIN RESULTS

The fact that the process is Gaussian, allows a useful explicit rewriting of the functional F in terms of the

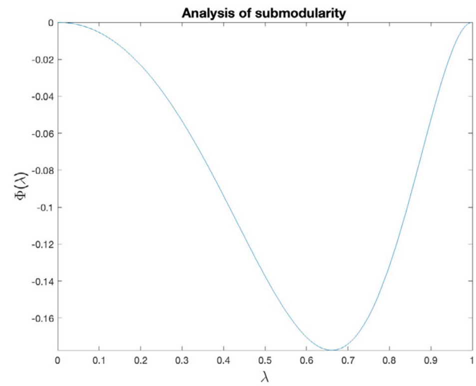


Fig. 2. Analysis of submodularity of variance reduction depending on model parameter λ for a line graph with 5 nodes. Study of the sign of function $\Phi(\lambda)$, given the subset selection $\mathcal{A} = \{3\}, \mathcal{B} = \{2, 3\}, \{k\} = \{4\}$.

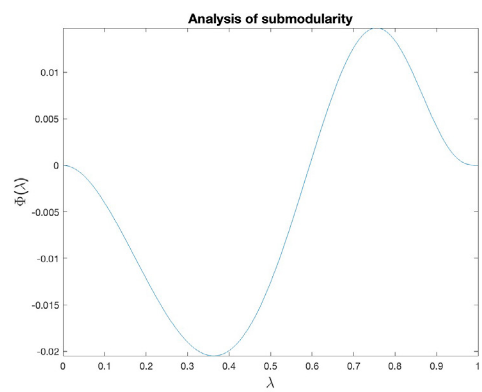


Fig. 3. Analysis of submodularity of variance reduction depending on model parameter λ over the cycle graph with 5 nodes. Study of the sign of function $\Phi(\lambda)$, given the subset selection $\mathcal{A} = \{3\}, \mathcal{B} = \{2, 3\}, \{k\} = \{4\}$.

matrices A and B, and the covariance matrix Σ. We define C as in (3) and for any subsets $\mathcal{K}, \mathcal{J} \subseteq \mathcal{V}$, we let $C_{\mathcal{K},\mathcal{J}}$ to denote the submatrix of C having rows in $\mathcal{K}$ and columns in $\mathcal{J}$.

Proposition 1. The functional $F(\mathcal{K})$ admits the following representation

$$F(\mathcal{K}) = (C\mathbb{1})'_{\mathcal{K}}(C_{\mathcal{K},\mathcal{K}})^{-1}(C\mathbb{1})_{\mathcal{K}}\,. \tag{7}$$

Proof. It is a standard result that the process $X_{-\mathcal{K}}$ conditioned to $X_{\mathcal{K}}$ is Gaussian

$$X_{-\mathcal{K}}|X_{\mathcal{K}} \sim N\left(\mu^{(\mathcal{K})}, C^{(\mathcal{K})}\right)$$

where

$$\mu^{(\mathcal{K})} = \mu_{-\mathcal{K}} + C_{-\mathcal{K},\mathcal{K}}(C_{\mathcal{K},\mathcal{K}})^{-1}(X_{\mathcal{K}} - \mu_{\mathcal{K}})$$
$$C^{(\mathcal{K})} = C_{-\mathcal{K},-\mathcal{K}} - C_{-\mathcal{K},\mathcal{K}}(C_{\mathcal{K},\mathcal{K}})^{-1}C_{\mathcal{K},-\mathcal{K}}\,.$$

Consequently, we have that

$$Y|X_{\mathcal{K}} \sim N\left(\mu^{(\mathcal{K})} + X_{\mathcal{K}}, \mathbb{1}'C^{(\mathcal{K})}\mathbb{1}\right)\,.$$

Therefore, the functional $F(\mathcal{K})$ in (5) takes the form

$$F(\mathcal{K}) = Var(Y) - Var(Y|X_{\mathcal{K}}) = \mathbb{1}'C\mathbb{1} - \mathbb{1}'C^{(\mathcal{K})}\mathbb{1}\,. \tag{8}$$

