RAPORT DE CERCETARE

CSC-N - COMMUNITY STRUCTURE AND DIFFUSION IN SOCIAL AND ECONOMIC NETWORKS.

GRANT CNCS – UEFISCDI, PN-II-RU-TE-2014-4-2332

DEZVOLTAREA DE NOI MODELE COMPUTAȚIONALE PENTRU DETECTAREA DE COMUNITĂȚI BAZATE PE TEORIA JOCURILOR ȘI INTELIGENȚĂ COMPUTAȚIONALĂ.

O1. Pregatirea dezvoltării de noi modele computaționale pentru detectarea de comunități bazate pe teoria jocurilor și inteligența computațională.
   A1. Analiza teoretică a jocului propus și a conceptelor de echilibre disponibile ca și posibile soluții pentru problema de detectarea structurii de comunități.

O2. Dezvoltarea de noi modele computaționale pentru detectarea de comunități bazate pe teoria jocurilor și inteligența computațională.
   A2. Dezvoltarea de noi operatori evolutivi potrivit pentru conceptele de echilibre de la O1A1; analiza complexității.
   A4. Comparații cu alte metode
   A5. Experimente pe rețele reale sociale și economice. Adaptare la specific de marketing.
   A6. Documentarea modelelor de difuzie de informație existente
   A7. Construirea platformei online pentru detectarea structurilor de comunități.
   A8. Diseminarea rezultatelor

REZULTATE PREVĂZUTE PENTRU LIVRARE:

O1:
- raport de cercetare;
- articole trimise spre publicare;
- pagina web;

O2:
- raport de cercetare;
- 3 articole indexate Web of Science;
- 10 participări la conferințe;
- o platformă web pentru detectarea structurilor de comunități
REZULTATE OBTINUTE:

LISTA DE LUCRARI PUBLICATE SAU TRIMISE SPRE PUBLICARE (ACTIVITATEA ASOCIATA ESTE INDICATA IN PARANTEZE +O2 A8):

1. Noemi Gasko, Rodica Ioana Lung, Mihai Suciu, EXTREMAL OPTIMIZATION AND NETWORK COMMUNITY STRUCTURE, The 7th International Conference on BIOINSPIRED OPTIMIZATION METHODS AND THEIR APPLICATIONS, Bled, Slovenia, 18-20 mai 2016 (O2, A1,A2,A3,A4)
4. Noémi Gaskó, Rodica Ioana Lung and Mihai Suciu, Community Detection in Bipartite Networks Using a Noisy Extremal Optimization Algorithm, Springer Series in Computational Intelligence (accepted, indexat ISI, ISDA 2016, O2, A1,A2,A3,A4)

Lucrari aflate in evaluare la reviste:

1. Lung, R.I., Suciu, M.A., Gasko, N., A Probabilistic Game Theoretic Approach to the Problem of Community Structure Detection in Complex Networks, in evaluare la revista PlosOne (zona rosie AIS) O1

PREZENTARI LA CONFERINȚE:


PARTICIPAȚI LA CONFERINȚE:
<table>
<thead>
<tr>
<th>Nr. Crt</th>
<th>Nume si prenume</th>
<th>Perioada</th>
<th>Localitate</th>
<th>conferinta</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Lung Rodica Ioana</td>
<td>3-6 iulie 2016</td>
<td>Poznan/Polonia</td>
<td>EURO 2016</td>
</tr>
<tr>
<td>2.</td>
<td>Gasko Noemi</td>
<td>3-6 iulie 2016</td>
<td>Poznan/Polonia</td>
<td>EURO 2016</td>
</tr>
</tbody>
</table>

PLATFORMA

www.econ.ubbcluj.ro/~rodica.lung/cscn
The community structure detection in networks is a graph clustering problem that attempts to identify groups of nodes that are highly interconnected to each other and sparsely connected outside the group. While the concept of community seems trivial, the automatic extraction of the community structure for simple unweighted and undirected networks has proven to be a computational challenge that requires truly adaptive and scalable methods.


One of the drawbacks of MNEO is that it requires an interval for the number of communities to search for, which may not always be available a-priori. A-MNEO addresses this issue by estimating the number of communities during the initialization process. Another drawback of MNEO comes from the periodic use of the network shifting mechanism, whether necessary or not. A-MNEO uses the network shifts adaptively, when the search stagnates for a period of time. Another adaptive feature introduced within A-MNEO is the updating mechanism of the number of nodes modified in one EO iteration, replacing the exponentially decreasing mechanism of MNEO. A-MNEO was tested on the well-known GN and LFR benchmarks¹ (Lancichinetti and Fortunato, 2009; Girvan and Newman, 2002).

