Social Clubs and Social Networks

Chaim Fershtman[1] Dotan Persitz[1]

[1]Coller School of Management
Tel Aviv University

September 13, 2025

Social Environments

- In modern life most social contacts are formed within social context.
- Sociologists refer to social contexts as social foci - “Social, psychological, legal, or physical entities around which joint activities are organized.” (Feld (1981)).
- Examples - family, gym, neighborhood, alumni, department, conference, interest group, workplace, scouts, army unit, synagogues, churches.
- Within the club, agents may connect with one another and form their social ties.
- Some literature: Simmel (1908/1955), Young and Larson (1965a,b), Kadushin (1966, 2011), Feld (1981), Granovetter (1983), Blau and Schwartz (1984), Rivera et al. (2010) and Jackson et al. (2012).

Social Environments and Social Networks

- Usually the structural formation of society is described as a sequential process - first social clubs are formed and then the social network emerges.
- Rivera et al. (2010, p. 106): “If networks are the fabric of inter-personal interaction, social foci are the looms in which they are woven”.
  - Stochastic non-strategic network formation models (Snijders et al. (2006) and Chandrasekhar and Jackson (2018)).
- Exceptions are Feld (1981) and McPherson et al. (2001) that point out a feedback from network contacts to (less costly) formation of new clubs.
- We suggest a model of **simultaneous** formation of the social network and the social clubs’ system.

Clubs are Linking Platforms

- To make our point as clear as possible we focus on clubs as linking platforms.
- One aspect of clubs we ignore: direct benefits (e.g. health effects of training in a gym). These benefits are the focus of the well-established literature on club theory.
- Another aspect of clubs we ignore: clubs and coalitions frequently serve as a binding agreement that constrains members' activities:
  - Myerson (1980), Slikker and Van den Nouweland (2001) and Arnold and Wooders (2015) discuss cooperative games with conference structure.
  - Caulier et al. (2013a,b, 2015) discuss coalitional networks.

The General Framework

- The Common Strategic Social Network Formation Setup:
  - Agents need to decide whether they form links with other agents .
  - Each link is costly.
  - Benefits from direct links as well as indirect links.
- The Social Clubs framework
  - Agents choose clubs.
    - Membership is a decision of the candidate only (we study other rules as well).
    - Membership is costly.
    - An agent may have multiple memberships.
  - Two agents that share a club are connected in the induced network by a weighted link.
  - The weight is determined by a function that depends on the characteristics of the two agents and the environment.
  - Indirect connections (through third parties) are endogenously depreciated.
  - The agents benefit from their position in the induced weighted network.

Why Should You Care?

- We believe that this framework captures a realistic characteristic of socialization that is missing from the current strategic network formation models.
- We will try to convince you that the model is operational.
- We will have novel insights on several important issues in the social networks literature:
  - The interpretation of weak ties
  - The prevalence of complete networks.
  - Observations of “too” high clustering.
  - The effect of club rules on social dynamics.

An Environment

- $N = \{1, ..., n_a\}$ is a finite set of identical agents.
- $S = \{1, ..., n_s\}$ is a finite set of clubs.
- The pair $\{i, s\}$ implies that individual i is affiliated with club s.
- $A^c \equiv \{\{i, s\} : i \in N, s \in S\}$ is the set of all possible affiliations.
- An environment is a triplet $G \equiv < N, S, A >$ where $A \subseteq A^c$.
- Additional Notation:
  - Individual i's affiliations: $S_G(i)$ ($s_G(i) \equiv |S_G(i)|$).
  - Club s's members: $N_G(s)$ ($n_G(s) \equiv |N_G(s)|$).
  - Additional affiliation: $G + \{i, s\}$.
  - Discarded affiliation: $G - \{i, s\}$.
  - Additional club: $G + m \equiv$ where $m \subseteq N$ and s is a vacant club.

