Fixed-Points of Social Choice: An Axiomatic Approach to Network Communities

[Extended Abstract]

Christian Borgs
Microsoft Research
borgs@microsoft.com

Jennifer Chayes
Microsoft Research
jchayes@microsoft.com

Adrian Marple
Stanford University
amarple@stanford.edu

Shang-Hua Teng
USC
shanghua@usc.edu

ABSTRACT

We address the fundamental question of what constitutes a community in a social network. Inspired by social choice theory [1], we start from an abstract social network framework, called *preference networks* [2]; these consist of a finite set of members and a vector giving a total ranking of the members in the set for each of them (representing the preferences of that member.)

Within this framework, we take an axiomatic approach to study the formation and structures of communities. Our study naturally involves two complementary approaches. In the first, we apply social choice theory and define communities indirectly by postulating that they are fixed points of a preference aggregation function obeying certain desirable axioms. In the second, we directly postulate desirable axioms for communities without reference to preference aggregation, leading to a natural set of eight community axioms.

These two approaches allow us to formulate and analyze community rules. We prove a taxonomy theorem that provides a complete characterization of the most comprehensive community rule and the most selective community rule consistent with all community axioms. This structural theorem is complemented with a complexity result: we show that while identifying a community by the selective rule is straightforward, deciding if a subset satisfies the comprehensive rule is coNP-complete. Our studies also shed light on the limitation of defining community rules solely based on preference aggregation. In particular, we show that many aggregation functions lead to communities which violate at least one of our community axioms. These include any aggregation function satisfying Arrow's independence of irrelevant alternative axiom as well as commonly used aggregation schemes like Borda count or generalizations thereof. Finally, we give a polynomial-time rule consistent with seven axioms and weakly satisfying the eighth axiom.

KDD '14 New York City, New York USA

Categories and Subject Descriptors

Theory of computation [**Algorithmic game theory and mechanism design**]: Social networks; Information systems [**World Wide Web**]: Web applications—*Social networks*; Information systems [**Information systems applications**]: Data mining—*Clustering*; Human-centered computing [**Collaborative and social computing**]: Social network analysis

General Terms

Network analysis, theory, community identification

Keywords

Community identification, social network analysis, axiomatization, community axioms, clustering, preference networks, social choice theory, network algorithms, sampling, preference aggregations, complexity theory

1. INTRODUCTION: FORMULATING PREFERENCES AND COMMUNITIES

A fundamental problem in network analysis is the characterization and identification of subsets of nodes in a network that have significant structural coherence. This problem is usually studied in the context of *community identification* and *network clustering*. Like other inverse problems in machine learning, this one is also conceptually challenging: There are many possible ways to measure the degree of coherence of a subset and many possible interpretations of affinities to model network data. As a result, various seemingly reasonable/desirable conditions to qualify a subset as a community have been studied in the literature [7, 11, 10, 6, 2, 5, 3, 9, 12]. The fact that there are an exponential number of candidate subsets to consider makes direct comparison of different community characterizations quite difficult.

Among the challenges in the study of communities in a social and information network are the following two basic mathematical problems:

Extension of individual affinities/preferences to community coherence: A (social) network usually represents the pairwise interactions among its members, while the notion of communities is defined over its larger subsets. Thus, to model the formation of communities, we need a set of *consistent* rules to extend the pairwise relations or individual preferences to community coherence.

Inference of missing links: Since networks typically are sparse, we also need methods to properly infer the missing links from the given network data.

In this paper, we take what we believe is a novel and principled approach to the problem of community identification. Inspired by the classic work in social choice theory [1], we propose an axiomatic approach towards understanding network communities, providing both a framework for comparison of different community characterizations, and relating community identification to well-studied problems in social choice theory [1]. Here, we focus on the problem of defining community rules and coherence measures from individual preferences presented in the input social/information network, but we think that this study will also provide the foundation for an axiomatic approach to the problem of inferring missing links. We plan to address this second problem in a subsequent paper, which will use this paper as input.

Through the lens of axiomatization, we examine both mathematical and complexity-theoretic structures of communities that satisfy a community rule or a set of community axioms. We also study the stability of network communities, and design algorithms for identifying and enumerating communities with desirable properties.

While the approach initiated here is conceptual, we believe it will ultimately enable a more principled way to choose among community formation models for interpretation of current experiments, and also suggest future experiments.

1.1 Preference Networks

Before presenting the highlights of our work, we first define an *abstract* social network framework which enables us to focus on the axiomatic study of community rules. This framework is inspired by social choice theory [1] and was first used in [2] in the context of community identification for modeling social networks with complete preference information. We will refer to each instance of this framework as a *preference network*. Below, let S_n denote the set of all permutations of $[1:n] = \{1,...,n\}$.

Definition 1 (Preference Networks). *For a set of n elements $V = \{1,...,n\}$, a* preference network $A = (V,\Pi)$ *is defined by $\Pi = \{\pi_1,...,\pi_n\}$, where $\pi_i \in S_n$ specifies the total ranking*[1] *of V in the order of i's preference: $\forall s,u,v \in V$, s prefers u to v, denoted by $u \succ_{\pi_s} v$, iff $\pi_s(u) < \pi_s(v)$.*

As argued in [2], a real-life social network may be viewed as sparse, observed social interactions of an underlying latent preference network. In this view, the communities of a preference network may be considered to be the ground truth set of communities in its observed social network.

1.2 Highlights of the Paper

Our main contribution is an axiomatic framework for studying community formation in preference networks, and mathematical, complexity-theoretic, and algorithmic investigation of community structures in this framework. Our work on axiomatization of network communities can be organized into two related parts: (1) communities as fixed points of social choice aggregation functions; and (2) communities via direct axiomatic characterization. In the second part, we specify eight axioms we would like the communities to obey, and find conditions under which such communities exist. In the first part, we specify social choice aggregation functions for which the communities will be fixed points; this first method allows for an "indirect" axiomatic characterization in that the aggregation functions themselves could be taken to obey axioms, which would then indirectly characterize the communities which result as fixed points.

Communities as fixed points of social choice

Our approach of starting from preference networks to study communities naturally connects community formation to social choice theory [1], which provides a theoretical framework for understanding the problem of combining individual preferences into a collective preference or decision. In this first part of our analysis, we use *preference aggregation functions* studied in social choice theory [1] to characterize communities by defining communities as fixed points of a preference aggregation function.

Since real-world voting schemes and preference aggregation functions do not always produce a total order, we will use the following notation in the definition below. Let $\bar{S}_n$ denote the set of all *ordered partitions* of $[1:n] = \{1,...,n\}$. For a $\sigma \in \bar{S}_n$, for $i,j \in [1:n]$, we use to denote that i is *strictly preferred to* j (that is, i and j belong to different partitions, and the partition containing i is ahead of the partition containing j in σ). In this case, we also say $j \prec_\sigma i$. We will use $i \succeq_\sigma j$ to denote that $i \succ_\sigma j$ or i and j are in the same partition. As usual, let S_n^* be the set of finite sequences with elements in S_n. A preference aggregation function is then defined to be an arbitrary function $F: S_n^* \rightarrow \bar{S}_n$. If (V,Π) is a preference network, and $S \subset V$, then the image $F(\Pi_S)$ — where $\Pi_S = \{\pi_s : s \in S\}$ — is called the *aggregated preference* of S.

Definition 2. (Communities as Fixed Points of Social Choice) *Let $F: S_n^* \rightarrow \bar{S}_n$ be a preference aggregation function. For a preference network $A = (V,\Pi)$, $S \subseteq V$ is a* community of A with respect to F *if and only if $u \succ_{F(\Pi_S)} v$, $\forall u \in S, v \in V - S$.*

The function $\mathcal{C}_F$ mapping A into the set of communities defined above is called the fixed point rule with respect to F*. If F is not specified, i.e., if there exists an F such that $\mathcal{C} = \mathcal{C}_F$, we call $\mathcal{C}$ simply* a fixed point rule.

