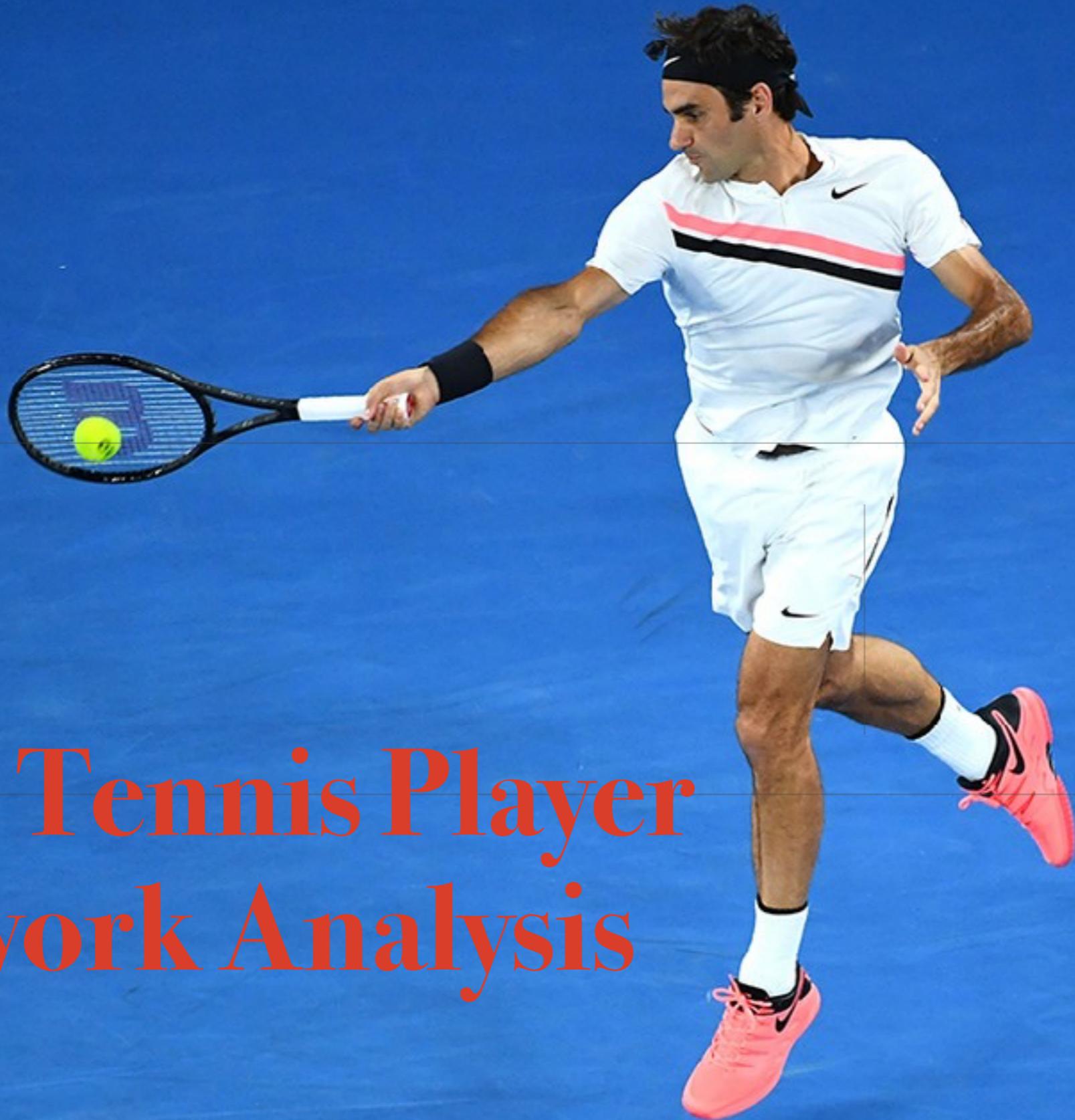


# ATP Tennis Player Network Analysis



Kuang-Yu Li  
30.01.2019

---

# Outline

---

- ❖ Introduction
- ❖ Structural Insight
- ❖ Centrality Insight
- ❖ Graph Model
- ❖ Conclusion

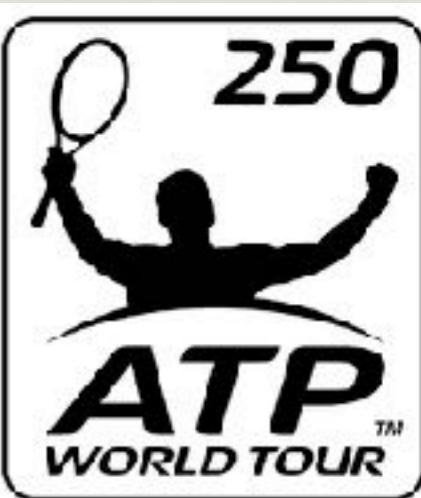
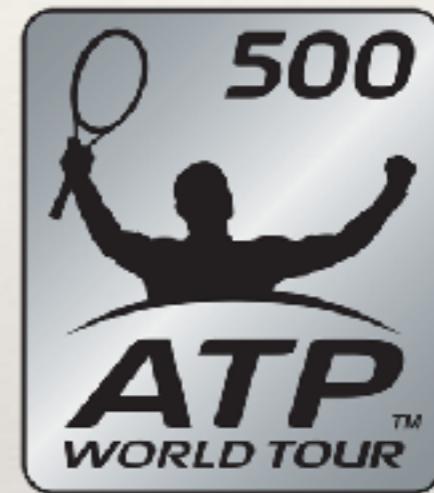
# ATP World Tour

- ❖ ATP (Association of Tennis Professionals) for Men Tennis
  - ❖ 1,814 active players
- ❖ ATP World Tour
  - ❖ 60+ tournaments and 2000+ matches in 2020
  - ❖ 5 main Categories
  - ❖ Single elimination
  - ❖ Draw is based on ranking





Tournament category	W	L	SL	QF	R16	R32	R64	R128	Q
Grand Slam	2000	1200	720	360	180	90	45	10	25
ATP World Tour Finals	+500 (1500 max)	+400 (1000 max)			(200 for each round robin match win) (600 max)				
Masters 1000	1000	600	360	180	90	45	10 (25)	(10)	25 (16)
500 Series	500	300	180	90	45	(20)			20 (10)
250 Series	250	150	90	45	20	(5)			12 (5)
ATP Challenger Tour Finals	+50	+30			(15 for each round robin match win)				
Challenger 125,000 +H	125	75	45	25	10				5
Challenger 125,000	110	65	40	20	9				5
Challenger 100,000	100	60	35	18	8				5
Challenger 75,000	90	55	33	17	8				5
Challenger 50,000	80	48	29	15	7				3
Challenger 35,000 +H	80	48	29	15	6				3
Futures 15,000 +H	35	20	10	4	1				
Futures 15,000	27	15	8	3	1				
Futures 10,000 +H	27	15	8	3	1				
Futures 10,000	10	10	6	2	1				
Other events									
Olympics (since 2016 no points are awarded)	750	450	340 270	135	70	35	5		



# Introduction: ATP Ranking

## ❖ ATP Ranking:

- ❖ known as 'world ranking'
- ❖ a 52-week rolling ranking
- ❖ renewed weekly
- ❖ active player only

## ❖ Other Ranking:

- ❖ Grand Slams
- ❖ GOAT points

Ranking	Move	Country	Player	Age	Points	Tourn Played	Points Dropping	Next Best
1	-	ESP	Rafael Nadal	33	10,235	13	1,200	0
2	-	CRO	Novak Djokovic	32	9,720	17	2,000	0
3	-	SUI	Roger Federer	38	6,590	17	180	0
4	-	RUS	Daniil Medvedev	23	5,960	24	180	180
5	-	AUT	Dominic Thiem	25	5,890	22	45	90
6	-	GRC	Stefanos Tsitsipas	21	5,375	27	720	90
7	-	GER	Alexander Zverev	22	3,345	25	180	45
8	-	ITA	Matteo Berrettini	23	2,870	24	10	45
9	-	ESP	Roberto Bautista Agut	31	2,630	23	360	45
10	-	FRA	Gael Monfils	33	2,555	22	45	0
11	-	BEL	David Goffin	29	2,555	27	90	45

# Grand Slam Ranking

## Open Era

Rank	Player	Total	Years	Australian Open	French Open	Wimbledon	US Open
1	瑞士 Roger Federer	20	2003–2018	6	1	8	5
2	西班牙 Rafael Nadal	19	2005–2019	1	12	2	4
3	塞尔维亚 Novak Djokovic	16	2008–2019	7	1	5	3
4	美国 Pete Sampras	14	1990–2002	2	0	7	5
5	瑞典 Björn Borg	11	1974–1981	0	6	5	0
6	美国 Jimmy Connors	8	1974–1983	1	0	2	5
	捷克 Ivan Lendl	8	1984–1990	2	3	0	3
	美国 Andre Agassi	8	1992–2003	4	1	1	2
9	美国 John McEnroe	7	1979–1984	0	0	3	4
	瑞典 Mats Wilander	7	1982–1988	3	3	0	1
11	瑞典 Stefan Edberg	6	1985–1992	2	0	2	2
	德国 Boris Becker	6	1985–1996	2	0	3	1
13	澳大利亚 Rod Laver	5	1968–1969	1	1	2	1
	澳大利亚 John Newcombe	5	1970–1975	2	0	2	1

# GOAT Point Ranking

Rank	Country	Name	Ranking GOAT Points
1	 SUI	Roger Federer	201
2	 SRB	Novak Djokovic	185
3	 ESP	Rafael Nadal	168
4	 GBR	Andy Murray	57
5	 ARG	Juan Martin Del Potro	10
5	 SUI	Stan Wawrinka	10
7	 GER	Alexander Zverev	7
8	 BUL	Grigor Dimitrov	6
8	 CAN	Milos Raonic	6
10	 AUT	Dominic Thiem	5
10	 JPN	Kei Nishikori	5
12	 USA	John Isner	4
13	 RUS	Daniil Medvedev	3
13	 CRO	Marin Cilic	3
15	 CRO	Ivo Karlovic	2
16	 FRA	Jo Wilfried Tsonga	1
16	 RSA	Kevin Anderson	1
16	 GRE	Stefanos Tsitsipas	1
16	 ESP	Tommy Robredo	1

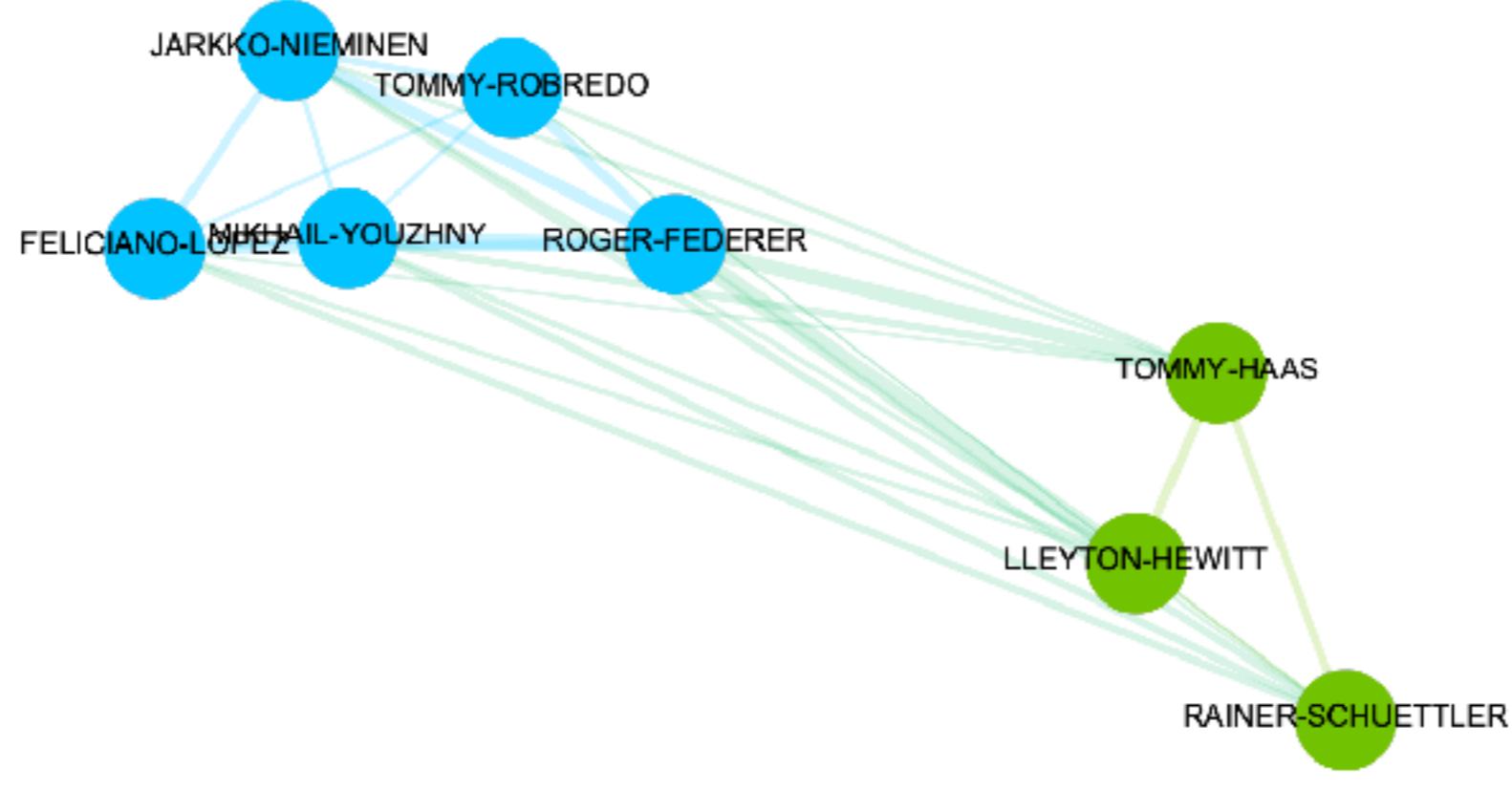
Can we evaluate a players trough network analysis?

# Data

- ❖ Real data collected from ATP match statistic from 1996~2016

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	tourney_year_id	tourney_id	tourney_slug	tourney_ur	tourney_round	order	match_order	winner_name	winner_play	winner_slug	loser_name	loser_player	loser_slug	winner_seed	loser_seed	match_score
25449	1998-414	32	hamburg	/en/scores/aRound of 32	5	15	Jens Kniboschil	k260	jens-kniboschil	Filip Dewulf	d228	filip-dewulf	Q		64 63	
25450	1998-414	32	hamburg	/en/scores/aRound of 32	5	15	Fabrice Santoro	s424	fabrice-santoro	Tommy Haas	h355	tommy-haas			64 64	
25451	1998-414	32	hamburg	/en/scores/aRound of 64	5	1	Felix Mantilla	m535	felix-mantilla	Guillaume Raoux	r186	guillaume-ra	9		64 (RET)	
25452	1998-414	32	hamburg	/en/scores/aRound of 64	5	2	Tommy Haas	h855	tommy-haas	Carlos Moya	m805	carlos-moya			10 62 75	
25453	1998-414	32	hamburg	/en/scores/aRound of 64	5	3	Michael Chang	c274	michael-chang	David Prinosil	p273	david-prinosil	11 WC	WC	64 36 61	
25454	1998-414	32	hamburg	/en/scores/aRound of 64	5	4	Tim Henman	h336	tim-henman	Jan Apell	a376	jan-apell	12 Q		63 62	
25455	1998-414	32	hamburg	/en/scores/aRound of 64	5	5	Sergi Bruguera	b350	sergi-bruguera	Jacobo Diaz	d291	jacobo-diaz	13 Q		62 61	
25456	1998-414	32	hamburg	/en/scores/aRound of 64	5	6	Jens Kniboschil	k260	jens-kniboschil	Magnus Norman	n250	magnus-nor	Q		14 63 60	
25457	1998-414	32	hamburg	/en/scores/aRound of 64	5	7	Oliver Gross	g297	oliver-gross	Nicolas Kiefer	k316	nicolas-kiefer	WC		15 36 62 61	
25458	1998-414	32	hamburg	/en/scores/aRound of 64	5	8	Goran Ivanisevic	i034	goran-ivanisi	Galo Blanco	b518	galo-blanco	16		46 64 63	
25459	1998-414	32	hamburg	/en/scores/aRound of 64	5	9	Karim Alami	a203	karim-alami	Martin Sinner	s367	martin-sinner	WC		75 62	
25460	1998-414	32	hamburg	/en/scores/aRound of 64	5	10	Hicham Arazi	a226	hicham-arazi	Andrea Gaudenzi	g254	andrea-gaudenzi			75 75	
25461	1998-414	32	hamburg	/en/scores/aRound of 64	5	11	Emilio Berlefe A	a190	emilio-berlefe	Marzio Martelli	m543	marzio-mart	Q		61 75	
25462	1998-414	32	hamburg	/en/scores/aRound of 64	5	12	Francisco Clavet	c252	francisco-cla	Marc Kevin Goellner	g252	marc-kevin-goellner			64 26 64	
25463	1998-414	32	hamburg	/en/scores/aRound of 64	5	13	Albert Costa	c378	albert-costa	Andrei Medvedev	m475	andrei-i-medvedev			63 16 76(3)	
25464	1998-414	32	hamburg	/en/scores/aRound of 64	5	14	Filip Dewulf	d228	filip-dewulf	Fernando Vicente	v195	fernando-vicente	Q		75 64	
25465	1998-414	32	hamburg	/en/scores/aRound of 64	5	15	Slava Dosedel	d198	slava-dosedel	Dominik Hrbaty	h477	dominik-hrbaty			60 62	
25466	1998-414	32	hamburg	/en/scores/aRound of 64	5	15	Nicolas Escude	e140	nicolas-escu	Julian Alonso	a272	julian-alonso			64 76(6)	
25467	1998-414	32	hamburg	/en/scores/aRound of 64	5	17	Wayne Ferreira	f196	wayne-ferreir	Jan Siemerink	s361	jan-siemerink			64 62	
25468	1998-414	32	hamburg	/en/scores/aRound of 64	5	18	Richard Fromber	f164	richard-fromber	Marcelo Filippini	f165	marcelo-filippini			63 67(7) 64	
25469	1998-414	32	hamburg	/en/scores/aRound of 64	5	19	Magnus Gustafsson	g182	magnus-gust	Boris Becker	b028	boris-becker			75 64	
25470	1998-414	32	hamburg	/en/scores/aRound of 64	5	20	Nicolas Lapentti	l290	nicolas-laper	Brett Steven	s321	brett-steven			61 61	
25471	1998-414	32	hamburg	/en/scores/aRound of 64	5	21	Thomas Muster	m099	thomas-mus	Albert Portas	p305	albert-portas			62 60	
25472	1998-414	32	hamburg	/en/scores/aRound of 64	5	22	Fabrice Santoro	s424	fabrice-santoro	Carlos Costa	c179	carlos-costa			63 63	

# Network

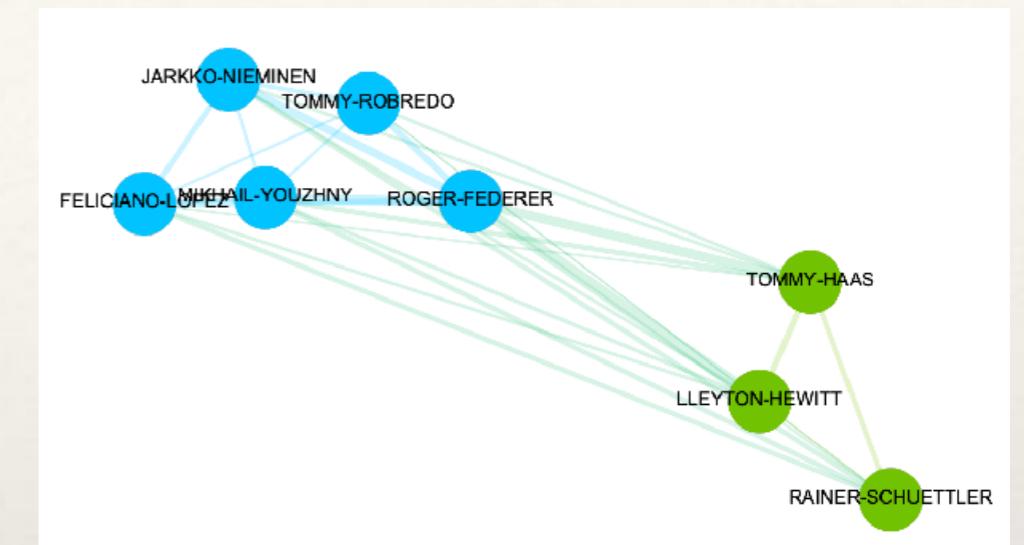


- ❖ Node: tennis player (order: 3,035)
- ❖ Edge: # of match fought between two players (size: 45,563)
- ❖ Type: undirected graph with weighted edge