The GN benchmark consists of networks having 128 nodes grouped in 4 communities of 32 nodes each. The degree of each node is $z_{in} + z_{out} = 16$, where $z_{in}$ is the number of edges linking each node to others in the same community and $z_{out}$ represents the number of edges linking each node with nodes from the other communities. Eight benchmark sets were randomly generated for different values of $z_{out} \in \{1, 2, ..., 8\}$. Each set contains 30 networks.

¹ generated using the code available at https://sites.google.com/site/andrealancichinetti/software
The LFR benchmark networks have 128 nodes, consisting of 6 sets of 30 networks with the mixing parameter (ratio between number of links outside the community and node degree) \( \mu \in \{0.1, 0.2, \ldots, 0.6\} \). The LFR parameters are: average vertex degree 20, maximum vertex degree 50, community size [10, 50].

The most challenging networks from the synthetic sets are the GN \( z_{out} = 8 \) and the LFR sets with 128 nodes and \( \mu \in \{0.5, 0.6\} \), because these are networks with the most unclear defined communities as the number of links a node has within its community is equal \( (z_{out} = 8, \mu = 0.5) \) or less \( (\mu = 0.6) \) than the number of links to nodes outside its community. Another feature that may increase the difficulty of the network is the difference between the number of links a node has in its community and the maximum number of links connecting it to another community: a difference of only 1 makes it difficult for an algorithm to correctly assign the node.

The results are evaluated using the normalized mutual information (NMI)\(^2\) proposed by Lancichinetti et al. (2009). The NMI can be used as an indicator when the correct community structure is available; a higher value of the NMI indicates a "better" solution; the maximum value of 1 indicates that the correct cover has been found. We compute the NMI of the individual having the best modularity value in A-MNEO population \( A \), and also the best NMI value found in the population.

Statistical comparisons are performed by using the Wilcoxon sign rank nonparametric test for NMIs reported by each considered method in 30 independent runs for each real network and on the 30 networks for each GN or LFR set. The Wilcoxon sign rank assesses if there is a significant difference between two sample medians: the null hypothesis that two samples come from the same population can be rejected with a level of significance \( \alpha = 0.05 \) if the computed \( p \)-value is smaller than 0.05.

**Comparison with other methods** The results obtained with A-MNEO are compared with those provided by different state of the art algorithms: *Louvain* (Blondel et al., 2008), *ModOpt* - Modularity optimization (Sales-Pardo et al., 2007), *Oslom* (Lancichinetti et al., 2011), *Infomap* (Rosvall and Bergstrom, 2008)\(^3\), and with *NoisyEO* (Lung et al., 2015), a similar extremal optimization based method that maximizes modularity.

The *Louvain* method (Blondel et al., 2008) maximizes the modularity gradually by starting with small communities detected by locally optimizing modularity, aggregating them as nodes in a new network, and repeating the local modularity optimization and aggregation until a maximum is found and a hierarchical structure is uncovered. *Louvain* was designed to deal with large networks.

The *ModOpt* algorithm (Sales-Pardo et al., 2007) goes beyond simply maximizing the modularity, by defining a basin of attraction for covers that are local maxima in the modularity landscape and using it to compute an affinity matrix for network nodes. By using simulated annealing the nested block diagonal structure of this matrix is revealed and further used to identify the hierarchical structure of a network. If a network has a

---

\(^2\)by using the code available at https://sites.google.com/site/andrealancichinetti/software

\(^3\)For Louvain, ModOpt, Oslom, and Infomap we used the source code available at https://sites.google.com/site/andrealancichinetti/software.
flat organization of nodes, \textit{ModOpt} acts as a standard community detection algorithm. \textit{Oslom} (Lancichinetti et al., 2011) optimizes repeatedly a score defined for clusters. This score is computed by using the probability that an external node has a given number of neighbors in a cluster in a random null model compared to the actual number of connections that the node has with the cluster. Optimized clusters are further analyzed in order to convey the optimum cover. \textit{Oslom} is capable to deal with different types of networks: unweighted and undirected, directed, weighed, hierarchical or dynamic. \textit{Infomap} (Rosvall and Bergstrom, 2008) uses a random walk to mimic the information flow within a network, based on the assumption that information flows more quickly and easily within a module. The random walk is encoded using Huffman coding; the community structure problem is then transformed into a coding problem that minimizes its expected description of the code. A deterministic greedy search algorithm followed by a heat-bath algorithm are used to solve the optimization problem and uncover the community structure. \textit{NoisyEO} (Lung et al., 2015) is another method that uses extremal optimization in a similar setting with A-MNEO to maximize the modularity. Within \textit{NoisyEO} the number of nodes modified in a EO iteration is linearly decreased until the middle of the search and set to one until the end, making the diversity induced by this mechanism dependent on the number of generations (in a similar manner with MNEO). \textit{NoisyEO} also uses a network shifting mechanism, in a similar manner with MNEO, with the only difference that all network modifications are random, following an uniform distribution, without preserving node degrees.

