Dynamic Ego Networks

David Hachen Omar Lizardo
Michael Penta Brandon Sepulvado
Joseph Workman

University of Notre Dame - iCeNSA

NetSci'14 Symposium: Temporal Networks, Human Dynamics and Social Physics, June 3, 2014

iCeNSA
Interdisciplinary Center for Network Science & Applications

Overview

Thinking about Social Networks Dynamically

- Evolution of a person's ego network
  - Size, clustering coefficient, composition
- Evolution of a social ties
  - formation, dissolution, strengthening, reciprocation
- Focus in this talk is on ego network size (degree)

Ego Network Size (Degree)

- important status marker (popular people)
- correlated with other measures of social status
- in organization settings indicative of centrality
- in diffusion studies critical is spreading of ideas, diseases, information
- Central questions
  - How stable is a person's degree?
  - How much between subject variability (inequality) is there in degree?
  - What factors explain the between subject and the within subject (over time) variability?
    - Preferential attachment
    - exogenous person-specific factors
  - How can we measure degree dynamically?

Dynamic Measures of Degree

- Tie formation and tie decay must be taken into account (Burt 2000)
- Recent paper by Miritello et al. (2013) shows how to use temporally fine-grained dynamic network data to measure various aspects of ego network size
- Construction of degree centrality indicators become dependent on the temporal window within which individuals are observed
- Each individual is observed on a sequential number of time units up until the last observation point $\{t_0, t_1, t_2, t_3 ... t_\omega\}$
  - Time units may be hours, days, weeks, months, etc.
  - Time window $\Omega = \{t_0, t_1, t_2, ... t_\omega\}$
  - Left and right censoring: Use time windows that are before the first and after the last time point in Ω to establish whether the first interaction in Ω indicates a new tie and the last one a deactivation

Three Measures of Degree

- In this dynamic context we can distinguish at least three empirically and conceptually distinct notions of vertex degree
  - **Cumulative Degree** $k_i(T_k)$: Count of the number of *unique contacts* that a given individual has interacted with up until a given time point T_k
  - **Effective Degree** $\kappa_i(t_k)$: The number of *active contacts* that an individual has the *potential* to interact with at a given time point t_k
  - **Active Degree** $\kappa_i^*(t_k)$: The number of *active contacts* than an individual *does in fact interact with* at a given time point t_k

Computing Cumulative, Effective and Active Degree

- **Effective Degree** κ_i: Beginning at t_0, $\kappa_i(t_1)$ is computed by adding the number of newly formed ties, $n_{\alpha,i}(t_1)$, and subtracting the number of deactivated ties $n_{\omega,i}(t_1)$
  - $\kappa_i(t_k) = \kappa_i(t_k - 1) + [n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)]$
  - When relatively stable $\overline{\kappa}_i$
  - Effective degree captures the size of the pool of people with whom an ego can interact
- **Cumulative Degree** $k_i(T_k)$: Total number of alters a person has interacted with up until time t_k
  - $k_i(t_k) = \sum_j a_{ij}(t_k)$
  - Where $a_{ij}(t_k) = 1$ when the ij tie appears for the first time in Ω
  - Cumulative Degree: $k_i(T_w) = \sum_{\mathbf{t_0}}^{t_w} k_i(t_k)$

Computing Cumulative, Effective and Active Degree

- **Active Degree** κ_i^*:
  - $\kappa_i^*(t_k) = \sum_j a_{ij}^*(t_k)$
  - Where $a_{ij}^*(t_k) = 1$ if i interacts with j at t_k
  - When normalized by effective degree, active degree gives the *proportion* of the total network that the individual activates at any given point in time:
  - $\kappa_i^{**}(t_k) = \frac{\kappa_i^*(t_k)}{\kappa_i(t_k)}$

Churn (Turnover)

