



Peter V. Marsden "Recent Developments  
in Network Measurements" in *Models and  
Methods in Social Network Analysis*

---

第7回社会ネットワーク分析輪読会  
2005年7月13日  
成蹊大学専任講師 山本晶



# 自己紹介

---

- 研究領域
  - 購買における対人影響、クチコミ、目利き、ヒット予測...
- 現在のテーマと悩み
  - 他の消費者の購買に影響を与える「インフルエンサー」の特定
    - アーカイブデータ(行動履歴データ)と質問調査データ
    - 質問調査票の表現？
    - データ収集？ サンプリング？ 第三章
    - 再現性？ 一般化？ 検定？



# 本の紹介と第二章の概略

---

- 書籍について

- Wasserman and Faust(1994)をアップデート
- More in-depth, up-to-date, more advanced

- 第二章について

- Marsden (1990) “Network Data and Measurement”の続編
  - 安田先生2003年講義@東大経済
- テーマはデータ収集に関する諸問題



# 2003年 「経営学文献購読」 社会ネットワーク分析文献リスト

- [5月7日] Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, 16, 435-463.
- [5月14日] Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.
- [5月21日] Burt, R. S. (1987). Social contagion and innovation, cohesion versus structural equivalence. *American Journal of Sociology*, 92, 1287-1335.
- [5月28日] Granovetter, M. S. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91, 481-510.
- [6月4日] Gulati, R. (1998). Alliances and networks. *Strategic Management Journal*, 19, 293-317.
- [6月11日] Gulati, R. (1999). Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strategic Management Journal*, 20, 397-420.
- [6月18日] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48, 147-160.
- [6月25日] Baker, W. E. (1990). Market networks and corporate behavior. *American Journal of Sociology*, 96, 589-625.
- [7月2日] Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42, 35-67.
- [7月9日] Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, 61, 674-698.
- [7月16日] Burt, R. S. (1988). The stability of American markets. *American Journal of Sociology*, 94, 356-395.
- [7月23日] White, H. (1981). Where do markets come from? *American Journal of Sociology*, 87, 517-547.
- [注意] 7月23日に予定されていた下記の論文は、授業の進捗状況と、参加者の興味と関心領域に配慮して、リストから外しました。  
Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly*, 45, 425-455.



# ネットワーク分析の設計

---

- Whole network (socio-centric network):
  - ネットワークの境界内の全行為者と紐帯が対象
- Ego-centric network:
  - 中心となる行為者 (ego) とそれを取り巻く行為者 (alter)、それらを結ぶ紐帯
  - しばしば母集団からサンプリングされる
- One mode data set:
  - 対象が一種類。Ex. 行為者間の関係
- Two-mode data set:
  - 対象が二種類。Ex. 行為者とイベントの関係



# ネットワーク分析の設計(続)

---

- Cognitive Social Structure (CSS) design
  - 複数の情報源から関係を測定
  - ネットワークに関する情報提供者間の認識の違いを把握できる
- Ego centricとwhole networkの関係
  - 前者は後者の一部である
  - “dense”に抽出すれば前者は後者に近づく



# ネットワークの境界

---

- どうやって決めるか？
  - Positional approach:
    - 分析対象の特徴。社員かどうか
  - Event based approach: 特定の活動への参加
    - 浜辺に期間内に3日以上現れたかどうか
  - Relational approach: social connectedness
    - “Expanding selection”. Fixed listからはじめて、Snowball samplingの結果ネットワーク内の行為者と複数のリンクが確認された場合のみ分析対象として追加
      - Fixed list: 分析者のネットワークに対する事前知識を利用して境界を決定
      - Expanding selection: 対象者のネットワークに対する知識を利用して境界を決定
    - Doreian and Woodard (1994)
      - 大量データから”k-core”基準を満たす行為者のみ抽出
      - Kの値は分析者が目的に応じて設定



# サーベイ調査

---

- アーカイブデータが無い場合、現実的なデータ収集方法
  - 日記調査などより軽負担
  - Self-reportであることに留意
- Whole networkの場合
  - 名簿作成 「想起」でなく「認知」が可能
  - ネットワーク内の全員にインタビュー。各々の関係について回答してもらう。
- Ego-centric networkの場合
  - 事前にalterは把握できない。ネットワークの境界は回答者の「想起」に依存
  - Alter間の関係や親密さについても回答者に質問





# サーベイ調査(続)

---

- ネットワークデータの取得方法
  - 二項判断 (sociometric choicesとも呼ばれる)
    - 閾値の問題あり
  - 紐帯の強さに関する順序評価(ordinal rating)
  - 順位付け(ranking)
    - Ratingとrankingには相関あり