# Network Construction

Raw Data: detailed match statistic

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	tourney_year_id	tourney_code	tourney_slug	tourney_ur	tourney_round	order	match_order	winner_name	winner_d	winner_s	loser_name	loser_player	loser_s	winner_seed	loser_seed	match_score	
25440	1998-414	32_hamburg	/en/scores/1Round of 32	5	15	Jens Klapisch	d250	ens-klapisch	filip-dewulf	d228	filip-dewulf	Q		64	63		
25450	1998-414	32_hamburg	/en/scores/1Round of 32	5	15	Fabrice Santoro	s424	fabrice-santoro	tommy-haas	h255	tommy-haas			64	64		
25451	1998-414	32_hamburg	/en/scores/1Round of 64	6	1	Felix Mantilla	m335	felix-mantilla	guillaume-raoux	r186	guillaume-raoux	9		64	(RET)		
25452	1998-414	32_hamburg	/en/scores/1Round of 64	6	2	Tommy Haas	h355	tommy-haas	carlos-moya	m605	carlos-moya			10	67	75	
25453	1998-414	32_hamburg	/en/scores/1Round of 64	6	3	Michael Chang	c274	michael-chang	baud-prinsell	p273	baud-prinsell	WC	WC	64	36	61	
25454	1998-414	32_hamburg	/en/scores/1Round of 64	6	4	Tim Henman	h383	tim-henman	jan-azel	a276	jan-azel	12	Q	63	62		
25455	1998-414	32_hamburg	/en/scores/1Round of 64	5	5	Sergi Bruguera	b350	sergi-bruguera	jacopo-diaz	d291	jacopo-diaz	13	Q	62	61		
25456	1998-414	32_hamburg	/en/scores/1Round of 64	5	6	Jens Klapisch	d260	ens-klapisch	Magnus Norman	n250	Magnus-Norman	C		14	63	60	
25457	1998-414	32_hamburg	/en/scores/1Round of 64	5	7	Oliver Gross	g297	oliver-gross	Nicolas Kiefer	k116	Nicolas-Kiefer	WC		15	35	67	61
25458	1998-414	32_hamburg	/en/scores/1Round of 64	5	8	Goran Ivanisevic	o334	goran-ivanisevic	Galo Blanco	b518	Galo-Blanco	16		45	64	63	
25459	1998-414	32_hamburg	/en/scores/1Round of 64	5	9	Karin Azari	a203	karin-azari	Martin Sinner	s267	Martin-Sinner	WC	WC	75	62		
25460	1998-414	32_hamburg	/en/scores/1Round of 64	5	10	Hicham Azaiez	a226	hicham-azaiez	Andrea Gaudenzi	g254	andrea-gaudenzi			75	75		
25461	1998-414	32_hamburg	/en/scores/1Round of 64	5	11	Emilio Berlejo A	a150	emilio-berlejo-a	Marco Martelli	m543	marco-martelli	Q		61	75		
25462	1998-414	32_hamburg	/en/scores/1Round of 64	5	12	Francisco Caver	c252	francisco-caver	Marco-Kevin Goellner	g252	Marco-Kevin-Goellner			64	25	64	
25463	1998-414	32_hamburg	/en/scores/1Round of 64	5	13	Albert Costa	c378	albert-costa	Anatoli Mihalev	m475	Anatoli-Mihalev			63	15	78	13
25464	1998-414	32_hamburg	/en/scores/1Round of 64	5	14	Filip Dewulf	d228	filip-dewulf	Fernando Vicente	v155	Fernando-Vicente	Q		75	64		
25465	1998-414	32_hamburg	/en/scores/1Round of 64	5	15	Slava Dosedze	d291	slava-dosedze	Dominik Hrbay	h277	Dominik-Hrbay			60	62		
25466	1998-414	32_hamburg	/en/scores/1Round of 64	5	16	Nicolás Escudé	e140	nico-as-escude	Julian Alarósz	a272	Julian-Alarósz			64	75	69	
25467	1998-414	32_hamburg	/en/scores/1Round of 64	5	17	Wayne Ferreira	f196	wayne-ferreira	Jan Siemerink	s361	Jan-Siemerink			64	62		
25468	1998-414	32_hamburg	/en/scores/1Round of 64	5	18	Richard Fairbank	f164	richard-fairbank	Marco Filippi	f165	Marco-Filippi			63	67	(7) 64	
25469	1998-414	32_hamburg	/en/scores/1Round of 64	5	19	Magnus Gustafsson	g182	magnus-gustafsson	Boris Becker	b628	Boris-Becker			75	64		
25470	1998-414	32_hamburg	/en/scores/1Round of 64	5	20	Nicolas Lapentti	g290	nico-as-lapentti	Bratt Steven	s421	Bratt-Steven			61	61		
25471	1998-414	32_hamburg	/en/scores/1Round of 64	5	21	Thomas Muster	m099	thomas-muster	Albert Portas	p205	Albert-Portas			62	60		
25472	1998-414	32_hamburg	/en/scores/1Round of 64	5	22	Frédéric Santoro	s424	frederic-santoro	Carlos Costa	c179	carlos-costo			63	63		



Input Data: only winner & loser

A	B	C
1 MICHAEL-STICH		NICKLAS-KULTI
2 JIM-COURIER		MICHAEL-STICH
3 MAGNUS-LARSSON		NICKLAS-KULTI
4 MARTIN-SINNER		JIM-COURIER
5 JIMMY-ARIAS		MICHAEL-STICH
6 FABRICE-SANTORO		NICKLAS-KULTI
7 PATRIK-KUHNEN		MAGNUS-LARSSON
8 PAUL-HAARHUIS		JIM-COURIER
9 MAGNUS-GUSTAFSSON		NICKLAS-KULTI
10 GILAD-BLOOM		MICHAEL-STICH
11 TODD-WOODBRIDGE		PATRIK-KUHNEN
12 JASON-STOLTBURG		JIMMY-ARIAS
13 SLOBODAN-ZIVOTINOVIC		MAGNUS-LARSSON
14 UDO-RIGLEWSKI		FABRICE-SANTORO
15 MARK-WOODFORDE		MARTIN-SINNER
16 BORIS-BECKER		MAGNUS-LARSSON
17 ALEXANDER-MRONZ		JIM-COURIER
18 HORST-SKOFF		JIMMY-ARIAS
19 SERGI-BRUGUERA		UDO-RIGLEWSKI
20 JOHN-FITZGERALD		MAGNUS-GUSTAFSSON
21 JEAN-PHILIPPE-FLEURIAN		MICHAEL-STICH
22 MARK-KOEVERMANS		MARTIN-SINNER
23 GUILLAUME-RAOUX		TODD-WOODBRIDGE
24 SANDON-STOLLE		GILAD-BLOOM

Data Parser

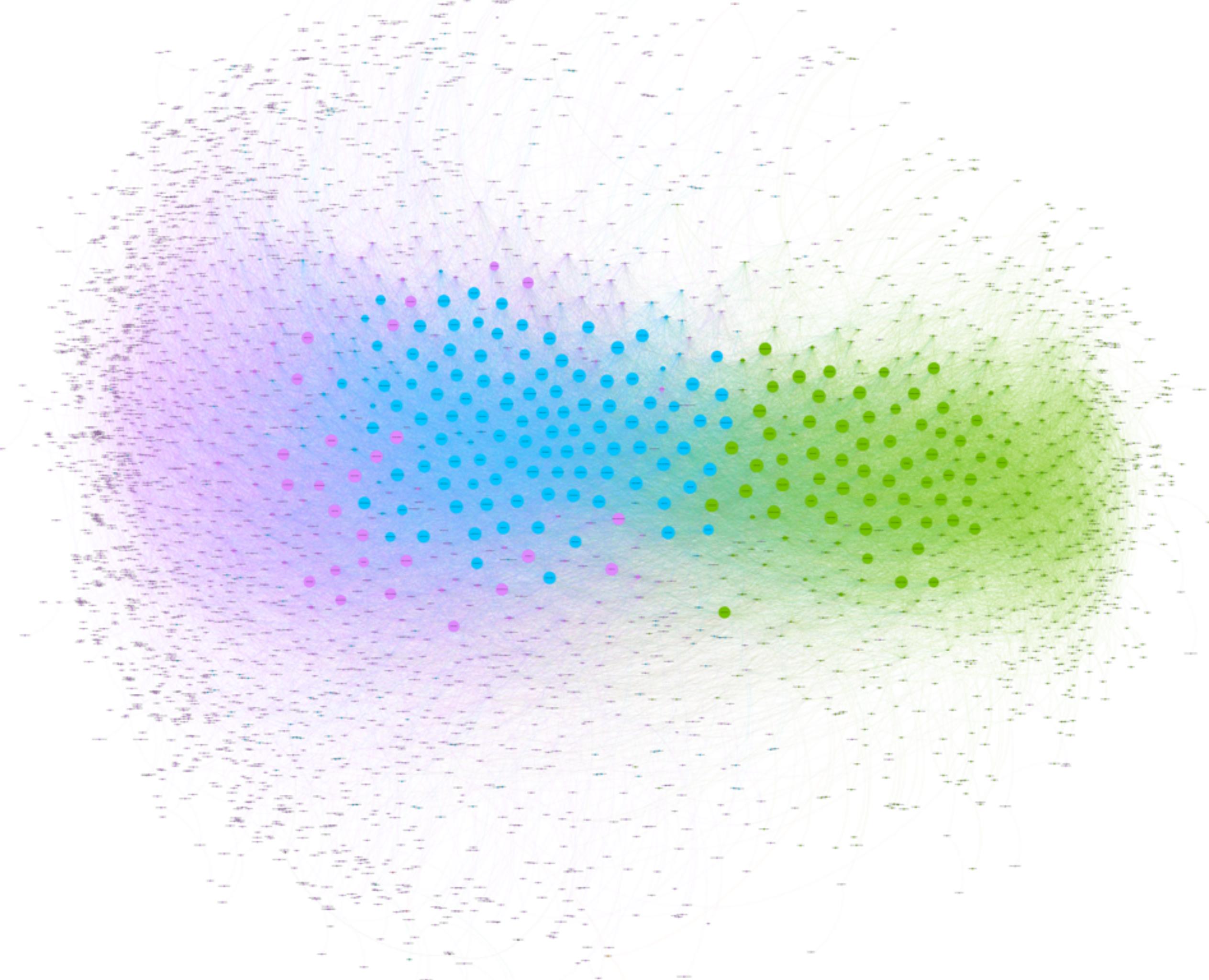
Network X

---

# My Expectation

---

- ❖ The graph is well connected with small diameter
- ❖ high clustering coefficient
- ❖ short average length
- ❖ one of the centrality can rank better



# Structural Insights

---

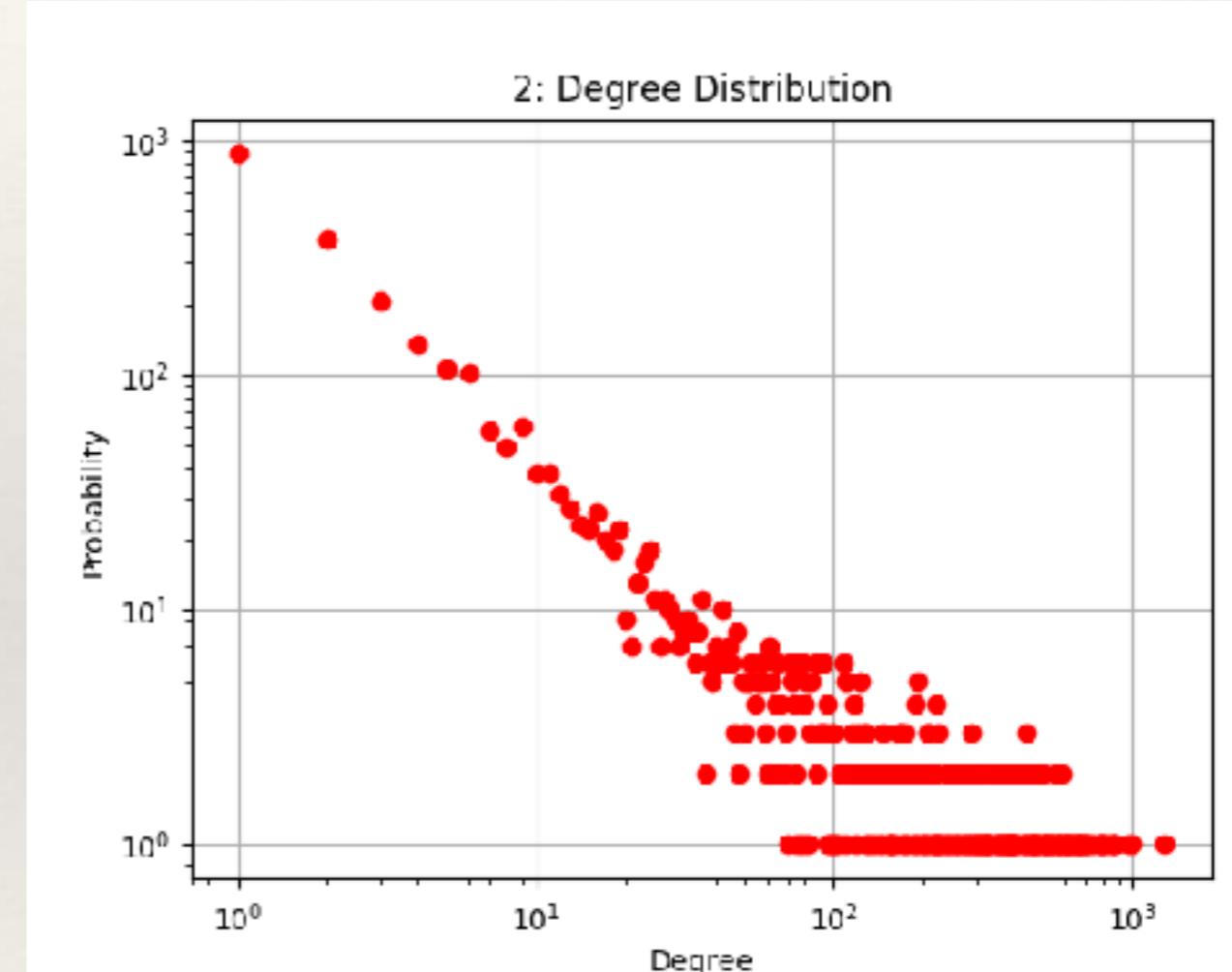
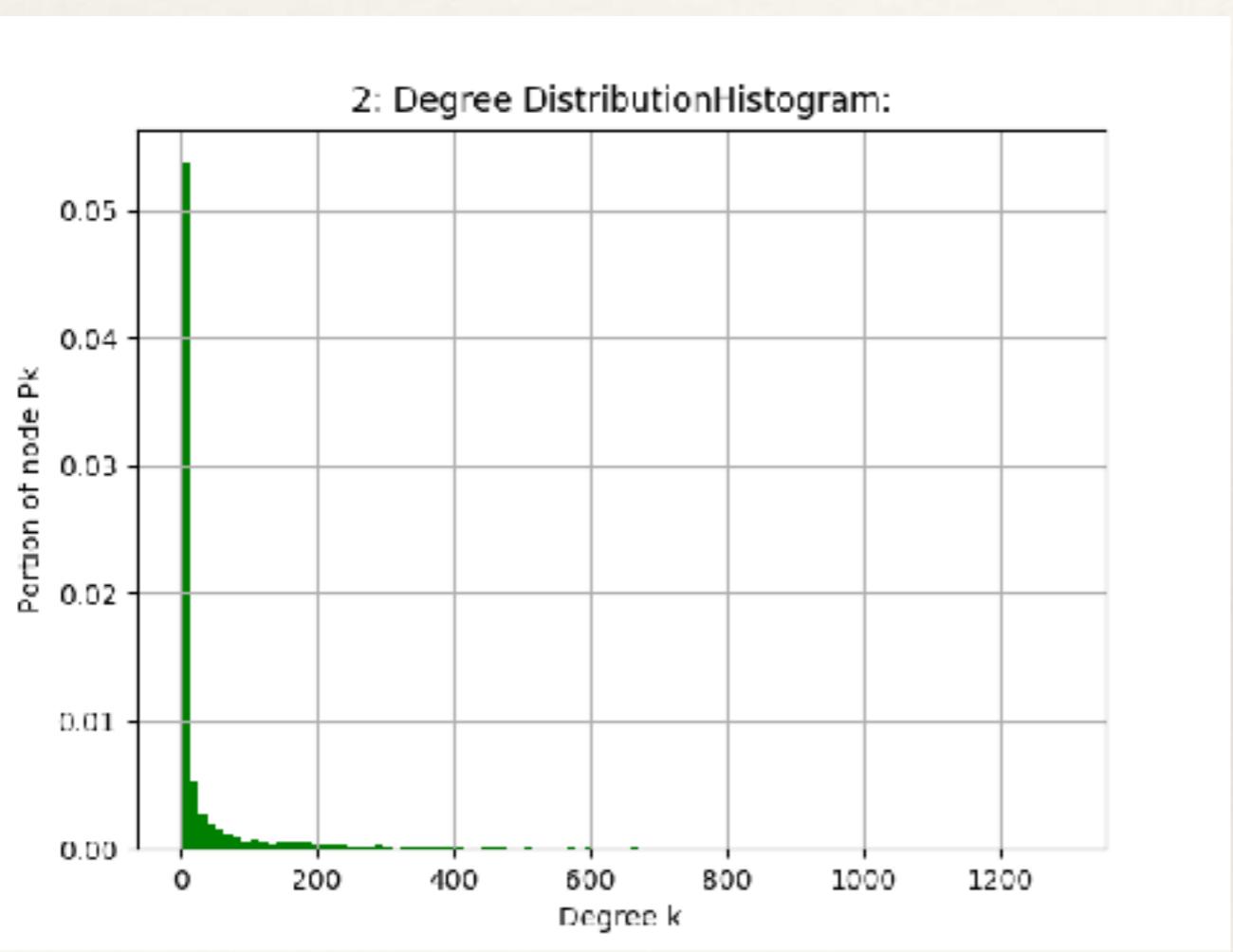
# Structural Metrics

---

- ❖ order: 3,035
- ❖ size: 45,563
- ❖ density: 0.009896
- ❖ diameter: 8
- ❖ radius: 4
- ❖ average path length: 3.263307
- ❖ average clustering coefficient: 0.004897
- ❖ transitivity: 0.350256
- ❖ number of triangle: 745,720
- ❖ number of clique: 12,066,625
- ❖ number of component: 1