- Effective Degree: Size of the pool of people with whom an ego can interact
- Churn captures the extent of change in who is in this pool of people
- When $\kappa_i(t_k)$ is fairly stable (because $\kappa_i(t_k) \simeq \bar{\kappa}_i$ and $n_{\alpha,i}(t_k) \simeq n_{\omega,i}(t_k)$) then churn is just the count of number of new (or deactivated) ties:
  - $C_i(t_k) = n_{\alpha,i}(t_k) = n_{\omega,i}(t_k)$
- When normalized by effective degree, gives us the average proportion of the effective ego-network that has changed in the time period:
  - $C_i^*(t_k) = \frac{C_i(t_k)}{\bar{\kappa}_i}$
- When there is churn, cumulative degree overstates effective degree

NetSense Study

- Two-year data collection effort (now extended to 4),funded by NSF and subsidized by Sprint (Striegel et al. 2013)
- $\sim$ 200 first-year students who began attending Notre Dame in the Fall of 2011 were given a smartphone and free voice/text/data plans
- Capture all communicative interactions: sender number, receiver number, time, date, communication form (voice, text)
- Random allocation of students to same-sex dorms
- Survey data: Prior to arrival students completed two surveys (and then subsequently every 4 months)
  - Demographic/Attitudes/Tastes/Behaviors survey
  - Network Survey: asked about their top 20 contacts

Data

- 161 students
  - $\sim$ 35 students dropped out of study before week 70
- 106 weeks: two academic years including summer
- $\sim$ 6 million communication events (voice or text)
- $\sim 55,000$ ties (weighted edges): communication between someone in the study and another person (possibly also in the study)
- $\sim 25,000$ mutual ties: at least 1 communication from i to j and from j to i
- Vast majority of ties and communication are with other people who are not in the NetSense study

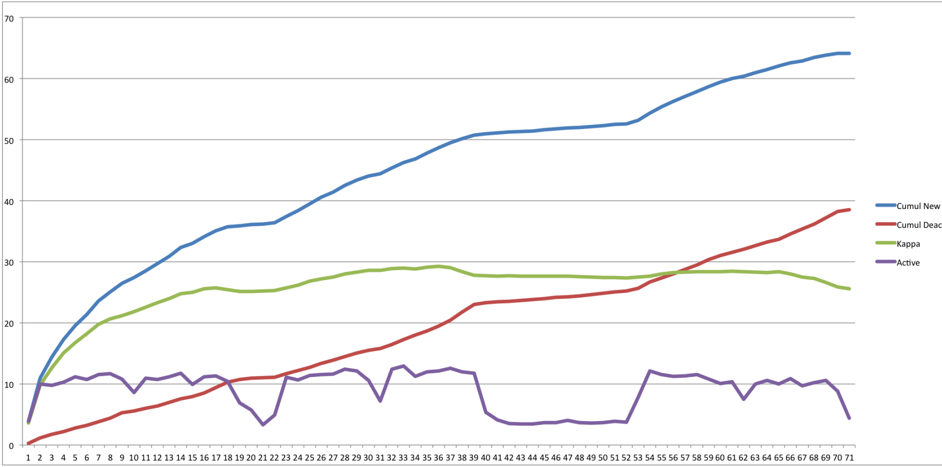
Types of Ties

- Communication events pick up activity with both pre-existing (prior to arrival at ND) ties and new ties
- **Pre-existing Ties**:
  - **Family ties** (3%): Identified from repeated name-generator surveys (top-20 contacts)
  - **Non-family ties** (5%): Ties to high school friends and others that remain active after arrival: Identified from repeated name-generator surveys (top-20 contacts)
- **New (Post-Arrival) ND Ties** (52%, $N \approx 13,000$): Ties formed after arrival on campus
  - No left-censoring: first time an interaction occurs is when the tie begins
- Residual Category: unknown type of ties (40%)

Time Window (Ω)

- Focus of the study is on the evolution of each student's ND ego-network during their first academic year (AY1),the summer, and the first semester of AY2, a 70 week period
- Remaining 36 weeks (Winter break 2012, Spring 2013 semester and Summer 2013) are used to determine if tie deactivates during the 70 week period
- If a tie is active during these 70 weeks but not active in the subsequent 36 weeks it is considered to have been deactivated
- Of the 13k post-arrival ties, 91% (N $\simeq 11.5k$) activated during the 70 week period
- Remainder activated after that and are not considered in this analysis
- Of the 11.5k ties activated, 62% (were *not* active in weeks 71-106, so these ties are considered to have been deactivated during the 70 week period
- No left censoring because all Post-Arrival ties begin after arrival at ND