# Ego-centric networkの調査方法

---

- Name generators
  - 回答者のalterを特定する質問項目
  - 自由想起によって得られ、ネットワークの境界を決定
    - 質問によってネットワークサイズの影響を受ける
  - GSS (General Social Survey) “discuss important matters?”
  - Single name generator: “core network”を抽出
  - Multiple name generator: 広義のサポートネットワークを抽出可能。しかし回答者負担大
    - 少なくともintimacyとactivityは測定するべきである (Burt 1997)
  - データ収集メディアの違い(紙・インタビュー・コンピュータ)
  - 妥当性テスト テスト・再テストによる確認
- Name interpreters
  - Ego-alter, alter-alterの関係に関するデータを得る質問項目



# 想起・認知・忘却

---

- Brewer and Webster (1999)、Brewer (2000)
  - 回答者はbest friend, close friend, other friendsを自由想起。つぎに名簿を見てリストに追加
  - 全体で1/5が忘却。親密さと忘却は反比例
  - 可能な場合は「想起」に加えて「認知」も含める
  - Multiple name generatorsによって忘却を防止
- 想起パターン
  - 社会グループ毎(家族、同級生、隣人、職場)にalterを認知
  - 回答者に社会グループに関するヒントを与えるとより完全なネットワークが得られる
  - 紐帯の強い順に想起



# Test-Retest Studies

---

- 再テストによって妥当性を確認
  - 再テストは1ヶ月以内
  - 多くの研究で75%のalterが再テストでも確認された
- ネットワーク規模の安定性
  - Rapkin and Stein (1989) 再テストでは規模縮小
  - Morgan, Neal, and Carder (1997) 2ヶ月ごとに7回調査
    - Core network(主に家族)は安定的
    - 周辺は不安定。24%のalterが一回のみ確認
    - ネットワークの特性はalterよりも安定的



# ネームジェネレータの意味と解釈

---

- GSS “discuss important matters”
  - Important matterはひとによって違う
    - 個人的な人間関係、健康、仕事、政治
  - 親密さ？ 接触頻度？ 役割？
  - “important matters” = “significant people”  
(Straits 2000)



# 文脈効果・インタビュアー効果

---

- 文脈効果
  - 直前の質問によってName generatorsの解釈が変わる
- インタビュアー効果
  - インタビュアーのトレーニング
  - コンピュータの利用



# ネームインタープリター

---

- 距離のあるalterに関する記憶は少ない
- Alterに関する質問は回答者(ego)にとって面倒で退屈
  - Alterの数に制限を置いたほうが(3 - 5人)ネットワーク密度などの指標が正確に測定できる



# ネットワーク規模の測定

---

- Killworth et al (1998)の”Scale up” method
  - $m/c=e/t$  を利用して $c$ を推定
    - $m$ : ego-centric network内のsubpopulation
    - $c$ : network size
    - $e$ : subpopulationの規模
    - $t$ : population
    - Subpopulationの例: Michaelという名前の人
  - Active network (相互認知、接触2年以内) の平均規模は108人 (フロリダ)、286人 (米国調査)
  - 日本人の平均規模
    - 129.8人 (電通総研、2003年)、272.84 ~ 329.35人 (辻ほか)
- Reverse small world (RSW) method
  - 500人のターゲットを提示 (職業、居住地など)
  - Alterがターゲットを知っているかどうかを質問
  - 平均規模は129人 (Jacksonville, Fl), 77人 (Mexico city)

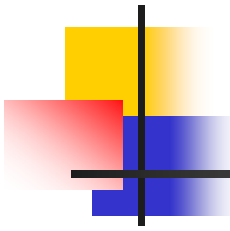




# ポジションジェネレーター

---

- (1) Name generatorで挙げたalterについて(2) name interpreterで掘り下げるのではなく、position generatorで直接聞く
  - 例： 次の15の職業についている友人、親戚、知人はいますか(Lin, Fu, and Hsung 2001)
  - Follow upでcontactの強さなど測定
  - 個々のalterを挙げさせるのではないため、調査時間がかからない
- Smith (2002) 異人種間の友人関係を調査
  - Position generator (W 42%, B 62%)
  - Global approach (W 24%, B 45%)
    - 友人の数を聞く。つぎにそのなかで異人種の友人数を聞く
  - Name generator (W 6%, B 15%)
    - 友人の名前を挙げてもらう。その友人の人種を聞く



# リソースジェネレーター

---

- Van der Gaag and Snijders (2004)が提案
  - 個人別社会資本を測定
  - 特定の能力のある知人を持っているかどうか
    - 自動車修理、文芸、高所得
  - Follow upで紐帯の数、質、強さを質問