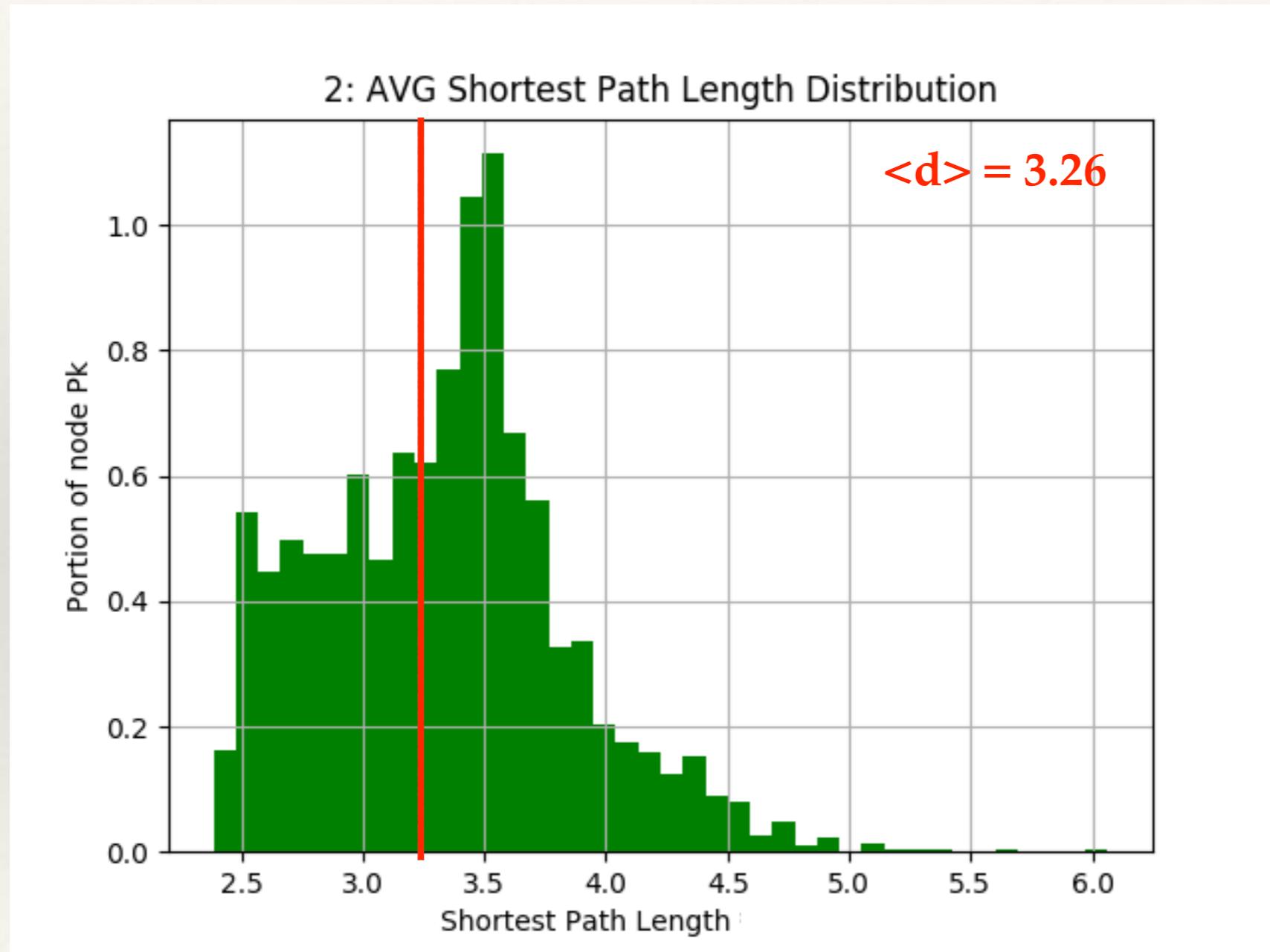
# Degree Distribution

- ❖ close to a scale-free network

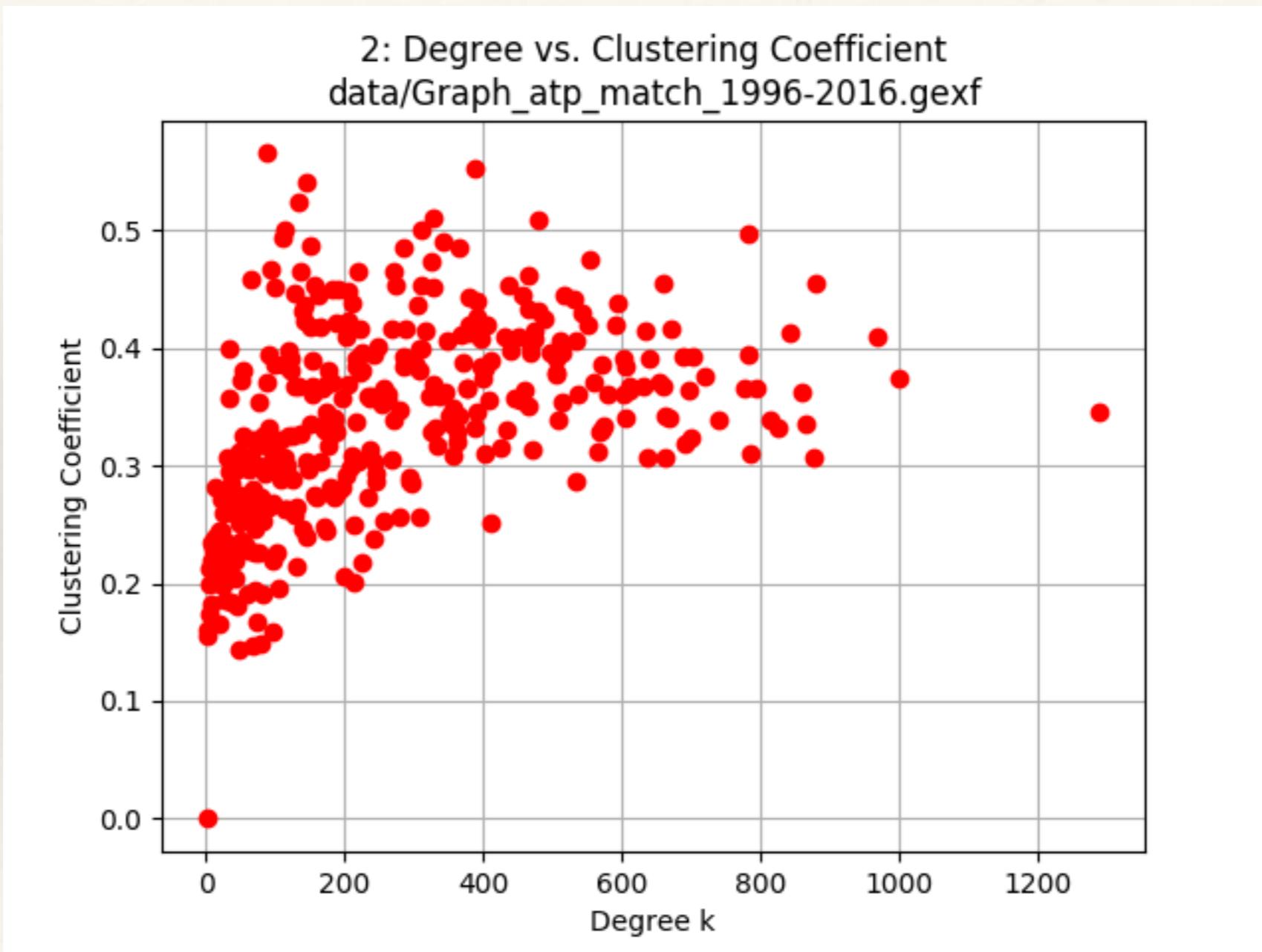


# Distance Distribution

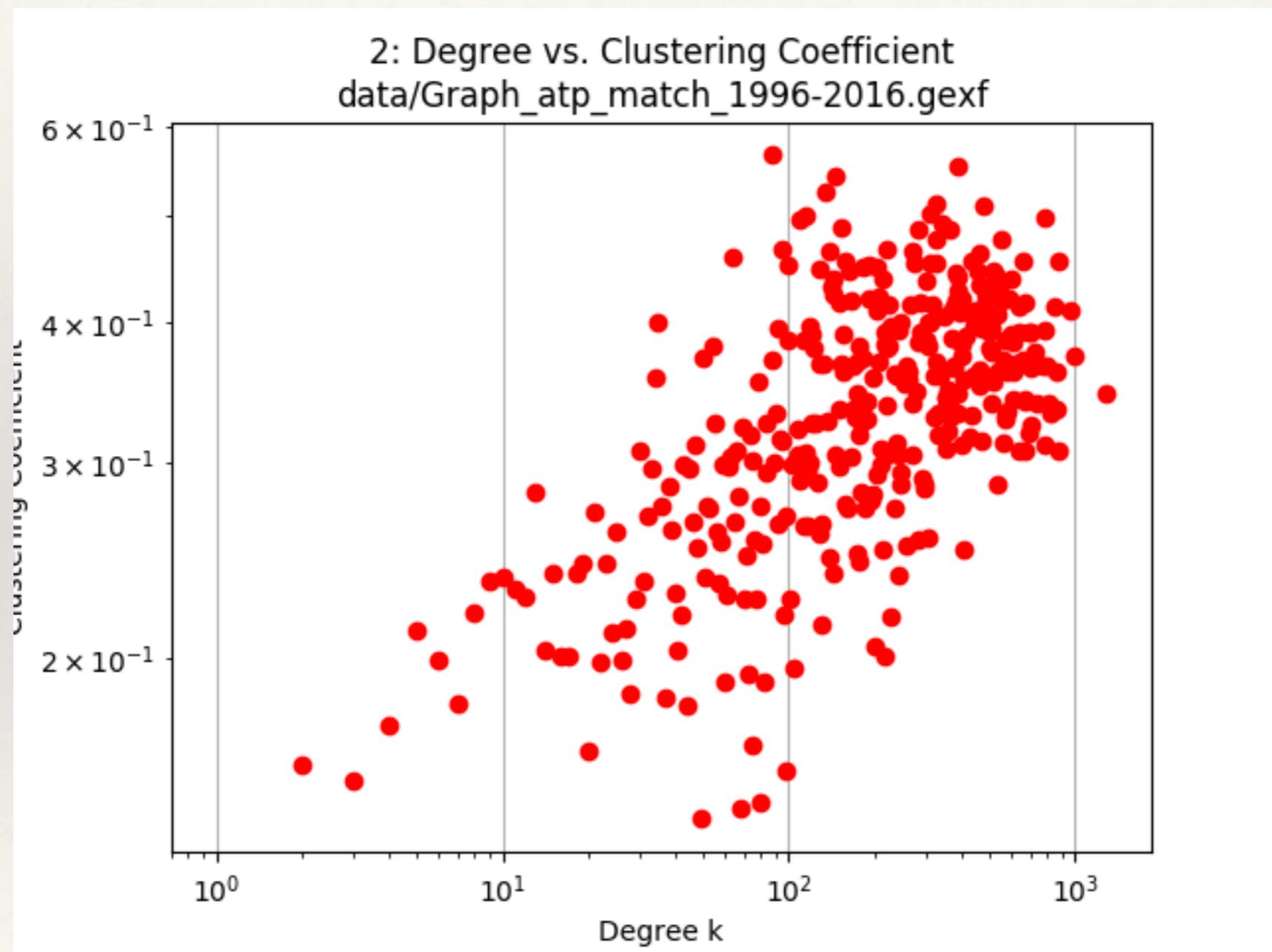
❖ 2

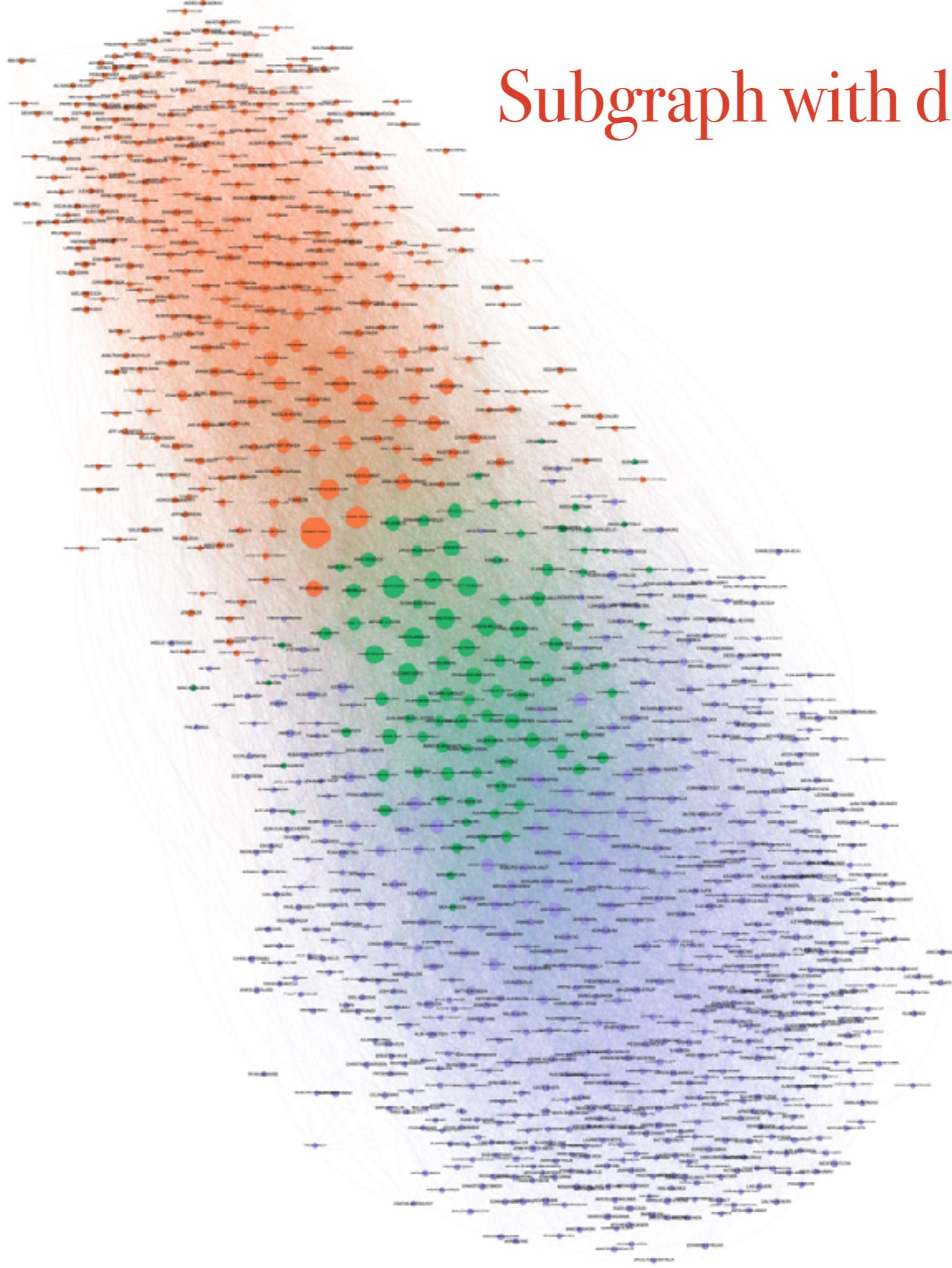


# Average Coefficient vs. Degree



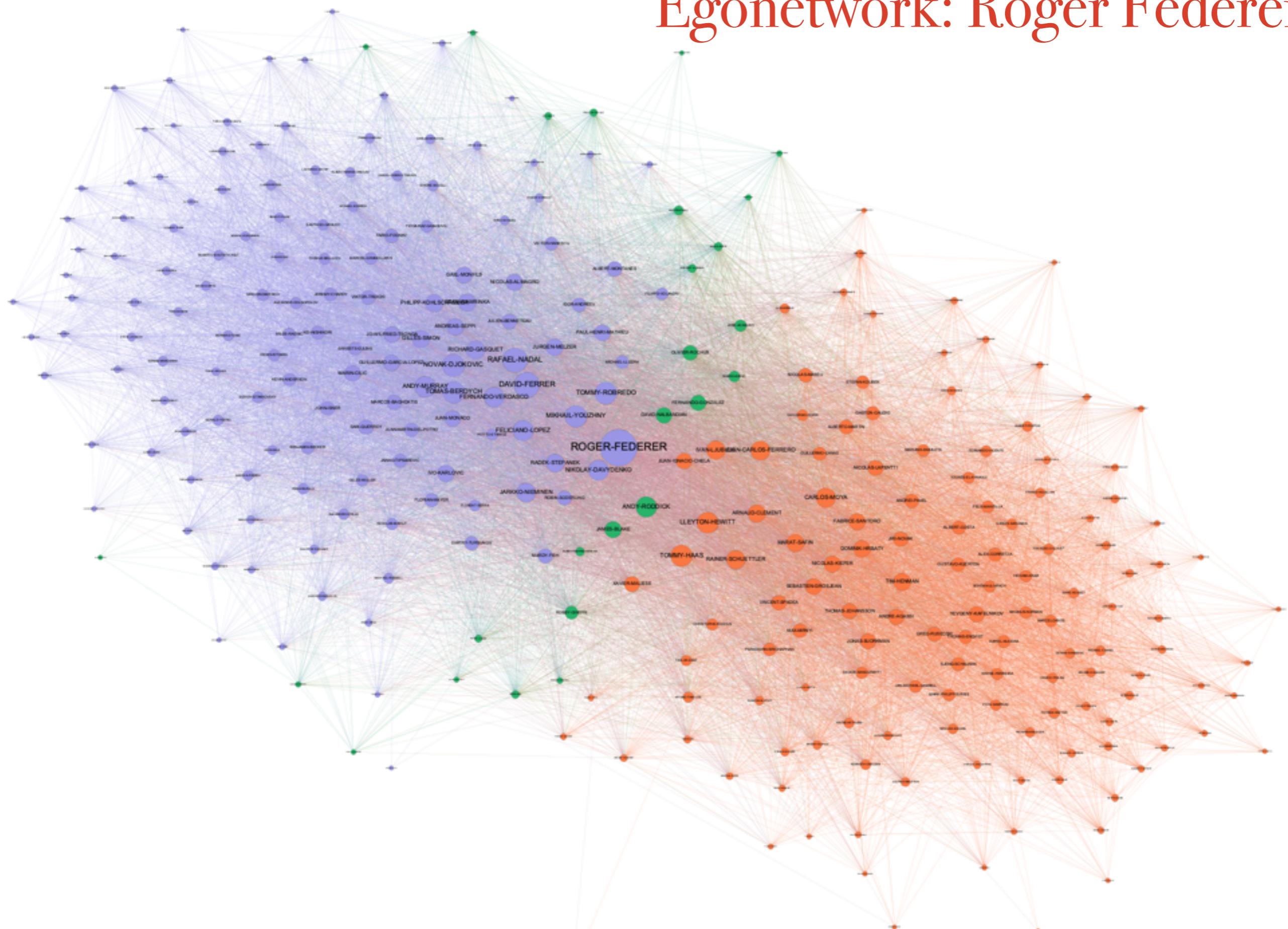
# Average Coefficient vs. Degree





# Subgraph with degree > 10

# Egonetwork: Roger Federer



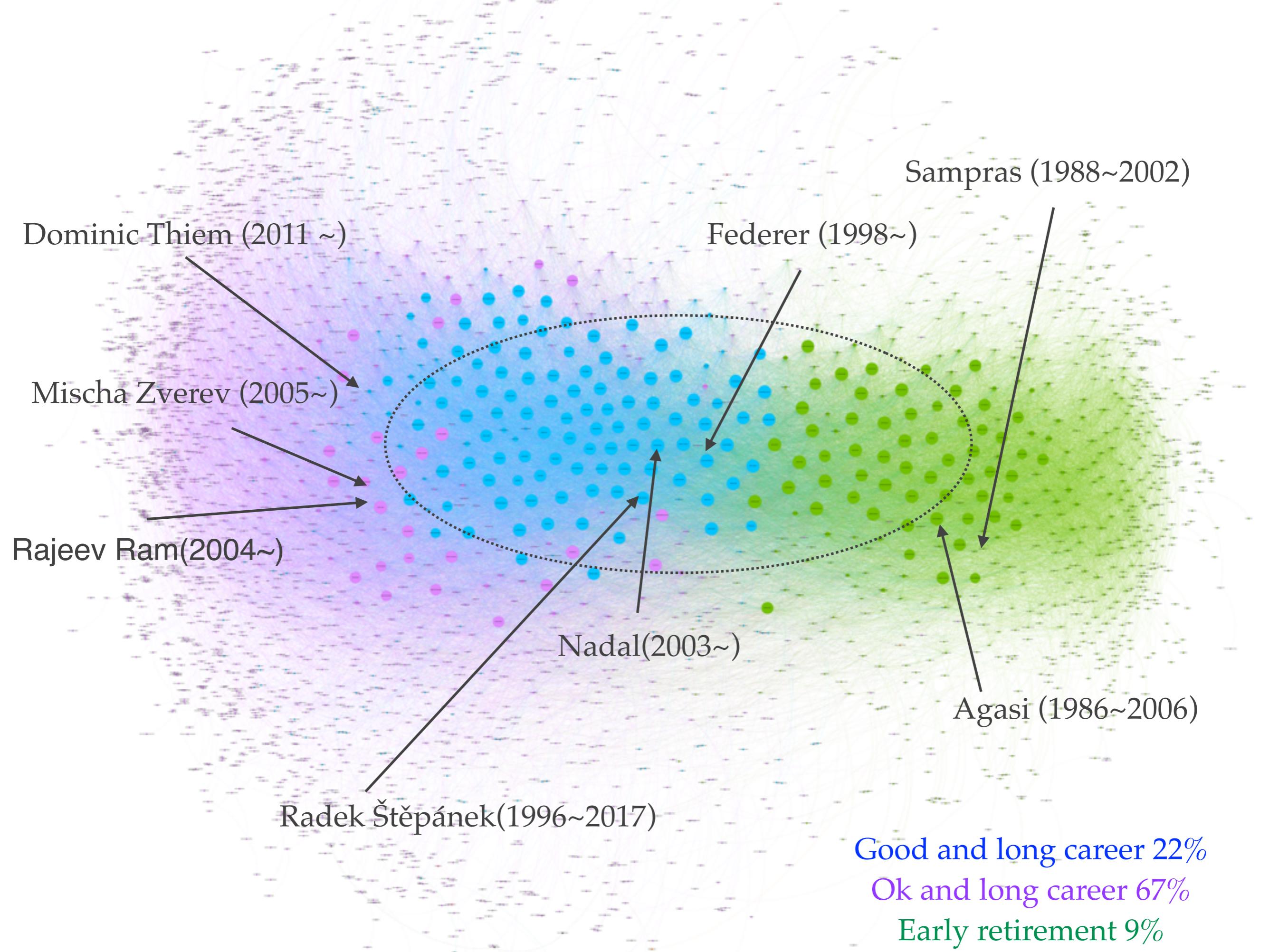
---

# Ego Network: Metrics

---

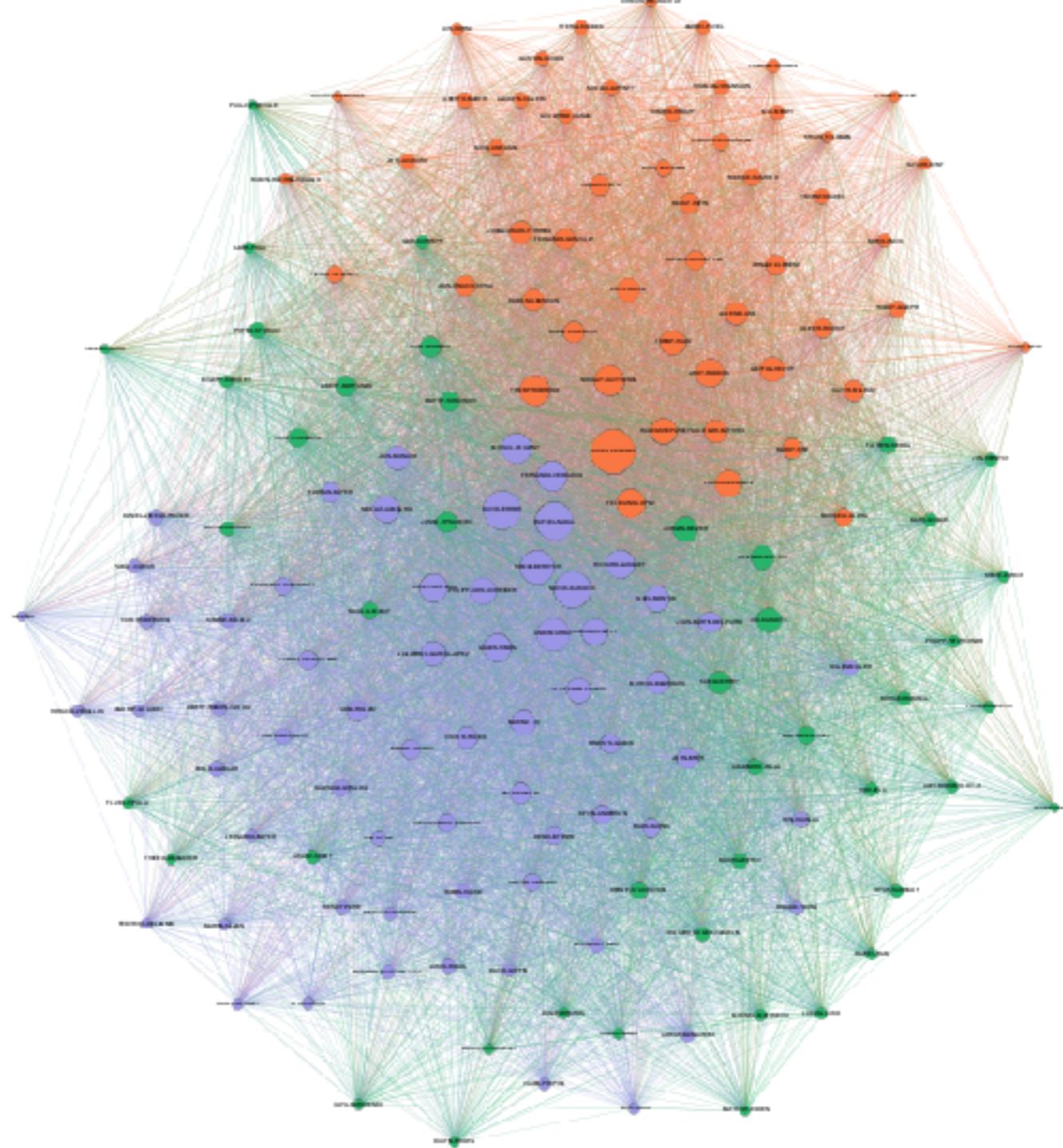
- ❖ order: 303
- ❖ size: 16,047
- ❖ density: 0.350731
- ❖ diameter: 2
- ❖ radius: 1
- ❖ average path length: 1.649269
- ❖ average clustering coefficient: 0.028295
- ❖ transitivity: 0.589713
- ❖ number of triangle: 404,647
- ❖ number of clique: 5,621,903
- ❖ number of component: 1