The Induced Weighted Network

- We are interested in the induced network where the nodes are the agents.
- Two agents are linked if and only if they share a club.
- The quality of a link may depend on various characteristics of the environment.
- Hence, we denote the weight of a link between two agents $i, i' \in N$ in G by $w(i, i', G) \in [0, 1]$.
- We denote the weighted network induced by Environment G and Weighting Function w by $g = < N, E_G, W_{G,w} >$.

Distance and Indirect Connections

- A path between i and i' in g is a non-empty subgraph p of g where the nodes are $\{x_1, x_2, x_3, \ldots, x_{l-1}, x_l\}$ (all distinct) and the edges are $\{x_1 x_2, x_2 x_3, \ldots, x_{l-1} x_l\}$, $x_1 = i$ and $x_l = i'$.

Definition (The Weight of a Path)

The weight of path $p = \{x_1, \ldots, x_l\}$ in the induced weighted network g is $WP_g(p) = \prod_{k=1}^{l-1} W(\{x_k, x_{k+1}\})$.

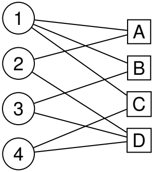
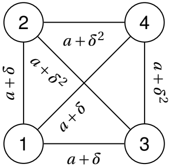
- Path p is a *shortest weighted path* between agents i and i' if there is no path p' between agents i and i' such that $WP_g(p') > WP_g(p)$.
  - The shortest path between two agents may be indirect even if they share a club in G (impossible in most network formation models).
- $d(i, i'|G, w)$ denotes the weight of a shortest path between agents i and i' in the induced weighted network g.
- $d(i, i'|G, w) = 0$ if there is no path between agents i and i'.

Preferences

- The agent benefits from her position in the network.
- Let $c \geq 0$ denote the homogeneous participation fee.
- The utility of Agent i from the Environment G and the weighting function w is:

$$u_i(G, w, c) = \sum_{k \in N, k \neq i} d(i, k | G, w) - s_G(i) \times c$$

Example

Environment	List of Clubs	Agents' Network	Utilities
	Club A: 1 2 Club B: 1 3 Club C: 1 4 Club D: 2 3 4		$u_1 = 3(a+\delta) - 3c$ $\forall i \in \{2,3,4\}$: **if** $\delta \geq \frac{1-a}{2}$: $u_i = (a+\delta) + 2(a+\delta)^2 - 2c$ **Otherwise:** $u_i = (a+\delta) + 2(a+\delta^2) - 2c$

The weighting function: $w(i, i', G)$ equals $a + \delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$.

Open Clubwise Stability

Open Clubwise Stability

An environment G is Open Clubwise Stable (OCS) for weighting function w and membership fees c if:

$$\forall s \in S, \forall i \in N_G(s) : u_i(G, w, c) \geq u_i(G - \{i, s\}, w, c) \qquad \text{(No Leaving)}$$

$$\forall s \in S, \forall i \notin N_G(s) : u_i(G, w, c) \geq u_i(G + \{i, s\}, w, c) \qquad \text{(No Joining)}$$

$$\forall m \subseteq N : \exists i \in m : u_i(G + m, w, c) > u_i(G, w, c) \Rightarrow \qquad \text{(No New Club)}$$

$$\exists j \in m : u_j(G + m, w, c) < u_j(G, w, c)$$

An environment G is ***Efficient*** if there is no other environment G' such that $\sum_{i \in N} u_i(G', w, c) > \sum_{i \in N} u_i(G, w, c)$.

The Club Congestion Model

McPherson and Smith-Lovin (1982) p. 884

One aspect of voluntary associations which is particularly crucial for the network of informal relations is the size of a given organization. Large organizations generate more potential acquaintances. One could argue that the contacts which occur in a larger organization are more "superficial" than those in smaller organizations.

Definition (The Club Congestion Model)

A club congestion function is a non-increasing function $h:\{2,3,\dots,n_a\}\to[0,1]$.