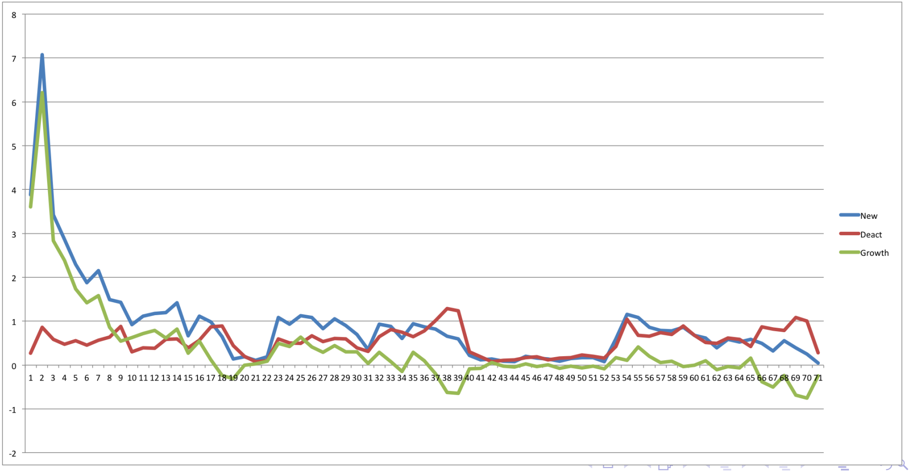
Mean Effective Degree, Active Degree, Cumulative Degree, and Cumulative Deactivated Ties (N = 161)



Effective Degree

- κ_i (effective degree) is fairly stable after an initial growth period
- All students begin with $\kappa_i(t_0) = 0$
- By t_{12} κ_i begins to stabilize for the typical student at around 25 and remains there through the Fall 2011 semester ($\overline{\kappa_i}(t_{13}, t_{18}) = 25.09$)
- κ_i increases at the beginning of the Spring 2012 semester and then stays fairly constant: $\overline{\kappa_i}(t_{26}, t_{36}) = 28.49$
- There is a decline in κ_i at end of the AY1 (weeks 37-39) to around 27 as some ties created in first year are not active in Fall 2012
- Throughout the Summer κ_i is fairly constant at around 27
- κ_i does not increase at beginning of Fall 2012 like it did at beginning of Spring 2012.
- Some decline in κ_i at end of Fall 2012 semester $\kappa_i(t_70) = 25.86$ as ties that were active are not active in the Spring semester and summer.

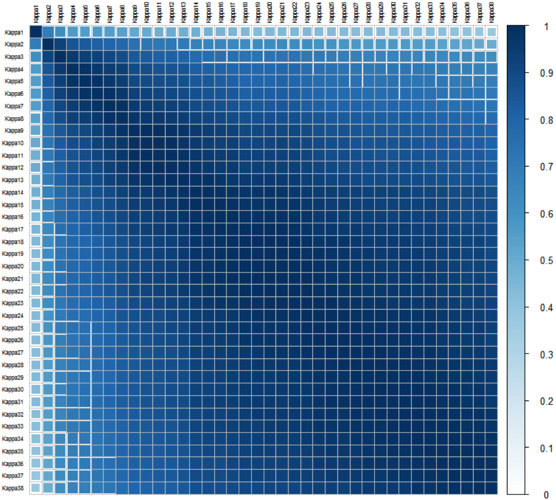
Mean Weekly Counts of Number of New Ties ($n_{\alpha,i}(t_k)$), Deactivated Ties ($n_{\omega,i}(t_k)$), and Growth in Number of Ties ($G_i(t_k)$), N = 161



Stability in Effective Degree

- Effective Degree is fairly stable because number of new ties $n_{\alpha,i}$ is similiar to number of deactivated ties $n_{\omega,i}$ after week 7
- Captured by the Growth count: $G_i(t_k) = n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)$
- High growth in first 5 weeks
- Growth evidences period effects
  - Very high during first 5 weeks, but note that deactivation levels are not negligible
  - Small negative growth (decline) at end of each semester
  - Small positive growth at beginning of each semester, especially Spring 2012