We now observe that

$$\mathbb{1}'C_{-\mathcal{K},-\mathcal{K}}\mathbb{1} = \mathbb{1}'C\mathbb{1} - \mathbb{1}'C_{-\mathcal{K},\mathcal{K}}\mathbb{1} - \mathbb{1}'C_{\mathcal{K},-\mathcal{K}}\mathbb{1} - \mathbb{1}'C_{\mathcal{K},\mathcal{K}}\mathbb{1}$$

and

$$\mathbb{1}'C_{-\mathcal{K},\mathcal{K}}\mathbb{1} = \mathbb{1}'(C\mathbb{1})_{\mathcal{K}} - \mathbb{1}'C_{\mathcal{K},\mathcal{K}}\mathbb{1}\,.$$

Substituting in (8) we obtain the thesis. □

In the special case when $|\mathcal{K}| = 1$, the form of the functional takes a particularly simple form. Denote

$$v_k = \sum_i M_{ik}$$

and

$$c_k = \sum_i M_{ki}^2 \sigma_{ik}^2\,.$$

Corollary 1. For $\mathcal{K} = \{k\}$, the functional F takes the following form

$$F(k) = \frac{(C\mathbb{1})_k^2}{C_{kk}} = \frac{(M\Sigma v)_k^2}{c_k}\,. \tag{9}$$

Analysing the functional in (9) we highlight that variables v and c_k correspond in particular to two known centrality measures, that are respectively the Bonacich centrality defined by Bonacich (1987) and the cycle centrality defined by Talamas and Tamuz (2017). These measures provide particular information on the graph.
The Bonacich centrality v_k for node $k \in \mathcal{V}$ coincides with the sum of all the paths in $\mathcal{G}$ that start at k: both cycles from k to k and outer paths from k to general node $i \in \mathcal{V}$. The cycle centrality c_k of node $k \in \mathcal{V}$ coincides instead with the weighted sum of the number of network cycles the agent is in. Indeed, if we rewrite the M matrix of the defined model through the power series expansion:

$$M = \left[\sum_{l=0}^{\infty} A^l\right] B$$

and if we assume equal openness degree λ for all the agents $i \in \mathcal{V}$, then

$$\begin{aligned} M_{ki}^2 &= (1-\lambda)^2 \left(\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} (A^l)_{ki}(A^m)_{ki}\right) \\ &= (1-\lambda)^2 \left(\sum_{r=0}^{\infty}\sum_{s=0}^{r} (A^s)_{ki}(A^{r-s})_{ki}\right)\,. \end{aligned} \tag{10}$$

5. EXAMPLES AND SIMULATIONS

Firstly, recall that given a graph $\mathcal{G}$, two agents i and j have the same role, and thus the same centrality measure, if there is a symmetry on $\mathcal{G}$ or if we can define a proper relabelling which maps i to j. Given this statement, we can neglect in our analysis the case of simple cycles where by symmetry all agents have the same role.
In the following examples, let the original beliefs be independent and identically distributed with variance $\sigma_i^2 = 1$ and Λ be such that $\Lambda_{ii} = \lambda$ for all $i \in \mathcal{V}$ and $\lambda \in (0,1)$ to analyse the effect of parameter λ on the selection.

Consider the star graph. In this case there are only two possible roles, fulfilled by the central agent and the external (or leaf) one. It can be proven analytically that, in this case, for any possible choice of the parameter λ, the optimal choice consists in the observation of the central node whichever is the centrality measure chosen.
The numerator of (9) can be found solving the linear system $Mv{=}b$, that is