Informally, this definition says that a community is a subset $S \subseteq V$ such that, when we aggregate the preferences of all its members, the resulting aggregated preference puts the members S as the top $|S|$ elements, In other words, under the aggregation function F, the members of the community "vote" for themselves. Thus, S is a *fixed point* of its aggregated preference. Definition 2 generalizes the following concept of self-determination of [2]:

Definition 3. (B^3CT Communities) *Let $A = (V,\Pi)$ be a preference network. For each $S \subseteq V$ and $i \in V$, let $\phi_S(i)$ denote the number of votes that member i would receive if each member $s \in S$ was casting a vote for each of*

[1] In broader settings, one may want to consider preferences that allow *indifference* or partially ordered preferences, or both. One may also model a social network by a cardinal affinity network that specify each member's preference by a weighted affinity vector. For simplicity of exposition, we first focus on preference networks. In Section 5, we discuss the possible extension of our framework.

its $|S|$ most preferred members according to its preference π_s, $\phi_S(i) = |\{s : (s \in S) \ \& \ (i \in \pi_s[1 : |S|])\}|$. Then, S is B^{3}CT-self-determined *if everyone in S receives more votes from S then everyone outside S.*

Note that the B^{3}CT voting function ϕ_S is an instance of preference aggregation. We will also refer to a community according to Definition 2 as a *F-self-determined community.* We are particularly interested in those aggregation functions that satisfy various axioms in social choice theory [1], since this enables us to utilize established social choice theory to study all conceivable self-determination community rules within one unified framework. For example, it allows us to reduce the fairness analysis for community formation to the fairness of preference aggregation functions.

Arrow's celebrated impossiblility theorem and subsequent work in social choice theory [1] point to both challenges and exciting opportunities for understanding communities in preference networks. Recall that Arrow's theorem states that for $n > 2$, no (strictly linear) preference aggregation function satisfies all of the following three axiomatic conditions: Unanimity, Independence of Irrelevant Alternatives, and Non-Dictatorship (see Section 2 for definitions.) On the other hand, preference aggregation function exists if one relaxes any of these conditions. For instance, the well-known Borda count is a unanimous voting method with no dictators.

In this paper, we will examine the impact of preference aggregation functions on the structure of the self-determined communities that they define as well as the limitation of formulating community rules solely based on preference aggregation. See more discussion below.

Communities via direct axiomatic characterization

In this second approach, we will use a more direct axiomatic characterization to study network communities. To this end, we use a set-theoretical *community function* as a means to characterize a community rule.

Definition 4 (Community Functions). *Let $\mathcal{A}$ denote the set of all preference networks over V. A* community function *is a function $\mathcal{C} : \mathcal{A} \longmapsto 2^{2^V}$, and a subset $S \subseteq V$ is a* community *in a preference network $A = (V, \Pi)$ defined by $\mathcal{C}$ if and only if $S \in \mathcal{C}(A)$.*

We use axioms to state properties, such as fairness and consistency, that a desirable community function should have. An example is the property – referred to as the *anonymity* – that the community function should be isomorphism-invariant, that is, for any $\sigma \in S_n$, for all preference network $A = (V, \Pi)$, if $S \subseteq V$ is a community in A, then $\sigma(S)$ should still be a community in the $A' = (V, \Pi')$ where $\pi'_{\sigma(i)} = \sigma \circ \pi_i$. Another example is the property of *monotonic characterization*: If S is a community in $A = (V, \Pi)$, then S should remain a community in every preference network $A' = (V, \Pi')$ satisfying: (1) for all $i \not\in S$, $\pi'_i = \pi_i$, and (2) for all $u, s \in S$ and $v \in V$, if $u \succ_{\pi_s} v$ then $u \succ_{\pi'_s} v$.

In Section 2, we propose a natural set of eight desirable community axioms. Six of them, including both examples above, provide positive characterization of communities. These axioms concern the consistency, fairness, and robustness of a community function, as well as the community structures when a preference network is embedded in a larger preference network. The other two axioms address the necessary stability and self-approval conditions that a community should satisfy.

Constructing and Analyzing Community Rules

While Definition 4 is convenient for the study of the mathematical structure of our theory, community identification is a computational problem as much as a mathematical problem. Thus, it is desirable that communities can be characterized by a constructive community function $\mathcal{C}$ that is:

- **Consistent**: $\mathcal{C}$ satisfies all (or nearly all) axioms;
- **Constructive**: Given $A = (V, \Pi)$, and $S \subseteq V$, one can determine in polynomial-time if $S \in \mathcal{C}(A)$.
- **Samplable**: One can efficiently obtain a random sample of $\mathcal{C}(A)$.
- **Enumerable**: One can efficiently enumerate $\mathcal{C}(A)$, for instance, in time $O(n^k \cdot |\mathcal{C}(A)|)$ for a constant k.

Our two axiomatic approaches allow us to formulate a rich family of community rules and analyze their properties. Using the fixed-point rule, we can define a constructive community function based on any polynomial-time computable aggregation function. Alternatively, we can use one or a set of axioms as a community rule. We can also define a community rule by the intersection of a fixed-point rule and a set of axioms. In this paper, we aim to *characterize the community rules that satisfy a set of "reasonable" axioms*, and address the basic questions: *Is there a aggregation function leading to a community rule satisfying this set of "reasonable" axioms? What is the complexity of the community rules based on these axioms?*

Structural and Complexity-Theoretic Results

Our main structural result is a taxonomy theorem that provides a complete characterization of the most comprehensive community rule and the most selective community rule consistent with all our community axioms. This result illustrates an interesting contrast to the classic axiomatization result of Arrow [1] and the more recent result of Kleinberg on clustering [4] that inspired our work. Unlike voting or clustering where the basic axioms lead to impossiblity theorems, the preference network framework offers a natural community rule, which we call the *Clique Rule,* that is intuitively fair, consistent, and stable, although selective[2]: S is a community according the Clique Rule iff each member of S prefers every member of S over every non-member. Indeed the Clique Rule satisfies all our axioms. Our analysis then leads us to a community rule which is consistent with all axioms – we call it the *Comprehensive Rule* – such that for any community rule $\mathcal{C}$ satisfying all axioms and all preference network A, $\mathcal{C}_{clique}(A) \subseteq \mathcal{C}(A) \subseteq \mathcal{C}_{comprehensive}(A)$.

We complement this structural theorem with a complexity result: we show that while identifying a community by the Clique Rule is straightforward, it is coNP-complete to determine if a subset satisfies the comprehensive rule.

Our studies also shed light on the limitations of formulating community rules solely based on preference aggregation. In particular, we show that many aggregation functions lead to communities which violate at least one of our community axioms. We give two impossibility-like theorems. (1) Any fixed-point rule based on commonly used aggregation schemes like Borda count or generalizations thereof – such as the B^{3}CT self-determination rule – is inconsistent with (at

[2]See Section 3 for more details.

least) one of our axioms. (2) For any aggregation function satisfying Arrow's independence of irrelevant alternative axiom, its fixed-point rule must violate one of our axioms.

Finally, using our direct axiomatic framework, we analyze the following natural constructive community function inspired by preference aggregations.

Definition 5 (Harmonious Communities). *$S \subseteq V$ is a* harmonious community *of $A = (V, \Pi)$ if $\forall u \in S, v \in V - S$, the majority of $\{\pi_s : s \in S\}$ prefer u over v.*

We will show that the harmonious community rule is consistent with seven axioms and weakly satisfies the eighth axiom. In addition, various stable versions of harmonious communities (see the discussion below) enjoy some degree of samplablility and enumerability.

Stability of Communities and Algorithms

In real-world social interactions, some communities are more stable or durable than others when people's interests and preferences evolve over time. For example, some music bands stay together longer than others. Inspired by the work of [2] and Mishra *et al* [6] towards modeling this phenomenon, we examine the impact of stability on the community structure.

To motivate our discussion, we first recall the main definition and result of [2]: For $0 \leq \beta < \alpha \leq 1$, a subset $S \subseteq V$ is an $(\alpha, \beta)-B^3CT$ *community* in $A = (V, \Pi)$ iff $\phi_S(u) \geq \alpha \cdot |S|$, $\forall u \in S$ and $\phi_S(v) < \beta \cdot |S|$, $\forall v \notin S$. It is shown in [2] that in any preference network, there are only polynomial many stable B^{3}CT communities when the parameters α, β are constants, and they can be enumerated in polynomial time, showing that the strength of community coherences have both structural and computational implications.

