Structural Stability of a Family of Spatial Group Formation Games

Chenlan Wang, *Graduate Student Member, IEEE*, Mehrdad Moharrami, Kun Jin, David Kempe, Paul Jeffrey Brantingham, and Mingyan Liu, *Fellow, IEEE*

***Abstract*—We introduce and study a group formation game in which individuals form groups so as to achieve high collective strength. This strength could be group identity, reputation, or protection, and is equally shared by all group members. The group's strength is derived from its access to resources possessed by its members, and is traded off against the geographic dispersion of the group; spread-out groups are costlier to maintain. We seek to understand the properties of stable groupings in such a setting. We define several types of equilibria, where a member wishing to join a new group requires the acceptance of that group, and may further require permission to leave its current group. We show that under natural assumptions on the group utility functions, some of these equilibria always exist, and that any sequence of improving deviations by agents (or subsets of agents in the same group) converges to an equilibrium. In characterizing the properties of these equilibria, We show that an "encroachment" relationship — which groups have members in the territory of other groups — always gives rise to a directed acyclic graph (DAG). We relate our model to observations of well-established groups in a real-world dataset.**

***Index Terms*—DAG, equilibrium, game theory, geographic dispersion, graph, group formation, resources, stability.**

I. Introduction

THE formation of groups and group membership plays an important role in human societies. Individuals form groups (or join existing groups) to benefit from shared interests or resources, reputation, protection, safety, monetary rewards, etc. [1]. For example, clubs are formed by individuals sharing similar interests, gangs often form to provide their members protection [2], and states form strategic alliances. The process by which groups are formed, and the groups' resulting stability, is of great interest and has been studied in many fields, including computer networks and wireless communications [3], [4], [5], social sciences [6], economics, and political sciences [7], [8]. Different models have been introduced to study such processes; including coalition formation games [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], agent-based modeling [21], [22], and signed network formation games [23], [24].

One of the key tradeoffs in the formation of groups is between adding more resources to a group (resource pooling) vs. making it less cohesive. In the examples above, when a club broadens its interests, it can add more members, but at the cost of less thematic cohesion. When a gang expands its territory, it can add more members (who add the ability to offer protection to others), but at the cost of less spatial cohesion, which makes it harder to coordinate actions. Similarly, when a strategic alliance comprises more countries, it can draw on more strength, but will suffer in less geographic cohesion. The big-picture question we investigate in this work is the following: *what are the effects of this tradeoff in terms of the group structures that one observes at equilibrium?*

To answer this question more formally, we introduce a group formation game (GFG)[1] in which the group members' utilities capture a trade-off between the combined resources of the group and the geographical dispersion. At a high level, our GFG is similar to hedonic games [12] studied within the family of coalition games. In the GFG we define (formal definitions are in Section III), agents exist in a Euclidean space. Each agent is endowed with a real-valued amount of individual resources. Groups form endogenously, and agents will form/join groups so as to maximize their utilities. Agents derive their utility from the utility of the group they belong to, plus optionally an individual component. The group utility is the same for all group members and is a linear combination of two terms: the strength of the group (which only depends on the group's members' resources and locations), and the negative effect of stronger groups. The strength of the group is assumed to be a monotone increasing function of the total resources of the group's members, and a monotone decreasing function of how "dispersed" the group is. We define and study several such notions, specifically, the volume, diameter, and surface area of the group, and study the game in equilibrium (Section IV) and structural properties of the groups formed at equilibrium (Section V).

Manuscript received 4 September 2023; revised 8 January 2024; accepted 22 February 2024. Date of publication 6 March 2024; date of current version 12 June 2024. This work was supported in part by NSF Under Grant CNS-1939006, Grant CNS-2012001, Grant ATD-2027277, and Grant CCF-1934986, and in part by ARO under Grant W911NF1810208. Recommended for acceptance by Dr. Z. Lyu. *(Corresponding author: Chenlan Wang.)*

Chenlan Wang, Kun Jin, and Mingyan Liu are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: chenlanw@umich.edu; kunj@umich.edu; mingyan@umich.edu).

Mehrdad Moharrami is with the Department of Computer Science, University of Iowa, Iowa City, IA 52242 USA (e-mail: moharami@umich.edu).

David Kempe is with the Department of Computer Science, University of Southern California, Los Angeles, CA 90007 USA (e-mail: david.m.kempe@gmail.com).

Paul Jeffrey Brantingham is with the Department of Anthropology, University of California, Los Angeles, Los Angeles, CA 90095 USA (e-mail: branting@ucla.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSE.2024.3370091, provided by the authors.

Digital Object Identifier 10.1109/TNSE.2024.3370091

[1] A preliminary version of this paper appeared in [25].

In choosing an appropriate equilibrium concept, we note that the settings described above typically allow agents to join a group only with the group's consent, whereas individuals may or may not leave a group unilaterally. We will primarily focus on the former case and discuss the latter in Section VII. We employ two notions of equilibria/stability that generalize the pairwise equilibria of [26]; the first is adapted from the original concept of the individually stable equilibrium by Drèze and Greenberg [20].

1) Individually Stable Equilibrium (ISE): each agent (weakly) prefers membership in its group over any group that would (weakly) prefer the agent to join that group. [2]
2) Strong Individually Stable Equilibrium (SISE): each subset of agents who are currently in the same group (weakly) prefers membership in their current group over joint membership in any group that would (weakly) prefer the entire subset to join.[3]

Several facts are worth noting about the notions of ISE and SISE: (i) ISE is different from the standard definition of Nash equilibrium [12], in which agents can join a group unilaterally, without the group's approval. (ii) This type of approval by groups is a natural requirement in many realistic scenarios; see, e.g., [2]. (iii) An alternative characterization of ISE is obtained by considering a normal-form representation of the game. Each agent chooses a subset of agents including itself as a strategy. When these strategies are consistent with each other (i.e., the sets form a partition into "cliques"), the result is the groups given by the strategies, with corresponding payoffs.[4] Otherwise, the result has utility $-\infty$ for all agents. Then, SISE can be viewed as equilibria under possible group deviations.

In Section IV, we show that under natural assumptions on the group utility functions, both an ISE and SISE always exist; this is done by showing that a better-response type of algorithm always finds such an equilibrium. Under a certain condition, we also introduce a polynomial-time algorithm that characterizes an SISE. These equilibria are in general not unique. Section V then focuses on the structural properties of the groups that form at equilibrium. We define the territory of a group as the convex hull of the individuals inside the group; our main focus is on understanding the relationship between the convex hulls of different groups, and to what extent they may overlap or contain each other, for different families of group utility functions. A particularly useful tool in characterizing the group structure is a graph capturing the territorial relationships among groups. We explore which graphs can be obtained at equilibria of the game in this way. We also show in Section VI how our model may be useful in explaining the territorial relationships of criminal gangs found in the Hollenbeck area in Los Angeles.

In Section VII, we discuss what happens when a group's utility also depends on weaker groups. We show that the existence of ISE and SISE is no longer guaranteed through a counter-example, but that by requiring an individual to obtain (its current group's) permission prior to departure, an equilibrium can be restored (a contractual version of ISE and SISE)[5].

II. Related Work

The literature most relevant to our work is on coalition formation games (CFG), which have been studied in the context of stable marriage, roommate assignment, and research team formation; see e.g., [11], [15], [16], [17], [18], [19], [27]. Coalition formation has also been used in the area of communication networks to construct distributed algorithms for performance improvement [3], [4], [5]. For example, using CFG results in higher performing multi-antenna systems in [3]. Coalition formation games were first introduced in [9] as transferable utility cooperative games formally given by $(\mathcal{N}, v, \mathcal{B})$, where $\mathcal{N}$ is the player set, v the value function associated with the coalitions, and $\mathcal{B}$ a given coalition structure (a partition of the player set $\mathcal{N}$). As the partition (or coalition/group structure) is given, this class of games primarily focuses on the allocation of values to achieve certain goals. Under certain conditions on the value function v, one could also formulate a game $(\mathcal{N}, v)$, where the goal is to find a desirable coalition structure $\mathcal{B}$, e.g., one that maximizes social welfare [10].

To study how stable coalition structures are formed, [12], [20] examine the coalition formation in the hedonic setting, where the utility of each agent depends only on the members of its own coalition. These games are called hedonic games (HG), given by $(\mathcal{N}, (\succ_i)_{i\in\mathcal{N}})$ [12], where $(\succ_i)_{i\in\mathcal{N}}$ is a profile of individual preferences. More explicitly, $\succ_i$ is the preference ranking of agent $i \in \mathcal{N}$ over the set $\{S \in 2^{\mathcal{N}} | i \in S\}$. Individuals have preferences over coalitions, and their preferences determine whether a stable partition/coalition structure exists. Hedonic games can also be formulated with structural constraints (graph-restricted), see, e.g., [17], [18], [19].

The equilibria considered in coalition formation games include core stability (CS), Nash stability (NS), individual stability (IS), and contractually individual stability (CIS). More discussion on some of these as they relate to the present study is given in subsequent sections. The task of finding a stable partition is in general NP-hard [28], which motivated studies on finding properties that can reduce complexity, see e.g., [15], [16], [11], [27].