Thus, A-MNEO is directly compared with two other methods that rely on modularity optimization, two stochastic methods that are also based (indirectly) on optimization, and with a similar existing EO based approach.

\textbf{Numerical results} The results obtained for the synthetic benchmarks are presented in Figure 1. For the GN and LFR benchmarks with 128 nodes and \( z_{out} \in \{1, 2, 3, 4\} \) and \( \mu \in \{0.1, 0.2, 0.3\} \) respectively, all methods reported NMI values of 1 for all networks, and therefore we did not represent the results. However, increasing \( z_{out} \) and \( \mu \), respectively, leads to differences between results. For the GN benchmark, \textit{Louvain} and \textit{Infomap} fail to identify the community structure starting with \( z_{out} = 7 \), while for \( z_{out} = 8 \) all methods except the EO based ones fail. With insignificant differences between A-MNEO and A-MNEO\(^{(m)}\), A-MNEO provides best results for this benchmark. Other methods in literature report an average NMI of approx. 0.4 for \( z_{out} = 8 \) (Gong et al., 2012; Pizzuti, 2008); Gong et al. (2013) report a maximum NMI value of approx. 0.6.

Moving to the LFR benchmark, it is clear that \( \mu = 0.5 \) and \( \mu = 0.6 \) are the most challenging sets for all methods. A-MNEO results are significantly better than those provided by \textit{Oslom} and \textit{Infomap}; while there is no statistical difference indicated in the matrix between A-MNEO and A-MNEO\(^{(m)}\), the results of A-MNEO\(^{(m)}\) cannot be considered statistically better than those provided by \textit{Louvain}, \textit{ModOpt} and \textit{NoisyEO}. In (Honghao et al., 2013) the results reported by an Ant colony optimization, compared with other methods, report average NMI values of 0.2 for \( \mu = 0.6 \).
Figure 1: Comparisons with other methods. Boxplots of NMI values obtained for the 30 networks in each set by each considered method. Wilcoxon $h$ values matrix illustrate the statistical significance of the differences in results for the seven methods: a black box corresponds to $p < 0.05$ and rejection of the null hypothesis that the two samples have the same median. Results obtained for the GN and LFR benchmarks (128 nodes).

Lung, R.I., Suciu, M.A., Gasko, N., A Probabilistic Game Theoretic Approach to the Problem of Community Structure Detection in Complex Networks, PlosOne, under evaluation. We explore the possibility of using a probabilistic version of the Nash ascendancy relation within one of the extremal optimization methods designed by us to identify the community structure of a network. The probabilistic ascendancy relation takes into account only a fraction of the network nodes when comparing two strategy profiles. We show that $p$–Nash non-ascended solutions are also Nash equilibria of the game. We address the practical concern that using the probabilistic relation may not yield results as good as the Nash ascendancy relation when maintaining the same number of iterations by using numerical experiments. Results show that there are very few significant differences in results between different values for the fraction of nodes and that the performance of $p$MNEO is significantly better than that of other state-of-art methods on the tested benchmarks.

Regarding the running time, the differences between different probability levels are low: an average decrease of 10% in running time when switching from $p = 100\%$ to $p = 25\%$. This result indicates that the Nash ascendancy relation, while indeed com-
Algorithm 1 \( k\text{EO}(s, s_{\text{best}}) \) iteration

1: For current configuration \( s \) evaluate \( u_i(s) \), the payoff function corresponding of node \( i \in \{1, \ldots, n\} \).
2: find the \( k \) worst components and replace them with a random value;
3: if \( (s \text{ is better} \) than \( s_{\text{best}}) \) then
4: \hspace{1em} set \( s_{\text{best}} := s \).
5: end if

\( ^{1} \text{Nash, Berge, Modularity, Community fitness, or Replace all;} \)

Computationally expensive, does not influence the running time of MNEO as much as other components of the method. However, the probabilistic version does offer a less computational expensive alternative to the Nash ascendancy relation, that can be used to enhance other heuristics that attempt to solve the community structure detection problem using game theoretic approaches.