Community



# My k-Core

❖  $k = 78$



# Implementation

```
while (tmpG.order() < cur_k) or (flag_first == False): # algo stops when least order >= cur_k

    flag_first = True
    tmpG = G.copy()

    # Step 1- remove all nodes from the original graph that have a degree smaller than k
    # and all the incident nodes*
    for i_node in list(G.nodes):
        if (tmpG.has_node(i_node) == True) and (tmpG.degree(i_node) < cur_k):
            nbr_nodes = tmpG.neighbors(i_node)
            tmpG.remove_node(i_node)

    # Step 2- remove nodes that have fewer than k neighbors
    for i_node in list(tmpG.nodes):
        if (tmpG.has_node(i_node) == True) and len(list((tmpG.neighbors(i_node)))) < cur_k:
            tmpG.remove_node(i_node)

    # Step 3- iterate until no remaining node has fewer than k neighbors
    for j_node in list(tmpG.nodes):
        if (tmpG.has_node(j_node) == True) and len(list((tmpG.neighbors(j_node)))) < cur_k:
            tmpG.remove_node(j_node)
            print('[iter_j]remove node %s with less %d neighbor ' % (j_node, cur_k))
            print('[iter_i]remove node %s with less %d neighbor ' % (i_node, cur_k))

    print('[cur_k: %d]' % cur_k)
    print('tmpG order: ' + str(tmpG.order()))
    print(list(tmpG.nodes))

    # Step 4- the remaining nodes form the k-core
    curG = tmpG
    cur_k = cur_k - 1
```

---

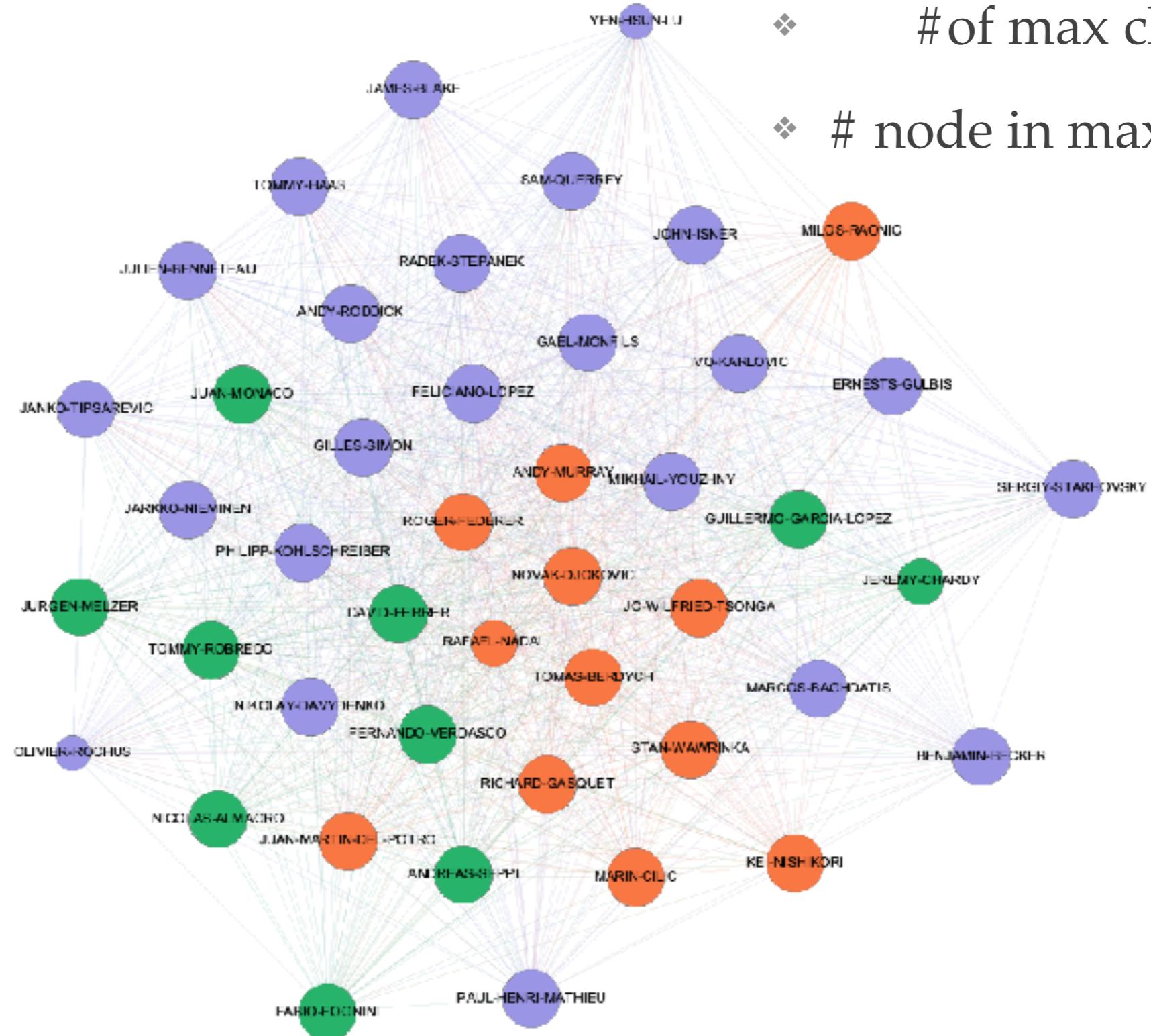
# My k-Core: Metrics

---

- ❖ order: 162
- ❖ size: 8954
- ❖ density: 0.686604
- ❖ diameter: 2
- ❖ radius: 2
- ❖ average path length: 1.313396
- ❖ average clustering coefficient: 0.039433
- ❖ transitivity: 0.744889
- ❖ number of triangle: 252,986
- ❖ number of clique: 6,800,942
- ❖ number of component: 1

# Maximal Clique

- ❖ # of max clique: 3
- ❖ # node in max clique: 43



---

# Maximal Clique

---

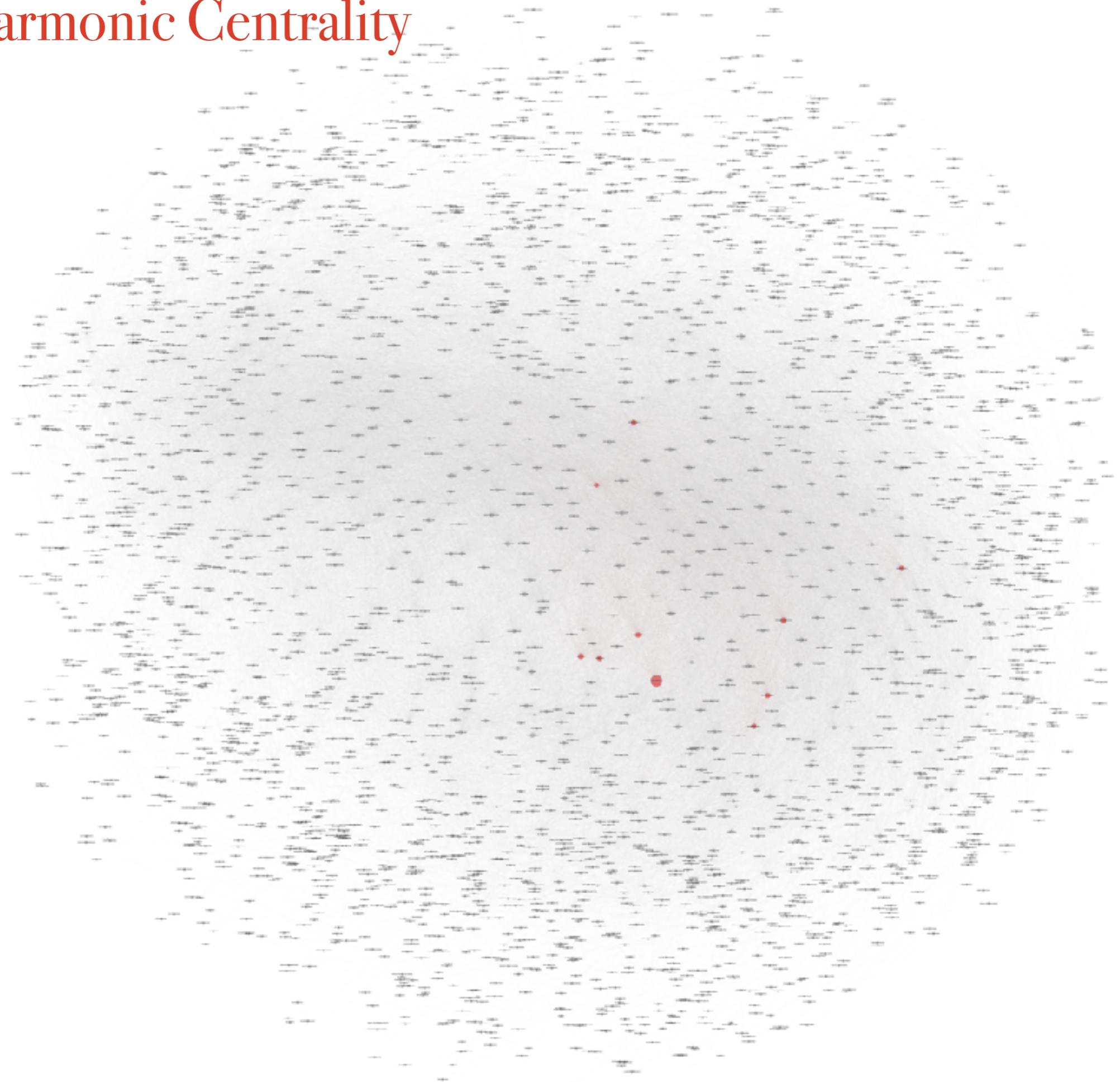
- ❖ ['TOMMY-HAAS', 'ROGER-FEDERER', 'FERNANDO-VERDASCO', 'TOMMY-ROBREDO', 'JARKKO-NIEMINEN', 'RAFAEL-NADAL', 'DAVID-FERRER', 'TOMAS-BERDYCH', 'FELICIANO-LOPEZ', 'JURGEN-MELZER', 'RICHARD-GASQUET', 'NIKOLAY-DAVYDENKO', 'IVO-KARLOVIC', 'MIKHAIL-YOUZHNY', 'JAMES-BLAKE', 'ANDY-RODDICK', 'RADEK-STEPANEK', 'JULIEN-BENNETEAU', 'ANDY-MURRAY', 'JUAN-MONACO', 'STAN-WAWRINKA', 'GILLES-SIMON', 'PHILIPP-KOHLSCHREIBER', 'JUAN-MARTIN-DEL-POTRO', 'NICOLAS-ALMAGRO', 'NOVAK-DJOKOVIC', 'ANDREAS-SEPPI', 'ERNESTS-GULBIS', 'GUILLERMO-GARCIA-LOPEZ', 'SAM-QUERREY', 'JANKO-TIPSAREVIC', 'JO-WILFRIED-TSONGA', 'PAUL-HENRI-MATHIEU', 'GAEL-MONFILS', 'MARCOS-BAGHDATIS', 'JOHN-ISNER', 'BENJAMIN-BECKER', 'SERGIY-STAKHOVSKY', 'KEI-NISHIKORI', 'MARIN-CILIC', 'MILOS-RAONIC', 'FABIO-FOGNINI', 'JEREMY-CHARDY']
- ❖ ['TOMMY-HAAS', 'ROGER-FEDERER', 'FERNANDO-VERDASCO', 'TOMMY-ROBREDO', 'JARKKO-NIEMINEN', 'RAFAEL-NADAL', 'DAVID-FERRER', 'TOMAS-BERDYCH', 'FELICIANO-LOPEZ', 'JURGEN-MELZER', 'RICHARD-GASQUET', 'NIKOLAY-DAVYDENKO', 'IVO-KARLOVIC', 'MIKHAIL-YOUZHNY', 'JAMES-BLAKE', 'ANDY-RODDICK', 'RADEK-STEPANEK', 'JULIEN-BENNETEAU', 'ANDY-MURRAY', 'JUAN-MONACO', 'STAN-WAWRINKA', 'GILLES-SIMON', 'PHILIPP-KOHLSCHREIBER', 'JUAN-MARTIN-DEL-POTRO', 'NICOLAS-ALMAGRO', 'NOVAK-DJOKOVIC', 'ANDREAS-SEPPI', 'ERNESTS-GULBIS', 'GUILLERMO-GARCIA-LOPEZ', 'SAM-QUERREY', 'JANKO-TIPSAREVIC', 'JO-WILFRIED-TSONGA', 'PAUL-HENRI-MATHIEU', 'GAEL-MONFILS', 'MARCOS-BAGHDATIS', 'JOHN-ISNER', 'BENJAMIN-BECKER', 'SERGIY-STAKHOVSKY', 'KEI-NISHIKORI', 'MARIN-CILIC', 'MILOS-RAONIC', 'FABIO-FOGNINI', 'OLIVIER-ROCHUS']
- ❖ ['TOMMY-HAAS', 'ROGER-FEDERER', 'FERNANDO-VERDASCO', 'TOMMY-ROBREDO', 'JARKKO-NIEMINEN', 'YEN-HSUN-LU', 'DAVID-FERRER', 'TOMAS-BERDYCH', 'FELICIANO-LOPEZ', 'IVO-KARLOVIC', 'JURGEN-MELZER', 'JULIEN-BENNETEAU', 'RADEK-STEPANEK', 'RICHARD-GASQUET', 'NIKOLAY-DAVYDENKO', 'GAEL-MONFILS', 'PAUL-HENRI-MATHIEU', 'ANDY-MURRAY', 'MIKHAIL-YOUZHNY', 'JAMES-BLAKE', 'ANDY-RODDICK', 'JUAN-MARTIN-DEL-POTRO', 'NICOLAS-ALMAGRO', 'JUAN-MONACO', 'SAM-QUERREY', 'JANKO-TIPSAREVIC', 'JOHN-ISNER', 'JO-WILFRIED-TSONGA', 'NOVAK-DJOKOVIC', 'ERNESTS-GULBIS', 'MARIN-CILIC', 'STAN-WAWRINKA', 'MARCOS-BAGHDATIS', 'ANDREAS-SEPPI', 'MILOS-RAONIC', 'JEREMY-CHARDY', 'GILLES-SIMON', 'PHILIPP-KOHLSCHREIBER', 'KEI-NISHIKORI', 'GUILLERMO-GARCIA-LOPEZ', 'FABIO-FOGNINI', 'BENJAMIN-BECKER', 'SERGIY-STAKHOVSKY']

# Centrality Insights

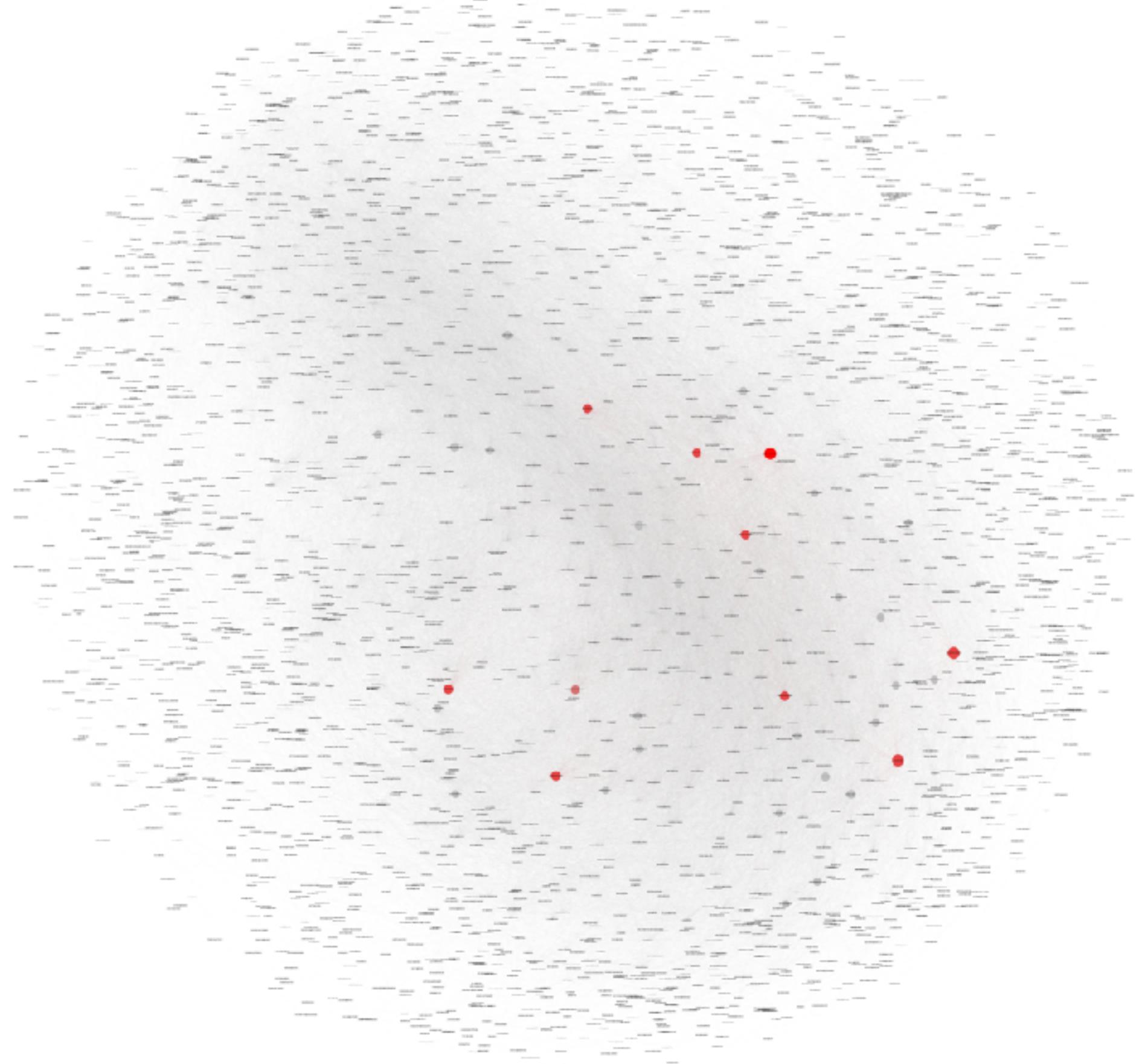
# Centrality Comparison $Hi^* Bi ki Ei$

#	Harmonic		Betweenness		Degree		Eigen	
1	JAN-HERNYCH	0.458	RAJEEV-RAM	0.016	TOMMY-HAAS	0.106	ROGER-FEDERER	0.234
2	RAJEEV-RAM	0.453	MISCHA-ZVEREV	0.016	ROGER-FEDERER	0.100	RAFAEL-NADAL	0.192
3	MICHAEL-BERRER	0.449	MICHAL-PRZYSIEZNY	0.016	MIKHAIL-YOUZHNY	0.098	NOVAK-DJOKOVIC	0.188
4	SIMONE-BOLELLI	0.449	ANDREY-GOLUBEV	0.012	RAINER-SCHUETTLER	0.098	DAVID-FERRER	0.173
5	ANDREY-GOLUBEV	0.449	SERGIY-STAKHOVSKY	0.012	JARKKO-NIEMINEN	0.098	ANDY-MURRAY	0.156
6	VICTOR-HANESCU	0.449	GEORGE-BASTL	0.012	LLEYTON-HEWITT	0.098	TOMAS-BERDYCH	0.153
7	BJORN-PHAU	0.448	MICHAEL-BERRER	0.012	FELICIANO-LOPEZ	0.097	MIKHAIL-YOUZHNY	0.132
8	MISCHA-ZVEREV	0.448	JAN-HERNYCH	0.011	TOMMY-ROBREDO	0.097	ANDY-RODDICK	0.130
9	TEYMURAZ-GABASHVILI	0.448	RUBEN-RAMIREZ-HIDALGO	0.011	RADEK-STEPANEK	0.095	FERNANDO-VERDASCO	0.128
10	RADEK-STEPANEK	0.446	BJORN-PHAU	0.010	CARLOS-MOYA	0.095	TOMMY-ROBREDO	0.128