Intertemporal Correlation of κ_i



Summary

- After an initial period of growth, effective degree is fairly stable
- Stability is the result of tie formation and tie dissolution levels being very similar
- Temporal pattern to tie formation and dissolution
  - Growth at beginning a semester
  - Decline at end of a semester
  - Stability during semesters and breaks
- Cumulative Degree overstates effective degree
  - By week 70, the typical student has interacted with about 64 other ND students
  - But has also deactivated 38 ties
  - Result $\kappa_i(t_{70}) \simeq 26$
  - Turnover rate of 146%
  - Implication: Though the size of people's social networks is fairly stable, there is a good deal of change in who is in a person's network

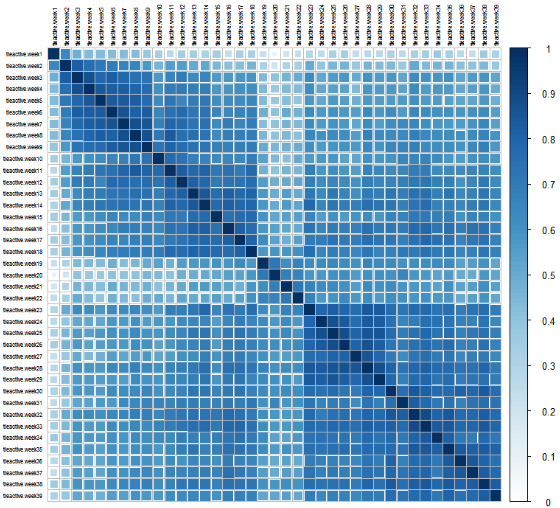
Churn (Turnover)

- Compute Churn during periods of relative stability in effective degree
- Normalize counts of new ties by effective degree to obtain indicator of churn, $C_i^*(t_k) = \frac{C_i(t_k)}{\bar{\kappa}_i}$
  - Weeks 13-18: $\mu_{C_i^*} = .24$
  - Weeks 26-35: $\mu_{C_i^*} = .33$
  - Weeks 56-65: $\mu_{C_i^*} = .25$
- No correlation with effective degree
- Small positive correlation turnover in Fall 2011 and Spring 2012 ($r = 0.26$), no correlations between turnover in Fall 2011 and Fall 2012, and in Spring 2012 and Fall 2012
- Churn is less stable and more episodic than degree

Active Degree

- $\kappa_i^*(t_k)$ shows strong period effects
- During breaks, $\kappa_i^*(t)$ drops as students communicate less with their ND friends
- $\kappa_i^*(t_k)$ picks up during semesters and remains fairly constant for the average student at around 11 students per week
- $\kappa_i^*(t_k)$ is less stable than $\kappa_i(t_k)$
- $\kappa_i^*(t_k)$ is on average much lower that $\kappa_i(t_k)$
- During the Spring 2012 semester the typical ND student interacted each week with between 40-50% of the ND contacts that they *could* interact with (e.g., $\kappa^{**}(t_{28}) = \frac{11.6}{27.5} = 0.42$)

Intertemporal Correlation of κ_i^*



Effective Degree Distributions

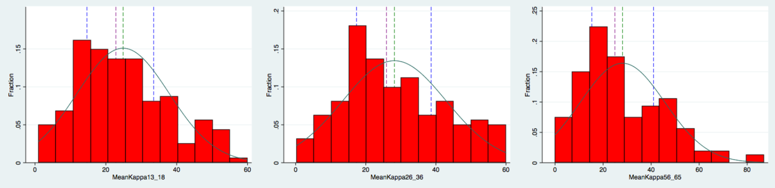
- Three relatively stable time periods: Fall 2011 (weeks 13-18), Spring 2012 (weeks 26-36), Fall 2012 (weeks 56-65)

			Correlations (Spearman)	
	Mean	St. Deviation	Weeks 26-36	Weeks 56-65
Weeks 13-18	25.09	13.84	0.94	0.74
Weeks 26-36	28.49	16	1	0.81
Weeks 56-65	28.35	18.7		

- Increasing variation could indicate preferential attachment
- However most of the variation is between subjects and not within subjects (i.e.overtime)