Mihai Suciu, Rodica Ioana Lung, Nomi Gask, Game theory, Extremal optimization, and Community Structure Detection in Complex Networks, GECCO ’16, July 20-24, 2016, Denver, CO, USA (A core conferences list, ISI)

Tested variants Within NoisyEO the search is guided in two phases of the \( k\text{EO} \) iteration: (i) nodes having the lowest fitness are selected and randomly re-assigned in other communities, and (ii) selection for survival: the decision regarding the replacement of \( s_{\text{best}} \) (Alg. 1, line 3) can be made by using any fitness function that evaluates a community structure. The effect of this decision is studied in this paper: how does the use of the game theoretic tools in this part influences the results compared with the use of the modularity or the community fitness (which actually defines the node payoff)? To answer this question five variants of NoisyEO that only differ in the fitness mechanism used for selection of survival in line 3, Alg. 1 are considered:

1. \textit{Nash}: the Nash ascendancy relation is used: if \( s \) Nash ascends \( s_{\text{best}} \), it will replace it;
2. \textit{Berge}: the Berge ascendancy relation is used: if \( s \) Berge ascends \( s_{\text{best}} \), it will replace it;
3. \textit{Modularity}: \( s \) replaces \( s_{\text{best}} \) if it has a higher modularity (eq. (??)); in this case the fitness of a node is also computed using the modularity:

\[
u_i^{(Q)}(s) = \sum_{j \in C_i} \left( A_{ij} - \frac{k_i k_j}{2m} \right)
\]

where \( A \) is the adjacency matrix, \( m \) is the total number of links, and \( k_i \) is the degree of node \( i \), and \( C_i \) is the community of node \( i \);
4. \textit{Community fitness}: \( s \) replaces \( s_{\text{best}} \) if it has a higher community score (eq. (??));
5. \textit{Replace all}: \( s \) replaces \( s_{\text{best}} \) always, regardless of any fitness measure;
Numerical experiments were performed with the five variants of NoisyEO and results were compared with four state-of-art methods: Louvain Blondel et al. (2008), OSLOM Lancichinetti et al. (2011), Infomap Rosvall and Bergstrom (2008), and ModOpt Sales-Pardo et al. (2007). On the small (128 nodes) synthetic benchmarks two variants stand out with best results that are significantly better than the others in most cases: Modularity (3) and Community fitness (4), both based on optimization, while results obtained when using the ascendency relation are almost as good, with values close to 1 for the networks with well defined community structures ($\mu < 0.5$, $z_{out} \leq 8$). For $\mu = 0.5$ results obtained using the Community fitness are statistically better than all others except Modularity, while for $\mu = 0.6$ there are no statistical differences between results. Compared with other methods (Figure 2), for GN $z_{out} = 7, 8$, and for LFR $\mu > 0.3$ they are better than ModOpt and Infomap and as good as the others.

For the larger, but with better defined structures, LFR S and B sets with 1000 nodes, when comparing results with those reported by other methods (Figure 2), we can see that the EO based approaches lack the precision provided by the rest, with very good results but having average values less than 1. Except ModOpt none of the methods is affected by the value of $\mu$.

Because in the final population the difference between the NMI value of the individual having the best modularity value and the individual having the best NMI value are significant for these networks. The differences between the two illustrate the well known draw-back of modularity: in most cases, the maximum value of modularity does not correspond to the real community structure. This limitation is also visible in Figure 3 where the results obtained by all methods on the real-world networks are presented: all methods that rely on modularity report similar results with the EO variants when reporting the NMI of the individual with best modularity value in the population (continuous colored lines); only OSLOM, which does not use modularity, reports better results on the dolphins and karate networks, and Infomap, which also does not use modularity, on karate. When considering the individuals with best NMI values in the final EO population the results are different for the dolphins and karate networks: apparently game based variants (Nash and Berge) and Replace all report better results, but statistical comparisons show no difference between these two EO variants and the others. Similarities between results obtained by all methods for the real-world networks and those obtained for the LFR set with 128 nodes and $\mu = 0.6$ suggest that this set may be closest to simulate real networks and that behavior of a method on this set may be an indicator of possible good performance on real-world networks.