The basic form of our game is a special case of the widely studied hedonic games mentioned above [12], [13], [20], [29], [30], and hence the existence of and convergence to equilibria follow from past work. However, we also go beyond past work in the following two ways: (1) We consider a version where the utility of weaker groups is affected by the composition of stronger groups, which generalizes the game to become non-hedonic. We show that similar arguments still apply in this case. (2) By focusing on more restricted utility functions defined on spatial locations, we are able to provide a more fine-grained understanding of the specific nature of equilibria, whereas past work on general hedonic games only established their existence.

Beyond coalition formation games, group formation has been studied computationally using Agent-Based Models

[2] This of course includes that the agent would not prefer to form a new (singleton) group by itself.

[3] Again, this includes that this subset of agents would not prefer to form a new group by themselves.

[4] Under this characterization, choosing a subset to form a group with is perhaps more aptly considered a "preference" rather than a "strategy" or "action"; for simplicity, we will use these terms interchangeably in this paper.

[5] This equilibrium concept is adapted to the *individually stable contractual equilibrium* in [20] by Drèze and Greenberg.

(ABM) [21], [22], in the context of group identity [31], using agent similarity [7], and by modeling inter-group conflict [23], [32], [33], [34].

Clustering studies the formation of groups based on similarities in agents' features, but typically in a non-strategic setting. The majority of clustering algorithms first pick cluster centers and then cluster points by assigning them to the closest center, e.g., K-*means* [35], *mean-shifting clustering* [36], and *clustering methods in* [37]. Using such "standard" clustering algorithms to analyze or explain observed group structure of human groups is called into question by the fact that none of these algorithms can produce overlapping cluster structure.[6] As we will see in Section VI, such structures do occur in reality.

Last but not least, how groups form has also been studied in the context of a network, called signed network formations (SNF). This is a well-studied area in the field of sociology and social psychology [32], where agents make pairwise decisions (e.g., whether to be friend or enemy with another, connected agent) that result in a network of signed edges. Hiller [23] is the first to model this as a game. There are many differences between a signed network game and our model, chief among which is the presence of a given network structure in the former and its absence in the latter. This means that in a signed network game an agent can only take action with respect to its neighbors on the network, whereas in our game an agent may join a group with any other agent.

III. The Group Formation Game (GFG)

A. The Model

Each of the $n \geq 2$ individuals/agents is indexed by $i \in \mathcal{N} = \{1, 2, \ldots, n\}$. Agent i is located at $x_i \in \mathbb{R}^d$ (where $d \geq 1$ is the dimension of the Euclidean space), and has positive scalar resources (abilities, skills, characteristics, etc.) $r_i > 0$. We write $\mathbf{x} = (x_1, \ldots, x_n)$ for the vector (technically, matrix) of all agents' locations, and $\mathbf{r} = (r_1, \ldots, r_n)$ for the vector of all agents' resources. Note, all symbols used in this work can be found in Table 3 in Appendix D.

We study group formation as a one-shot game in which each individual chooses the agents with whom she wants to be in the same group. Thus, agent i's strategy/action/preference space is $\mathcal{A}_i = \{A \subseteq \mathcal{N} | i \in A\}$; we will denote agent i's action by $a_i \in \mathcal{A}_i$. The set of all joint action profiles is $\mathcal{A} = \times_{i=1}^{n} \mathcal{A}_i$, and a joint action profile is denoted by $\mathbf{a} = (a_1, a_2, \ldots, a_n)$. The profile of the actions of all agents except i is $\mathbf{a}_{-i}$. In particular, we can write $\mathbf{a} = (a_i, \mathbf{a}_{-i})$.

We are interested only in profiles under which the groups form a *disjoint partition* of the individuals, and in which the actions chosen by agents are consistent. We say that the action profile $\mathbf{a}$ is *feasible* if and only if $a_i = \{j \in \mathcal{N} | a_j = a_i\}$ for all $i \in \mathcal{N}$; in words, if all agents that i wants to be in a group with also want to be in the *same* group.

A feasible profile $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ partitions the players $\mathcal{N}$ into $m = m(\mathbf{a}) \leq n$ disjoint *groups*. We index these groups as $G_1(\mathbf{a}), G_2(\mathbf{a}), \ldots, G_m(\mathbf{a})$; we will discuss a specific useful indexing scheme in Section III-C. In this and other notation, we omit the dependence on $\mathbf{a}$ for legibility when it is clear from the context. When the preference profile $\mathbf{a}$ is feasible, and the indexing of resulting groups has been fixed, we use $\sigma_i = \sigma_i(\mathbf{a})$ to denote the index of the (unique) group k such that $i \in G_k(\mathbf{a})$. We use the same notation for subsets of a group, i.e., for $S \subset G_k(\mathbf{a})$ we set $\boldsymbol{\sigma}_S = k$.[7] In that case (when the groups are clear from the context), we often consider $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{a})$ itself to be the strategy (or group affiliation) profile.

As discussed in the introduction, the reasons motivating individuals to form groups include seeking protection, pooling resources, and gaining reputation, among others. Within this context, we introduce our main assumption on the utility functions.

Assumption 1: The utility of a group depends on its own members' resources and locations as well as on the composition of other groups as follows:

$$U_G(\mathbf{r}, \mathbf{x}) = f(R_G, D_G) - \sum_{G': f(R_{G'}, D_{G'}) > f(R_G, D_G)} h(f(R_G, D_G), f(R_{G'}, D_{G'})), \quad (1)$$

where f represents the strength of a group: it is strictly increasing in R_G, a measure of the resources the group possesses, and strictly decreasing in D_G, a measure of the dispersion of the group membership; $h(y, z)$ is monotone decreasing in y and increasing in z, and non-negative whenever $z \geq y$.

Such utility functions capture situations where groups are negatively impacted by those that are stronger; h thus captures the *negative externality* which stronger groups impose on weaker groups. The assumptions on f and h imply that a group with higher strength also has a larger utility and vice versa:

$$U_G(\mathbf{r}, \mathbf{x}) > U_{G'}(\mathbf{r}, \mathbf{x}) \text{ iff } f(R_G, D_G) > f(R_{G'}, D_{G'}), \quad (2)$$

a fact that we will frequently exploit in subsequent analysis.

If $h \equiv 0$, then the groups' utilities are independent of one another and depend only on their own memberships, which is precisely the hedonic form of the utility function:

$$U_G(\mathbf{r}, \mathbf{x}) = f(R_G, D_G). \quad (3)$$

Hedonic utilities are thus a special case of the utilities given in (1).

For simplicity of presentation, we will frequently write $f(R_G, D_G)$ as f_G and $U_G(\mathbf{r}, \mathbf{x})$ as U_G.

The assumptions on $f(R_G, D_G)$ are placed to encode that the strength of a group (1) increases in the resources available to the group and (2) decreases as the group members are more "spread out." Specifically, we will consider $R_G = \sum_{i \in G} r_i$ as the total resources available to the group. We will consider a generic D_G to measure how spread out the group is, referred to as the group's *coverage*. Some natural examples are:

1) The Chebyshev distance $D_G = \max_{i,j \in G} \|x_i - x_j\|_\infty$.
2) The *volume* of the convex hull of $\{x_i | i \in G\}$.
3) The *surface area* of the convex hull of $\{x_i | i \in G\}$; this captures a notion of the "border" of the group.

[6]This can be observed by recalling that the Voronoi cells for any point set are disjoint.

[7]If the members of S are not all in the same group, then $\boldsymbol{\sigma}_S$ is undefined.

We will assume that the group utility is shared by all members of the group [8] and constitutes their individual utility; i.e., the utility of agent $i \in G := G_{\sigma_i}(\mathbf{a})$ is

$$u_i(\mathbf{a}) = U_G(\mathbf{r}, \mathbf{x}). \tag{4}$$

This assumption says that a group's combined strength is enjoyed equally by all its members, referred to as the common ranking property [29]. For notational convenience, we define the strength of an empty group as $f(R_\emptyset, D_\emptyset) = -\infty$ (and thus $U_\emptyset = -\infty$).

B. Equilibrium and Stability

Next, we define the notions of equilibria we study. At a high level, our goal is to capture stability against deviations by individuals or groups. Importantly, such deviations to another group are possible only when the group accepts these new member(s). We call partitions that are stable against deviations *acceptance equilibria*. We define two such equilibria, depending on whether we are considering individual deviation or group deviation.

Individually Stable Equilibrium (ISE): A group affiliation profile $\boldsymbol{\sigma}^*$ (and its corresponding strategy profile $\mathbf{a}$) is an *Individually Stable Equilibrium* (ISE) if and only if no agent can benefit from joining a group that would accept her. Formally, $\mathbf{a}$ is an ISE if and only if for all agents i, $G = G_{\sigma_i^*}(\mathbf{a})$, and any group $G' = G_k(\mathbf{a})$ (including $G' = \emptyset$) with $k \neq \sigma_i^*$, at least one of the following two inequalities holds:

$$U_G(\mathbf{r}, \mathbf{x}) \geq U_{G' \cup \{i\}}(\mathbf{r}, \mathbf{x}) \tag{5}$$

$$U_{G'}(\mathbf{r}, \mathbf{x}) > U_{G' \cup \{i\}}(\mathbf{r}, \mathbf{x}). \tag{6}$$

The first inequality states that agent i is weakly better off in her current group than by joining G' (and hence would prefer not to deviate); the second inequality states that group G' is better off without having agent i join, and hence prefers not to accept i. By including $G' = \emptyset$ and recalling that $U_\emptyset = -\infty$, this definition also captures the fact that i would not prefer to establish a group by herself. Note that since all members of a group have the same utility, a group's approval is equivalent to approval by any member of the group.