In Section 4, we consider several stability conditions in our axiomatic community framework. In one direction, we examine the structure of the communities (defined by a fixed-point community rule) that remain self-determined even after a certain degree of perturbation in its members' preferences. In this context, for example, we can reinterpret the B^{3}CT-stability defined above as follows: A subset $S \subseteq V$ is an $(\alpha, \beta)-B^3CT$ *community* in a preference network A if it remains self-determined when $|S| \cdot (\alpha - \beta)/2$ members of S make arbitrary changes to their preferences. In the other direction, we consider some notions of stability derived directly from the social-choice based community framework where members of a community separate themselves from the rest. We can further use the separability as a measure of the community strength and stability to capture the intuition that stronger communities are also themselves more integrated. As a concrete example, we show in Section 4 that there are a quasi-polynomial number of stable harmonious communities for all these notions of stability. This result demonstrates that there exists a constructive community function that essentially satisfies all our axioms, whose stable communities are quasi-polynomial-time samplable and enumerable.

To fit our submission in 10 pages, we use Proposition to state properties whose proofs we will omit. A full version of our paper will be on Arxiv shortly after the KDD deadline.

2. COHERENT COMMUNITIES: AXIOMS

In this section, we define our eight core axioms, give a more formal treatment of social choice axioms and examine several properties of community rules and the relations these have with each other.

2.1 Axioms for Community Functions

For the following definitions, fix a community function $\mathcal{C}$ and a preference network $A = (V, \Pi)$.

Axiom 1 (Anonymity (A)). *For any permutation $\sigma : V \to V$, $S \in \mathcal{C}(A) \Longleftrightarrow \sigma(S) \in \mathcal{C}(V, \sigma(\Pi))$.*

A staple axiom, Anonymity states that labels should have no effect on a community function.

Axiom 2 (Group Stability (GS))). *Let $S \in \mathcal{C}(A)$. For all $G \subset S$, $G' \subseteq V - S$ and all tuples of bijections $(f_i : G \to G', i \in S - G)$, there exists $s \in S - G$, $u \in G$ such that $u \succ_{\pi_s} f_s(u)$.*

This axiom provides a type of game theoretic stability. In plain words, it states that no subgroup in a community can be replaced by a comparable group of non-members that are clearly preferred by the remainder of the community members. For instance, if the subgroup is of size 1, this means that there is no outsider that is universally preferred to this member, excluding that member's own opinion. On the other end of the spectrum, if the subgroup is all but one person, then group stability states that there must be someone from that member's top choices, and thus represents a type of individual rationality condition.

Axiom 3 (Self-Approval (SA)). *Let $S \in \mathcal{C}(A)$. For all $G' \subseteq V - S$ and all tuples of bijections $(f_i : S \to G', i \in S)$, there exists $s, u \in S$, such that $u \succ_{\pi_s} f_s(u)$.*

SA generalizes the intuition that a singleton should be a community only if that member prefers herself to everyone else. It uses the same partial ordering of groups as in GS.

Axiom 4 (Monotonicity (Mon)). *Let $S \subset V$. If Π' is such that for all $s, u \subseteq S$, and $v \in V$, $u \succ_{\pi'_s} v \Rightarrow u \succ_{\pi_s} v$, then $S \in \mathcal{C}(V, \Pi') \Longrightarrow S \in \mathcal{C}(V, \Pi)$.*

With Monotonicity if a member of a community gets promoted without negatively impacting other members, then that subset must remain a community. Thus this axiom reflects the fact that a high position means a greater affinity towards that person.

Axiom 5. (Coherence Robustness of Non-Members (CRNM)) *Let $S \subset V$. If Π' is such that for all $u, s, t \in S, v, w \notin S$, both $\pi'_s(u) = \pi_s(u)$ and $v \succ_{\pi'_s} w \Longleftrightarrow v \succ_{\pi'_t} w$, then*

$$S \in \mathcal{C}(V, \Pi') \Longrightarrow \; S \in \mathcal{C}(V, \Pi)$$

Axiom 6. Coherence Robustness of Members (CRM). *Let $S \subset V$. If Π' is such that for all $s, t, u, w \in S, v \notin S$, both $\pi'_s(v) = \pi_s(v)$ and $u \succ_{\pi'_s} w \Longleftrightarrow u \succ_{\pi'_t} w$, then*

$$S \in \mathcal{C}(V, \Pi') \Longrightarrow \; S \in \mathcal{C}(V, \Pi)$$

The two Coherence Robustness Axioms acknowledge the fact that if community members agree about their preferences about either members or non-member they are less likely to be a community.

Axiom 7 (World Community (WC)). $V \in \mathcal{C}(A)$

Axiom 8 (Embedding (Emb)). *Let V' be a set of additional voters and let Π' be an arbitrary preference profile over $V \cup V'$, s.t. for all $i, j \in V$, $\pi'_i(j) = \pi_i(j)$. Then $\mathcal{C}((V, \Pi)) = \mathcal{C}((V \cup V', \Pi')) \; \cap \; 2^V$.*

In other words, if we embed a preference network in a larger preference network in a manner in which all members of V prefer members of V over all members of V' and then look only at the communities which are subsets of V, the set of communities is the same as the original one.

2.2 Properties of Social Choice Axioms

Before we begin to talk about the properties induced by social choice axioms, we look at the properties that fixed point rules have without any further assumptions. To this end, we will define two properties of a community rule $\mathcal{C}$.

PROPERTY 1 (Independence of Outside Opinions (IOO)). *Let $\Pi, \Pi' \in S_n^n$ be two preference profiles such that $\forall s \in S$, $\pi'_s = \pi_s$. Then $S \in \mathcal{C}((V, \Pi')) \Leftrightarrow S \in \mathcal{C}((V, \Pi))$.*

IOO simply states that the preferences of outsiders cannot influence whether or not a subset is a community.

PROPERTY 2. (Null Community (NC)) $\forall A \in \mathcal{A}$, $\emptyset \in \mathcal{C}(A)$.

This property is a kind of dual to the WC Axiom, although without as strong an intuition as to what it means.

It turns out that these two Properties (and one of our Axioms) are always satisfied by any fixed point rule.

PROPOSITION 1. *All fixed point rules satisfy* Independence of Outside Opinions*,* Null Community*, and* World Community*.*

Turning now to social choice axioms, we must first formally define the axioms informally described in Section 1.2.

SOCIAL CHOICE AXIOM 1 (Unanimity (U)). *An aggregation function, F, satisfies* Unanimity *if for all preference profiles, $\Pi \in S_n^*$, and all pairs of candidates, $\{i, j\} \subseteq [n]$,*

$$\pi(i) > \pi(j) \forall \pi \in \Pi \Rightarrow F(\Pi)(i) > F(\Pi)(j).$$

The question then is: what properties capture the intuition behind Pareto Efficiency and how do they relate to this social choice axiom? To answer this, we define the following two properties of a community rule $\mathcal{C}$.

PROPERTY 3 (Pareto Efficiency (PE)). *For all $A \in \mathcal{A}$ and $S \in \mathcal{C}(A)$, if $u \in S$ and $v \notin S$ then there is a $s \in S$ such that $u \succ_{\pi_s} v$.*

PROPERTY 4 (Clique (Cq)). *For all $A \in \mathcal{A}$, if $u \succ_{\pi_s} v$ for all $u, s \in S$ and $v \notin S$, then $S \in \mathcal{C}(A)$.*

Pareto Efficiency is a negative property that states that subsets where a non-member that is Pareto preferred to a member should not be a community. In contrast, Clique is a positive Property, in that it states that a completely self-loving group (i.e. cliques) must be a community. As mentioned, both are implied by Unanimity.

PROPOSITION 2. *Let F be an aggregation function that satisfies* Unanimity*. $\mathcal{C}_F$ satisfies the* PE *and* Cq *Properties.*

SOCIAL CHOICE AXIOM 2 (Non-Dictatorship (ND)). *Given an aggregation function, an election is dictatorial if there exists a voter such that the aggregate preference is identical to that voter's preferences. An aggregation function satisfies* Non-Dictatorship *if no election is dictatorial.*

Instead of showing properties implied by ND as we did with Unanimity, we do the inverse, and show that a dictatorship violates some of our axioms.

PROPOSITION 3. *Let F be a preference aggregation function. If $\mathcal{C}_F$, the fixed point rule with respect to F, satisfies* GS *or* A*, then F satisfies* Non-Dictatorship*.*

The last of the three social choice axioms, Independence of Irrelevant Alternatives, simply states that the aggregate relation between any two pairs of candidates should not depend on the preferences for any other candidate.