An interesting aspect of these results is related to the Replace all: even when no selection is used (the offspring always replaces the parent), results are similar to the game based approaches. The explanation may be that the extremal optimization procedure does select the worst nodes and modifies them randomly, and this selection is made based on the node fitness computed using $u_i(s) = f(c_i) - f(c_i \{i\})$, the same with the

5 Using the source code available at https://sites.google.com/site/andrealancichinetti/software, last accessed May, 2015
Figure 2: Synthetic benchmarks. Comparisons with other methods. Color error-bars (purple and blue) represent the best and worst EO variant, Community fitness and Replace all, respectively. Black error bars correspond to results obtained by the other state-of-art methods (Louvain, ModOpt, Oslom, and Infomap).

Figure 3: Results obtained on the real-world networks: continuous color lines represent average NMI values of the individuals reported by the method and dotted lines the best NMI value in the final population. Black lines represent results obtained by other methods.
node payoff of game $\Gamma$. It can be inferred that the search is actually directed by the inner mechanism of the extremal optimization algorithm which randomly replaces the worst components in each individual, while the selection for survival plays only a small refinement role.

Mihai Suciu, Rodica Ioana Lung, Noemi Gasko, About Nash Equilibrium, Modularity Optimization, and Network Community Structure Detection, GAMES 2016, the 5th World Congress of the Game Theory Society, Maastricht University 24-28 iulie 2016. This article explores the possibility of using the Nash equilibrium concept in the context of the network community detection problem. While it seems appealing to model the community structure by considering nodes as rational agents that choose the community that best fits their interests, there is still a big gap that has to be filled when constructing the game and in considering the approximation method. In this paper two possible payoff functions are considered. Results are controversial: one of them is better on the synthetic benchmark and the other yields better results on the real-world networks. However, compared with results offered by state-of-art methods in the literature, the tested approach appears better; further work consists in considering more realistic payoff functions, inspired from real-world models of community structure formation in order to further explore the possibility of using game theoretic concepts to approach this problem.

Rodica Ioana Lung, Noemi Gasko, Mihai Suciu, Network community structure detection and the Berge-Zhukovskii equilibrium, 28th European Conference on Operational Research, Poznan, 3-6 iulie, 2016. Figure 4 presents mean and standard errors for the results reported by BEO ($\epsilon = 0.001$) and Infomap, OSLOM, and Modularity Optimization in order to assess if these results are in fact comparable with those obtained by other methods. According to the Wilcoxon sum rank test, BEO results are signifi-

Figure 4: Mean and standard error of the mean of NMI for BEO and the other algorithms for GN, LFR and real networks.

Figure 4: Mean and standard error of the mean of NMI for BEO and the other algorithms for GN, LFR and real networks.
BEO results are better than those obtained by Infomap and Modularity Optimization. The Wilcoxon sum rank test was performed for all tested values of \( \epsilon \).

These results indicate the potential of this approach: they show that the concept of \( \epsilon \)-Berge-Zhukovskii equilibrium may be used to approach the community structure detection problem with results just as good and even better than those provided by other state-of-art methods. Many implications arise from here, of both practical and theoretical nature. From a practical point of view further experiments should be performed to study the behavior of search methods that attempt to compute the Berge-Zhukovskii equilibrium for the community structure detection game. From a theoretical point of view the challenge is to prove that the Berge-Zhukovskii equilibrium of this game indeed represents a network cover.

**Nomi Gask, Rodica Ioana Lung and Mihai Suciu, Community Detection in Bipartite Networks Using a Noisy Extremal Optimization Algorithm, Springer Series in Computational Intelligence (accepted, indexat ISI, ISDA 2016)** In this paper we are exploring the use of a community structure detection algorithm designed for unweighted and undirected networks for finding the structure of bipartite networks. By simply replacing the Newman modularity with the Barber modularity we find promising results, but to improve the method more attention to the specificities of bipartite networks has to be paid. The problem also requires attention: benchmarks with known structures are needed to evaluate the performance of different methods and allow quantitative comparisons, as a higher modularity may not always represent a better structure.

**Noemi Gasko, Rodica Ioana Lung, Mihai Suciu, EXTREMAAL OPTIMIZATION AND NETWORK COMMUNITY STRUCTURE, The 7th International Conference on BIOINSPIRED OPTIMIZATION METHODS AND THEIR APPLICATIONS, Bled, Slovenia, 18-20 mai 2016** A comparative analysis of four variants of extremal optimization updating procedures for the community structure detection problem is performed in this paper. The results show that the use of an adaptive method of setting the number of nodes to be randomly reassigned each iteration is beneficial; however, differences between tested variants are not significant enough to enable us to draw a conclusion regarding the best variant for the tested problems. Only when using an exponential rule, results are worse than the other EO variants, but even in those situations, they are very good.