Strong Individually Stable Equilibrium (SISE): A group affiliation profile $\boldsymbol{\sigma}^*$ (and its corresponding strategy profile $\mathbf{a}$) is a *Strong Individually Stable Equilibrium (SISE)* if no *subset of agents* from the same group can be better off by deviating together to another group that would accept them. Formally, $\mathbf{a}$ is an SISE if and only if for every pair of groups $G = G_k(\mathbf{a}), G' = G_{k'}(\mathbf{a})$ (again, allowing for $G' = \emptyset$) and every subset $S \subseteq G$ of agents, at least one of the following two inequalities holds:

$$U_G(\mathbf{r}, \mathbf{x}) \geq U_{G' \cup S}(\mathbf{r}, \mathbf{x}), \tag{7}$$

$$U_{G'}(\mathbf{r}, \mathbf{x}) > U_{G' \cup S}(\mathbf{r}, \mathbf{x}). \tag{8}$$

The first inequality states that agents in S weakly prefer staying in G over deviating to join G', the second that agents in G' prefer not to accept the additional members S. Again, by including $G' = \emptyset$, we capture that the members of S do not prefer to form a new group. Notice that it is possible to express this condition concisely by heavily exploiting the fact that all members of S obtain the same utility.

C. States and the Ordering of Groups and States

To examine how agents change their group affiliations and how the process may converge to an equilibrium, we will think of action profiles as *states* and deviations as *transitions* between states.

Consider a state $\mathbf{a}$, with $m = m(\mathbf{a})$ denoting the number of groups in this state. Assign indices $1, 2, \ldots, m$ to these groups in descending order of their group utility, i.e., the m groups $G_1, G_2, \ldots, G_m$ are such that $U_{G_1}(\mathbf{r}, \mathbf{x}) \geq U_{G_2}(\mathbf{r}, \mathbf{x}) \geq \cdots \geq U_{G_m}(\mathbf{r}, \mathbf{x})$. Equivalently, we have $f(R_{G_{\sigma_1}}, D_{G_{\sigma_1}}) \geq f(R_{G_{\sigma_2}}, D_{G_{\sigma_2}}) \geq \cdots \geq f(R_{G_{\sigma_n}}, D_{G_{\sigma_n}})$. With a slight abuse of notation, we will also use the simpler $f_i(\mathbf{a}) := f(R_{G_{\sigma_i}}, D_{G_{\sigma_i}})$. The agents can be similarly indexed by $1, 2, \ldots, n$ such that $u_1(\mathbf{a}) \geq u_2(\mathbf{a}) \geq \cdots \geq u_n(\mathbf{a})$, Since agents from the same group have the same utility, the ordering can be chosen without loss of generality such that agents from the same group appear consecutively.

Note that the indexing of groups and agents in two different states $\mathbf{a}, \mathbf{a}'$ may be very different, even when the two states share common groups. This observation will become particularly relevant when we obtain $\mathbf{a}'$ from $\mathbf{a}$ by the deviation of one agent or a subset of agents: the fact that the utilities of the affected groups (resp., agents) may have changed can result in a different ordering of the groups and agents.

Each state $\mathbf{a}$, based on the above ordering, is associated with a vector $\boldsymbol{\Psi}(\mathbf{a}) = (f_1(\mathbf{a}), f_2(\mathbf{a}), \ldots, f_n(\mathbf{a}))$ which allows us to order the states lexicographically: a state $\mathbf{a}$ is ranked higher than (or appears before) $\mathbf{a}'$ in lexicographical order iff there exists an index $k \in \{1, \ldots, n\}$ such that:

$$f_k(\mathbf{a}) > f_k(\mathbf{a}'), \text{ and } f_{k'}(\mathbf{a}) = f_{k'}(\mathbf{a}'), \text{ for all } k' < k. \tag{9}$$

In this case, we write $\boldsymbol{\Psi}(\mathbf{a}) \succ \boldsymbol{\Psi}(\mathbf{a}')$. Notice that it is possible to have $\mathbf{a} \neq \mathbf{a}'$ while $\boldsymbol{\Psi}(\mathbf{a}) = \boldsymbol{\Psi}(\mathbf{a}')$.

IV. Existence of Equilibria and Convergence

In this section, we prove that every instance of the GFG has at least one SISE and thus at least one ISE (since any SISE is an ISE). We in fact establish a much stronger result: that any updating dynamics under which agents (resp., subsets of agents of the same group) always strictly improve their utility converges to an ISE (resp., SISE). Thus a best or better response algorithm always converges. Similar improvement algorithms in the context of hedonic games can be found in [12], [20], [29].

A. Convergence

Algorithm 1 presents a generic (asynchronous) improvement update algorithm. We will show that this algorithm converges to an SISE. In Algorithm 1, IR_S is the set of strictly improving responses for agents in S (including singletons), i.e., the set of groups that strictly improve the agents' utility over their current state, and would accept these agents as members. IR_S is only defined for subsets of agents that are currently in the same

[8] Adding an agent-specific term (which does not depend on the group utility or the action profile) to this utility function will not impact the subsequent analysis.

Algorithm 1: Improvement Update.

```
Input: σ^(0), a partition of 𝒩
t ← 0
while there exists a set S with IR_S ≠ ∅ do
    for all S ⊂ 𝒩 do
        if S ⊂ G_k for some k then
            IR_S ← {k' | U_{G_{k'}∪S}(r, x) > U_{G_{σ_S^(t)}}(r, x) and
            U_{G_{k'}∪S}(r, x) ≥ U_{G_{k'}}(r, x)}
        else
            IR_S ← ∅
        end if
    end for
    if there exists a set S with IR_S ≠ ∅ then
        Let S be arbitrary such that IR_S ≠ ∅
        Let k' ∈ IR_S be arbitrary
        Obtain σ^(t+1) from σ^(t) by moving agents in S to
        G_{k'} and leaving other memberships unchanged
    end if
    t ← t + 1
end while
```

group. Since both S and their new group are chosen arbitrarily, Algorithm 1 represents an improvement updating procedure, when sets of agents belonging to the same group move together. Restricting S in Algorithm 1 to be singletons, we obtain the same results for dynamics in which only individuals change their strategies to improve their utility.

Lemma 1 below shows that every state transition under Algorithm 1 results in a higher-ranked new state according to the lexicographical ordering. Thus, using lexicographic rank as a potential function[9] immediately implies Theorem 1, that the dynamics converges to an SISE, which proves the existence of an SISE, and thus an ISE.

Lemma 1: In Algorithm 1, for each iteration t, the new state $\mathbf{a}^{(t+1)}$ (corresponding to $\boldsymbol{\sigma}^{(t+1)}$) strictly precedes the state $\mathbf{a}^{(t)}$ (corresponding to $\boldsymbol{\sigma}^{(t)}$) in the lexicographical order of states. That is,

$$\boldsymbol{\Psi}(\mathbf{a}^{(t+1)}) \succ \boldsymbol{\Psi}(\mathbf{a}^{(t)}). \quad (10)$$

Proof: When a set of agents S deviates from G_k to another group $G_{k'}$, the only groups whose *strength* may be affected are G_k and $G_{k'}$; notice that even in the non-hedonic setting, while the externalities imposed by these two groups on other groups (via the h function) may change other groups' *utilities*, their *strengths* f_G only depend on their own members. Also recall that the lexicographic ordering is defined with respect to the *strengths* (not utilities) of the groups' members.

Let ℓ^- and ℓ^+ denote the indices of the groups $G_k \setminus S$ and $G_{k'} \cup S$ in the new utility[10]-based ordering after the deviation; so $G'_{\ell^-} = G_k \setminus S$ and $G'_{\ell^+} = G_{k'} \cup S$. Let $\hat{u}$ be the maximum utility of $G_k, G_{k'}, G'_{\ell^-}, G'_{\ell^+}$ in their respective groupings. For any group $G \neq G_k, G_{k'}$, the composition of the group, and hence its strength, stays unchanged. Hence, the elements in $\boldsymbol{\Psi}(\mathbf{a}^{(t)})$ and $\boldsymbol{\Psi}(\mathbf{a}^{(t+1)})$ corresponding to members of such groups G are identical.

Next, we analyze the strengths of the groups whose composition *does* change. By the definition of IR_S, we have

$$U_{G'_{\ell^+}}(\mathbf{r}, \mathbf{x}) > U_{G_k}(\mathbf{r}, \mathbf{x}) \quad (11)$$

$$U_{G'_{\ell^+}}(\mathbf{r}, \mathbf{x}) \geq U_{G_{k'}}(\mathbf{r}, \mathbf{x}). \quad (12)$$

Noticing that at least one of $U_{G'_{\ell^+}}$ and $U_{G'_{\ell^-}}$ must equal $\hat{u}$, we now distinguish the two corresponding cases.