SOCIAL CHOICE AXIOM 3. (Independence of Irrelevant Alternatives) *An aggregation function F satisfies* Independence of Irrelevant Alternatives (IIA) *if for all preference profiles, $\Pi, \Pi' \in S_n^m$ such that for candidates a and b we have that for all i, $a \succ_{\pi_i} b \Leftrightarrow a \succ_{\pi'_i} b$, then $a \succ_{F(\Pi)} b \Leftrightarrow a \succ_{F(\Pi')} b$.*

This axiom is often considered the strongest of the three as it says that the aggregate preference between two candidates does not even depend on the preferences voters have between either of the two and some other candidate. We will demonstrate this strength by proving an impossibility result involving modest assumptions about the fixed point rule of an aggregation function that satisfies IIA.

THEOREM 1. *There is no aggregation function $F : S_n^* \longrightarrow \bar{S}_n$ that satisfies* IIA *such that the fixed point rule with respect to F satisfies the* Clique *Property and the* Group Stability *Axiom.*

PROOF. Assume that F satisfies IIA, and the resulting fixed point rule $\mathcal{C}_F$ satisfies Cq and GS. We will first show that $\mathcal{C}_F$ must satisfy PE.

Let $S \subset V$ s.t. $1 < |S| < |V|$. In the following preference profiles, Π, Π', and Π'', we assume that every member of S has the same preference, π, π', and π'' respectively. First, let π rank all members of S above non-members. By the Cq, $S \in \mathcal{C}_F(A)$ and thus $\forall s \in S, v \notin S, s \succ_{F(\Pi)} v$. By IIA, this means that if $s \in S$ is unanimously preferred to $v \notin S$, then s must be strictly preferred to v in the aggregate preference.

Now let π' be the same as π only with the least preferred member of S, s', and the most preferred non-member, v', switched in rank. By GS, S is not a community. Furthermore, by the partial Unanimity property described above, the only way this is possible is if $v' \succeq_{F(\Pi')} s'$. Applying this property again yields the following two statements:

$$\forall s \in S - \{s'\}, s \succ_{F(\Pi')} s' \qquad \forall v \notin S \cup \{v'\}, v' \succ_{F(\Pi)} v$$

By symmetry and IIA, this means this for any two members or two non-members if one is unanimously preferred to the other, then it must be strictly preferred in aggregate preference.

Finally, consider π'' where v' is switched with the second lowest ranked member, s''. Because s'' must be strictly preferred to s' in the aggregate preference, by the above additional partial Unanimity property, v must be strictly rather than weakly preferred to s' in the aggregate preference. Thus, by symmetry and IIA again, if a non-member, $v \notin S$, is unanimously preferred to a member $s \in S$, v must be strictly preferred to s in the aggregate preference. Put together, these three partial Unanimity properties, constitutes Unanimity.

Since F satisfies both IIA and Unanimity, by Arrow's Impossibility Theorem [1] it must be a dictatorship. However, by Proposition 3, F must satisfy Non-Dictatorship. $\square$

2.3 Additional Properties of Axioms

Here we state some additional properties of interest that community rules (not necessarily fixed point rules) have when they satisfy one or more of our main axioms.

PROPOSITION 4. *If $\mathcal{C}$ satisfies the* WC *and* Emb *Axioms, then $\mathcal{C}$ also satisfies the* Cq *and* NC *Properties.*

PROPOSITION 5. *Any community rule that satisfies one of* Mon, CRM, *or* CRNM *satisfies the* IOO *Property.*

PROPERTY 5 (Outsider Departure (OD)). *A community rule $\mathcal{C}$ satisfies the* Outsider Departure *Property if for a given preference network $A = (V, \Pi)$, community $S \in \mathcal{C}(A)$, and outsider $v \notin S$, we have that $S \in \mathcal{C}(V - \{v\}, \Pi|_{V-\{s\}})$.*

PROPOSITION 6. *If $\mathcal{C}$ satisfies the* Mon *and* Emb *Axioms it also satisfies the* OD *Property.*

PROPOSITION 7. *If a community rule satisfies the* GS *and* SA *Axioms it also satisfies the* PE *Property.*

3. COMMUNITIES RULES

We now examine community rules through the lens of our axiomatic framework. In Section 3.1, we focus on a what we call weighted fixed-point rules, starting with the B^{3}CT community function from [2]. We show that it violates both Axioms Mon and GS. The violation of the monotonicity axiom was initially something of a surprise and somewhat counter-intuitive to us. We find it is also illustrative regarding the potential subtlety of community rules. We then show that the fixed-point community rule based on any Borda-count-like voting function is inconsistent with either the Group Stability axiom or the Clique property. This impossibility result and Theorem 1 illustrate some basic limitations of fixed-point rules.

Next, we analyze community rules that are formulated directly in terms of our axioms. In Section 3.2, we prove a Taxonomy Theorem regarding the spectrum of the axiom-conforming community rules. We present our complexity result in Section 3.3. Finally, we study the properties of the harmonious community function in Section 3.4.

3.1 Weighted Fixed Point Rules

This section focuses on a class of community rules that lie in between general fixed point rules and the B^{3}CT community rule, which we call *weighted fixed point rules*. First, we will look at some of the properties of the B^{3}CT rule as a particular case of a weighted fixed point rule.

THEOREM 2. *The B^{3}CT community rule, $\mathcal{C}_{B^3CT}$, does not satisfy* Mon *or* GS. *It satisifies all other axioms.*

PROOF. We only include the analysis for MON and GS. See the full arxiv version for others. Let $V = [1 : 6]$ and $\Pi = (\pi_1, ..., \pi_6)$ be

$$\pi_1 = [142356], \quad \pi_2 = [523416], \quad \pi_3 = [631425]$$
$$\pi_4 = [456123], \quad \pi_5 = [156423], \quad \pi_6 = [165423]$$

$S = \{1, 2, 3\} \in \mathcal{C}_{B^3CT}(V, \Pi)$. However, in violating Axiom Mon, $S = \{1, 2, 3\}$ is no longer a B^{3}CT community with the following preferences: $\pi'_1 = [124356], \pi'_2 = [234516], \pi'_3 = [314625]$, and $(\pi'_4, \pi'_5, \pi'_6) = (\pi_4, \pi_5, \pi_6)$, which satisfies the condition of Axiom Mon. Let $T = (1, 5, 6) \in \mathcal{C}_{B^3CT}(V, \Pi)$, $G = \{5, 6\} \subset T$ and $G' = (2, 4) \subset V - T$. As member 1 prefers 2 to 5 and 4 to 6, T does not satisfy GS. □

We now analyze the fixed point rule defined by the family of aggregation functions such as Borda count and B^{3}CT voting that derive a cardinal social preference from ordinal individual preferences. Let $W = (w^1, ..., w^n)$, $w_i \in \mathbb{R}^n_+$. For a set of preferences $\pi_1, ..., \pi_k \in S_n$, let

$$\phi^W_{(\pi_1,...,\pi_k)}(i) = \sum_{s=1}^{k} w^k_{\pi_s(i)}.$$

In B^{3}CT, w^k is the vector of k ones followed by $(n-k)$ zeros, while Borda count uses $w^k = (n, n-1, ..., 1)$ for all k.

DEFINITION 6 (WEIGHTED FIXED POINT RULE). *For a set of vectors $W = (w^1, ..., w^n \in \mathbb{R}^n_+)$, $S \in \mathcal{C}_W$ if and only if $\min_{u \in S} \phi^W_{\Pi(S)}(u) > \max_{v \notin S} \phi^W_{\Pi(S)}(v)$.*

PROPOSITION 8. *Weighted fixed-point rules satisfy Axiom* Anonymity.

THEOREM 3. (IMPOSSIBILITY OF WEIGHTED AGGREGATION SCHEMA) *Weighted Fixed Point Rules are inconsistent with either the* Group Stability *Axiom or the* Clique *Property.*

PROOF. Let $A = (V, \Pi)$ be a preference network, $S \subset V$ a subset with $k = |S|$, and $\mathcal{C}_W$ a weighted fixed point rule. First notice that for $i \leq k$ and $j > k$ we must have $w^k_i > w^k_j$ otherwise it will violate the Clique Property. We will with only slight loss of generality assume that the weights of vector w^3 are weakly decreasing. Throughout the course of this proof, we will use the fact that a preference profile violates Group Stability (GS), and the loss of generality will mean that when $i \succ_\pi j$ for the purposes of GS it must satisfy one of two conditions. Either $\pi(i) \leq k$ and $\pi(j) > k$, in which case no assumption needs to be made, or it must be proven that $w^k_{\pi(i)} = w^k_{\pi(j)}$, in which case positions $\pi(i)$ and $\pi(j)$ can be transposed without affecting our assumption.