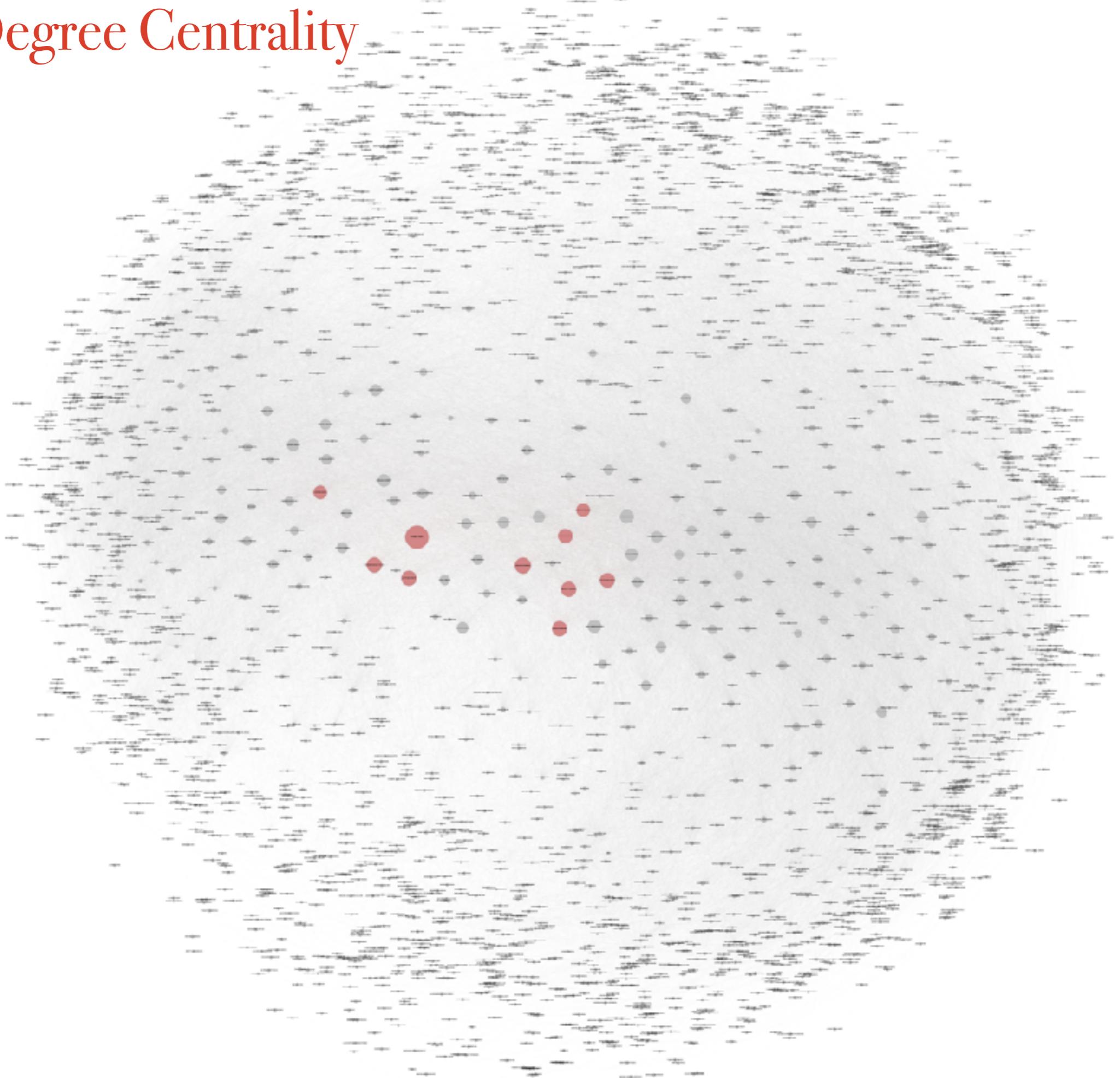
# Harmonic Centrality



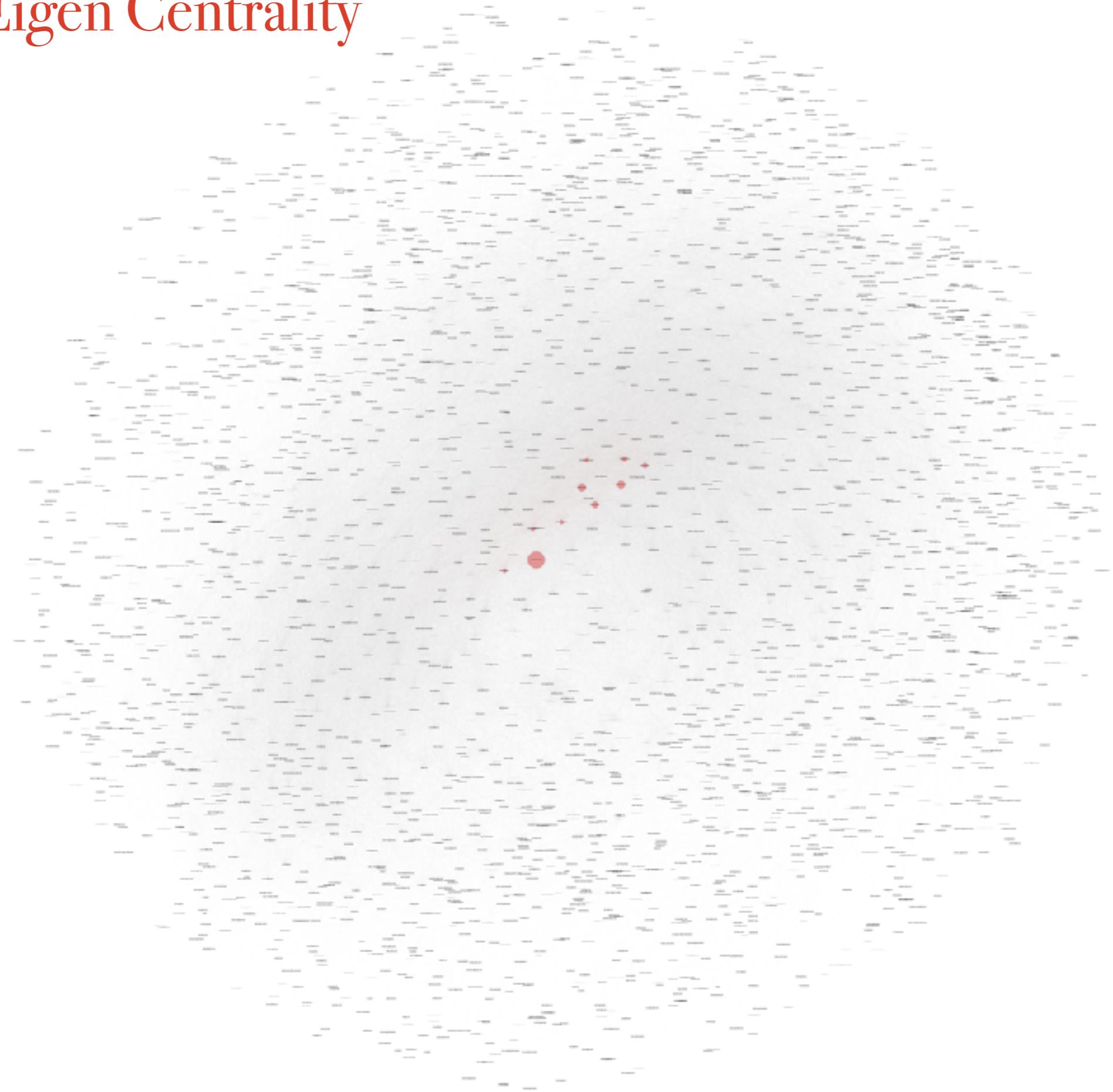
# Betweenness Centrality



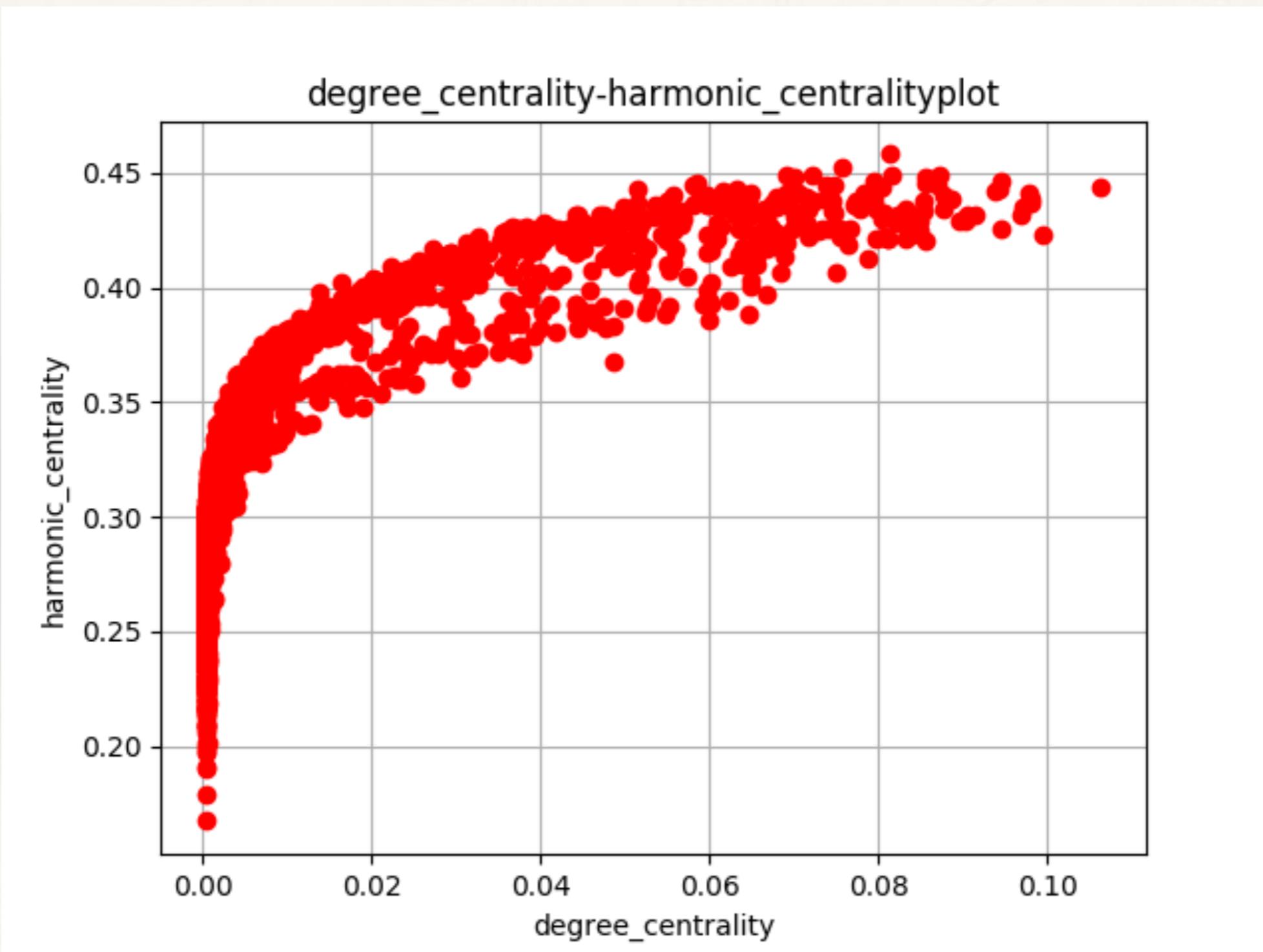
# Degree Centrality



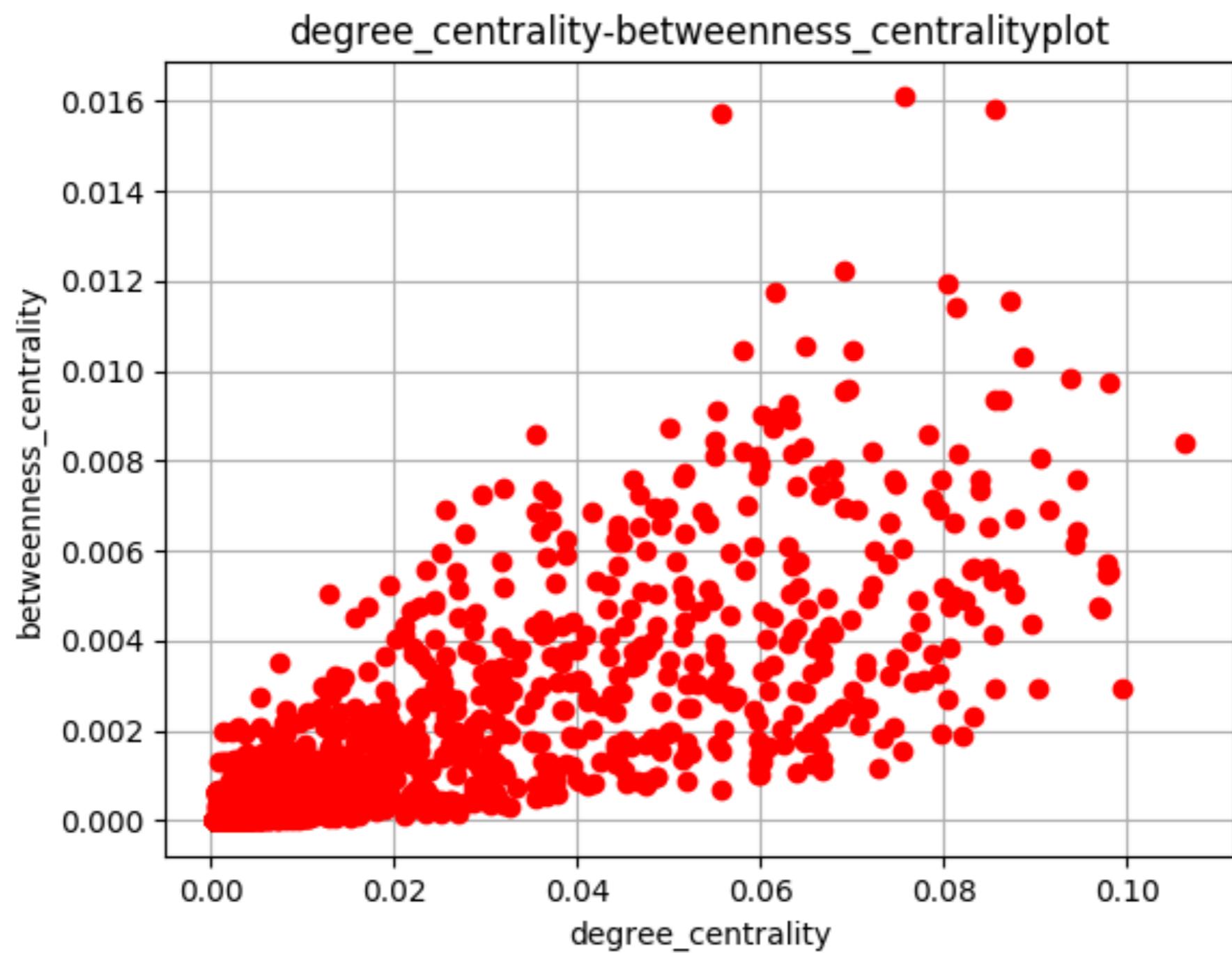
# Eigen Centrality



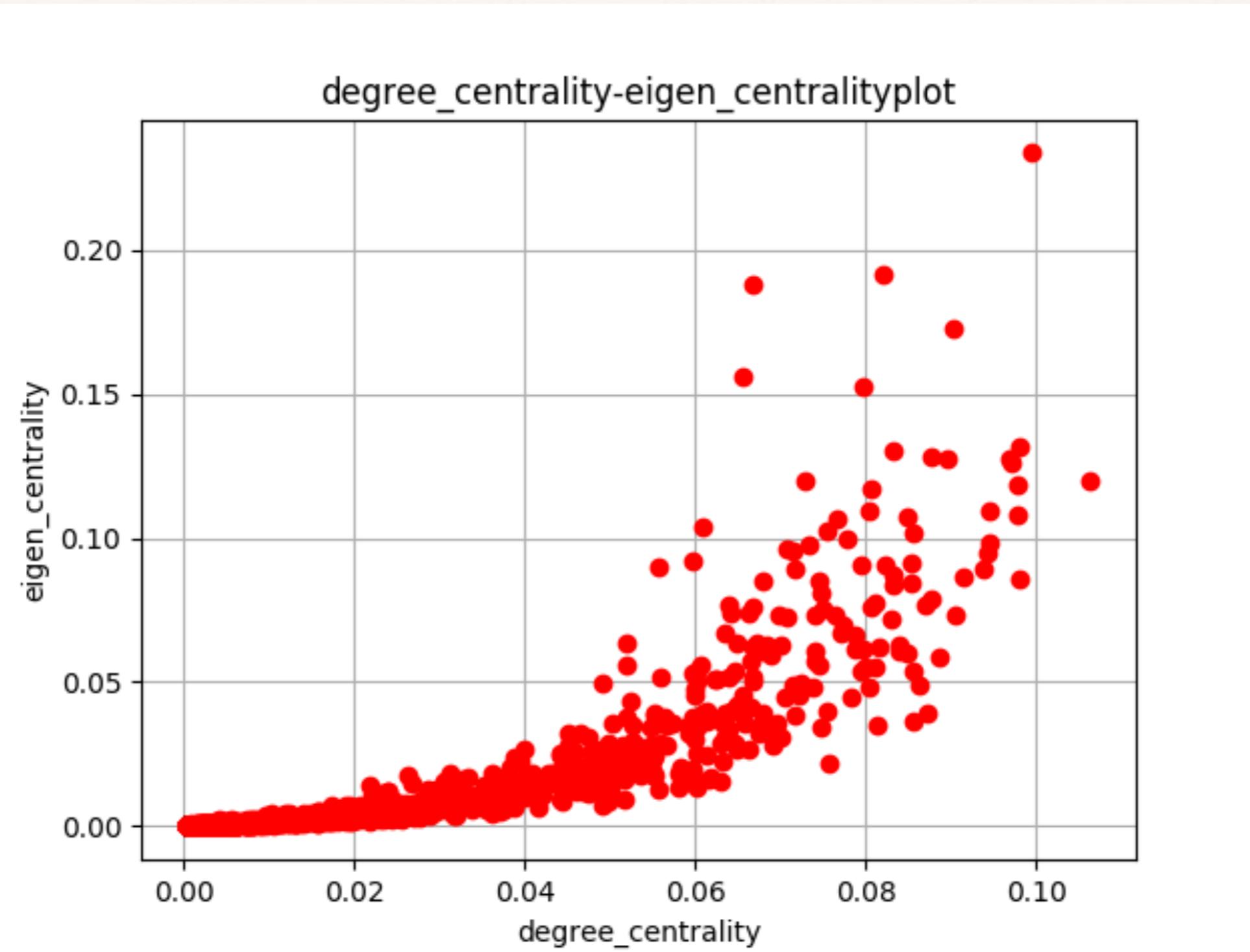
# $Hi^* - ki$ Plot



# *Bi – ki* Plot



# *Ei-ki* Plot



# Katz Centrality and Page Rank

# Katz Centrality

#	Beta = 0.1 Alpha = 0.9( $1/\lambda$ )		Beta = 0.2		Beta = 0.5		Beta = 1	
1	ROGER-FEDERER	0.199	ROGER-FEDERER	0.199	ROGER-FEDERER	0.199	ROGER-FEDERER	0.199
2	RAFAEL-NADAL	0.162	RAFAEL-NADAL	0.162	RAFAEL-NADAL	0.162	RAFAEL-NADAL	0.162
3	NOVAK-DJOKOVIC	0.157	NOVAK-DJOKOVIC	0.157	NOVAK-DJOKOVIC	0.157	NOVAK-DJOKOVIC	0.157
4	DAVID-FERRER	0.149	DAVID-FERRER	0.149	DAVID-FERRER	0.149	DAVID-FERRER	0.149
5	ANDY-MURRAY	0.133	ANDY-MURRAY	0.133	ANDY-MURRAY	0.133	ANDY-MURRAY	0.133
6	TOMAS-BERDYCH	0.131	TOMAS-BERDYCH	0.131	TOMAS-BERDYCH	0.131	TOMAS-BERDYCH	0.131
7	MIKHAIL-YOUZHNY	0.117	MIKHAIL-YOUZHNY	0.117	MIKHAIL-YOUZHNY	0.117	MIKHAIL-YOUZHNY	0.117
8	ANDY-RODDICK	0.115	ANDY-RODDICK	0.115	ANDY-RODDICK	0.115	ANDY-RODDICK	0.115
9	TOMMY-ROBREDO	0.114	TOMMY-ROBREDO	0.114	TOMMY-ROBREDO	0.114	TOMMY-ROBREDO	0.114
10	FERNANDO-VERDASCO	0.113	FERNANDO-VERDASCO	0.113	FERNANDO-VERDASCO	0.113	FERNANDO-VERDASCO	0.113

No difference in Katz Centrality with different  $\beta$

# Katz Centrality vs Reality

#	Beta 0.1 Alpha = 0.9( $1/\lambda$ )	Won GS	Won Title	GOAT point*	Active Year	Highest Ranking
1	Roger Federer	20	103	201	1998~	1(2004.02)
2	Rafael Nadal	19	84	168	2001~	1(2008.08)
3	Novak Djokovic	16	77	185	2003~	1(2011.07)
4	David Ferrer	0	27	124	2000~2019	3(2013.07)
5	Andy Murray	3	46	57	2005~	1(2016.11)
6	Tomas Berdych	0	13	32	2002~2019	4(2015.05)
7	Mikhail Youzhny	0	10	27	1999~2019	8(2008.01)
8	Andy Roddick	1	32	33	2000~2012	1(2003.11)
9	Tommy Robredo	0	12	32	1998~	5(2006.08)
10	Fernando Verdasco	0	7	22	2001~	7(2009.04)

Katz Centrality is almost aligned with GOAT points rank!