1) If $U_{G'_{\ell^+}} \geq \hat{u}$, then at least $|G'_{\ell^+}|$ agents achieve utility at least $\hat{u}$, whereas previously, at most the $|G_{k'}|$ agents of $G_{k'}$ could have had utility $\hat{u}$ (or possibly less) by Inequality (12), while all agents of G_k had strictly lower utility by Inequality (11). Thus, strictly more agents now have utility at least $\hat{u}$ (and none of their utilities decreased).
2) If $U_{G'_{\ell^+}} < \hat{u}$, then by Inequality (12) and Inequality (11), we also have that $U_{G_k} < \hat{u}$ and $U_{G_{k'}} < \hat{u}$. Thus, $\hat{u} = U_{G'_{\ell^-}}$, so all agents in G'_{ℓ^-} saw a strict increase in utility to the new largest value $\hat{u}$, which strictly exceeds all previous utilities of agents in $G_k \cup G_{k'}$. As a result, the number of agents of utility $\hat{u}$ or more has strictly increased, while no agent with utility at least $\hat{u}$ saw a decrease in utility.

Now we observe that a strict increase in *utility* must coincide with a strict increase in *strength*; this is because the utility is composed of a strength term and an externality term, but the externalities with respect to other groups cannot have changed unless the group's strength has increased (recall the relation between utility and strength in (2)). Because out of all agents $G_k \cup G_{k'}$, the number of agents with highest utility (and thus strength) has strictly increased, the vector of agents' strengths is now lexicographically higher.

Therefore, in all cases, the new vector $\boldsymbol{\Psi}(\mathbf{a}^{(t+1)})$ ranks lexicographically before $\boldsymbol{\Psi}(\mathbf{a}^{(t)})$ (i.e. $\boldsymbol{\Psi}(\mathbf{a}^{(t+1)}) \succ \boldsymbol{\Psi}(\mathbf{a}^{(t)})$). □

Theorem 1: Algorithm 1 converges to an SISE in a finite number (at most $(n/\ln(n+1))^n$) of steps. In particular, an SISE always exists, and thus an ISE always exists.

Proof: By Lemma 1, each iteration results in a new state with higher-ranked $\boldsymbol{\Psi}(\mathbf{a})$; in particular, no state can be repeated. When the dynamics terminate, no subset of agents S from any existing groups can make an improvement move. This is exactly the equilibrium condition given in (7) and (8). Since no state can be repeated, and the total number of states is at most the total number of partitions of n items, i.e., the Bell Number $B_n \leq (n/\ln(n+1))^n$, the algorithm must converge to an SISE in the given finite time. □

We next introduce a second algorithm that will find an SISE that contains the group with the highest strength among all possible subsets of $\mathcal{N}$. We will refer to such an SISE as an HSISE (Highest-ranked Strong Individually Stable Equilibrium).

The arbitrary tie breaking in Algorithm 2 could lead to different lower-utility groups over the remaining iterations and thus different $\boldsymbol{\Psi}(\cdot)$ vectors. If tie breaking is never invoked, then the resulting HSISE is the highest-ranked state according to

[9] A similar potential function based on individual utility in the hedonic setting appeared in [30].

[10] Recall that the ordering by utilities and strengths is the same.

Algorithm 2: Finding an HSISE.

1: **Input:** $\mathcal{N}_1 = \{1, 2, \ldots, n\}$
2: $k \leftarrow 1$
3: **while** $\mathcal{N}_k \neq \emptyset$ **do**
4: Pick $G_k \in \text{argmax}_{G \subseteq \mathcal{N}_k} f(R_G, D_G)$ that has the largest number of agents, breaking ties arbitrarily
5: Set $\sigma_i = k$ for all $i \in G_k$
6: $\mathcal{N}_{k+1} \leftarrow \mathcal{N}_k \setminus G_k$
7: $k \leftarrow k + 1$
8: **end while**

the lexicographical order. In either case, the HSISE may not be unique as detailed further below.

Theorem 2: Algorithm 2 finds an HSISE.

Proof: Starting with $|\mathcal{N}_1| = n$, the sequence $|\mathcal{N}_k|, k = 1, 2, \ldots$ is strictly decreasing; thus, the algorithm terminates after a finite number of iterations. Consider group G_k identified during the k-th iteration. No members of G_k will be able to join any group $G_{k'}$, $k' < k$, and strictly increase their utility, for otherwise this contradicts the optimality of $G_{k'}$ during the k'-th iteration. Similarly, no member of G_k is willing to join any group $G_{k'}$, $k' > k$, for otherwise this contradicts the optimality of G_k during the k-th iteration. Thus the output of the algorithm is an SISE. Notice that G_1 is by construction the most powerful group among all possible subsets, making the output an HSISE.

Proposition 1: Assume that all agents are embedded as $x_i \in \mathbb{R}^d$, and the group strength is of the form $f(R_G, D_G)$, where $D_G = \max_{i,j \in G} \|x_i - x_j\|_\infty$. Then, Algorithm 2 runs in time $\mathcal{O}(n^{2d+2})$.

Proof: For any vector $x \in \mathbb{R}^d$ (in particular, those corresponding to agents), let $x(i)$ denote the i-th coordinate. For any group G, let $b_G^-(i) = \min\{x(i) | x \in G\}, b_G^+(i) = \max\{x(i) | x \in G\}$ be the smallest and largest entries in coordinate i of members of G.

Now consider Step 4 of Algorithm 2. Instead of searching over all $2^{|\mathcal{N}_k|}$ subset of agents, it suffices to focus on groups $G = \{x \in \mathcal{N}_k | b^-(i) \le x(i) \le b^+(i) \text{ for all } i = 1, \ldots, d\} \subseteq \mathcal{N}_k$ characterized by values $b^-(i), b^+(i), i = 1, \ldots, d$. This is because for any group G, we could consider the corresponding $b_G^-(i), b_G^+(i)$, and then add any agent $x \in \mathcal{N}_k$ who was previously not in G. This would strictly increase the resources R_G without increasing D_G, and hence strictly increase $f(R_G, D_G)$. In particular, if G is part of an HSISE, it must include all agents from $\mathcal{N}_k$ satisfying all of the inequalities.

Because we can focus on such groups, it suffices to search over all possible $(2\,d)$-tuples of values $b^-(i), b^+(i)$, of which there are at most $n^{2\,d}$, since each is a corresponding coordinate of an agent. Thus, Step 4 can be implemented in time $\mathcal{O}(n^{2d+1})$ (including computing R_G and D_G), and the complexity of the algorithm is $\mathcal{O}(n^{2d+2})$.

B. Non-Uniqueness of Equilibria

While Theorems 1 and 2 guarantee the existence of an SISE, the SISE may not be unique. This also implies that there may not be a unique ISE. This is shown in the following simple example.

Example 4.1: Consider three agents on a line, with locations $x_1 = 0$, $x_2 = 0.6$, and $x_3 = 1.2$, and resources $r_1 = 2$, $r_2 = 1$, and $r_3 = 2$. The group utility function is $U_G = (\sum_{i \in G} r_i)/(1 + \max_{i,j \in G} |x_i - x_j|)$.

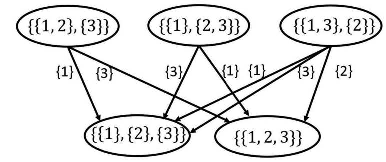


Fig. 1. State transition graph for a three-agent group formation game. Each edge shows an improving unilateral deviation by an agent, whose index is shown as the edge label.

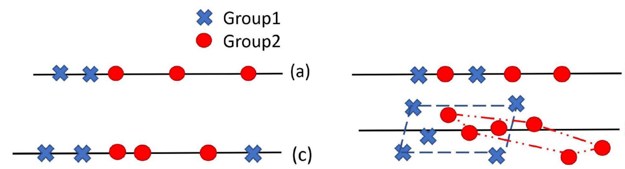


Fig. 2. Two groups that are (a) non-overlapping, (b) mutually encroaching, (c) nested, in 1D, and (d) one encroaching on another in 2D.

Fig. 3. Group $\{1, 2, 4\}$ (in blue) and Group $\{3, 5, 6\}$(in red) are mutually encroaching.

There are five different states in total: $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1\}, \{2, 3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{1, 2, 3\}\}$. The unilateral deviation from any state to another state is shown as the state transition graph in Fig. 1 . Simple calculations show that a single large group and groups of isolated individuals are both SISEs. (On the other hand, in partitions with two groups, such as $\{\{1, 2\}, \{3\}\}$, there is always an agent who prefers to deviate, resulting in either the partition $\{\{1\}, \{2\}, \{3\}\}$ or $\{\{1, 2, 3\}\}$.)

V. Structural Properties

Having established the existence of equilibria, we now turn to their properties. In particular, we are interested in the combinatorial structure of the overlap between different groups' "territories." We formally define the following:

Definition 1: The *territory* of a group G, denoted by X_G, is the convex hull of its members' locations:

$$X_G = \left\{ \sum_{i \in G} \lambda_i x_i \,\middle|\, \sum_{i \in G} \lambda_i = 1 \text{ and } \lambda_i \ge 0 \text{ for all } i \right\}.$$

Individuals inside the territory of group G, denoted by T_G, are defined as $T_G = \{i \in \mathcal{N} | x_i \in X_G\}$. The *area/volume* of X_G is denoted by $|X_G|$.