Consider the following scenario:

$$V = \{a, b, c, d, e\}, \ S = \{a, b, c\}$$
$$\pi_a = [abcde], \ \pi_b = [abcde], \ \pi_c = [adebc]$$

By our assumption we know $w^3_1 \geq w^3_2 \geq w^3_2 > w^3_4 \geq w^3_4$, which implies $a \succ_{F_W(\Pi)} b \succeq_{F_W(\Pi)} c \succ_{F_W(\Pi)} e$ and $b \succ_{F_W(\Pi)} d$. The only way S can not be a community (since it violates GS) then is if $d \succeq_{F_W(\Pi)} c$. Notice that this implies that we cannot have both $w^3_2 = w^3_3$ and $w^3_4 = w^3_5$.

Now consider a modified preference profile:

$$\pi_a = [abdce], \ \pi_b = [abdce], \ \pi_c = [caebd]$$

In this profile c and d have been swapped and c promoted slightly over a in π'_c. Thus we have $a \succ_{F_W(\Pi')} b$, $b \succeq_{F_W(\Pi')} d \succ_{F_W(\Pi')} e$, and $b \succeq_{F_W(\Pi')} c$. Since Π' also violates GS we must have either $b \sim_{F_W(\Pi')} d$ or $c \sim_{F_W(\Pi')} d$. The former, however, implies $w^3_2 = w^3_3$ and $w^3_4 = w^3_5$ and is hence a contradiction. Therefore the latter must be true. Notice that this implies the following:

$$w^3_1 + 2w^3_4 = 2w^3_3 + w^3_5 \leq w^3_2 + 2w^3_4$$

However, because $w^3_1 \geq w^3_2$ this means in actuality $w^3_1 = w^3_2$.

This brings us to the final preference profile:

$$\pi_a = [abdce], \ \pi_b = [dcabe], \ \pi_c = [cbaed]$$

Note the Π'' still violates GS (and respects our slight loss of generality since $w^3_1 \geq w^3_2$). Weight from π'_b has changed from promoting d over c to indifference in π''_b. Because $F_W(\Pi')$ was indifferent between d and c this means that we now have $c \succ_{F_W(\Pi'')} d$. We also have $b \succ_{F_W(\Pi'')} d$ using the same argument from Π', and it is easy to see that $a \succ_{F_W(\Pi'')} d$ and in $F_W(\Pi'')$ everything is preferred to e. Together this implies that S is a community, which is a contradiction. $\square$

3.2 Taxonomy of Community Rules

We start with perhaps the simplest rule for communities that satisfies the Clique Property.

RULE 1 (CLIQUE RULE ($\mathcal{C}_{clique}$)). *$S \subseteq V$ is a community of $A = (V, \Pi)$, if and only if $\forall u, s \in S$, $v \notin S$, $u \succ_{\pi_s} v$.*

PROPOSITION 9. *$\mathcal{C}_{clique}$ satisfies all Axioms.*

However, the clique rule has the following structural feature, which essentially rules out any non-trivial overlap of communities. "Real-world" communities could have non-trivial overlaps among themselves. The Clique Rule appears to be too restrictive.

PROPOSITION 10. *For any preference network A, if $S_1, S_2 \in \mathcal{C}_{clique}(A)$, then either $S_1 \cap S_2 = \emptyset$ or $S_1 \subset S_2$, or $S_2 \subset S_1$.*

Are there more inclusive rules consistent with all axioms that admit overlapping communities? To address this question, we consider rules defined by community axioms.

RULE 2 (AXIOM BASED COMMUNITY RULES). *For*

$$X \in \{\mathsf{A},\ \mathsf{GS},\ \mathsf{SA},\ \mathsf{Mon},\ \mathsf{CRM},\ \mathsf{CRNM},\ \mathsf{WC},\ \mathsf{Emb}\}$$

let $\mathcal{C}_X$ denote the community rule: for a preference network $A = (V, \Pi)$, $S \in \mathcal{C}_X(A)$ if and only S enjoys Axiom X.

We can extend this community formulation to any community property. We can also take intersection of two rules: For any community functions $\mathcal{C}_1$ and $\mathcal{C}_2$, we let $\mathcal{C} = \mathcal{C}_1 \cap \mathcal{C}_2$ denote the community function that for all preference network A, $\mathcal{C}(A) = \mathcal{C}_1(A) \cap \mathcal{C}_2(A)$.

Our Taxonomy Theorem is a consequence of the following fundamental lemma.

LEMMA 1. (INTERSECTION LEMMA: GS AND SA) *For $X \in$ {A, Mon, CRM, CRNM, WC, Emb}, if $\mathcal{C}$ satisfies Axiom X, then $\bar{\mathcal{C}} = \mathcal{C} \cap \mathcal{C}_{GS} \cap \mathcal{C}_{SA}$ satisfies Axioms X,* GA *and* SA.

PROOF. $\mathcal{C}_{GS}$ and $\mathcal{C}_{SA}$ are both consistent with A, WC, and Emb, thus if $\mathcal{C}$ satisfies Axiom $X \in$ {A, WC, Emb}, then $\bar{\mathcal{C}}$ remains consistent with Axiom X. To see $\bar{\mathcal{C}}$ satisfies Axiom Mon, choose Π' such that, for all $u, s \in S$ and $v \in V$, $u \succ_{\pi_s} v \implies u \succ_{\pi'_s} v$. We need to show that if $S \in \bar{\mathcal{C}}(A)$ then $S \in \bar{\mathcal{C}}((V, \Pi'))$. Suppose this is not the case, then either (1) $S \notin \mathcal{C}_{GS}((V, \Pi'))$ or (2) $S \notin \mathcal{C}_{SA}((V, \Pi'))$. In Case (1), there exists $G \subset S$, $G' \subset V - S$, $|G| = |G'|$, and bijections $(f_s : S \to G' | s \in S - G)$ such that $\forall s \in S - G, \forall u \in G$, $u \prec_{\pi'_s} f_s(u)$. Then by Mon, we have $u \prec_{\pi_s} f_s(u)$, which shows $S \notin \mathcal{C}_{GS}(A)$. Case 2 can be similarly proved.

Suppose $\mathcal{C}$ satisfies Axiom CRM. Consider a Π' specified by Axiom CRM, which is defined by a permutation σ of S, such that $\forall u_1, u_2, s \in S$, $u_1 \succ_{\pi'_s} u_2$ iff $u_1 \succ_\sigma u_2$. We need to show that if $S \in \bar{\mathcal{C}}((V, \Pi'))$ then $S \in \bar{\mathcal{C}}(A)$. Suppose this is not the case, then either (1) $S \notin \mathcal{C}_{GS}(A)$ or (2) $S \notin \mathcal{C}_{SA}(A)$.

In Case (1), there exists $G \subset S$, $G' \subset V - S$, $|G| = |G'|$, a set of bijections $(f_s : G \to G', s \in S - G)$, such that $\forall s \in S - G, u \in G$, $u \prec_{\pi_s} f_s(u)$. Let $T \subset S$ be the set of $|G|$ least preferred elements by σ. We now show that there exists bijections $(f'_s : T \to G', s \in S)$ such that $\forall s \in S - T, u \in T$, $u \prec_{\pi'_s} f'_s(u)$, which would imply that $S \notin \mathcal{C}_{GS}((V, \Pi'))$.