Numerical results also show that extremal optimization may be very powerful in addressing the problem of community structure detection. Its main drawback, however, arises from the fact that random the computational time required by the iterative random changes makes this approach less efficient for large networks. On the other hand, this method proved very efficient for small networks with less visible community structures. Further work consists in finding the means to improve its scalability while maintaining its efficiency in dealing with ambiguous community structures.
Figure 5: Community structure of the averaged whole brain functional connectivity network.

Figure 6: Community structure of the anterior and posterior salience network in case of (A) healthy subjects, (B) cannabis users without ADHD, (C) subjects with childhood diagnoses of ADHD who does not use cannabis, (D) subjects with childhood diagnoses of ADHD who regularly use cannabis.

Lung, R. I., Suciu, M., Meszlnyi, R., Buza, K., & Gask, N. (2016, September). Community Structure Detection for the Functional Connectivity Networks of the Brain. In International Conference on Parallel Problem Solving from Nature (pp. 633-643). Springer International Publishing (A core conferences list, ISI) The analysis of brain functional connectivity networks from the community structure point of view can offer important information about the structure and functioning of the brain. The brain networks are relatively small, with very unclear structure, not detected by existing algorithms. In this paper we propose a game theoretic approach capable to identify strong connections in these networks and construct community structures that can offer relevant knowledge about the functioning of the brain. Moreover, the emergent field of neuromarketing proposes the analysis of the brain’s reaction to various stimuli in connection to advertisements and exposure to different products. The method proposed in this paper can be used to analyze functional connectivity networks of the brain constructed from such data and offer more insights into possible reactions. Thus the developed method opens the door to use community structure detection algorithm for analyzing sensitive marketing data.

Neuromarketing is an emerging interdisciplinary field that uses tools of neuroscience
in marketing applications by studying the reaction of individuals to different marketing approaches. In this project we are interested in analyzing functional magnetic resonance imaging (fMRI data) by using the method proposed in the above paper to gain a deeper understanding about the functioning of the brain and its reaction to various stimuli. We find this application challenging, as it strongly relates to the second objective of the project, that of studying information diffusion in networks.

**Information Diffusion in Social Networks** In this phase we performed a thorough documentation of existing literature on information diffusion in social networks by using the Scopus (www.scopus.com) database. We found several major approaches to information diffusion.

According to Pei and Makse (2013) the network diffusion models can be divided into two main classes: independent interaction models and threshold models.

Within independent interaction models it is assumed that an infected node transmits to its neighboring nodes the information/infection with a certain probability and that such probabilities across the network are independent. Examples are the susceptible-infected-recovered (SIR), susceptible-infected-susceptible (SIS), and the Bass model.

Threshold models take into account the possibility that a node accepts an information only if a fraction of its neighbors have already done so. We mention here the Linear Threshold Model, the General Threshold Model, the standard rumor model, the voter model, the strategic games model, etc.

An important component of diffusion models is also the identification of influential nodes/individuals in the network either by considering various centrality measures. A further step is to find sets of influential nodes.

Zhang et al. (2016) points out the lack of a common ground of the field. In his extensive review classifies dynamic spreading models into: threshold models, cascading models and epidemic models.

Related to our project, we were particularly interested in works relating the community structure of a network with modeling information diffusion. We found two mainstream approaches:

- The first one uses information diffusion methods to identify the community structure of a network (Shen et al., 2010; Chen, 2011; Hajibagheri et al., 2013; Fan et al., 2015; Jeub et al., 2015; Salnikov et al., 2016);
- The second one, which is also in the scope of our project, is to use the insights provided by knowing the community structure of the network to model and optimize information diffusion in the network (Anshelevich et al., 2014; Jiang et al., 2015; Vega-Oliveros and Berton, 2015; Halappanavar et al., 2016; Curato and Lillo, 2016)

**Game theory** offers efficient tools for modeling information diffusion (Jackson and Zenou, 2015) and the implication of game theory in modeling information diffusion is mentioned in many articles. However, there are very few works that connect community
structure detection methods, game theory and information diffusion models (Krishnamurthy, 2014; Zhang et al., 2016). One reason for this may be that equilibria concepts provided by game theory are often intractable and thus considered useless by the practitioners community; it is our assumptions that by providing methods of incorporating such concepts as alternates to optimization solution concepts used within diffusion models would enhance these models while also bridging the gap between game theory and practical applications.

References


