Robustness against Frustration in Community Detection

Lucas Lopes Felipe
UFRJ
Rio de Janeiro, Brazil

Konstantin Avrachenkov
INRIA
Sophia Antipolis, France

Daniel Sadoc Menasché
UFRJ
Rio de Janeiro, Brazil

ABSTRACT

We reinterpret the Leiden algorithm for community detection through a game-theoretic lens, modeling it as a hedonic game where nodes act as rational agents. This perspective uncovers an equivalence between CPM optimization and Nash equilibria and introduces agent-level notions of frustration. The resolution parameter γ balances conflicting objectives, and robustness arises when these are naturally aligned. We prove convergence guarantees and demonstrate empirically that robustness improves accuracy under noise.

Keywords

Leiden algorithm, hedonic games, community detection

1. INTRODUCTION

Community detection is key for network science, with applications spanning social network analysis, epidemiology, and distributed systems [8, 6]. The Leiden algorithm [11] has emerged as a state-of-the-art method for partitioning networks into disjoint communities by optimizing the Constant Potts Model (CPM) [10]. While its empirical success is well-documented [10, 11], there is still room for developments in its theoretical understanding and interpretability remains essential.

In this work, we propose a game-theoretic framework for community detection that reinterprets the Leiden algorithm's local-move phase as a hedonic game [2, 4], where nodes act as rational agents aiming to maximize a personalized utility. Each agent balances two competing objectives: maximizing the number of neighbors (friends) and minimizing the number of non-neighbors (strangers) within its community. These objectives are weighted by a resolution parameter γ that acts as a conflict arbitrator. This lens provides a fresh interpretation of the CPM objective, which traditionally lacks an agent-level notion of conflict or frustration.

Our approach introduces an explicit treatment of **frustration**, a concept adapted from statistical physics (e.g., in the Ising model), to characterize scenarios where no community assignment simultaneously satisfies all agent-level preferences [9, 3]. Our analysis focuses on **individual utility changes**, showing that they align with global CPM improvements. In this setting, **robust nodes**—those that remain optimal across all $\gamma \in [0, 1]$—emerge as frustration-free, as their conflicting objectives naturally align.

Table 1: Hedonic games, Ising Model, and communities

Hedonic Game Theory	**Ising Model (Physics)**	**Community Detection**
Players (agents)	**Spins** (magnetic particles)	**Nodes** in a network
Coalition (group of agents)	Aligned spins (same state)	Community (cluster of nodes)
Individual utility, $\varphi_i^\gamma = (1-\gamma)d_i^k - \gamma\hat{d}_i^k$	Local energy	Tradeoff: internal links vs. internal non-links
Nash Equilibrium (no agent can improve utility by moving)	Ground state (no local change reduces energy)	Locally optimal partition
Frustration: No coalition satisfies all preferences	**Frustration**: energy cannot be minimized for all edges	**Conflicting objectives**: best community differs by criterion
Potential Game: global potential increases with utility	**Hamiltonian**: global energy decreases with better alignment	**CPM** minimizes internal non-links and maximizes links
Totally robust agent: best group for all γ	**Unfrustrated spin**: aligned with all neighbors	**Robust node**: community is optimal across $\gamma \in [0, 1]$
Best-response dynamics	Spin flips (e.g., Glauber dynamics)	Node relocation (e.g., Leiden)
Stable partition = strategic equilibrium	Low-energy configuration	Robust and stable community structure

Table 1 summarizes this analogy across hedonic games, the Ising model, and community detection, showing how equilibrium concepts, dynamics, and robustness unify under a shared framework.

2. FORMAL RESULTS

We establish that the CPM corresponds to a hedonic potential game, transforming the problem into a game framework, and enabling a naturally distributed approach. We prove that CPM local maximizers—partitions where no node can improve its utility by unilaterally switching communities—are equilibria of this hedonic game [1]. This equivalence enables game-theoretic analysis of the Leiden algorithm's local-move phase. We demonstrate that the local-move phase of the Leiden algorithm converges to an equilibrium in pseudo-polynomial time [5], specifically $O(cV^2)$ steps, where V is the number of nodes and $\gamma = b/c$ (where $b, c \in \mathbb{N}$).