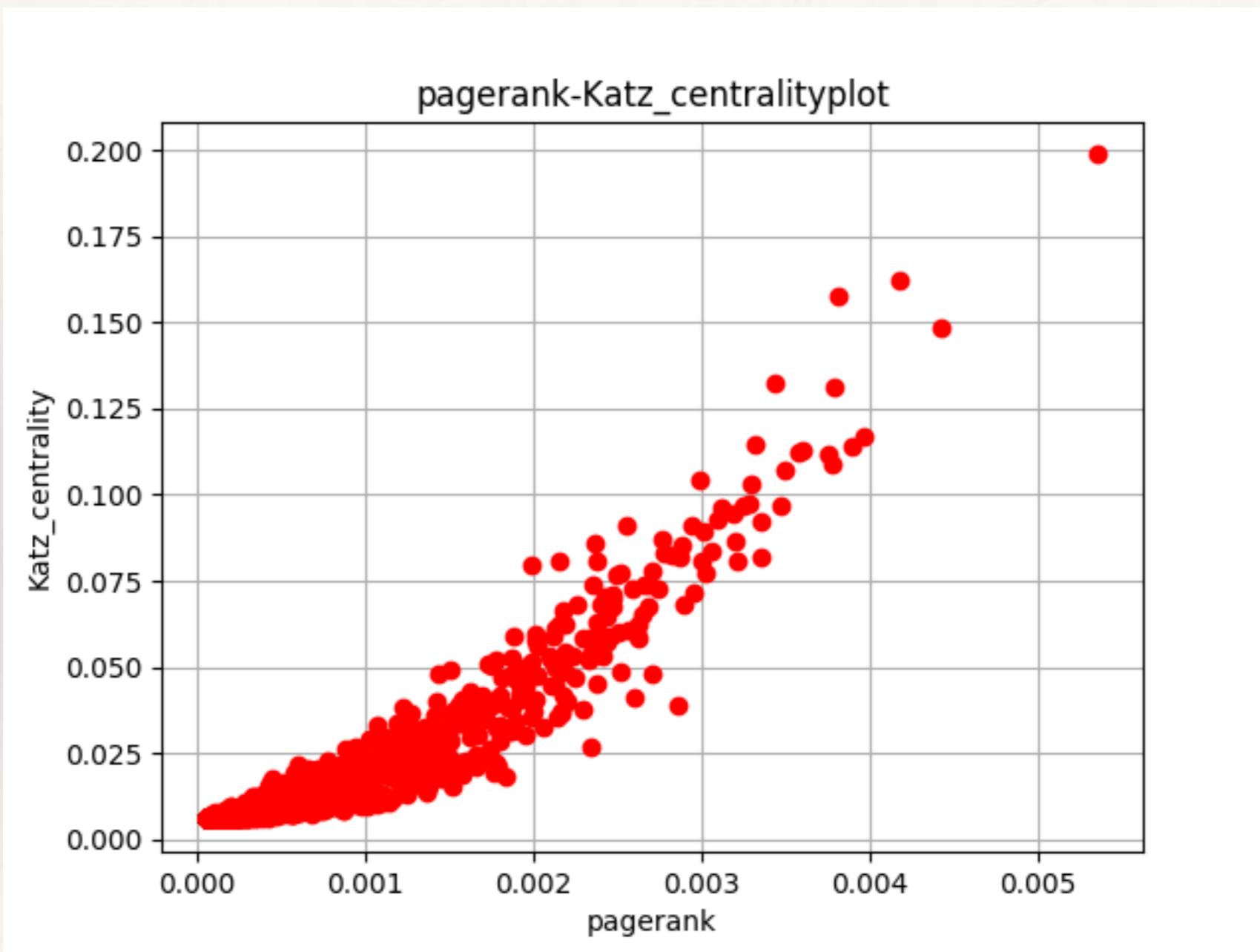
# Page Rank

#	alpha = 0.85	
1	ROGER-FEDERER	0.00536
2	DAVID-FERRER	0.00442
3	RAFAEL-NADAL	0.00417
4	MIKHAIL-YOUZHNY	0.00396
5	TOMMY-ROBREDO	0.00389
6	NOVAK-DJOKOVIC	0.00381
7	TOMAS-BERDYCH	0.00379
8	TOMMY-HAAS	0.00377
9	FELICIANO-LOPEZ	0.00375
10	FERNANDO-VERDASCO	0.00360

# Katz vs Page Rank

#	Katz Centrality		Page Rank	
1	ROGER-FEDERER	0.199	ROGER-FEDERER	0.00536
2	RAFAEL-NADAL	0.162	DAVID-FERRER	0.00442
3	NOVAK-DJOKOVIC	0.157	RAFAEL-NADAL	0.00417
4	DAVID-FERRER	0.149	MIKHAIL-YOUZHNY	0.00396
5	ANDY-MURRAY	0.133	TOMMY-ROBREDO	0.00389
6	TOMAS-BERDYCH	0.131	NOVAK-DJOKOVIC	0.00381
7	MIKHAIL-YOUZHNY	0.117	TOMAS-BERDYCH	0.00379
8	ANDY-RODDICK	0.115	TOMMY-HAAS	0.00377
9	TOMMY-ROBREDO	0.114	FELICIANO-LOPEZ	0.00375
10	FERNANDO-VERDASCO	0.113	FERNANDO-VERDASCO	0.00360

# Scatter Plot: Katz-Page Rank



# Page Rank vs Reality

#	Beta 0.1 Alpha = 0.9( $1/\lambda$ )	Won GS	Won Title	GOAT point*	Active Year	Highest Ranking
1	Roger Federer	20	103	201	1998~	1(2004.02)
2	David Ferrer	0	27	124	2000~2019	3(2013.07)
3	Rafael Nadal	19	84	168	2001~	1(2008.08)
4	Mikhail Youzhny	0	10	27	1999~2019	8(2008.01)
5	Tommy Robredo	0	12	32	1998~	5(2006.08)
6	Novak Djokovic	16	77	185	2003~	1(2011.07)
7	Tomas Berdych	0	13	32	2002~2019	4(2015.05)
8	Tommy Haas	0	15	58	1996~2017	2(2002.05)
9	Feliciano Lopez	0	7	22	1997~	12(2015.03)
10	Fernando Verdasco	0	7	22	2001~	7(2009.04)

Page rank is somehow deviated from GOAT points rank!

---

# Observation

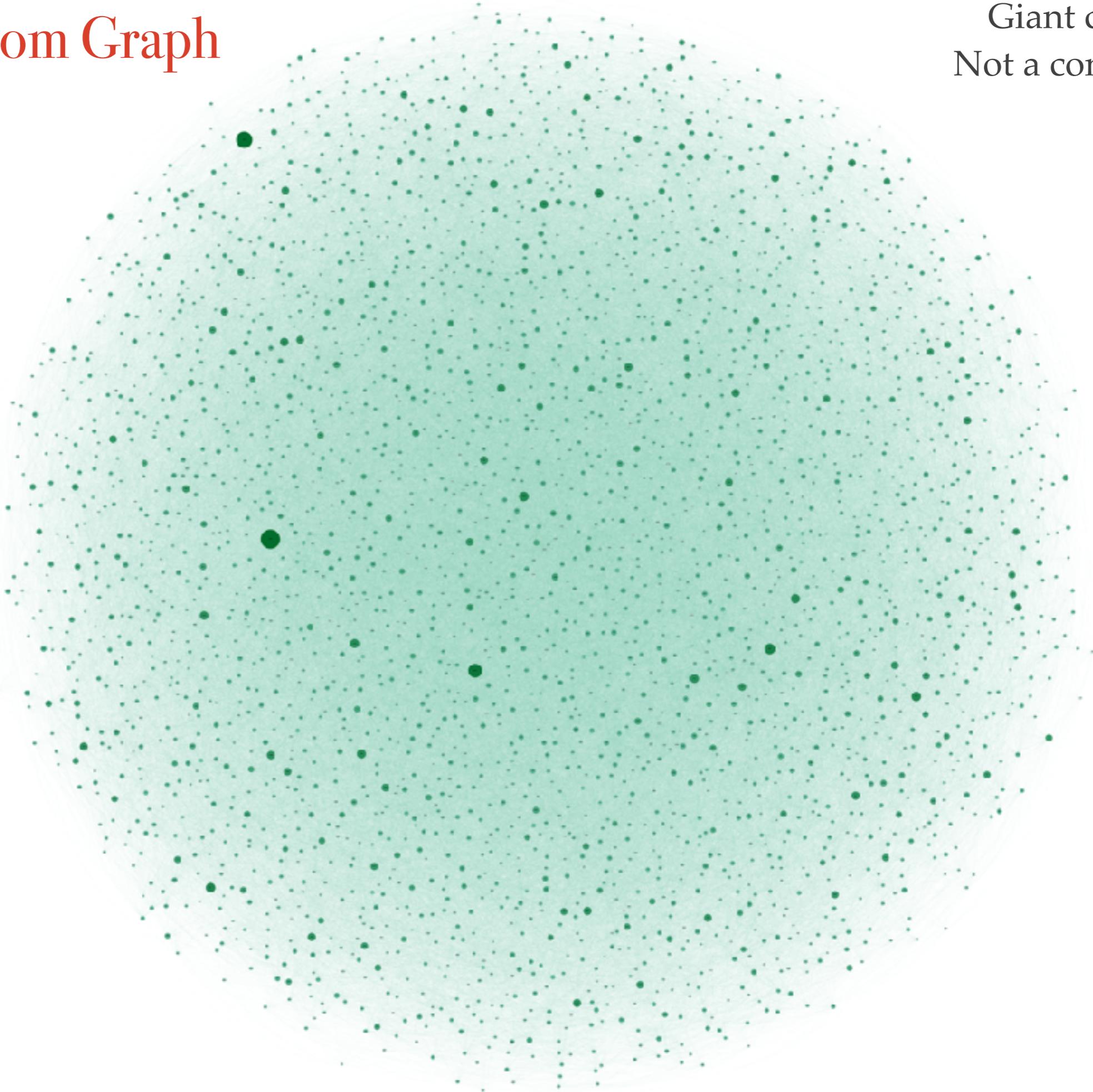
---

- ❖ Harmonic and Betweenness Centrality is uncorrelated to evaluation player
- ❖ Degree and Eigen Centrality is highly correlated to the evaluation player, almost aligned with ATP ranking
- ❖ **The quality of a player is reflected by his opponent**

# Graph Model

# Random Graph

Giant component  
Not a complete graph



---

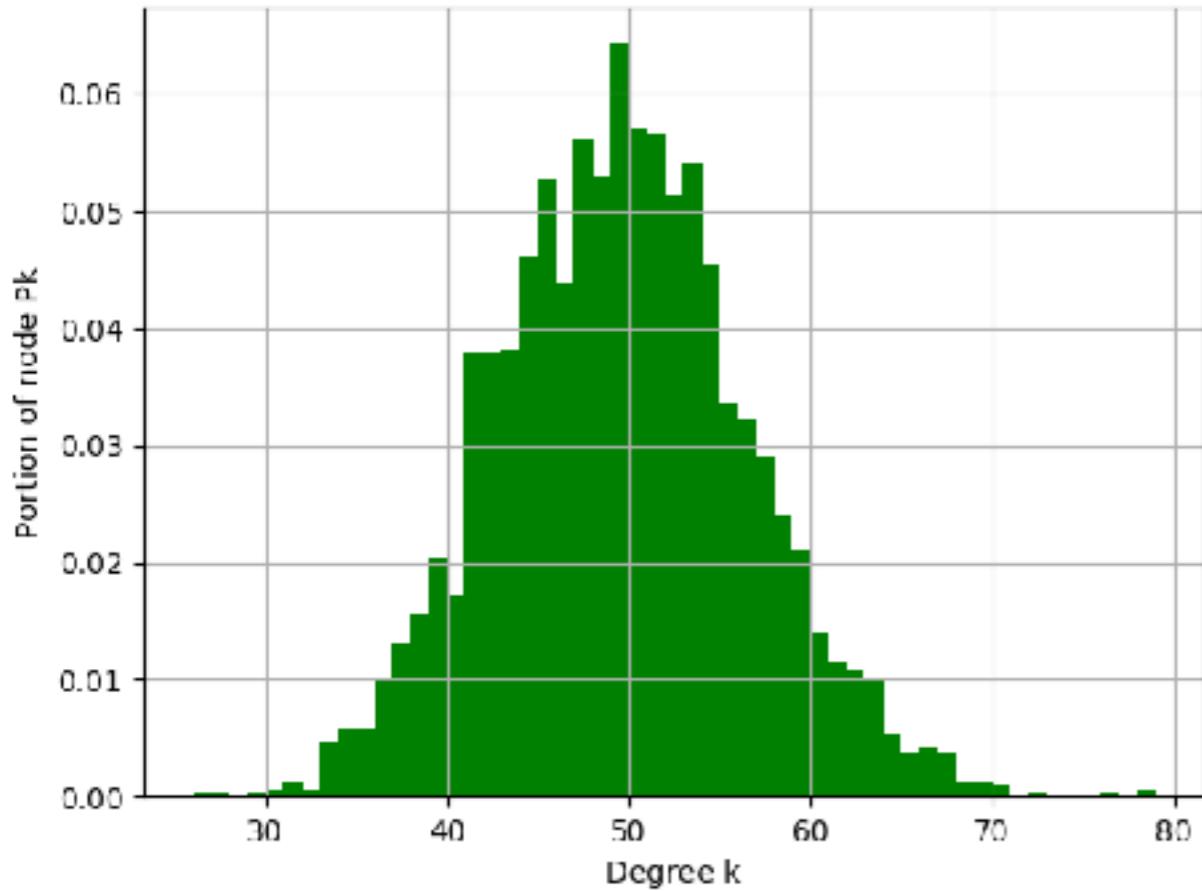
# Random Graph: Metrics

---

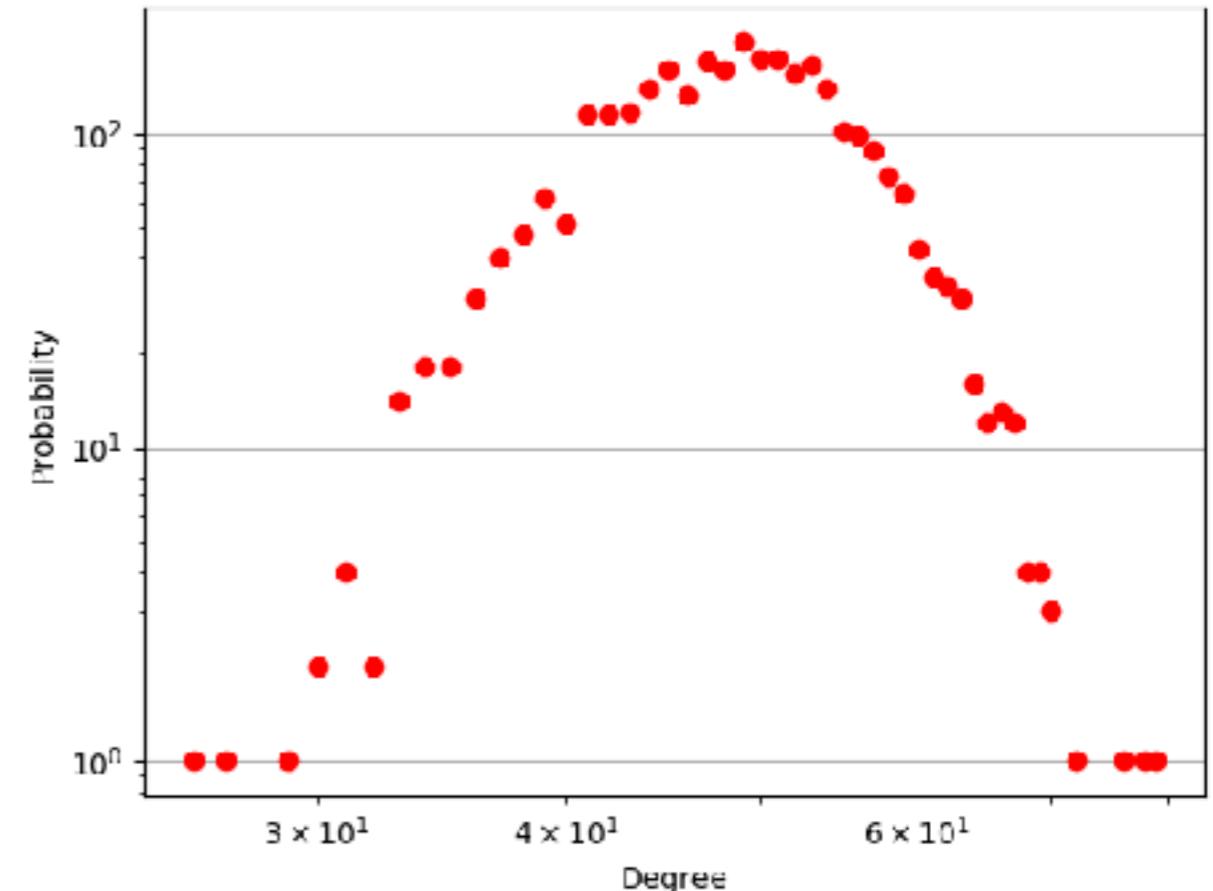
- ❖ average degree: 49.154
  - ❖ order: 3035
  - ❖ size: 74591
  - ❖ density: 0.016201
  - ❖ diameter: 3
  - ❖ radius: 3
  - ❖ average path length: 2.427481
  - ❖ average clustering coefficient: 0.016153
  - ❖ transitivity: 0.016151
  - ❖ number of triangle: 19729
  - ❖ number of clique: 53269
  - ❖ number of component: 1
- ❖ N node size = 3035,
  - ❖ Probability =  $\langle k \rangle / (N - 1) = 0.016$
  - ❖  $\langle k \rangle = 49.075$
  - ❖ Supercritical:  $\langle k \rangle = 49.154 > \ln 3035 = 8$

# Degree Distribution

4a: Degree Distribution Histogram:

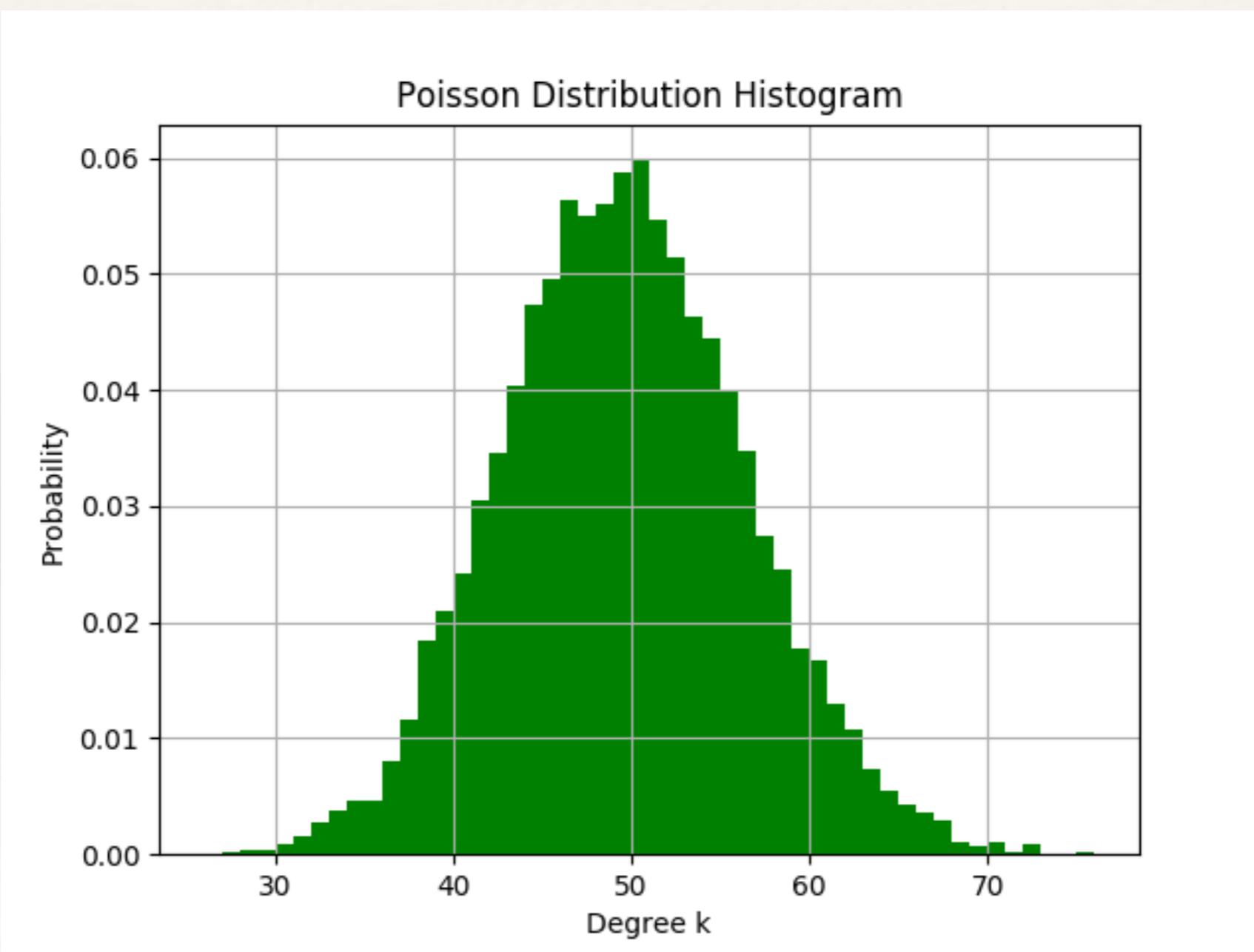


4a: Degree Distribution

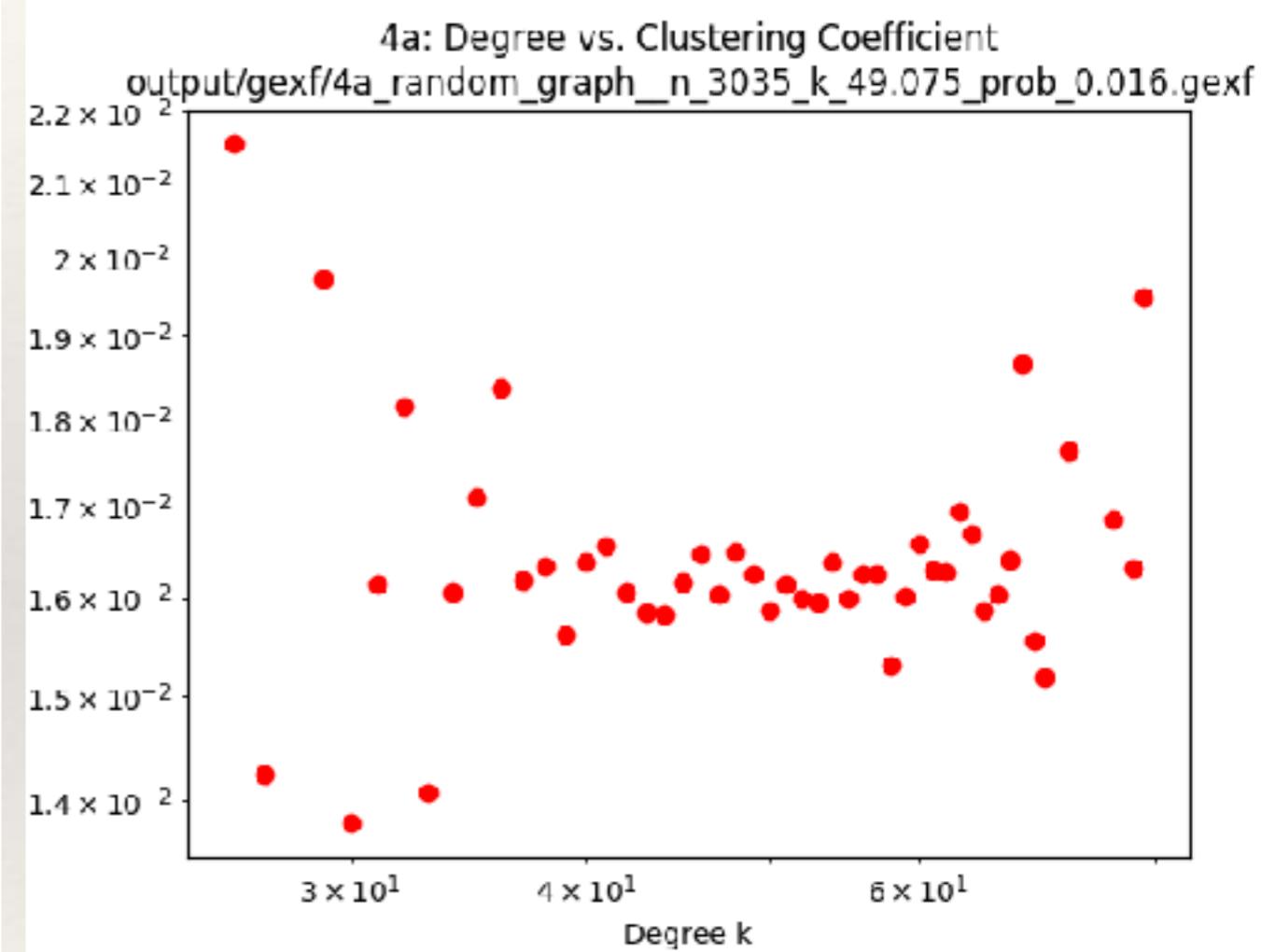
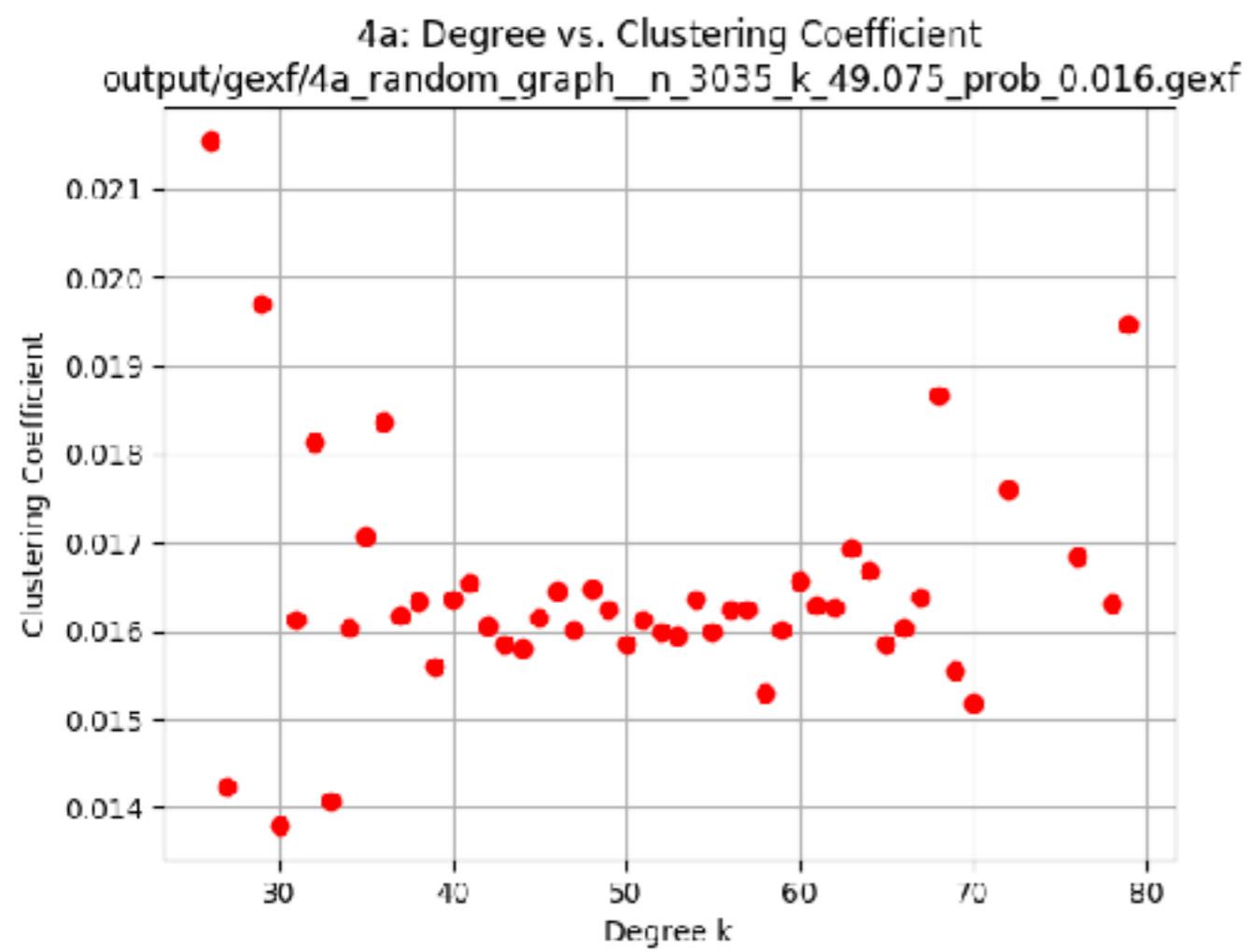


# Poisson Distribution

- ❖  $\lambda = 49$



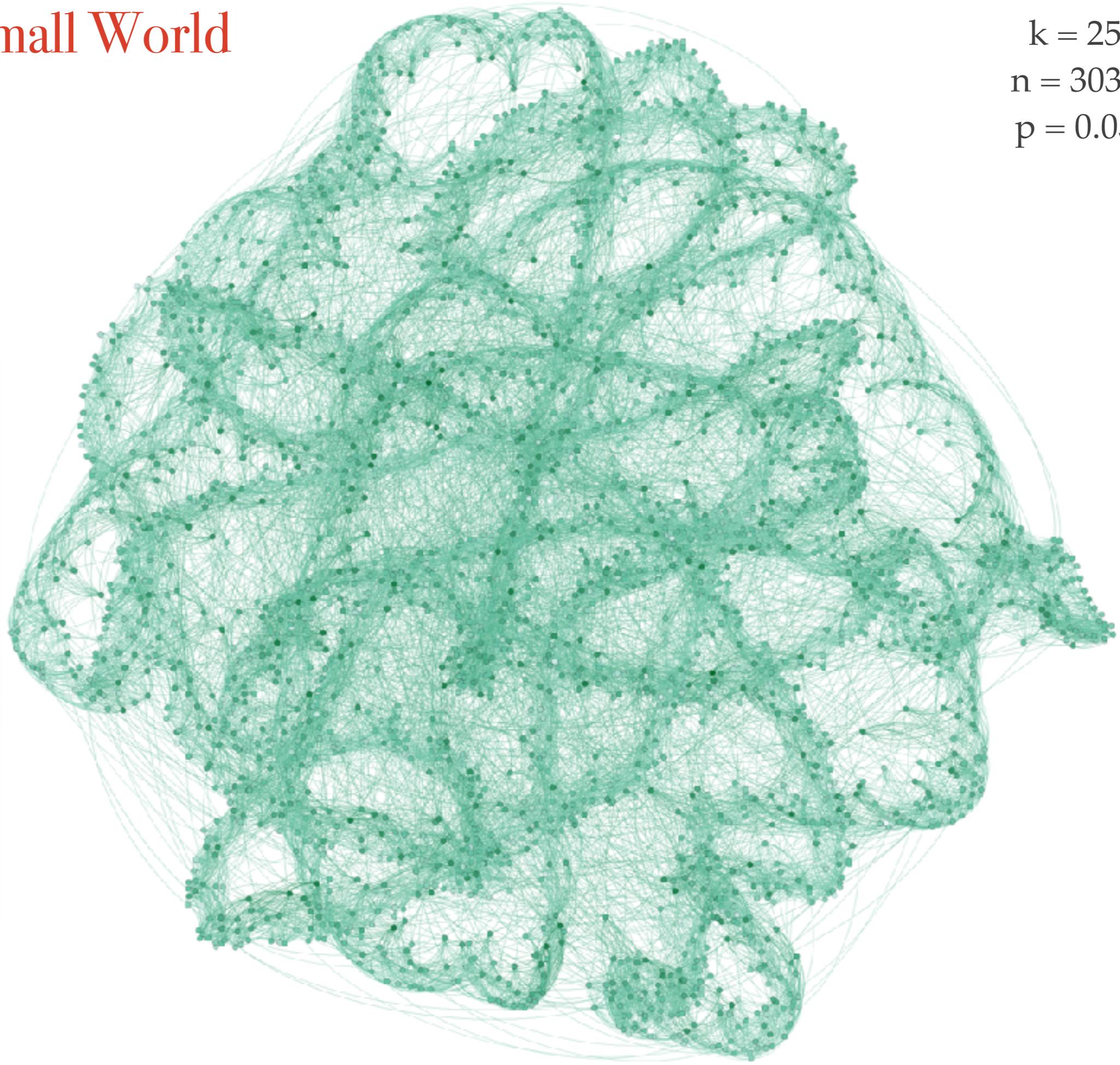
# Clustering vs. Degree



# Small World

n = 3035, k = 25							my network
p	0.001	0.01	0.05	0.1	0.5	1	
small wordness	0.062	0.128	<b>0.156</b>	0.149	0.032	0.002	0.051
average path length	11.348	5.430	3.953	3.530	2.894	2.829	3.263
clustering coefficient	0.714	0.696	0.618	0.528	0.094	0.007	0.166

# Small World

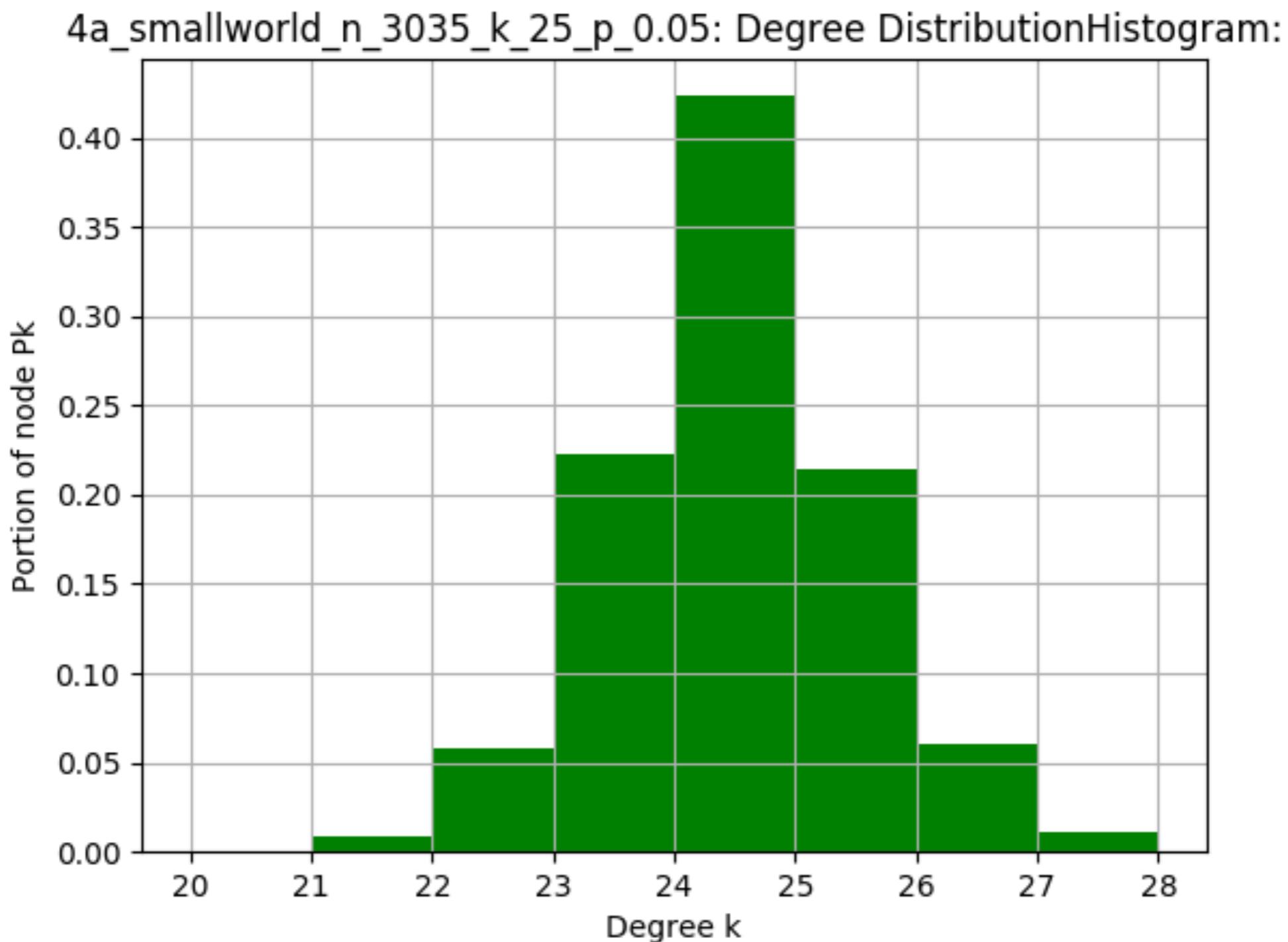


$k = 25$   
 $n = 3035$   
 $p = 0.05$

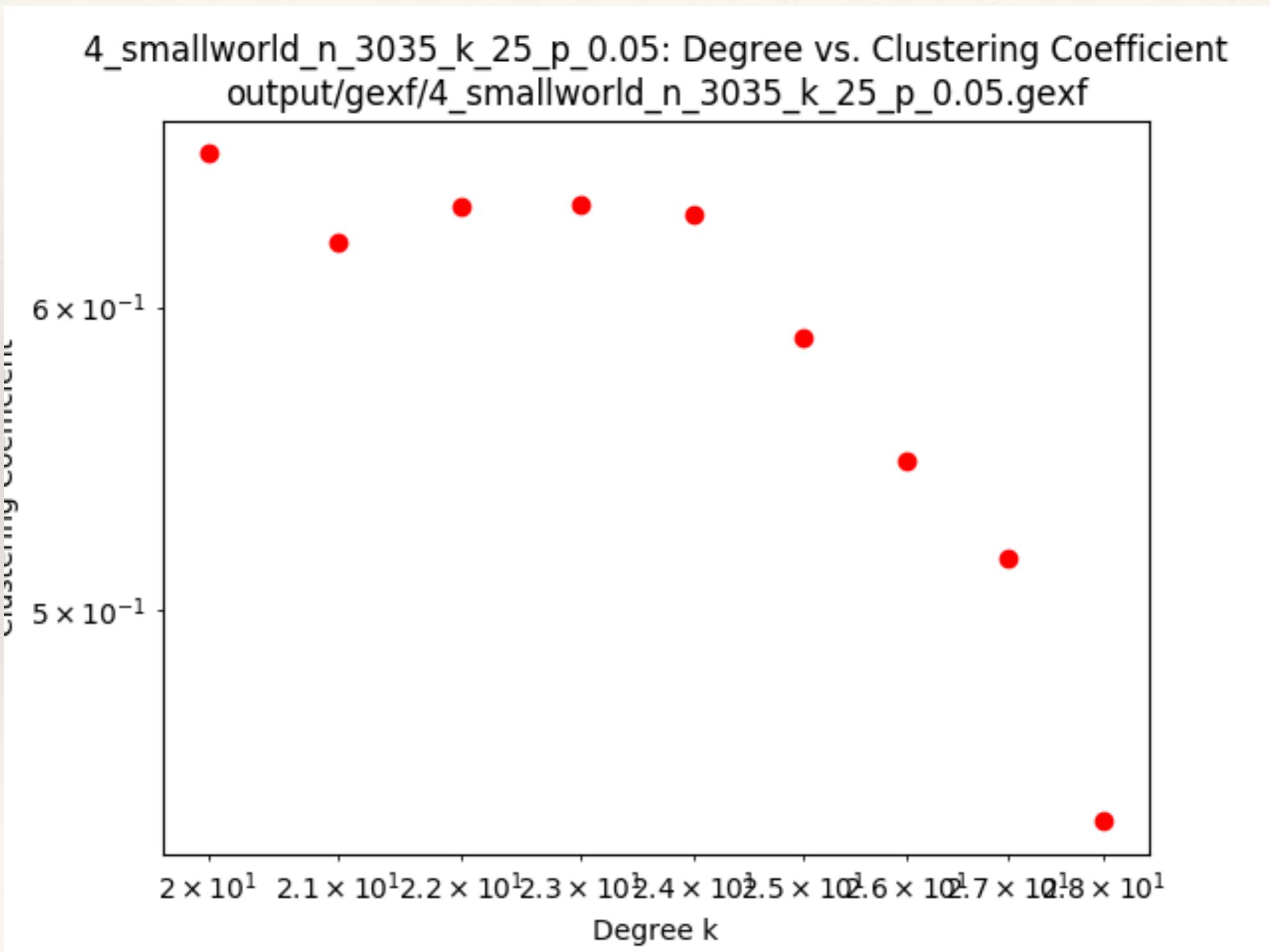
# Small World: Metric

- ❖ order: 3035
  - ❖ size: 36420
  - ❖ density: 0.007910
  - ❖ diameter: 6
  - ❖ radius: 5
  - ❖ average path length: 3.953023
  - ❖ average clustering coefficient: 0.618777
  - ❖ transitivity: 0.616572
  - ❖ number of triangle: 172503
  - ❖ number of clique: 13260
  - ❖ number of component: 1
- ❖ N node size = 3035
  - ❖ Probability = 0.05
  - ❖  $k = \langle k \rangle / 2 = 25$