Let us denote T by $T = \{t_1, ..., t_{|T|}\}$ such that $t_i \prec_\sigma t_{i+1}$. Fix an $s \in S - T$, and let us denote G by $G = \{g_1, ..., g_{|T|}\}$ such that $g_i \prec_{\pi_s} g_{i+1}$, and denote G' by $G' = \{g'_1, ..., g'_{|T|}\}$ such that $g'_i \prec_{\pi_s} g'_{i+1}$. Because f_s is a bijection between G and G', we have $\forall u \in G, u \prec_{\pi_s} f_s(u) \Longrightarrow g_i \prec_{\pi_s} g'_i$. In other words, $\pi_s(g_i) > \pi_s(g'_i)$. We define f'_s by mapping t_i to g'_i. Note that the positions of the preferences rankings of S as a set are the same in π'_s and π_s. Because T is the set of $|G|$ least preferred elements of S, we have $\pi'_s(t_i) > \pi_s(g_i)$. It then follows $\pi'_s(g'_i) = \pi_s(g'_i)$ that $\pi'_s(t_i) > \pi_s(g_i) > \pi_s(g'_i) = \pi_s(g'_i)$. Thus, $t_i \prec_{\pi'_s} g'_i$, and consequently, $S \notin \mathcal{C}_{GS}((V, \Pi'))$. Case (2) can be similarly proved. Thus, $\bar{\mathcal{C}}$ satisfies Axiom CRM. We can similarly prove CRNM. By definition, $\mathcal{C} \cap \mathcal{C}_{GS} \cap \mathcal{C}_{PE}$ satisfies GS and SA. $\square$

RULE 3. (COMPREHENSIVE COMMUNITY RULE) *For a preference network $A = (V, \Pi)$, $S \subseteq V$ is a community according to $\mathcal{C}_{comprehensive}$ iff S satisfies both* GS *and* SA *Axioms.*

THEOREM 4. (TAXONOMY) *$\mathcal{C}_{comprehensive}$ satisfies all Axioms. Moreover, for any community function $\mathcal{C}$ that satisfies all Axioms, for every preference network $A = (V, \Pi)$*

$$\mathcal{C}_{clique}(A) \subseteq \mathcal{C}(A) \subseteq \mathcal{C}_{comprehensive}(A). \quad (1)$$

PROOF. $\mathcal{C}_{all}(A) = 2^V$ satisfies Axioms A, Mon, CRM, CRNM, WC and Emb. As $\mathcal{C}_{comprehensive} = \mathcal{C}_{all} \cap \mathcal{C}_{GS} \cap \mathcal{C}_{SA}$, the theorem then follows from the Intersection Lemma. $\square$

3.3 Complexity of Community Rules

Theorem 4 shows that $\mathcal{C}_{comprehensive}$ and $\mathcal{C}_{clique}$ are the most inclusive and the most selective function, respectively, that satisfies all axioms. While it is very easy to determine whether a subset in a preference network satisfies Property Clique, in this section, we demonstrate that $\mathcal{C}_{comprehensive}$ is highly "non-constructive" by showing that the decision problem for determining whether a subset in a preference network satisfies Axiom Self-Approval or Group Stability is coNP-complete. Our reduction also provides examples of preference networks derived from 3-SAT instances.

THEOREM 5. *It is* coNP-complete *to determine whether $S \subset V$ satisfies Axiom* SA *in a given $A = (V, \Pi)$.*

PROOF. We reduce 3-SAT to this decision problem: Suppose $\mathbf{c} = (c_1, \ldots, c_m)$ is a 3-SAT instance with Boolean variables $\mathbf{x} = (x_1, \ldots, x_n)$ (i.e. $c_j = \{u_j, v_j, w_j\} \subset \cup_{i=1}^n \{x_i, \bar{x}_i\}$). We define a preference network as follows: (1) $V = A \cup B \cup D \cup X$ has $m + n + m + 2n$ members, where $A = \{a_1, \ldots, a_m\}$, $B = \{b_1, \ldots, b_n\}$, $D = \{d_1, \ldots, d_m\}$, and $X = \{x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$. The distinguished subset will be $S = A \cup B$, and for convenience we will denote its complement as $U = D \cup X$. (2) As we will focus on subset S here we only define the preferences of members in S. The preferences of U can be chosen arbitrarily. (i) Member b_i has preference $D \succ A \succ \{x_i, \bar{x}_i\} \succ \{b_i\} \succ X - \{x_i, \bar{x}_i\} \succ B - \{b_i\}$, where preferences between elements of each set can be chosen arbitrarily. (ii) Member a_j has preference $c_j \succ \{a_j\} \succ D \cup X - c_j \succ B \cup A - \{a_j\}$, where again preferences between elements of each set are arbitrary. Intuitively, members of A are used to enforce clause consistency (i.e. make sure each clause is satisfied) and members of B are there to enforce variable consistency (no variable to both true and false at the same time). Subsets of X naturally constitute an assignment of the variables, and D provides necessary padding in order to apply Self-Approval.

We now show that S does not satisfy Axiom Self-Approval if and only if the 3-SAT instance is satisfiable.

Suppose $Y = \{y_1, \ldots, y_n\}$ where $y_i \in \{x_1, \bar{x}_i\}$ is a satisfying assignment for the 3-SAT instance. Let $G' = Y \cup D$. Now consider the bijection, f, where $f(a_j) = d_j$ and $f(b_i) = y_i$. It is not hard to see that for all $s \in S$ and all i, $f(s) \succ_{\pi b_i} s$. All that is left is to find similar bijections for each a_j. First, note that for a_j all bijections f_j trivially satisfy $f_j(s) \succ_{\pi a_j} s$ where $s \in B \cup A - \{a_j\}$, since this set is ranked at the bottom of π_{a_j}. Therefore it is sufficient to show that there exists an element of G' that a_j prefers to itself. This happens so long as on of the literals from its clause is in G', which must be true by the fact that Y is a satisfying assignment.

Suppose $G' \subset U$ is a witness that S does not satisfy Axiom Self-Approval. We note the following: (1) $D \subset G'$ otherwise any b_i will have a member of A that cannot be mapped to a more preferred member of G'. (2) Let $Y = X \cap G'$. Then $|Y| = n$ by the above fact. (3) $\{x_i, \bar{x}_i\} \cap G' \neq \emptyset$ by b_i's preference, and by the pigeonhole principle the literals of Y are consistent (i.e. $\{x_i, \bar{x}_i\} \nsubseteq Y$). (4) $c_j \cap Y \neq \emptyset$ by c_j's preferences. Therefore the variable assignment implied by Y is a satisfying assignment for the 3-SAT instance. It follows that this problem is in coNP by Proposition 11 below. $\square$

Proposition 11. *For any $\pi \in S_n$ and disjoint sets G and G' of $[1; n]$ with $|G| = |G'|$, there exists a bijection $f : G \longrightarrow G'$ such that for all $g \in G$, $f(g) \succ_\pi g$ if and only for all $i \in [1 : |G|]$, $G'[i] \succ_\pi G[i]$, where $G[i]$ denote the i^{th} ranking element of G according to π.*

Theorem 6. *It is* coNP-complete *to determine whether $S \subset V$ satisfies Axiom* GS *in a given $A = (V, \Pi)$.*

Proof. We reduce the problem of determining whether a subset satisfies Axiom Self-Approval, which is coNP-complete to this decision problem: Let $A = (V, \Pi)$ be a preference network and $S \subset V$ be the subset in question. We define a new preference network, A' and subset S' of that network as follows: (1) $V' = \tilde{S} \cup V$, where $\tilde{S}$ is a disjoint copy of S. (2) $S' = S \cup \tilde{S}$. (3) Π' is defined as follows: (i) For $\tilde{s} \in \tilde{S}$, $\pi'_{\tilde{s}}$ ranks all of $\tilde{S}$ first, then adopts the same ranking as π_s. (ii) For $s \in S$, π'_s ranks all of S' first, then ranks $V - S$ arbitrarily. (iii) For $v \in V - S$, π'_v is arbitrary.

We now show that S' satisfies GS if and only if S satisfies SA. Assume S does not satisfy SA. Thus there is a $G' \subseteq V - S$ and a set of bijections $(f_i : S \to G', i \in S)$ such that for all $s, s' \in S$, $f(s') \succ_{\pi_s} s'$. Now choose $G = S$, it follows that for all $\tilde{s} \in \tilde{S}$, $s' \in S$ we have $f(s') \succ_{\pi'_{\tilde{s}}} s'$, by the definition of Π'. Therefore S' does not satisfy GS.

Assume S' does not satisfy GS. Thus there is a $\emptyset \neq G \subset S'$ and $G' \subseteq V' - S' = V - S$ and a set of bijections $(f_i : G \to G', i \in S - G)$ such that for all $s \in S - G$, $s' \in S$, $f(s') \succ_{\pi_s} s'$. Note the following: (1) $G' \cap \tilde{S} = \emptyset$ since all members of $\tilde{S}$ are preferred to all outsiders by everyone in S'. (2) $S \subseteq G'$ since all members of S prefers all members of S' before all outsiders. Hence $G' = S$ and it is trivial to apply the above argument to show that S does not satisfy SA. $\square$

3.4 Properties of Harmonious Communities

It remains open if there exists a *natural*, and *constructive* community function, that satisfies all eight axioms. Below, we shows that the harmonious community rule weakly satisfies Axiom GS in addition to satisfies all other seven axioms. We first parameterize Group Stability. For $0 < \delta \leq 1$, $\mathcal{C}$ satisfies the δ-Group Stability property if the following is true for all preference network $A = (V, \Pi)$ and $S \in \mathcal{C}(A)$.