Given a club congestion function h, the weight of a link between two agents $i,i'\in N$ is $w_h(i,i',G)=\max_{s\in S_G(i)\cap S_G(i')} h(n_G(s))$.

Externalities

- In the baseline model, unilateral actions, such as leaving or joining a club, may "shorten" or "break" shortest paths of other agents.
- Incorporating congestion introduces a new type of externality: Unilateral actions may also affect the quality of some links.
- For example, if Agent j joins a club with which Agent i is also affiliated, the quality of some links that Agent i maintains may change - either by making some paths shorter or by reducing the weight of some links due to stronger congestion.
- Therefore, unilateral actions may have positive or negative externalities.
- But, forming a new club never imposes negative externalities.
- Therefore, if G is efficient then it satisfies the condition of "No New Club Formation".

Two Useful Environments

Let $S_p = \{s | n_G(s) > 0\}$ denote the set of populated clubs.

Definition (m-complete)

G is an m-complete environment ($m \in \mathbb{N}$, $n_a \geq m \geq 2$) if:

$$\forall i, i' \in N : |S_G(i) \cap S_G(i')| = 1.$$

$$\forall s \in S_p : n_G(s) = m$$

Definition (m-star)

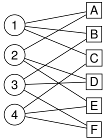
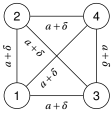
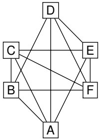
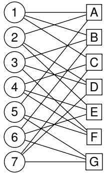
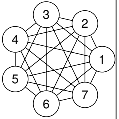
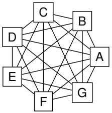
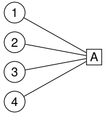
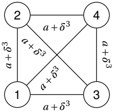
G is an m-star environment ($m \in \mathbb{N}$, $n_a \geq m \geq 2$) if:

$$\forall s \in S_p : n_G(s) = m$$

$$\exists i \in N \quad \text{such that} \quad \forall s', s'' \in S_p : N_G(s') \cap N_G(s'') = \{i\}.$$

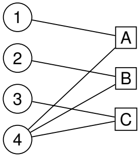
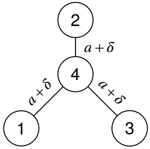
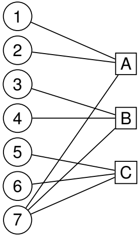
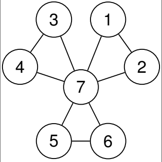
$$\forall j \in N \backslash \{i\} : s_G(j) = 1.$$

m-complete

	Environment	List of Clubs	Network	Clubs' Network	Utilities
All Paired or 2-Complete (n=4)		Club A: 1 2 Club B: 1 3 Club C: 1 4 Club D: 2 3 Club E: 2 4 Club F: 3 4			$\forall i \in \{1,2,3,4\}$: $u_i = 3(a+\delta) - 3c$
3-Complete (n=7)		Club A: 1 2 5 Club B: 1 3 6 Club C: 1 4 7 Club D: 2 3 7 Club E: 2 4 6 Club F: 3 4 5 Club G: 5 6 7	The weights are all $a+\delta^2$		$\forall i \in \{1,\ldots,7\}$: $u_i = 6(a+\delta^2) - 3c$
Grand Club or 4-Complete (n=4)		Club A: 1 2 3 4		A	$\forall i \in \{1,2,3,4\}$: $u_i = 3(a+\delta^3) - c$

The weighting function: $w(i,i',G)$ equals $a+\delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$.

m-star

	Environment	List of Clubs	Network	Utilities
2-Star (n=4)		Club A: 1 4 Club B: 2 4 Club C: 3 4		$\forall i \in \{1,2,3\}$: $u_i = (a+\delta)+2(a+\delta)^2 - c$ $u_4 = 3(a+\delta) - 3c$
3-Star (n=7)		Club A: 1 2 7 Club B: 3 4 7 Club C: 5 6 7	 The weights are all $a+\delta^2$	$\forall i \in \{1,\dots,6\}$: $u_i = 2(a+\delta^2)+4(a+\delta^2)^2 - c$ $u_7 = 6(a+\delta^2) - 3c$

The weighting function: $w(i,i',G)$ equals $a+\delta^{m-1}$ where m is the size of the smallest club i and i' share, $\delta \in (0,1)$, $a \in [0,1)$ and $a+\delta \in (0,1)$.