	Mean	St. Deviation	Between Subject	Within Subject	Intraclass Correlation (rho)
Weeks 13-18	25.09	13.9	13.78	1.81	0.984
Weeks 26-36	28.49	16.05	15.91	2.06	0.984
Weeks 56-65	28.35	18.84	18.68	2.48	0.983

Effective Degree Distributions



Change in Effective Degree

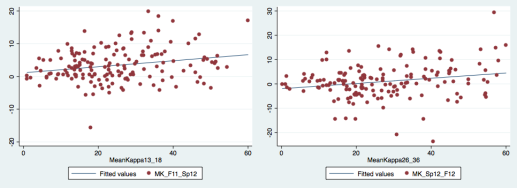


Figure: Change in Effective Degree by Degree at Beginning of Time Period, Fall 2011-Spring2012 and Spring 2012-Fall 2012

Summary

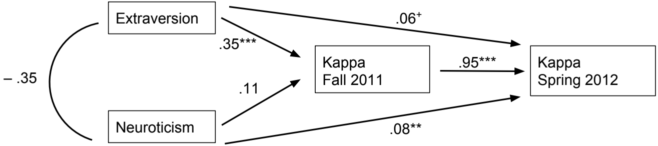
- Small positive effect of degree at time t on change in degree from time t to $t+1$
- Some evidence that low degree person's effective degree declines, higher degree person's degree increases
- However there is substantial variation in effective degree at time $t+1$ that is not accounted for by degree at time t
- What accounts for this variation between subjects that emerge in the Fall 2011 semester and persist through the next year?

Predicting Effective Degree

- Exogenous factors: demographic, background, activities, health, attitudes, aspirations, personality
- Measured prior to arriving at ND
- Path Model: Exogenous factors –> $\overline{\kappa_i}(t_{13}, t_{18})$ –> $\overline{\kappa_i}(t_{26}, t_{36})$
  - Alternatively can predict change scores

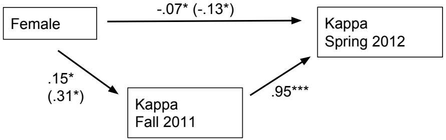
Extroversion

- Extroversion: 5 point scale (8 items), $\mu = 2.62$, $\sigma = .61$
- Extroversion strong predictor of effective degree
  - Net of other factors, a 1 SD increase in Extroversion yields .35 of a SD increase in $\overline{\kappa_i}(t_{13}, t_{18})$ ($\sim$ 6.7 friends for a 1 unit difference in Extroversion)
  - Net of Extroversion's effect on $\overline{\kappa_i}(t_{13}, t_{18})$ there is no direct effect of Extroversion on $\overline{\kappa_i}(t_{26}, t_{36})$
  - Because Extroversion is negatively correlated with Neuroticism, and Neuroticism has a small positive effect on $\overline{\kappa_i}(t_{26}, t_{36})$, the total effect of Extroversion on $\overline{\kappa_i}(t_{26}, t_{36})$ (.32) is less than the sum of its direct effect (.06) and its indirect effect through $\overline{\kappa_i}(t_{13}, t_{18})$ (.33)



Gender

- Women have .31 SD of $\overline{\kappa_i}(t_{13}, t_{18})$ (~ 4) more ND ties than men
- However men catch up so that by Spring women have .16 SD of $\overline{\kappa_i}(t_{26}, t_{36})$ (~ 2) more ND ties than men



Effective Degree

- In the NetSense data, differences in effective degree emerge within the first 13 weeks
- After week 13 a student's effective degree is fairly stable
- Most of the variance in their effective degree is between students and not over time
- Differences between students persist through the Spring 2012 and Fall 2102 semesters given the high correlation between effective degree in the Fall and Spring semesters
- The most important predictor of initial degree (and, therefore, subsequent degree) is Extroversion
  - The effective degree distribution is generated by differences among persons in their tendency to be social
  - Evidence for a small amount of preferential attachment: Change in effective degree is a function of initial level of effective degree
  - However most of the variance (inequality) in degree occurs early on and then persists