Algorithm 1: Better-Response Algorithm

```
Input  : G = (V, E); γ; Initial partition π.
Output: Equilibrium partition π.
changed ← true;
while changed do
    changed ← false;
    foreach node i ∈ V do
        Let C* ← arg max_C φ_i^γ(C);
        if C_i ≠ C* then
            Move node i to C*;
            changed ← true;
        end
    end
end
return π
```

Notation. We represent a network as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ the set of edges. A *partition* $\pi = \{\mathcal{C}_1, \ldots, \mathcal{C}_K\}$ groups nodes into disjoint communities. Each node $i \in \mathcal{V}$ belongs to a community $\mathcal{C}_k = \pi(i)$. The number of neighbors (edges) that node i has within community $\mathcal{C}_k$ is denoted by d_i^k, while $\hat{d}_i^k$ denotes the number of non-neighbors (nodes in $\mathcal{C}_k$ that are not adjacent to i). The resolution parameter $\gamma \in [0, 1]$ balances the preference for internal links (d_i^k) and internal non-links ($\hat{d}_i^k$). The individual utility of node i with respect to community $\mathcal{C}_k$ is given by $\varphi_i^\gamma(\mathcal{C}_k) = (1-\gamma)d_i^k - \gamma\hat{d}_i^k$. The global potential of a partition is denoted by $\Phi^\gamma(\pi)$ and satisfies $\Phi^\gamma(\pi) = -\mathcal{H}_{\text{CPM}}(\pi)$, where $\mathcal{H}_{\text{CPM}}$ is the Constant Potts Model Hamiltonian. We denote the change in utility when node i moves from community $\mathcal{B}$ to $\mathcal{A}$ as $\Delta\varphi_i^{\mathcal{BA}}$, and define a partition to be a γ-equilibrium if no node has incentive to deviate under resolution γ. A partition is said to be *robust* over an interval $[\gamma_0, \gamma_1]$ if it is a Nash equilibrium for all γ in that interval. A node is *robust* if it simultaneously maximizes d_i^k and minimizes $\hat{d}_i^k$ in its current community.

Theorem 2.1 (**Pseudo-Polynomial Time**). *When $\gamma = b/c$ is rational with $b, c \in \mathbb{N}$, the better-response dynamics of Algorithm 1 converges to an equilibrium partition in at most $O(cV^2)$ steps, thus achieving pseudo-polynomial time complexity.*

Proof sketch. The key insight is that the CPM-based potential function $\Phi^\gamma(\pi)$ is bounded between $-V^2$ and V^2 and takes values of the form U/c for integer U, due to the rationality of $\gamma = b/c$. Since each better-response move by a node strictly increases the potential by at least $1/c$, and the total range of potential values is at most $2V^2$, the number of improving steps is bounded by $2cV^2$. Consequently, the algorithm converges in pseudo-polynomial time, where the pseudo-polynomiality arises from the dependence on c (a numeric value encoded with $\log c$ bits), rather than on the bit-length of the input.

To assess partition stability, we propose two criteria: a strict multi-objective criterion requiring nodes to maximize neighbors and minimize non-neighbors simultaneously, and a relaxed single-objective criterion using a γ-weighted utility function, which the Leiden algorithm optimizes. A node is "robust" if it satisfies the strict criterion, and partition robustness is the fraction of such nodes.

Theorem 2.2 (**Robustness**). *Given a γ-equilibrium partition π^γ that holds as equilibrium for $\gamma = \gamma_0$ and for $\gamma = \gamma_1$, with $\gamma_1 > \gamma_0$, then π^γ is an $[\gamma_0, \gamma_1]$-robust equilibrium, denoted as $\pi^{[\gamma_0,\gamma_1]}$.*

Proof sketch. The proof relies on the linearity of the utility difference $\Delta\varphi_i^{\mathcal{BA}}(\gamma)$ with respect to the resolution parameter γ. Suppose node i belongs to community $\mathcal{B}$ and considers moving to community $\mathcal{A} \neq \mathcal{B}$. If π^γ is an equilibrium for both γ_0 and γ_1, then $\Delta\varphi_i^{\mathcal{BA}}(\gamma_0) \leq 0$ and $\Delta\varphi_i^{\mathcal{BA}}(\gamma_1) \leq 0$. Since the function is linear in γ, it follows that $\Delta\varphi_i^{\mathcal{BA}}(\gamma) \leq 0$ for all $\gamma \in [\gamma_0, \gamma_1]$. Thus, node i has no incentive to deviate within this interval. Repeating this argument for all nodes and all community pairs proves that π^γ is an $[\gamma_0, \gamma_1]$-robust equilibrium.