We are particularly interested in the possible cases of $i \in T_G$ while $i \notin G$. For the remainder of this section, we will adopt the following additional assumption.

Assumption 2: The group coverage function D_G depends only on the group's territory in the following sense: if $i \in T_G$, then $D_{G \cup \{i\}} = D_G$.

In Section III-A, we described the maximum pairwise distance, the volume of the convex hull, and the surface of the convex hull as three natural examples of group coverage. We note that all these examples satisfy Assumption 2.

A. Types of Structures in ISE and SISE

We define four types of overlap (or lack thereof) between pairs of groups.

Definition 2: Let G, G' be two groups.

- G and G' are *non-overlapping* if their territories are disjoint, i.e., $X_G \cap X_{G'} = \emptyset$.
- G *encroaches on* G' if G has at least one member in the territory of G', i.e., $G \cap T_{G'} \neq \emptyset$.
- G and G' are *mutually encroaching* if $G \cap T_{G'} \neq \emptyset$ and $G' \cap T_G \neq \emptyset$.
- G and G' are *nested* if all members of G are located within the territory of G', i.e., $G \subseteq T_{G'}$.

The four types of relationships are illustrated in Fig. 2. Note that non-mutual (or one-way) encroachment between groups in one dimension also means one group is nested inside the other (Fig. 2(c)); this need not be the case in higher dimensions (Fig. 2(d)).

It turns out that mutually encroaching structures cannot exist in any SISE (Corollary 1) but can exist in an ISE when the utility functions are non-hedonic (Example 5.1).

Example 5.1: Consider six agents on a line, with locations $\mathbf{x} = (0, 0.01, 0.51, 0.61, 1, 1.01)$ and resources $\mathbf{r} = (0.95, 0.95, 0.2, 0.1, 0.8, 0.8)$. Consider the non-hedonic group utility function $U_G = f_G + \sum_{G': f_{G'} > f_G} (f_{G'} - f_G)$, where $f_G = (\sum_{i \in G} r_i)/(e^{\max_{i,j \in G} |x_i - x_j|})$. It can be easily verified that the partition $\{\{1, 2, 4\}, \{3, 5, 6\}\}$, as shown in Fig. 3 , is an ISE since no single agent can successfully deviate. In particular, Agent 3 (resp. 4), currently located inside the territory of the blue (resp. red) group can always increase the *strength* of the blue (resp. red) group by joining. However, doing so will increase the strength of the group which the agent leaves even more, by making it much more cohesive. As a result, the group the agent leaves will overtake her new group in strength, and the non-hedonic utility function will impose a steep penalty on the agent's new group. As a result, the agent will not be accepted by her new target group in this game. Thus, the given grouping is in fact an equilibrium with a mutually encroaching structure.

While we have seen that mutually encroaching structures can occur in ISEs under non-hedonic utilities, focusing on either stronger equilibria (SISEs) or hedonic utilities makes encroachments disappear, as we will show below. In fact, this will follow directly from the following proposition, stating that no group can encroach or nest in another group with weakly higher utility.

Proposition 2: There is no SISE in which a group encroaches on or nests in another group with weakly higher utility.

Proof: We prove the contrapositive. Let $\boldsymbol{\sigma}^*$ be a group affiliation profile with two groups G and G' such that $G \cap T_{G'} \neq \emptyset$ (so G encroaches on G' or is nested in it) and $f_{G'} \geq f_G$. Thus, $U_{G'} \geq U_G$.

Let $i \in G \cap T_{G'}$. Since $i \in T_{G'}$, by Assumption 2, we have $D_{G'} = D_{G' \cup \{i\}}$. Because $R_{G' \cup \{i\}} > R_{G'}$, this implies that $f_{G' \cup \{i\}} > f_{G'} \geq f_G$. We now distinguish two cases:

- If $f_{G' \cup \{i\}} \geq f_{G \setminus \{i\}}$, then $U_{G' \cup \{i\}} - U_{G'} \geq f_{G' \cup \{i\}} - f_{G'} > 0$ and $U_{G' \cup \{i\}} - U_G \geq f_{G' \cup \{i\}} - f_G + h(f_{G'}, f_G) > 0$. Thus, agent i has an acceptable and beneficial deviation to join group G'.
- Otherwise, $f_{G' \cup \{i\}} < f_{G \setminus \{i\}}$, so $U_{G \setminus \{i\}} - U_G \geq f_{G \setminus \{i\}} - f_G + h(f_{G'}, f_G) > 0$. Thus, the set of agents $G \setminus \{i\}$ has a beneficial deviation to leave the current group G and set up their own group $G \setminus \{i\}$.

Hence, in both cases, there is a set of agents who have an acceptable and beneficial deviation to another group. Thus, $\boldsymbol{\sigma}^*$ is not an SISE. □

As an immediate corollary, we obtain the following:

Corollary 1: No SISE can contain two mutually encroaching groups.

Proof: For any two groups G, G', at least one (say, G) has weakly higher utility. By Proposition 2, G' will not encroach on G in any . So mutual encroachment cannot happen in an SISE.□

Summarizing the insights from Proposition 2 and Corollary 1, the following group structures may occur in an SISE:

- Non-overlapping groups: this occurs when groups are far from each other; thus, no agent or subset of agents has any incentive to deviate.
- Nested Structure: this can occur when a group with high resources is located within a much weaker group. Members of the inner group have no incentive to deviate, whereas members of the outer group will not be accepted by the inner group.
- Non-nested one-way encroaching structure: two groups may overlap as shown in Fig. 2(d), where if the red group has higher utility, its agents may not want to deviate to the blue group.

We saw that under SISEs, weaker groups will not encroach on or nest in stronger groups; hence, mutual encroachment does not occur in SISEs. We next show that in the special case of hedonic utility functions, i.e., $U_G(\mathbf{r}, \mathbf{x}) = f(R_G, D_G)$, the same group structures are also ruled out in AEs.

Proposition 3: When utilities are hedonic, there is no ISE in which a group encroaches on or nests in another group that has higher utility.

Proof: We prove the contrapositive. Consider a group affiliation profile $\boldsymbol{\sigma}^*$ in which two groups G and G' with $f_{G'} \geq f_G$ have $G \cap T_{G'} \neq \emptyset$, so G encroaches on or nests inside G'. The assumption that $f_{G'} \geq f_G$ also implies $U_{G'} \geq U_G$.

Let $i \in G \cap T_{G'}$, and consider a deviation of i to G'. Then, $R_{G' \cup \{i\}} > R_{G'}$, and because $i \in T_{G'}$, by Assumption 2, we have $D_{G'} = D_{G' \cup \{i\}}$. Hence, $f_{G' \cup \{i\}} > f_{G'} \geq f_G$. This also implies $U_{G' \cup \{i\}} \geq U_G$ by the assumption of hedonic utilities. Thus, agent i has an acceptable and beneficial deviation to group G', so $\boldsymbol{\sigma}^*$ is not an ISE. □

By observing that of any two groups G, G', at least one must have weakly higher utility, we again obtain a direct corollary ruling out mutual encroachment:

Corollary 2: Under hedonic utility functions, there is no ISE in which two groups are mutually encroaching.

B. ISEs, SISEs, and Their Graphical Representations

We now address the question of what types of structure can emerge globally, and how they may be characterized via *encroachment graphs*.

Definition 3: Given an ISE or SISE $\mathbf{a}^*$ and its group partition $G_1, \ldots, G_m$, we define the directed *encroachment graph* $\mathcal{G}(\mathbf{a}^*)$

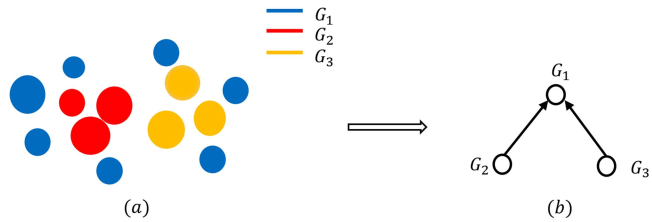


Fig. 4. (a) Shows agents and their resources, and a group affiliation forming an ISE (or SISE). Each colored dot is an agent from the corresponding group; the larger the dot, the more resources the agent has. The center of the dot gives the location of the agent. (b) Shows the corresponding directed encroachment graph.

as follows: the m nodes $V(\mathbf{a}^*)$ are the groups $G_1, \ldots, G_m$, and there is a directed edge from G_i to G_j if and only if G_i encroaches on G_j.

Example 5.2: Fig. 4 illustrates this definition with a concrete example.

As discussed in Section V-A, in any SISE, only a group with higher strength can encroach on a group with lower strength. It follows that the encroachment graph corresponding to any SISE is acyclic, as formally stated below.

Proposition 4: For every SISE $\mathbf{a}^*$, the encroachment graph $\mathcal{G}(\mathbf{a}^*)$ is acyclic.

Similarly, when the utility function is hedonic (i.e., $U_G(\mathbf{r}, \mathbf{x}) = f(R_G, D_G)$), in any ISE, only groups with higher strength can encroach on groups with lower strength. Again, this implies that the encroachment graph corresponding to any hedonic ISE is acyclic.