Implications

- Cumulative, Effective, and Active Degree conceptually and empirically distinct in a dynamic ego-network context
- Effective Degree captures the size of the set of alters an ego *could* interact with at a give time
- Cumulatative Degree captures the size of the set of alters an ego *has* interacted with up to some point in time
- Active Degree captures the size of the set of alters an ego *does* interacted with in a given time period

Implications

- Effective Degree best captures the notion of ego network size:
  - It is stable at relatively long time-scales
  - Closest conceptually to Dunbar's fixed inter-individual "ceiling" of sociability (Roberts et al. 2009)
- Cumulative Degree indexes a person's "history" of sociability but not their capacity to maintain ties nor their relative extent of tie activation
- Active degree captures the "strength of engagement" dimension
  - Gap between active and effective degree a useful index of the relative prevalence of non-decayed, weaker ties
  - Operational way of differentiating the inner-layer "support clique" from the active network "outer layer" (Hill and Dunbar 2003)
- Churn captures turnover in a person's effective network
  - It is independent of effective degree
  - It is less stable, more episodic, and less predictable
  - High churn could be indicative of a core-periphery ego network structure

Thanks!

Burt, Ronald S (2000). "Decay Functions". In: *Social Networks* 22.1, pp. 1–28. ISSN: 03788733. DOI: 10.1016/S0378-8733(99)00015-5. URL: http://linkinghub.elsevier.com/retrieve/pii/S0378873399000155.

Gould, Roger V (2002). "The origins of status hierarchies: A formal theory and empirical test". In: *American Journal of Sociology* 107.5, pp. 1143–1178. ISSN: 0002-9602 ER.

Hill, RA and RIM Dunbar (2003). "Social network size in humans". In: *Human nature* 14.1, pp. 53–72. URL: http://link.springer.com/article/10.1007/s12110-003-1016-y.

Miritello, Giovanna et al. (2013). "Limited communication capacity unveils strategies for human interaction". In: *Scientific reports* 3.April. arXiv: arXiv:1304.1979v1. URL: http://www.nature.com/srep/2013/130606/srep01950/full/srep01950.html.

Roberts, Sam G B et al. (2009). "Exploring variation in active network size: Constraints and ego characteristics". In: *Social Networks* 31.2, pp. 138–146.

Striegel, Aaron et al. (2013). "Lessons learned from the netsense smartphone study". In: *Proceedings of the 5th ACM workshop on HotPlanet*. ACM, pp. 51–56.

Operational criteria for classification as post-arrival tie

1. Named as ND student in ego-network surveys
2. Contact is shared by two unrelated NetSense study participants
3. Phone number of contact associated with ND email address
4. Phone number appear in the address books of unrelated NetSense study participants

Discriminant and Criterion Validity of Different Measures

	k_i	$\bar{\kappa}_i$	κ_i^*	G_i	Deact.	C_i	κ_i^{**}
Women	2.469	4.993*	3.499*	-0.0721	4.198*	-0.0217	0.0638*
	(1.21)	(2.32)	(3.71)	(-1.23)	(2.48)	(-0.41)	(3.52)
ND First	1.353	2.220	1.786*	0.0226	0.811	-0.0239	0.0475*
	(0.62)	(0.96)	(1.79)	(0.36)	(0.45)	(-0.36)	(2.32)
Outgoing	7.887*	10.38*	3.616*	0.0538	6.595*	-0.0644	-0.0446*
	(4.13)	(5.35)	(4.20)	(0.78)	(3.32)	(-1.11)	(-2.50)
Constant	19.97*	22.15*	8.710*	0.179*	15.66*	0.880*	0.391*
	(9.59)	(10.41)	(9.71)	(2.70)	(9.47)	(14.97)	(21.87)
N	156	156	156	156	156	156	156

* $p < 0.05$

Extroversion Scale

- I am someone who...
  - is talkative
  - is reserved
  - tends to be quiet
  - has an assertive personality
  - is sometimes shy, inhibited
  - is outgoing, sociable

Neuroticism Scale

- I am someone who...
  - is depressed, blue
  - is relaxed, handles stress well
  - can be tense
  - worries a lot
  - is emotionally stable, not easily upset
  - can be moody
  - remains calms in tense situations
  - gets nervous easily