$$\frac{(I-A')(I-A)b}{(1-\lambda)^2} = \mathbb{1}\,.$$

Thus, given $i = 1$ the central node,

$$b_1 = \frac{\lambda^2 + (m+1)\lambda + m}{m(\lambda+1)^2},$$

with $m = n - 1$, and for all $i \geq 2$

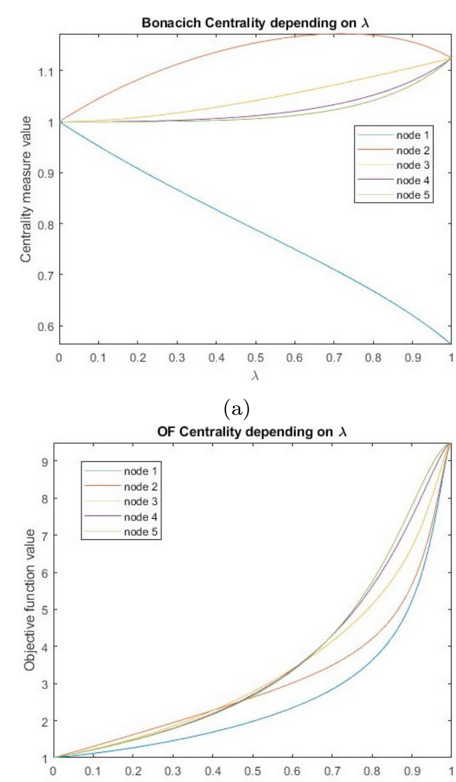
$$b_i = \frac{m(1 + m\lambda^2) + \lambda(1+m)}{m(\lambda+1)}\,.$$

The cycle centrality c_k can be computed using (10) obtaining:

$$\begin{aligned} c_1 &= \frac{m+\lambda^2}{m(1+\lambda)^2} \\ c_i &= \frac{(m-1)(1-\lambda^2)^2 + 1 + m\lambda^2}{m(1+\lambda)^2} \quad \forall i \geq 2 \end{aligned}$$

Now, comparing the expression of $F(k)$ for the central and the leaf node we retrieve that for every possible choice of the parameters the centrality of the central node is always greater then the one of the leaves.

Let now examine the line graph. For this case problem, simulations have been done with $N = 5$ and $N = 9$. As shown in Fig. 4 the Bonacich centrality, regardless of λ value, is higher for nearly extremal nodes (e.g., in the test cases node 2). The value of the proposed centrality measure on the contrary depends on parameter λ: the greater is λ, the more central is the selected node.



(a)

(b)

Fig. 4. Comparison of centrality measures for network roles depending on the choice of λ for a line graph with 9 nodes: (a) Bonacich Centrality v_k, (b) proposed functional $F(k)$.

Finally, consider the barbell graph in Fig. 5. Due to the graph symmetry, the three possible roles on this network are played by agents 1,2,4.

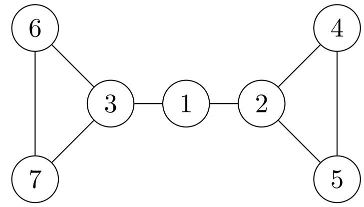


Fig. 5. Barbell graph with 7 nodes.

In Fig. 6, we represent, for various centrality measures, the values of centrality of the various agents as functions of the parameter λ. We observe that Bonacich centrality reaches the maximum with node 2, while cycle centrality with node 4, both regardless of the choice on λ. Conversely, the best node to observe according to the proposed centrality measure $F(k)$ depends on the value of λ and it could be node 1 or 2. For smaller values of λ (higher stubbornness level), the choice coincides with the one related to Bonacich centrality. Instead, for λ greater then a certain threshold $\bar{\lambda}$ the minimization of cycle centrality prevails in the ratio and the optimal choice is node 1.