Theorem 2.3 (**Full Robustness**). *An equilibrium partition π^γ, which 1) holds for $\gamma = 0$, and 2) in which all communities have the same number of nodes, is an equilibrium for any $\gamma \in [0, 1]$.*

Proof sketch. The proof shows that if a node has no incentive to move under $\gamma = 0$, and all communities are equally sized, then it also has no incentive to move under $\gamma = 1$. Specifically, under $\gamma = 0$, equilibrium requires that a node has more neighbors in its current community $\mathcal{B}$ than in any other community $\mathcal{A}$, i.e., $d_i^{\mathcal{B}} \geq d_i^{\mathcal{A}}$. Under $\gamma = 1$, the condition becomes $n_{\mathcal{B}} - (d_i^{\mathcal{B}} + 1) \leq n_{\mathcal{A}} - d_i^{\mathcal{A}}$, where $n_{\mathcal{A}} = n_{\mathcal{B}}$ due to the balanced partition. This inequality is satisfied whenever $d_i^{\mathcal{B}} \geq d_i^{\mathcal{A}}$, showing that equilibrium at $\gamma = 0$ implies equilibrium at $\gamma = 1$. Applying Theorem 2.2, this establishes robustness for all $\gamma \in [0, 1]$.

Our theoretical results confirm that the Leiden algorithm's local-move phase efficiently converges to a hedonic game equilibrium [5]. By leveraging pseudo-polynomial time complexity, multi-agent systems can efficiently scale, harnessing the power of distributed local moves.

3. EXPERIMENTAL RESULTS

Experimentally, we evaluate this framework using synthetic networks from the Symmetric Assortative Planted Partition Model (SAPPM) [7], testing efficiency, robustness, and accuracy in a community tracking scenario where initial partitions are perturbed by noise [5].

We find that partitions with higher robustness correlate with higher-quality solutions. In the SAPPM, we consider networks with varying connection probabilities (p) and difficulty factors ($\lambda = q/p$), and indicate that robustness effectively filters stable partitions among multiple equilibria.

Samples for confidence intervals. Each of the five heatmaps in Figure 1 comprises 10×10 cells, each cell corresponding to a pair o parameters (p, λ). To ensure confidence in the values expressed in each cell of the heatmap, we collect samples for the same parameters while varying the seed argument of the NetworkX SBM function.[1] In particular, the value reported in each cell of Figure 1 is an average across 100 network samples. This sufficed for the purpose of producing 95% confidence intervals of length less than 0.01.

Each heatmap corresponds to a different number of communities ($K \in \{2, 3, 4, 5, 6\}$), with $V = 1020$ fixed, and number of nodes per community $N = 1020/K$. The darker the

[1] `https://networkx.org/documentation/stable/reference/generated/networkx.generators.community.stochastic_block_model.html`

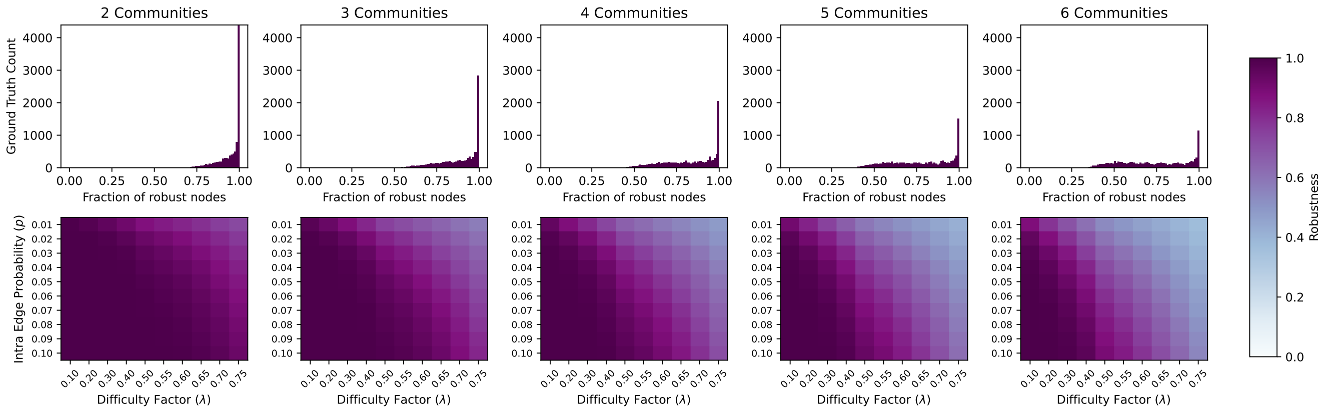

Figure 1: Robustness of the ground-truth partition varies as a function of the difficulty factor. Top row: Histograms showing the distribution of the fraction of robust nodes for various parameter settings. Bottom row: Heatmaps where each cell represents the average fraction of robust nodes for a given pair of connection probability p and difficulty factor λ.