Property 6. δ-Group Stability (δ-GS) *$\forall G \subset S$, $G' \subset V - S$ s.t. $0 < |G| = |G'| \leq \delta \cdot |S|$, it must be for all tuples of bijections $(f_i : G \to G', i \in S - G)$ there exists $s \in S - G$, $u \in G$ such that $u \succ_{\pi_s} f_s(u)$.*

Theorem 7. *The harmonious community function satisfies Axioms* A, SA, Mon, Emb, WC, CRM, *and* CRNM. *It also satisfies the* $1/2$-Group Stability *property.*

Proof. Let $\mathcal{H}$ denote the harmonious community function. We only include our proofs for Axioms CRNM and GS. See the full arxiv version for others. To see $\mathcal{H}$ is consistent with Axiom CRNM, consider a Π' specified by Axiom CRNM. Π' has property that there exists a permutation σ of $V - S$, such that $\forall i, j \in V - S$ and $\forall s \in S$, $i \succ_{\pi'_s} j \Leftrightarrow i \succ_\sigma j$. Let $v^* = \sigma(1)$, the most preferred element of σ. If $S \in \mathcal{H}((V, \Pi'))$ then for all $u \in S$, majority of $\Pi'(S)$ prefer u to v^*. Note that $\forall s \in S$ and $\forall v \in V - S$, $v \succ_{\pi_s} u$ implies $v^* \succ_{\pi'_s} u$. Thus, majority of $\Pi(S)$ prefer u to v, and $S \in \mathcal{H}((V, \Pi))$. We can similarly show that $\mathcal{H}$ satisfies Axiom CRM. The set T in the proof of Theorem 2 is also an example that $\mathcal{H}$ violates Axiom GS. We now show $\mathcal{H}$ satisfies the (1/2)-Group Stability property. We consider a pair $G \subset S$ and $G' \subset V - S$ satisfying $0 < |G| = |G'| \leq |S|/2$, and bijections $(f_s : G \to G' | s \in S)$. For each $u \in G$, because $\forall s \in S - G$, majority of S prefer u to $f_s(u)$, (who is not a member of S) and $|G| \leq |S|/2$, there must be $u \in G$ and $s \in S - G$ such that $u \succ_{\pi_s} f_s(u)$. $\square$

We conclude this subsection by taking a step back to compare the fixed-point community rules that we have discussed so far: Borda voting, B^3CT voting, and the harmonious rule. While all three have their own appealing simplicity and intuition and all satisfy Axioms A, SA, Emb, WC, CRM, and CRNM, there are significant differences with respect to Axioms Mon and GS, and the Outsider Departure property.

Outsider Departure: A harmonious community S remains a harmonious community when any outsider $v \notin S$ leaves the systems as the departure does not alter any pairwise preferences. However, the departure of an outsider can increase the votes of other outsiders 0 to $|S|$, causing the B^3CT community rule volatile when the preference network is evolving. The fixed-point rule of the Borda count rule is less sensitive to the outsiders' departure, but nevertheless does not satisfy this property.

Monotonicity: The harmonious rule satisfies Axiom Mon. The other two only satisfy a weaker version of monotonicity.

Group Stability: The subset T in the proof of Theorem 2 is a community according to all these three community rules. But T violates GS because 1 prefers outsiders over 5 and 6, even though 5 and 6 have the highest preference of 1: Element 1 is an *"arrogant" member* of its community. All aggregation functions satisfying Unanimity seem to be prone to existence of "arrogant" members. The harmonious rule satisfies the stability of majority subgroup, although the stability of the minority subgroup may not be guaranteed. The fixed-point rule of Borda count and B^3CT voting essentially have no guarantee on the group stability.

Small World: The harmonious rule also enjoys the following interesting property:

$S \in \mathcal{H}((V,\Pi)) \Leftrightarrow \forall v \subseteq V - S, S \in \mathcal{H}(S \cup \{v\}, \Pi|_{S\cup\{v\}})$.

4. STABILITY OF COMMUNITIES

In this section, we consider several stability measures and their impact to community structures.

4.1 Two Stability Models

We first study the structure of self-determined communities that remain self-determined even after a certain degree of changes in its members' preferences.

Definition 7 (Preference Perturbations). *Let Π and Π' be two preferences over $V = [1:n]$ and $S \subseteq V$. For $0 \le \delta \le 1$, we say Π' is a* δ-perturbation of Π with respect to S *if* $\max_{v\in V} |\{i \in S: \ \pi_i(v) \neq \pi'_i(v)\}| \le \delta|S|$.

We observe the following: For any $A = (V,\Pi)$, if $S \in \mathcal{C}_{B^3CT}(A)$ and is stable under δ-perturbations, then $\exists \alpha \ge \delta$ such that S is an $(\alpha, \alpha-\delta)$-B^3CT community. Conversely, if $S \subset V$ is an (α,β)-B^3CT community, then it is stable under $(\alpha-\beta)/2$-perturbations. Thus, the main result of [2] can be restated as: there are at most $n^{O(1/\delta)}$ B^3CT communities that are stable under δ-perturbations.

Axioms Coherence Robustness (CRM and CRNM) uses a family of *property-perserving* preference perturbations: Π' is a *membership-preserving perturbation of* Π *with respect to* S if $\forall s \in V$, $\{\pi_s(i) : i \in S\} = \{\pi'_s(j) : j \in S\}$ (which also implies $\{\pi_s(u) : u \notin S\} = \{\pi'_s(v) : v \notin S\}$). We can combine property-preserving conditions with the δ-perturbations to further refine the stability studies of community functions. The following is an illustrating example.

Theorem 8. *For any $A = (V,\Pi)$, the number of B^3CT communities that are stable under membership-preserving, δ-perturbations of Π is polynomial in $n^{1/\delta}$.*

Proof. We first show that if a B^3CT community S is stable under membership-preserving, δ-perturbations of Π, then either (1) S is an $(\alpha, \alpha-\delta)$-B^3CT community for some $\alpha > \delta$, or (2) $\exists s \in S$ such that $\pi_s[1:|S|] = S$. Let $u^* = \operatorname{argmin}\{\phi_S(u) : u \in S\}$ and $\alpha = \phi_S(u^*)/|S|$. We prove if S is not an $(\alpha,\alpha-\delta)$-B^3CT community, then $\pi_s[1:|S|] = S$ for some $s \in V$. This assumption implies $\phi_S(v^*)/|S| \ge \alpha - \delta$ where $v^* = \operatorname{argmax}\{\phi_S(v) : v \in V - S\}$. Let $T = \{s \in S : \pi_s(v^*) < |S|\}$. Then, $|T| = \phi_S(v^*) \ge (\alpha-\delta)|S|$. We know that $|T| < \alpha|S|$, since S is a B^3CT community. For each $s \in S - T$, if $\pi_s[1:|S|] \neq S$, then $\exists \pi'_s$ that agrees with π_s on S and $\pi'(v^*) \le |S|$. Thus, if $\pi_s[1:|S|] \neq S, \forall s \in S - T$, then using $\alpha|S| - |T| \le \delta|S|$ of them, we can find a membership-preserving, δ-perturbations Π' of Π such that S is not a B^3CT community in (V,Π'), contradicting with S is stable under membership-preserving, δ-perturbations of Π – there must be $s \in V$ with $\pi_s[1:|S|] = S$. The theorem then follows from [2]. $\square$

We can also strengthen the concept of fixed points in our social choice based community framework.

Definition 8. (δ-Strong Fixed Points) *Let $F: S_n^* \to \bar{S}_n$ be an aggregation function, and $\delta \in [0:1]$. Then, $S \in \mathcal{C}_F((V,\Pi))$ is* δ-strong *if for $\forall T \subseteq S$ such that $|T| \ge (1-\delta)\cdot|S|$, $u \succ_{F(\Pi_T)} v$, $\forall u \in S, v \in \bar{S} = V - S$.*

The δ-strong fixed points of an aggregation function interpolates between communities identified by the clique rule ($\delta = 1$) and self-determination ($\delta = 0$). Our goal is to understand the influence of F and δ ($0 < \delta < 1$) to the structure of the δ-strong F-self-determined communities.