# CSSデータ

---

- Krackhardt (1987) ネットワーク内のダイアド関係についての判断データ
  - 単一の観察者の判断。network内の行為者であることが多い
  - ダイアドを構成する二人の行為者によるlocally aggregated structure of judgment
  - ダイアドについての全員の判断consensus structure
- 質問調査票によりデータ収集
  - 紐帯のチェックリストを提示(二値)
  - 行列を提示。ペアを識別してもらう
  - Outgoing tieの強さの閾値について質問
  - ネットワーク内で親密な相手をランク付けしてもらう
- 事例
  - 21人の職場の同僚の調査。400の紐帯について、friendshipかadviceかの二値判断をしてもらう(Krackhardt 1987)
  - 二次元空間に行為者を配置してもらうFreeman (1994)
  - 二値、trichotomous, rankingで結果に変化なし(Batchelder 2002)
    - 二値で十分



# 情報提供者バイアス

- Kumbasar, Romney, and Batchelder (1994)
  - CSSの個人判断と全体のコンセンサスを比較
  - 個人の判断結果のほうが個人が中心的な位置と認識
  - 半数以上が次数中心性で自分が上位と考える "ego bias"
  - 情報提供者と隣接するalterと、情報提供者が含まれない紐帯を比較
    - 隣接するalterとの関係のほうが密度、互惠性、推移性ともに高い  
情報提供者の周辺で局地的密度が高くなる
    - 局地的バランスへの認知的プレッシャー
  - 情報提供者は「グループ」や「バランス」のスキーマによって関係を「創造」したり「みすごし」たりする
- Krackhardt and Kilduff (1999)
  - 近距離のalter間と遠距離のalter間で互惠性、推移性が高い
  - 遠距離の場合、バランススキーマを使用して記憶を補完
  - Status schema。ステータスの高い人が情報提供者の記憶で重要な位置を占める。ネットワーク内で突出した地位を占めるとして報告される



# 情報提供者の正確性と能力

---

- Bernard, Killworth, and Sailer (BKS;1981)
  - 認知データと行動データの乖離を指摘
    - Stated preference (アンケートデータ)とrevealed preference (POS data)
  - 自己申告の限界
  - 多くの後続研究を排出
- Freeman, Romney, and Freeman (1987)
  - 長期的な関係については正確な報告



# 観察データと認知データの対応

- Kashy and Kenny (1990)
  - 入次数が高い(認知される回数が多い)行為者は、観察によっても高い相互作用が確認された
  - 出次数の高さと観察された相互作用の対応は低かった
    - 乖離の原因は認知の際の閾値にある
- Freeman and Webster (1994)
  - 観察データと認知データには対応あり
  - ただし、認知データのほうが単純な構造。クラスタ数が少ない
  - グループスキーマの介在により、ダイアド関係ではなくカテゴリ所属として情報が蓄積される
- 認知データはネットワークの測定に有用である
- 社会的紐帯の観察自体が困難である



# 情報提供者の能力に関する調査

---

- 信頼できる情報提供者は、「正確である (accurate)」 (Romney and Weller 1984)
  - 認知データが他の情報提供者と類似しているような情報提供者のデータは、集計された観察データと合致する
- Romney, Weller and Batchelder (1986)
  - 情報提供者のCSS dataが局所構造やコンセンサス構造と合致する = 「能力がある (competent)」
  - 中心性の高い情報提供者のほうが能力が高い
  - 近接性が能力の源泉である (Bondonio 1998)



# 情報提供者の活用方法

---

- 情報提供者全員に自己申告させるかわりに、情報提供者の数人にCSS的な調査を実施することによってwhole networkを特定
  - 能力の高い情報提供者の認知データとコンセンサス構造の間に対応がある場合有効
  - 情報提供者のスクリーニングには中心性などの指標が有効





# アーカイブデータ

---

- 近年活発に利用されている
  - 墓碑広告の位置を利用して投資銀行のステータスを測定
  - 特許の引用データからイノベーションの評価指標を開発
  - キケロの手紙からローマ社会を分析
  - 論文の引用から研究領域の構造を分析
- アーカイブデータの妥当性
  - 測定された紐帯と研究テーマの紐帯の対応に依存
  - 論文の引用は敬意を表するため、反論のため、手法を特定するためなど多様な紐帯の可能性あり
  - Editorial policyの介在
  - ウェブサイトのリンク構造。リンク基準は多様
  - 名称変更、略称
  - アーカイブデータから得られるネットワークは媒体に大きく依存
    - E-mailの送受信関係から得られる人間関係は全人間関係ではない



# 観察データ

---

- 初期においては重要な情報源。近年は利用した研究が少ない
  - 測定困難
    - ビデオやテープ 人間の特定困難
    - 対象が大人数 多数の観察者が必要
    - 小型マイクで録音し、音声認識を利用してダイアド間のコミュニケーションを測定



# 結論

---

- 1990年代以降、質問法によるネットワークデータの収集が発展
  - Ego-centric networkの測定に関するより深い理解が深まる
  - 認知プロセスとバイアスに関する理解が深まる
  - 質問調査の妥当性・信頼性の課題は残る
  - アーカイブ・データの有効利用と質の確保