color, the higher the fraction of fully-robust nodes. The heatmaps reveal several important insights:

Robustness decreases as K grows. As the number of communities (K) increases, the robustness of the ground truth partition tends to decrease. This is expected since the chance that the community of a given node is the community where this node has the highest number of neighbors and fewest non-neighbors diminishes as K increases. To illustrate that point, consider two extreme cases. The grand coalition ($K = 1$) is always fully robust, as nodes have no alternative communities to transition to. A set of singletons ($K = V$), in contrast, is never fully robust, as under the criterion of maximizing neighbors, nodes can benefit by transitioning to a community with an additional node.

Effect of parameters p and λ. The robustness is higher for lower values of λ and higher values of p. Indeed, we observe a gradual decline of robustness as λ grows. For $K = 2$, the robustness remains high for a large range of p and λ. As K increases, the robustness heatmap starts to show a gradual decline where higher values of λ and lower values of p result in lower robustness. This occurs because lower intra-community connection probabilities (p) and higher inter-community connection probabilities ($q = p\lambda$) make communities less distinct and harder to detect.

Robust partitions and ground truth. Although not all the ground truth partitions correspond to an equilibrium, we observe that a significant fraction of those partitions does correspond to fully robust equilibria. These are the cells marked with dark blue in Figure 1, with robustness equal one. In those cases, nodes have no incentive to transition across communities for any value of γ. For the cases where a fraction of nodes has incentive to deviate, such incentive holds for any γ.

In community tracking, we simulate network evolution by perturbing ground-truth partitions with noise levels ($\eta \in \{0.1, 0.25, 0.5, 0.75, 1.0\}$). Compared against baselines — spectral clustering, Leiden (full-fledged), one-pass improvement, and a mirror strategy (also referred to as zero-order hold (ZOH) in the realm of signal processing and control theory) — our hedonic-game-based method (Leiden Phase 1) achieves high accuracy, measured via the Rand index. For $\eta < 0.5$, it recovers ground truth with a Rand index above 0.9, outperforming spectral clustering (0.85) and simpler baselines, especially under moderate noise.

4. CONCLUSION AND FUTURE WORK

Our hedonic framework enhances the Leiden algorithm's interpretability and utility in dynamic settings, such as social media or biological networks, where communities evolve. The distributed algorithm suits decentralized environments, while the robustness filter improves solution quality without excessive computation. Future work could extend this to overlapping communities using multiple game initializations or explore non-greedy node movement strategies to escape local optima, potentially integrating evolutionary game theory for further refinement. **Further reading:** The complete version of this work is available at [5].

5. REFERENCES

[1] Bogomolnaia, A., and Jackson, M. O. The stability of hedonic coalition structures. *Games and Economic Behavior 38*, 2 (2002), 201–230.

[2] Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D. *Handbook of computational social choice*. Cambridge University Press, 2016.

[3] Cataudella, V., Coniglio, A., de Arcangelis, L., and di Liberto, F. Cluster formulation for frustrated spin models. *Physica A: Statistical Mechanics and its Applications 192*, 1-2 (1993), 167–174.

[4] Dreze, J. H., and Greenberg, J. Hedonic coalitions: Optimality and stability. *Econometrica: Journal of the Econometric Society* (1980), 987–1003.

[5] Felipe, L. L., Avrachenkov, K., and Menasché, D. S. From Leiden to Pleasure Island: identifying robust communities through hedonic games, 2025. `https://tinyurl.com/hedonicgames`.

[6] Girvan, M., and Newman, M. E. Community structure in social and biological networks. *PNAS 99*, 12 (2002), 7821–7826.

[7] Holland, P. W., Laskey, K. B., and Leinhardt, S. Stochastic blockmodels: First steps. *Social Networks 5*, 2 (1983), 109–137.

[8] Newman, M. E. The structure and function of complex networks. *SIAM review 45*, 2 (2003), 167–256.

[9] Reichardt, J., et al. Statistical mechanics of community detection. *Phys. Rev. E 74* (2006), 016110.

[10] Traag, V. A., et al. Narrow scope for resolution-limit-free community detection. *Physical Review E 84*, 1 (2011), 016114.

[11] Traag, V. A., et al. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports 9*, 1 (2019), 1–12.