Efficiency

- Efficiency is hard.
- We characterize the efficient networks among those with homogeneous club size.

Definition (m-uniform environment)

G is an **m-uniform environment** ($m \in \{2, \ldots, n_a\}$) if $\forall s \in S_p : n_G(s) = m$.

- Denote the set of all m-Uniform environments with n agents by $\mathcal{G}_n^m$.

Efficiency Result

Efficient m-Uniform environments

Let $m \in \{2, \ldots, n_a\}$. For every club congestion function $h(\cdot)$ and m-Uniform Environment $G' \in \mathcal{G}^m_{n_a}$:

1. Let $c \in [0, (m-1)(h(m) - h^2(m)))$ and let G be an m-Complete Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.
2. Let $c \in ((m-1)[h(m) - h^2(m)], (m-1)h(m) + \frac{(n_a-m)(m-1)}{m} h^2(m)]$ and let G be an m-Star Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.
3. Let $c \geq (m-1)h(m) + \frac{(n_a-m)(m-1)}{m} h^2(m)$ and let G be the Empty Environment. Then, $\sum_{i=1}^{n_a} u_i(G, h, c) \geq \sum_{i=1}^{n_a} u_i(G', h, c)$.

From Efficiency to Stability

- Note that the result is independent of the specific club congestion function.
- Motivation to study when are those architectures stable.
- We did.
- But today I will show only the implications.

Very Low Membership Fees

- Recall: Spanning super environments of the All Paired environment are the only OCS when $c = 0$.
- Stability of All Paired Environments: the only sensible deviation is replacing a club membership with an indirect connection.
- The All Paired environment is the unique OCS if and only if

$$\min\left\{h(2) - h^2(2), h(2) - h(3)\right\} \geq c > 0$$

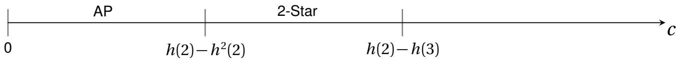
- In this range the All Paired Environment is also the unique efficient environment.

Weak Ties

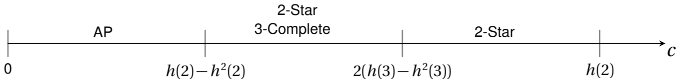
- Since the seminal Granovetter (1973) the concept of “weak ties” has become central to social networks literature.
- Weak ties are usually interpreted as direct links:
  - Infrequently used contacts (mainly in the literature on the role of networks in labor markets).
  - Carries low weight (mainly in the literature on the strategic formation of weighted networks and its various applications).
- Indirect connections which are frequently highlighted in the strategic network formation literature, are not referred to as weak ties.
- We consider both direct connections via a congested large club and depreciated indirect connections through memberships in small clubs as weak ties.

The Emergence of Weak Ties

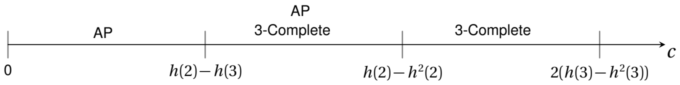
Congestion is the stronger friction - $h(3) < h^2(2)$:



Depreciation and congestion are equal - $h(3) = h^2(2)$:



Depreciation is the stronger friction - $h(3) > h^2(2)$:



Cliques

- The standard strategic theory of network formation restricts clubs to be of size 2.
  - Therefore, the main trade-off is costly direct connections versus depreciated indirect connections.
  - In particular, cliques can occur only if the linking costs are low.
- We endogenize the club size.
  - Another force should be taken into account - congestion.
  - For example, cliques can occur in high costs environment if congestion is less of a friction compared to indirect connection depreciation.