Proposition 12. *For any $\delta \in (0,1)$, the number of δ-strong B^3CT communities in any preference network is $n^{O(1/\delta)}$.*

4.2 Stable Harmonious Communities

Applying stability notions of Sections 4.1 we define the following types of stable harmonious communities.

Definition 9 (Stable Harmonious Communities). *For $\delta \in [0:1/2]$, a subset S is a* δ-stable harmonious community *in $A = (V,\Pi)$ if $\forall u \in S, v \in V - S$, at least $(1/2+\delta)$-fraction of $\{\pi_s : s \in S\}$ prefer u over v. For $\delta \in [0:1]$, S is a* δ-strong harmonious community *in A if $\forall u \in S, v \in V - S$ and $T \subseteq S$ such that $|T| \ge (1-\delta)\cdot|S|$, majority of $\{\pi_s : s \in T\}$ prefer u over v,*

Proposition 13. . *If S is a δ-strong harmonous community then S is a $\delta/2$-stable harmonious community.*

With a simple probabilistic argument, we can prove:

Theorem 9. *$\forall\delta \le 1/2$, the number of δ-stable harmonious communities in any preference network is $n^{3\log n/\delta^2}$.*

Proof. Let S be a δ-stable harmonious communities. For any $T \subseteq S$, we say T *identifies* S for all $u \in S$ and $v \in V - S$, majority of T prefer u to v. We now show that $\exists T \subset V$ of size $3\log n/\delta^2$ that identifies S. To this end, we consider a sample $T \subset S$ of $k = 3\log n/\delta^2$ randomly chosen elements (with replacements). We analyze the probability that T identifies S. Let $T = \{t_1, ..., t_k\}$, and for each $u \in S$ and $v \in V - S$, let $x_i^{(u,v)} = [u \succ_{\pi_{t_i}} v]$, where $[B]$ denotes the indicator varable of an event B. Then T identifies S iff $\sum_{i=1}^k x_i^{(u,v)} > k/2, \forall j \in S, v \in V - S$. We now focus on a particular (u,v) pair and bound $\Pr\left[\sum_{i=1}^k x_i^{(u,v)} \le k/2\right]$. We first note that

$$\mathbf{E}\left[\sum_{i=1}^k x_i^{(u,v)}\right] = \sum_{i=1}^k \mathbf{E}\left[x_i^{(u,v)}\right] \ge \left(\frac{1}{2}+\delta\right)\cdot k.$$

By a standard use of the Chernoff-Hoeffding bound,

$$\begin{aligned}
&\Pr\left[\sum_{i=1}^k x_i^{(u,v)} \le k/2\right] \\
&\quad \le \quad \Pr\left[\sum_{i=1}^k x_i^{(u,v)} \le (1+2\delta)^{-1}\mathbf{E}\left[\sum_{i=1}^k x_i^{(u,v)}\right]\right] \\
&\quad \le \quad e^{-\frac{(2\delta)^2(1/2+\delta)k}{2}} \le \frac{1}{n^3}.
\end{aligned}$$

If T fails to identify S, then there exists $(u \in S, v \in V - S)$ such that $\sum_{i=1}^k x_i^{(u,v)} \le k/2$. As there are at most $|S||V -$

$S| \leq n^2$ such (u,v) pairs to consider, by the union bound,

$$\begin{aligned}\Pr[T \text{ identifies } S] &\geq 1 - \sum_{u \in S, v \in V-S} \Pr\left[\sum_{i=1}^{k} x_i^{(u,v)} \leq k/2\right] \\ &> 1 - 1/n.\end{aligned}$$

Thus, if S is a δ-stable harmonious communities, then either $|S| \leq 3\log n/\delta^2$, or it has a subset T of size $3\log n/\delta^2$ that identifies it. We can thus enumerate all δ-stable harmonious communities by enumerate all (T,t) pairs, where T ranges from all multi-subsets of V of size $3\log n/\delta^2$ and $t \in [1:n]$ and check if T can identify a set of size t. $\square$

5. REMARKS

While the results of this paper are conceptual and are built on the abstract framework of preference networks, we hope this study is a significant step towards developing a rigorous theory of community formation in social and information networks, and in particular will be used to inform and choose among other approaches to community identification which have been developed. Below we discuss a few short-term research directions that may help to expand our understanding in order to make more effective connection with community identification in networks that arise in practice.

Preferences Models

We have based our community formation theory on the *ordinal* concept of utilities used in social choice and modern economic theory [1]. The resulting preference network framework, like that in the classic studies of voting and stable marriage, enables our axiomatic approach to focus on the conceptual question of network communities rather than the more practical question of community formation in an observed social network. To better connect with the real-world community identification problem, we need to loosen both the assumption of strict ranking and the assumption of complete preference information.

With simple modifications to our axioms, we can extend our entire theory to a preference network $A = (V, \Pi)$ that allows *indifferences*, i.e., Π is given by n ordered partitions $\{\pi_1, ..., \pi_n\} : \pi_i \in \bar{S}_n$. This extension enables us to partially expand our results to affinity networks. Recall an affinity network $A = (V, W)$ is given by n vectors $W = \{w_1, ..., w_n\}$, where w_i is an n-place non-negative vectors. We can extract an ordinal preference $\pi_i \in \bar{S}_n$ from the cardinal affinities by sorting entries in w_i – elements with the same weight are assigned to the same partition.

Although this conversion may lose some valuable affinity information encoded in the numerical values, it offers a path for us to apply our community theory – even in its current form – to network analysis. For example, as suggested in [2], given a social network $G = (V, E)$, we can first define an affinity network $A = (V, W)$ where w_i is the personalized PageRank vector of vertex i, and then obtain an preference network (V, Π) where $\pi \in \bar{S}_n$ ranks vertices in V by i's PageRank contributions to them.

Theoretically, we would like to extend our work to preference networks with partially ordered preferences as a concrete step to understand community formation in networks with incomplete or incomparable preferences. Like our current study, we believe that the existing literature in social choice – e.g., [8] – will be valuable to our understanding. We think an axiomatic community approach to preference networks with partially ordered preferences together with an axiomatization theory of personalized ranking in a network may offer us new understanding of how to address the two basic mathematical problems – extension of individual affinities/preferences to community coherence and inference of missing links – for studying communities in a social and information network. As this part of community theory becomes sufficiently well developed, well-designed experiments with real-world social networks will be necessary to further enhance this theoretical framework.

Structures, Algorithms, and Complexity

Our taxonomy theorem provides the basic structure of communities in a preference network, while the coNP-Completeness result illustrates the algorithmic challenges for community identification in addition to community enumeration. On the other hand, our analysis of the harmonious rule and the work of [2] seem to suggest some efficient notion of communities can be defined.

However, it remains an open question if there exists a natural and constructive community rule that (1) satisfies all axioms, (2) allows overlap communities, and (3) whose stable communities are polynomial-time samplable and enumerable.

6. REFERENCES

[1] K. J. Arrow. *Social Choice and Individual Values.* Wiley, New York, 2nd edition, 1963.

[2] M. F. Balcan, C. Borgs, M. Braverman, J. T. Chayes, and S.-H. Teng. Finding endogenously formed communities. In *SODA*, pages 767–783, 2013.

[3] J. Hopcroft, O. Khan, B. Kulis, and B. Selman. Natural communities in large linked networks. In *KDD*, 2003.

[4] J. Kleinberg. An impossibility theorem for clustering. In *NIPS*, 2002.

[5] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 2009.

[6] N. Mishra, R. Schreiber, I. Stanton, and R. Tarjan. Finding strongly-knit clusters in social networks. *Internet Mathematics*, 2009.

[7] M. E. J. Newman. Modularity and community structure in networks. *PNAS*, 2006.

[8] M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Aggregating partially ordered preferences: Impossibility and possibility results. *TARK*, 10, 2005.

[9] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi. Defining and identifying communities in networks. *PNAS*, 2004.

[10] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *PNAS*, 105(4):1118–1123, Jan. 2008.

[11] D. A. Spielman and S.-H. Teng. A local clustering algorithm for massive graphs and its application to nearly-linear time graph partitioning. *SIAM J. Comput.*, 42(1):1–26, 2013.

[12] K. Voevodski, S.-H. Teng, and Y. Xia. Finding local communities in protein networks. *BMC Bioinformatics*, 2009.