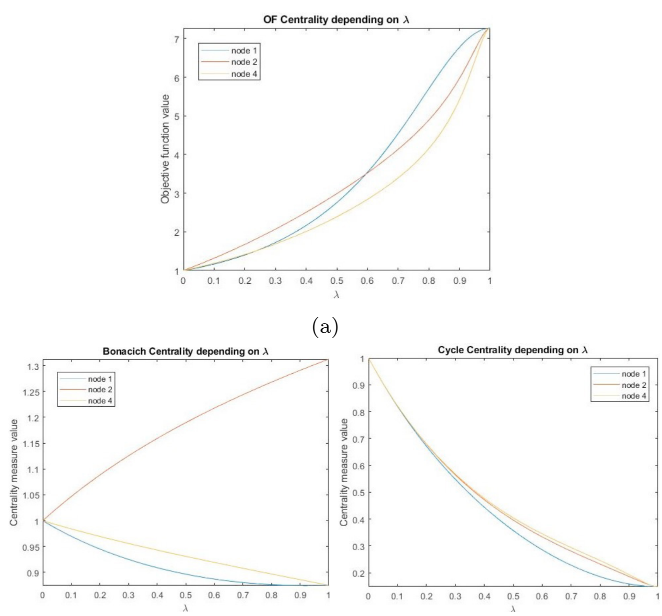


Fig. 6. Comparison of centrality measures for network roles depending on the choice of λ for the barbell graph chosen: (a) proposed functional $F(k)$, (b) Bonacich centrality v_k, (c) cycle centrality c_k.

This shows that, differently than with other notions of centrality, the ranking of nodes with respect to this new OF centrality depends not only on the network but also on agents' openness to social interaction λ.

Finally, we compare the optimal performance obtained by choosing the agent k^* of highest OF centrality with the performance obtained by choosing instead the agent k^B of highest Bonacich centrality. For the barbell graph in Fig. 5 with $\lambda = 0.8$, we obtain $F(k^*) = 5.70$ and $F(k^B) = 4.89$, with an initial total varince of $Var(Y) = 7.20$. This numerical result underlines the non negligible improvement that our new centrality allows to obtain.

6. CONCLUSIONS AND FUTURE DIRECTIONS

In this work we have analysed a subset selection problem for an Opinion Dynamics model. We have chosen as objective function of the maximization problem the variance reduction of the aggregate information Y. We have highlighted how our functional $F(k)$ can be interpreted in terms of known centrality measures and in particular in terms of paths and cycles of the original graph $\mathcal{G}$.
Moreover, we have shown, through some examples, that $F(k)$ is not submodular in our case study and thus the problem gets more complicate since state-of-the-art results cannot be applied. A natural follow-up of our work would be more in-depth analyses of the submodularity aimed to study possible correlations between topological properties of the graph and functional sumodularity.
Finally, preliminary simulations on low dimension graphs led to a node hierarchy which is function of the model parameters and such behaviour deviates from the one related to known centrality measures chosen as benchmark. Future works include so also test of the proposed centrality measure on a wider selection of high dimensional social networks.

REFERENCES

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5), 1403–1417.

Batlle, P., Bruna, J., Fernandez-Granda, C., and Preciado, V.M. (2022). Adaptive test allocation for outbreak detection and tracking in social contact networks. *SIAM Journal on Control and Optimization*, 60(2), S274–S293.

Bonacich, P. (1987). Power and centrality: a family of measures. *American Journal of Sociology*, 5, 1170–1182.

Das, A. and Kempe, D. (2008). Algorithms for subset selection in linear regression. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, 45–54. Association for Computing Machinery.

Das, A. and Kempe, D. (2018). Approximate submodularity and its applications: Subset selection, sparse approximation and dictionary selection. *Journal of Machine Learning Research*, 19(3), 1–34.

Friedkin, N. and Johnsen, E. (1990). Social influence and opinions. *Journal of mathematical Sociology*, 15(3-4), 193–205.

Krause, A., McMahan, H.B., Guestrin, C., and Gupta, A. (2008a). Robust submodular observation selection. *Journal of Machine Learning Research*, 9(93), 2761–2801.

Krause, A., Singh, A., and Guestrin, C. (2008b). Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.*, 9, 235–284.

Lauritzen, S.L. (1996). *Graphical Models*. Oxford University Press.

Ma, Y., Garnett, R., and Schneider, J. (2013). Σ - optimality for active learning on gaussian random fields. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Natarajan, B.K. (1995). Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24, 227–234.

Nemhauser, G., Wolsey, L., and Fisher, M. (1978). An analysis of approximations for maximizing submodular set functions—i. *Mathematical Programming*, 14, 265–294.

Talamas, E. and Tamuz, O. (2017). Network cycles and welfare. *SSRN Electronic Journal*.

Wainwright, M.J. and Jordan, M.I. (2008). Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1(1–2), 1–305.