Proposition 5: When utilities are hedonic, for every ISE $\mathbf{a}^*$, the encroachment graph $\mathcal{G}(\mathbf{a}^*)$ is acyclic.

When the utilities are hedonic, a converse can be obtained. In Appendix A, we show that for a wide class of hedonic utility functions, every DAG can arise as an ISE of a game with suitably chosen agent locations and resources, even in just two dimensions.

VI. Observations From a Real-World Scenario

This section examines a real-world scenario of the co-existence of rival groups. We look closely at criminal street gangs co-located in the Hollenbeck area east of Los Angeles [38], [39], [40], [41]. We highlight the nesting and encroaching relationships observed among these groups, and use our model to reason about the relative strengths of these groups. The analysis relies on two related datasets, which we describe below.

A. A Picture of the Hollenbeck Gangs

The first dataset includes anonymized Field Interview cards collected by the Los Angeles Police Department [42], [43]. Each card records gang affiliations of individuals present at a non-criminal stop as well as time and location information (e.g., name of a road), which can be geocoded.

The second dataset is a territorial map of gangs in the Hollenbeck area [44]. In Fig. 5, each group is shown in a different color and given a numerical index. The dataset also contains the area of each gang territory and the estimated membership size in each gang. These data are used under UCLA IRB Protocol #21-001700.

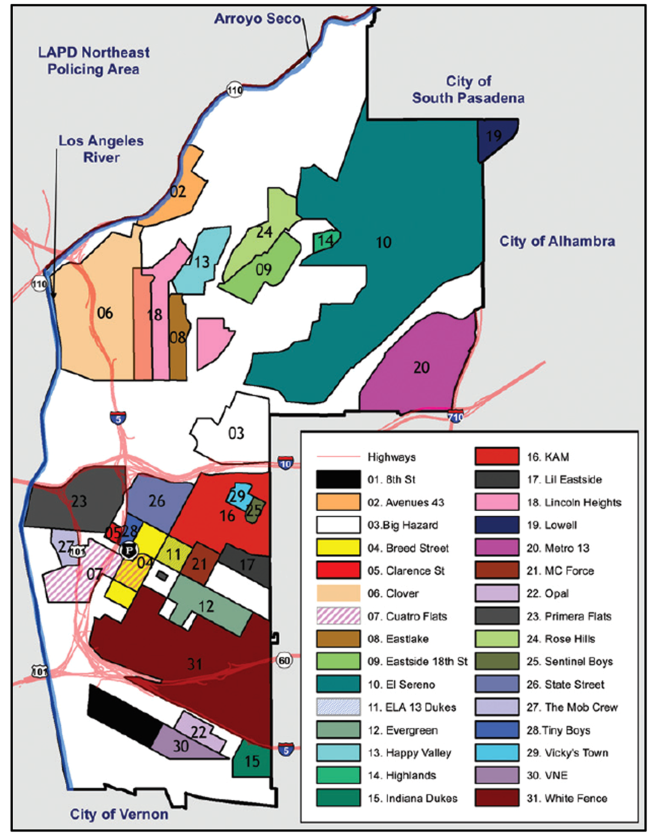


Fig. 5. Gang territories in LA's Hollenbeck area [42]. The area is bordered by the LA River, the Pasadena Freeway, and regions with no rival gangs.

B. Analysis

We start with the assumption that the gangs recorded in the above datasets are arranged spatially in a stable or equilibrium condition [45]. This is a fairly reasonable assumption given that some of these groups have existed for decades without much territorial change [41], [42].

With this assumption in mind, we observe in Fig. 5 that most groups have well-defined territories that provide a clear separation between them and their neighbors. Yet, there are also two distinct instances of *nesting*: groups 25 and 29 are both nested within group 16. There are also a number of instances of apparent encroachment: group 12 encroaches on 31, group 14 encroaches on 10, etc.

We recognize that encroachment may be the result of imprecision in how gang territories are drawn [41]. We are also aware that physical barriers, such as freeways, might serve as hard boundaries that exclude territorial encroachment that might otherwise occur [45].

While the process of group formation was not observed (and we do not have information on actual utility functions), the existence of overlapping territorial structures may be interpreted in light of our model. In particular, whereas much of the literature on gang spatial organization has focused on the largest (and most violent gangs) [42], [41], [46], our model draws attention to some of the smallest gangs in the community (i.e., groups 25 and 29).

In what follows, we explore how our model explains such formations, knowing that such explanations may not be unique.

TABLE I
RANKING OF NESTING GANG GROUPS UNDER DIFFERENT STRENGTH FUNCTIONS AND PROXY CHOICES

Strength function	Choice of R_G	G_{16}	G_{25}	G_{29}
$\frac{R_G}{e^{D_G}}$	n^{FI}	16	6	2
	n^{app}	16	6	2
$\frac{R_G}{{D_G}^2}$	n^{FI}	26	4	6
	n^{app}	26	2	4
$\frac{R_G}{D_G}$	n^{FI}	22	6	8
	n^{app}	19	2	6
$\frac{R_G}{\sqrt{D_G}}$	n^{FI}	18	16	15
	n^{app}	14	6	9
$\frac{R_G}{\log D_G}$	n^{FI}	11	25	23
	n^{app}	7	18	15

TABLE II
RANKING OF ENCROACHING GANG GROUPS UNDER DIFFERENT STRENGTH FUNCTIONS AND PROXY CHOICES

Strength function	Choice of R_G	G_{14}	G_{10}	G_{12}	G_{31}
$\frac{R_G}{e^{D_G}}$	n^{FI}	17	21	19	31
	n^{app}	17	21	19	31
$\frac{R_G}{{D_G}^2}$	n^{FI}	3	30	18	28
	n^{app}	8	31	19	28
$\frac{R_G}{D_G}$	n^{FI}	4	28	19	24
	n^{app}	15	30	22	18
$\frac{R_G}{\sqrt{D_G}}$	n^{FI}	7	13	21	12
	n^{app}	26	17	24	7
$\frac{R_G}{\log D_G}$	n^{FI}	15	2	17	3
	n^{app}	31	4	23	1

In particular, we are interested in what our model has to say about the relative strength/power of groups in a nesting or encroaching relationship, assuming these groups are in steady state.

We start by extracting from the first dataset (field interview cards) the number of interviewees belonging to group (gang) k, denoted by n_k^{FI}. We denote the estimated/approximate number of agents in group k by n_k^{es} and the area of group k in square feet by D_k^{sf}; both are from the second dataset, the territorial map.

These are then used as proxies for a group's resources and coverage, respectively. Specifically, we use a group's membership size (two options: n_k^{FI} or n_k^{es}) to approximate its resources (R_G); this is because we have no other information that would distinguish the members, so we simply ascribe one unit of resource to each. Similarly, we use a group's territorial area D_k^{sf} as its group coverage (D_G).

We consider a number of strength functions that are increasing in R_G and decreasing in D_G; these are given in the first column of Table I. The strength functions are generally of the form $f_{G_k} \propto \frac{n_k}{g(D_k^{\text{sf}})}$; here, $g(\cdot)$ is an increasing function ranging from exponential to logarithmic, and $n_k \in \{n_k^{\text{FI}}, n_k^{\text{es}}\}$ is one of the proxies for group size/resources.

For each combination, we then compute the strength of all groups; we order the groups in descending order of their computed strength. Table I shows the rankings of Groups 16, 25, and 29, the three groups in a nesting relationship. Table II shows the rankings of Groups 14 and 10, and Groups 12 and 31, the two pairs of groups in an encroaching relationship.

Our main observation here is that with the exception of the last strength function choice, where the cost of a group as a function of its area coverage grows logarithmically (the slowest among the functions considered), Groups 25 and 29 have higher strength (and thus utility) than Group 16.

Recall that our model implies that at an SISE (or an ISE under hedonic utility functions), if one group is nested within another, then the former must have higher group utility than the latter (Propositions 2 or 3). This is exactly what we are observing with all but the last function choice. There is empirical evidence to suggest that the utility in question is the group's "reputation," which corresponds to its "strength" or "power" [41], [46], [47]. Our model-based argument is that a much smaller gang that is completely encircled by a much larger gang must be able to hold its own (i.e., without disintegrating, disbanding, or weakening over time). The smaller gang needs to be stronger in reputation, finances, cohesion, and ideology, than the larger gang, or it would be absorbed or eliminated [40], [41].

Similar observations can be made on the encroaching groups: Group 14 has higher utility than Group 10 except for the last two functions, where the area coverage grows slower. This is again consistent with our model, which indicates that at equilibrium, if one group encroaches on another, the former must have higher group utility than the latter (Propositions 2 or 3).

The observation that a small group surrounded by a hostile larger group must be *stronger* may at first seem counter-intuitive. However, notice that it has parallels in history, such as Switzerland during World War II, and Israel post World War II. Both of these countries are relatively small, and are/were at the given time surrounded by hostile countries. The survival of Switzerland has been attributed to its military, democracy, and economic strength [48]. These are attributes obviously not captured by our model, but it is interesting to see that the same phenomenon emerges in a much simpler model.