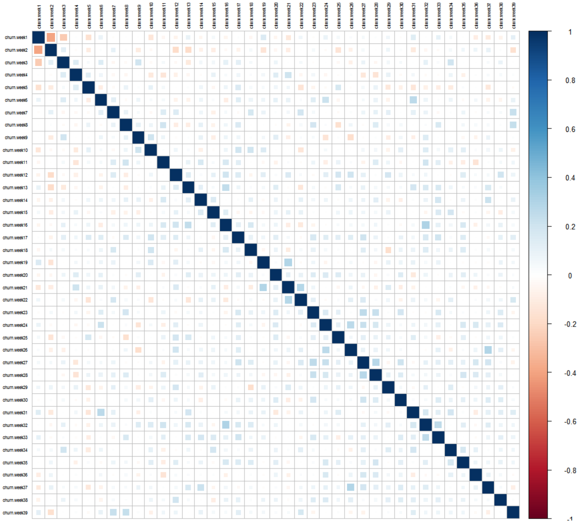
Effective Degree Growth

- $G_i(t_k) = n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)$
- Changes in $\kappa_i(t_k)$ at the ego level are solely a function of the difference between the number of ties formed $n_{\alpha,i}(t_k)$ and number of ties deactivated $n_{\omega,i}(t_k)$
- This implies that: $\kappa_i(t_k) = \kappa_i(t_k - 1) + G_i(t_k)$
- For any given time period Δ that begins at t_a and ends at t_w:
  1. Whenever $\sum_{t_a}^{t_w} n_{\alpha,i}(t) \neq \sum_{t_a}^{t_w} n_{\omega,i}(t)$, we should observe fluctuations in κ_i within persons over time
  2. Whenever $\sum_{t_a}^{t_w} n_{\alpha,i}(t) \simeq \sum_{t_a}^{t_w} n_{\omega,i}(t)$ we should observe little variation in κ_i within persons over time
  3. If (2) holds, then $\kappa_i(t_k) \simeq \bar{\kappa}_i$, where $\bar{\kappa}_i = \frac{1}{T} \sum_{t_a}^{t_w} \kappa_i(t)$ and T is the number of time points in Δ

Academic Year Time Periods

- Period 1: Formative Period (weeks 1-12, includes week 10 which is Fall Break)
- Period 2: Fall 2011 Semester (weeks 13-18)
- Period 3: Winter 2011-12 Break (weeks 19-22)
- Period 4: Beginning Spring Semester 2012 (weeks 23-25)
- Period 5: Spring Semester 2012 (weeks 26-36)
- Period 6: Summer Break (weeks 37-54)
- Period 7: Fall Semester 2012 (weeks 55-71)
- Period 8: Winter 2012-13 Break (week 72-75)
- Period 9: Spring Semester 2013 (weeks 76-92)
- Period 10: Summer 2013 (weeks 93-107)

Intertemporal Correlation of C_i^*



Effective Degree Distributions

- Three relatively stable time periods: Fall 2011 (weeks 13-18), Spring 2012 (weeks 26-36), Fall 2012 (weeks 56-65)
- Fall: $\mu_{\overline{\kappa_i}(t_{13},t_{18})} = 24.2$, $\sigma_{\overline{\kappa_i}(t_{13},t_{18})} = 12.5$
  - Between subject: $\sigma^{b}_{\overline{\kappa_i}(t_{13},t_{18})} = 12.35$
  - Within subject: $\sigma^{w}_{\overline{\kappa_i}(t_{13},t_{18})} = 1.70$
  - $\rho = .98$
- Spring: $\mu_{\overline{\kappa_i}(t_{26},t_{36})} = 26.9$, $\sigma_{\overline{\kappa_i}(t_{26},t_{36})} = 14.0$
  - Between subject: $\sigma^{b}_{\overline{\kappa_i}(t_{26},t_{36})} = 13.86$
  - Within subject: $\sigma^{w}_{\overline{\kappa_i}(t_{26},t_{36})} = 1.98$
  - $\rho = .98$
- $r_{\overline{\kappa_i}(t_{13},t_{18}),\overline{\kappa_i}(t_{26},t_{36})} = .94$