The Individual Congestion Model

- The observation: Club affiliations require attention and time.
- Agents with a thin portfolio of affiliations are able to pay attention to each of their memberships and to form high quality connections with other members.
- Agents who are members of many clubs, possess many weak direct relations since they devote little attention to each of their memberships.

Definition (The Individual Congestion Model)

An individual congestion function is a non-increasing function $b : \{1, 2, \ldots, n_s\} \to [0, 1]$.
Given an individual congestion function b, the weight of a link between two agents $i, i' \in N$ in Environment G is,

$$w_b(i, i', G) = b(s_G(i)) \times b(s_G(i'))$$

Alternative Club Rules

- Very specific club rules are implicit in the model.
- Open Clubwise Stability assumes that membership is open for all agents (as long as the participation fees are paid).
- Alternative club acceptance rules:
  - Membership requires approval of (a subset of) the existing club members.
  - Membership quotas.
  - Membership criteria.
  - Exclusivity rules.
- Other implicit specifications - leaving rules, rules for forming new clubs, coordination (within and between agents).
- Assumption: all clubs in a specific environment have the same, exogenously given, rules.

Closed Clubwise Stability

Closed Clubwise Stability

An Environment G is Closed Clubwise Stable if the following conditions hold:

1. No Leaving:
   $\forall s \in S, \forall i \in N_G(s) : u_i(G, w, c) \geq u_i(G - \{i, s\}, w, c)$.
2. No New Club Formation:
   $\forall m \subseteq N : \exists i \in m : u_i(G + m, w, c) > u_i(G, w, c) \Rightarrow$
   $\exists j \in m : u_j(G + m, w, c) < u_j(G, w, c)$.
3. No Joining: $\forall s \in S, \forall i \notin N_G(s) :$
   $u_i(G, w, c) \geq u_i(G + \{i, s\}, w, c) \quad OR$
   $\exists j \in N_G(s) : u_j(G, w, c) > u_j(G + \{i, s\}, w, c)$.

- OCS implies CCS.
- These two rules induce different dynamics.

Clustering

- Real life networks are highly clustered: The probability of two individuals who share a common neighbor to be connected is much higher than expected if connections were random.
- The literature attributes the high clustering to one of two explanations:
  - “Preference for transitivity”: Attraction is based on the “network” properties of the individuals.
  - “Homophily”: Attraction is based on “non-network” properties of the individuals.
- Recent literature attempts to provide econometric tools for estimating network formation models that incorporate these explanations. Mainly concerned with homophily on unobservables (e.g. Goldsmith-Pinkham and Imbens (2013), Mele (2017), Graham (2015, 2016))
- Another concern is that neglecting to account for self-selection into social contexts leads to an over-estimation of the importance of these factors (Rivera et al. (2010) and Miyauchi (2016)).

Our Model and Clustering

- Our setting provides a third explanation.
- In our framework, a network must exhibit high clustering since the individual's neighbors form a tightly knit group.
- we propose clubs as linking platforms rather than individuals' linking preferences as the fundamental that drives high clustering.

Take Home

- Novel framework for strategic network formation of undirected weighted networks.
- Agents choose affiliations and benefit from their position in the underlying network.
- The framework provides insights that are absent from link formation models.

Thanks


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
DANKSCHEEN
BIYAN
SHUKRIA
TINGKI
YAQHANYELAY
TASHAKKUR ATU
SUKSAMA
EKHMET
GRAZIE
MEHRBANI
MAAKE
KOMAPSUMNIDA
PALDIES
GOZAIMASHITA
EFCHARISTO
JUSPAXAR
BOLZİN
MERCI