The directionality of the other encroaching pair, Group 12 and Group 31, is not clear-cut from the map; this is also seen from the rank comparison. If we define a group's territory as the smallest rectangle or the convex hull containing all agents in the group, then Group 12 and Group 31 are effectively mutually encroaching. Recall that under our model, such a structure is possible only in an ISE with non-hedonic utility function. We unfortunately do not have sufficient ground-truth information to verify the actual utility function or the form of the equilibrium. However, these field observations motivate us to consider extensions to the model, discussed below.

VII. FURTHER EXTENSIONS OF THE MODEL

Our game model assumed that the utility of a group only depends on its own members and on groups with higher strength. We now consider a modification under which its utility depends on both stronger and weaker groups, and examine to what extent the results obtained so far continue to hold.

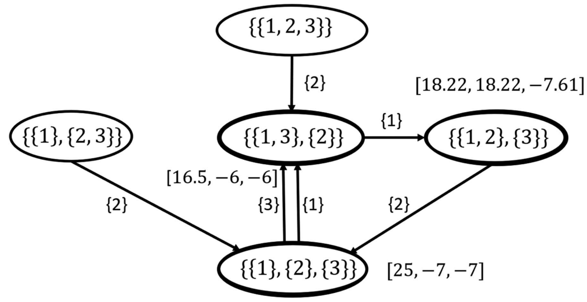


Fig. 6. State transition diagram of how three agents form groups. The utility vectors $[u_{[1]}, u_{[2]}, u_{[3]}]$ are shown in the graph for the three states where there is a cycle.

Our model departed from the more standard hedonic model by allowing a group's utility to depend not only on its own members, but also on stronger groups. Here, we consider what happens if we generalize the model further, and allow a group's utility to depend on the strengths of both stronger and weaker groups. Then, the utility function assumes the following general form:

$$U_G(\mathbf{r}, \mathbf{x}) = f(R_G, D_G) - \sum_{G': f(R_{G'}, D_{G'}) \neq f(R_G, D_G)} h(f(R_G, D_G), f(R_{G'}, D_{G'})). \tag{13}$$

Generalizing the assumptions previously placed on h for the case of dependence only on stronger groups, we now assume that $h(y, z)$ satisfies the following natural properties.

- $h(y, z)$ is monotone decreasing in y and increasing in z: a weaker impacted group or a stronger impacting group results in the impacted group suffering more harm.
- if $h(y, z) > h(y', z')$, then $h(z, y) < h(z', y')$: If a particular strength differential results in more harm to the impacted group, then it results in more benefit to the impacting group.
- $h(y, z) = 0$ if and only if $y = z$: two groups impact each other's utility if and only if their strengths are not the same.

In Appendix C, we show that the preceding conditions imply $h(y, z) > 0$ if and only if $y < z$, i.e., groups benefit from weaker groups and are harmed by stronger groups.

Under this class of utility functions, an ISE is no longer guaranteed to exist. In fact, below, we will show that this occurs for arguably the simplest non-trivial impact function h, namely $h(y, z) = z - y$.

Example 7.1: Consider three agents $\mathcal{N} = \{1, 2, 3\}$ on a line with locations $\mathbf{x} = [0, 0.04, 0.29]$ and resources $\mathbf{r} = [1, 9, 1]$. For notational convenience, we write f_G instead of $f(R_G, D_G)$. Consider the group utility function $U_G(\mathbf{r}, \mathbf{x}) = f_G + \sum_{G': G' \neq G}(f_G - f_{G'})$, where $f_G = \sum_{i \in G} r_i \,/\, e^{\max_{i,j \in G} \|x_i - x_j\|}$.

There are five different states in total, labeled the same way as in Example 4.1: $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1, 2\}, \{3\}\}$, $\{\{1\}, \{2, 3\}\}$, $\{\{1, 3\}, \{2\}\}$, $\{\{1, 2, 3\}\}$. The state transition diagram, driven by unilateral deviations, is given in Fig. 6 , where a directed edge label indicates the deviating agent(s). We see that regardless of the state, there always exists at least one agent who has a beneficial deviation and will be accepted by the destination group. In particular, states $\{\{1\}, \{2\}, \{3\}\}$, $\{\{1, 3\}, \{2\}\}$, and $\{\{1, 2\}, \{3\}\}$, form a cycle. As a result, there is no ISE in this example.

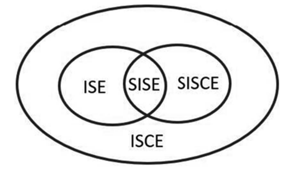


Fig. 7. Relationship among ISE, SISE, ISCE and SISCE.

The main driver of this behavior is that a powerful group under the given utility function would prefer as many weaker groups as possible to exist. In particular, agent 2 would not accept a singleton agent as a partner, but will "steal" agent 3 from agent 1, as this does not alter the number of weaker groups. Once 3 is partnered with 2, 2 can do even better by forming a singleton group again, leaving two weaker groups.

In the given example, if 3 could prevent 2 from leaving the group $\{2, 3\}$, a stable outcome would exist again. This suggests a different notion of equilibrium — which we term *individually stable contractual equilibria* — in which departures also need group approval. Similarly, a strong version of such equilibira can be defined. (Details can be found in Appendix B.) One can show, using ideas similar to those used in showing the existence of ISE and SISE (proofs can be found in Appendix C), that for utility functions of the form given in (13), a contractual equilibrium always exists. The relationship among these equilibria is summarized by a Venn diagram in Fig. 7.

Furthermore, mutually encroaching structures such as the one shown in Fig. 2(b) can become stable; such structures are indeed observed in the real-world Hollenbeck data, as discussed in Section VI (e.g., Groups 6 and 18, and Groups 4 and 7).

VIII. Conclusion and Future Work

We studied the existence and characteristics of different types of equilibria in a group formation game in which agents benefit from being part of a well-resourced and cohesive group. We primarily investigated two types of equilibria, ISE and SISE, in this game and explored their various structural properties. In particular, we showed that each SISE, or ISE under certain conditions, can be represented by a DAG, and conversely, that under mild conditions on the utility functions, every DAG can arise as an ISE of a suitably defined instance of the group formation game. We also connected our model to a real-world scenario that exhibits similar group structures to the ones the model predicts.

Inspired by observations of additional group structures in the real-world scenario, we introduced a modified group utility function in which groups are impacted by all other groups. In addition, we discussed further equilibrium notions, introducing contractual equilibria on top of ISE and SISE. While the assumption that D_G is a function of only X_G is natural, when it does not hold, some DAGs may not arise as equilibrium encroachment graphs; this would be a direction for future research. In addition, one can naturally extend the model to contain *hierarchical* structure, i.e., groups of groups, whereby an agent obtains utility from groups at each level that it belongs to. A deeper analysis of such models may shed interesting light on the emergence and stability of hierarchical organizations.

From an algorithmic viewpoint, it will be interesting to study what minimal changes a principal might perform on the locations or resources of agents in order to achieve the formation of a particular group structure.

Acknowledgment

The authors thank an anonymous reviewer for suggesting the real-world examples of Switzerland and Israel.

References

[1] T. Sandler and J. T. Tschirhart, "The economic theory of clubs: An evaluative survey," *J. Econ. Literature*, vol. 18, no. 4, pp. 1481–1521, 1980.

[2] J. Gravel, B. Allison, J. West-Fagan, M. McBride, and G. E. Tita, "Birds of a feather fight together: Status-enhancing violence, social distance and the emergence of homogenous gangs," *J. Quantitative Criminology*, vol. 34, no. 1, pp. 189–219, 2018.

[3] W. Saad, Z. Han, M. Debbah, and A. Hjorungnes, "A distributed coalition formation framework for fair user cooperation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4580–4593, Sep. 2009.

[4] Y. Zhang and M. Guizani, *Game Theory for Wireless Communications and Networking*. Boca Raton, FL, USA: CRC Press, 2011.

[5] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 77–97, Sep. 2009.

[6] L. Goette, D. Huffman, and S. Meier, "The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups," *Amer. Econ. Rev.*, vol. 96, no. 2, pp. 212–216, 2006.

[7] I. Milchtaich and E. Winter, "Stability and segregation in group formation," *Games Econ. Behav.*, vol. 38, no. 2, pp. 318–346, 2002.

[8] G. Hollard, "On the existence of a pure strategy nash equilibrium in group formation games," *Econ. Lett.*, vol. 66, no. 3, pp. 283–287, 2000.

[9] R. J. Aumann and J. H. Dreze, "Cooperative games with coalition structures," *Int. J. Game Theory*, vol. 3, pp. 217–237, 1974.

[10] H. Aziz and B. De Keijzer, "Complexity of coalition structure generation," 2011, *arXiv:1101.1007*.

[11] M. Slikker, "Coalition formation and potential games," *Games Econ. Behav.*, vol. 37, no. 2, pp. 436–448, 2001.

[12] A. Bogomolnaia and M. O. Jackson, "The stability of hedonic coalition structures," *Games Econ. Behav.*, vol. 38, no. 2, pp. 201–230, 2002.

[13] S. Banerjee, H. Konishi, and T. Sönmez, "Core in a simple coalition formation game," *Social Choice Welfare*, vol. 18, no. 1, pp. 135–153, 2001.

[14] D. Ray and R. Vohra, "Coalition formation," in *Handbook of Game Theory With Economic Applications*, vol. 4. Amsterdam, Netherlands: Elsevier, 2015, pp. 239–326.

[15] H. Aziz, F. Brandt, and H. G. Seedig, "Computing desirable partitions in additively separable hedonic games," *Artif. Intell.*, vol. 195, pp. 316–334, 2013.

[16] H. Aziz, F. Brandt, and H. G. Seedig, "Stable partitions in additively separable hedonic games," in *Proc. 10th Conf. Auton. Agents MultiAgent Syst.*, 2011, pp. 183–190.

[17] A. Igarashi and E. Elkind, "Hedonic games with graph-restricted communication," in *Proc. 15th Conf. Auton. Agents MultiAgent Syst.*, 2016, pp. 242–250.

[18] A. Igarashi, "Coalition formation in structured environments," in *Proc. 16th Conf. Auton. Agents MultiAgent Syst.*, 2017, pp. 1836–1837.

[19] M. Hoefer, D. Vaz, and L. Wagner, "Dynamics in matching and coalition formation games with structural constraints," *Artif. Intell.*, vol. 262, pp. 222–247, 2018.

[20] J. H. Dreze and J. Greenberg, "Hedonic coalitions: Optimality and stability," *Econometrica: J. Econometric Soc.*, vol. 48, pp. 987–1003, 1980.

[21] A. J. Collins, C. V. M. Cornelius, and J. A. Sokolowski, "Agent-based model of criminal gang formation," in *Proc. Agent-Directed Simul. Symp.*, 2017, pp. 1–10.

[22] A. J. Collins and E. Frydenlund, "Strategic group formation in agent-based simulation," *Simulation*, vol. 94, no. 3, pp. 179–193, 2018.

[23] T. Hiller, "Friends and enemies: A model of signed network formation," *Theor. Econ.*, vol. 12, no. 3, pp. 1057–1087, 2017.

[24] P. Cisneros-Velarde and F. Bullo, "Signed network formation games and clustering balance," *Dyn. Games Appl.*, vol. 10, no. 4, pp. 783–797, 2020.

[25] C. Wang, M. Moharrami, K. Jin, D. Kempe, P. J. Brantingham, and M. Liu, "Structural stability of a family of group formation games," in *Proc. IEEE 60th Conf. Decis. Control*, 2021, pp. 3080–3085.

[26] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks," *J. Econ. Theory*, vol. 71, no. 1, pp. 44–74, 1996.

[27] T. Ui, "A shapley value representation of potential games," *Games Econ. Behav.*, vol. 31, no. 1, pp. 121–135, 2000.

[28] T. Sandholm, K. Larson, M. Andersson, O. Shehory, and F. Tohmé, "Coalition structure generation with worst case guarantees," *Artif. Intell.*, vol. 111, no. 1/2, pp. 209–238, 1999.

[29] J. Farrell and S. Scotchmer, "Partnerships," *J. Quart. J. Econ.*, vol. 103, no. 2, pp. 279–297, 1988.

[30] B. Caskurlu and F. E. Kizilkaya, "On hedonic games with common ranking property," *Annu. Math. Artif. Intell.*, pp. 1–19, 2023.

[31] P. Dev, "Group identity in a network formation game with cost sharing," *J. Public Econ. Theory*, vol. 20, no. 3, pp. 390–415, 2018.

[32] F. Heider, "Attitudes and cognitive organization," *J. Psychol.*, vol. 21, no. 1, pp. 107–112, 1946.

[33] D. Cartwright and F. Harary, "Structural balance: A generalization of heider's theory," *Psychol. Rev.*, vol. 63, no. 5, pp. 277–293, 1956.

[34] J. A. Davis, "Clustering and structural balance in graphs," *Hum. Relations*, vol. 20, no. 2, pp. 181–187, 1967.

[35] J. MacQueen et al., "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probability*, 1967, pp. 281–297.

[36] K. Fukunaga and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 32–40, Jan. 1975.

[37] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[38] G. Tita, J. K. Riley, G. Ridgeway, A. F. Abrahamse, and P. Greenwood, *Reducing Gun Violence: Results From an Intervention in East Los Angeles*. Santa Monica, CA, USA: RAND Press, 2004.

[39] G. E. Tita and S. M. Radil, "Spatializing the social networks of gangs to explore patterns of violence," *J. Quantitative Criminology*, vol. 27, no. 4, pp. 521–545, 2011.

[40] P. J. Brantingham, G. E. Tita, M. B. Short, and S. E. Reid, "The ecology of gang territorial boundaries," *Criminology*, vol. 50, no. 3, pp. 851–885, 2012.

[41] P. J. Brantingham, M. Valasik, and G. E. Tita, "Competitive dominance, gang size and the directionality of gang violence," *Crime Sci.*, vol. 8, no. 1, 2019, Art. no. 7.

[42] Y. van Gennip et al., "Community detection using spectral clustering on sparse geosocial data," *SIAM J. Appl. Math.*, vol. 73, no. 1, pp. 67–83, 2013.

[43] K. Faust and G. E. Tita, "Social networks and crime: Pitfalls and promises for advancing the field," *Annu. Rev. Criminol.*, vol. 2, no. 1, pp. 99–122, 2009.

[44] M. A. Valasik, *Saving the World, one Neighborhood at a Time: The Role of Civil Gang Injunctions at Influencing Gang Behavior*. Irvine, CA, USA: University of California, Irvine, 2014.

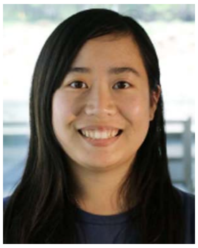
[45] L. M. Smith, A. L. Bertozzi, P. J. Brantingham, G. E. Tita, and M. Valasik, "Adaptation of an ecological territorial model to street gang spatial patterns in los angeles," *Discrete Continunous Dynamical Syst.*, vol. 32, no. 9, pp. 3223–3244, 2012.

[46] A. V. Papachristos, "Murder by structure: Dominance relations and the social structure of gang homicide," *Amer. J. Sociol.*, vol. 115, no. 1, pp. 74–128, 2009.

[47] S. H. Decker, "Collective and normative features of gang violence," *Justice Quart.*, vol. 13, pp. 243–264, 1996.

[48] S. P. Halbrook, *Target Switzerland: Swiss Armed Neutrality in World War II*. Cambridge, MA, USA: Da Capo Press, 2009.

[49] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge, U.K.: Cambridge Univ. Press, 2010.



Chenlan Wang (Graduate Student Member, IEEE) received the B.Eng. degree in electronic and information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2015, the M.S. degree in electrical and computer engineering in 2017 from the University of Michigan, Ann Arbor, MI, USA, where she is currently working toward the Ph.D. degree in electrical and computer engineering. Her research interests include game theory and social networks.

Mehrdad Moharrami received the B.Sc. degree in mathematics and electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2014, the M.Sc. degree in electrical engineering and the M.Sc. degree in mathematics from the University of Michigan, Ann Arbor, MI, USA, in 2017 and 2019, respectively, and the Ph.D. degree in electrical engineering in Winter 2020, for which he was awarded the Rackham Predoctoral Fellowship for an outstanding dissertation. He is an Assistant Professor in computer science with the University of Iowa, Iowa City, IA, USA. He was a TRIPODS Postdoctoral Research Fellow with the University of Illinois at Urbana Champaign, Champaign, IL, USA. His research interests include Markov decision processes, reinforcement learning, and random graph models for economics, learning, and computation.

Kun Jin received the M.Sc. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, MI, USA, in 2020, and the B.Sc. degrees in electronics and information science, and economics from Peking University, Beijing, China, in 2017. He is currently a Ph.D. student with the Department of Electrical and Computer Engineering, University of Michigan. His research interests include algorithmic game theory, algorithmic mechanism design, and AI FATE.

David Kempe received the Ph.D. degree in computer science from Cornell University, Ithaca, NY, USA, in 2003. He has been with the Faculty in the Computer Science Department, USC since the Fall of 2004, where he is currently a Professor. His research interests include computer science theory and the design and analysis of algorithms, with a particular emphasis on social networks, algorithms related to online and other learning models, and game-theoretic and pricing questions. He was the recipient of the NSF CAREER Award, VSoE Junior Research Award, ONR Young Investigator Award, a Sloan Fellowship, and an Okawa Fellowship, in addition to several USC mentoring awards and Best Paper awards.

Paul Jeffrey Brantingham received the Ph.D. degree in anthropology from the University of Arizona, Tucson, AZ, USA, in 1999. He is currently a Professor with the Department of Anthropology, University of California, Los Angeles, Los Angeles, CA, USA. His research interests include modeling and measuring the fundamental dynamics of crime patterns and the role of neutral processes in human cultural evolution.

Mingyan Liu (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland, College Park, College Park, MD, USA, in 2000. She is currently a Professor and the Peter and Evelyn Fuss Chair of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI, USA. Her research interests include sequential decision and learning theory, game theory and incentive mechanisms, with applications to large-scale networked systems, cybersecurity, and cyber risk quantification. She is a Member of the ACM.