# Degree Distribution



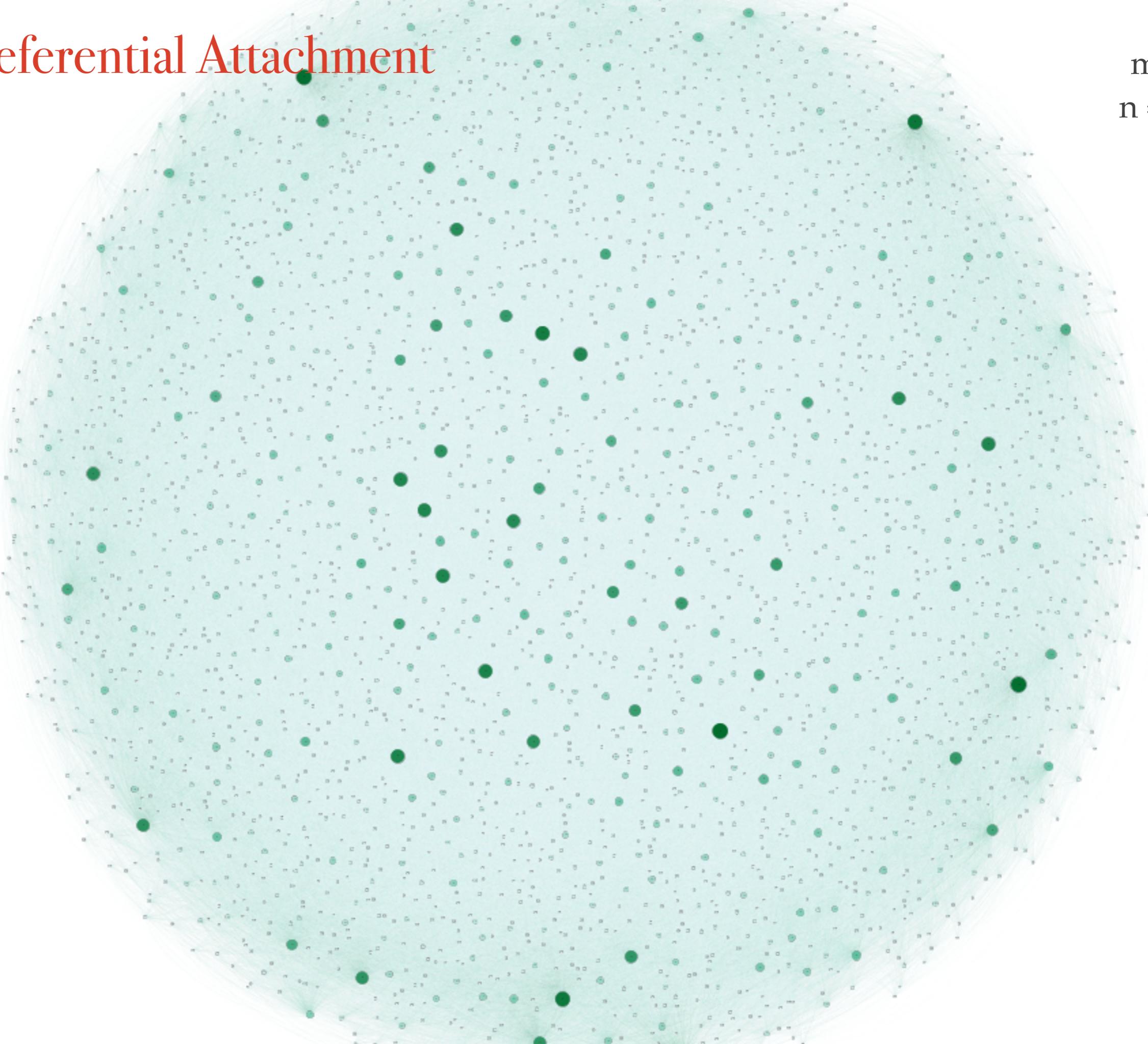
# Clustering vs. Degree



# Preferential Attachment

$m = 25$

$n = 3035$

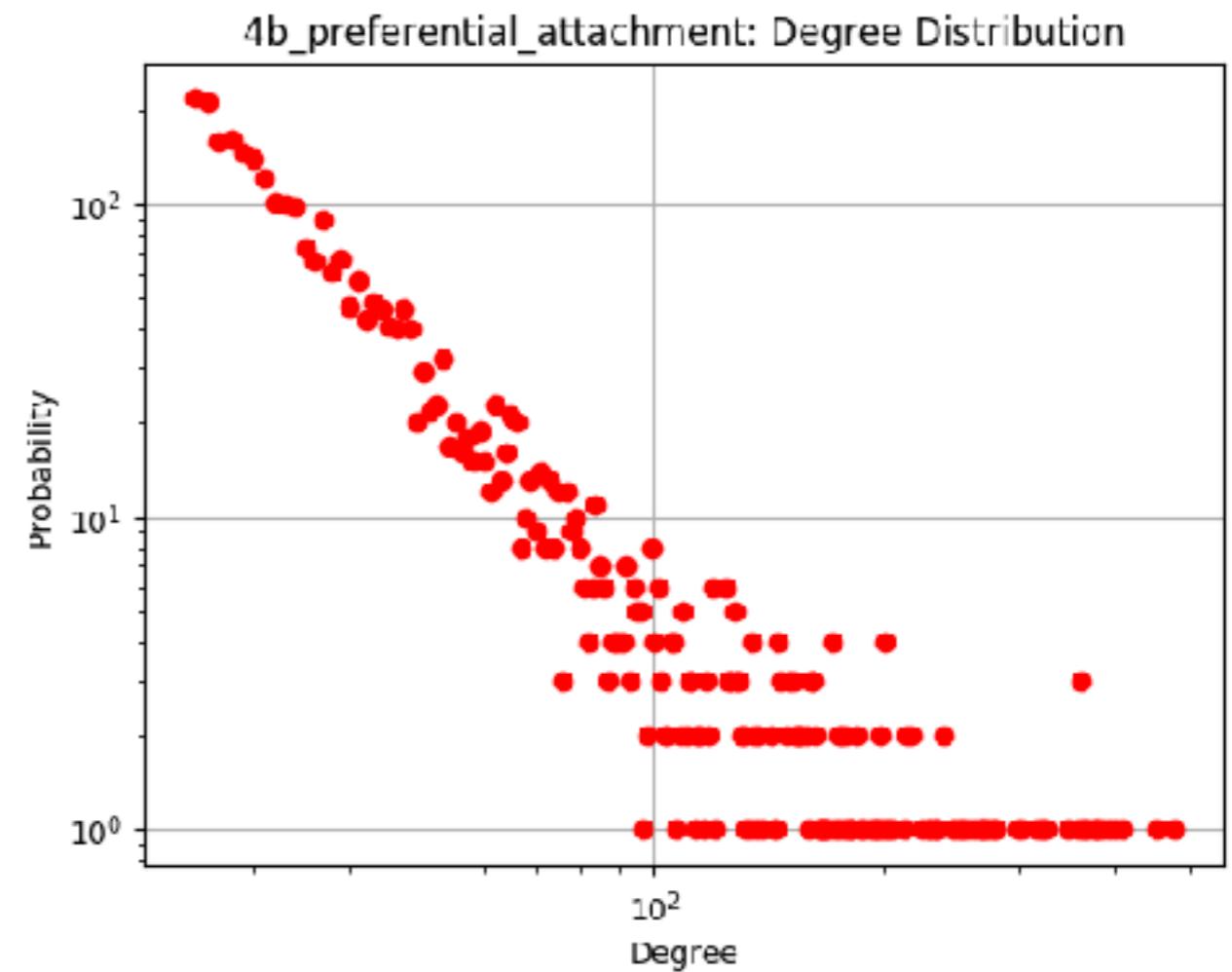
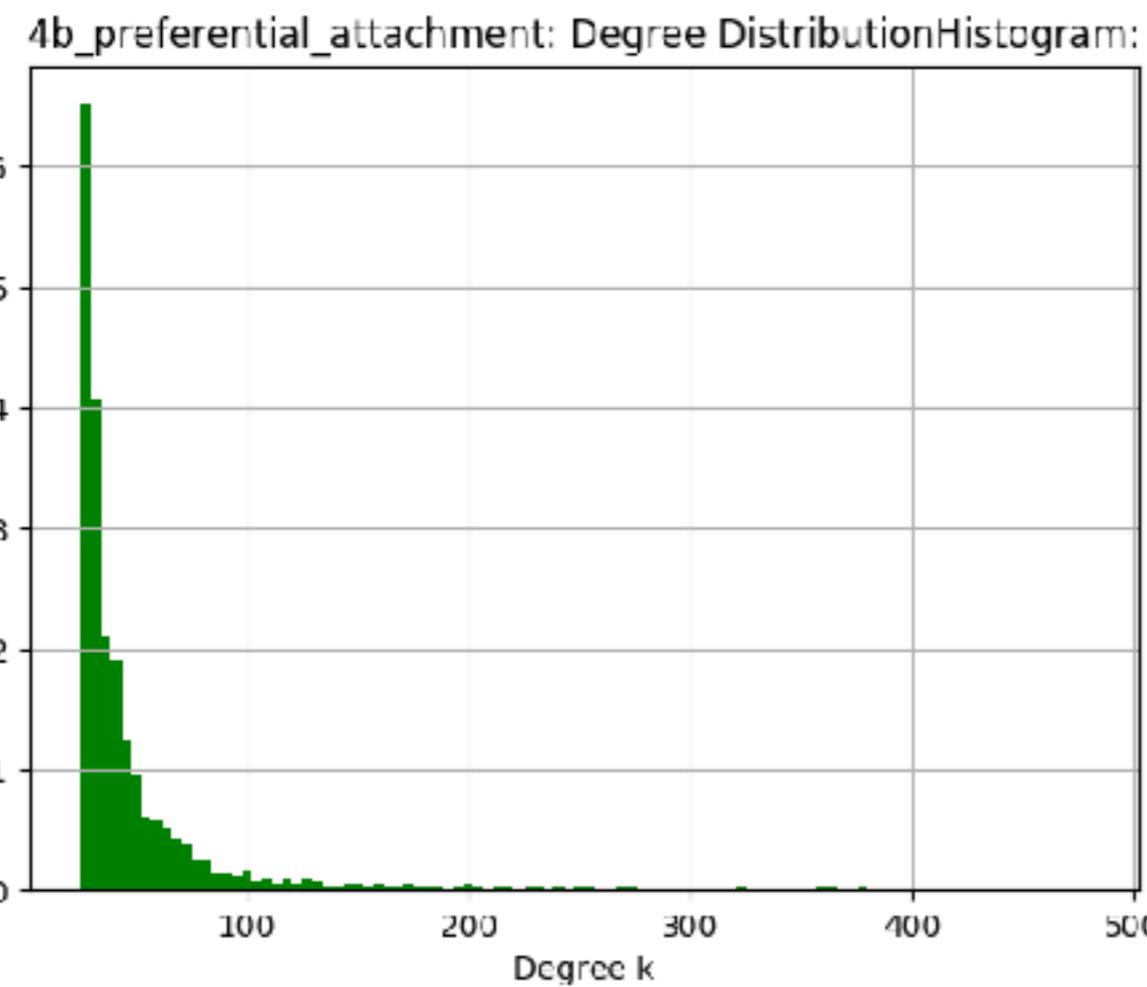


# Preferential Network: Metrics

---

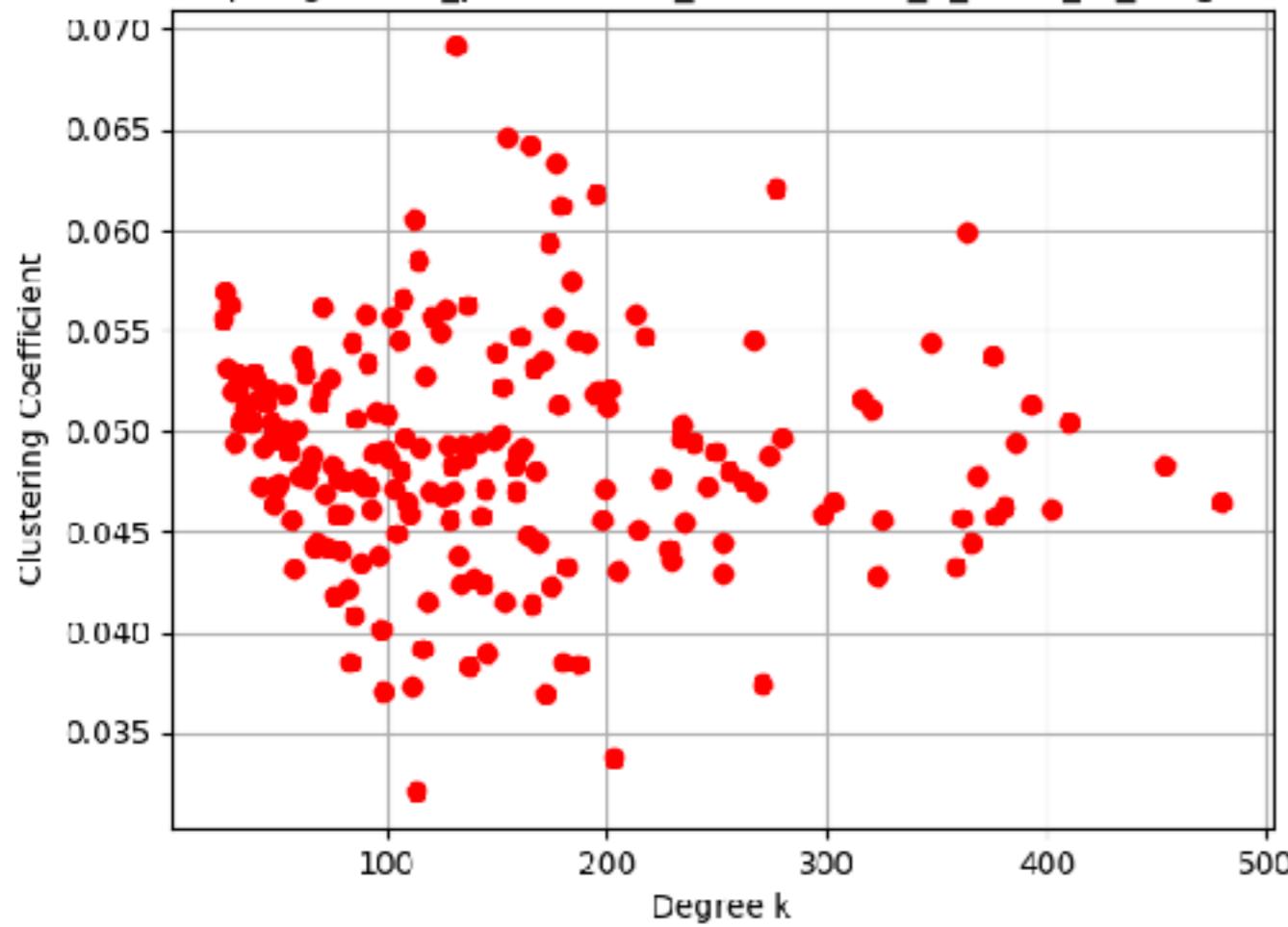
- ❖ average degree: 49.58
  - ❖ order: 3035
  - ❖ size: 75250
  - ❖ density: 0.016344
  - ❖ diameter: 3
  - ❖ radius: 2
  - ❖ average path length: 2.355214
  - ❖ average clustering coefficient: 0.051652
  - ❖ transitivity: 0.049413
  - ❖ number of triangle: 108398
  - ❖ number of clique: 80531
  - ❖ number of component: 1
- ❖ N node size = 3035,
  - ❖  $m = 25 \sim \langle k \rangle / 2$
  - ❖ Supercritical:  $\langle k \rangle = 49.154 > \ln 3035 = 8$

# Degree Distribution

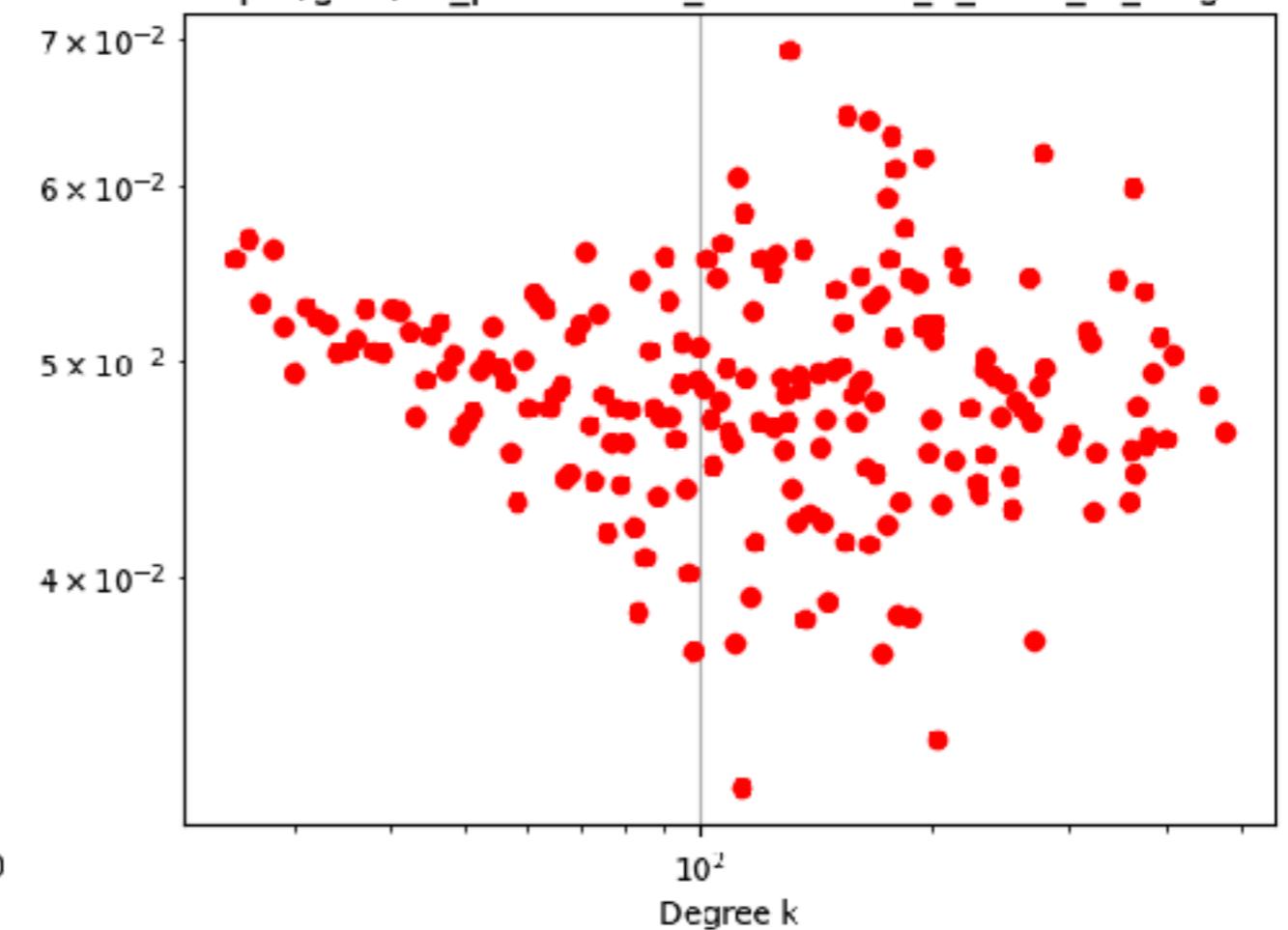


# Clustering vs. Degree

4b\_preferential\_attachment: Degree vs. Clustering Coefficient  
output/gexf/4b\_preferential\_attachment\_n\_3035\_m\_25.gexf



4b\_preferential\_attachment: Degree vs. Clustering Coefficient  
output/gexf/4b\_preferential\_attachment\_n\_3035\_m\_25.gexf

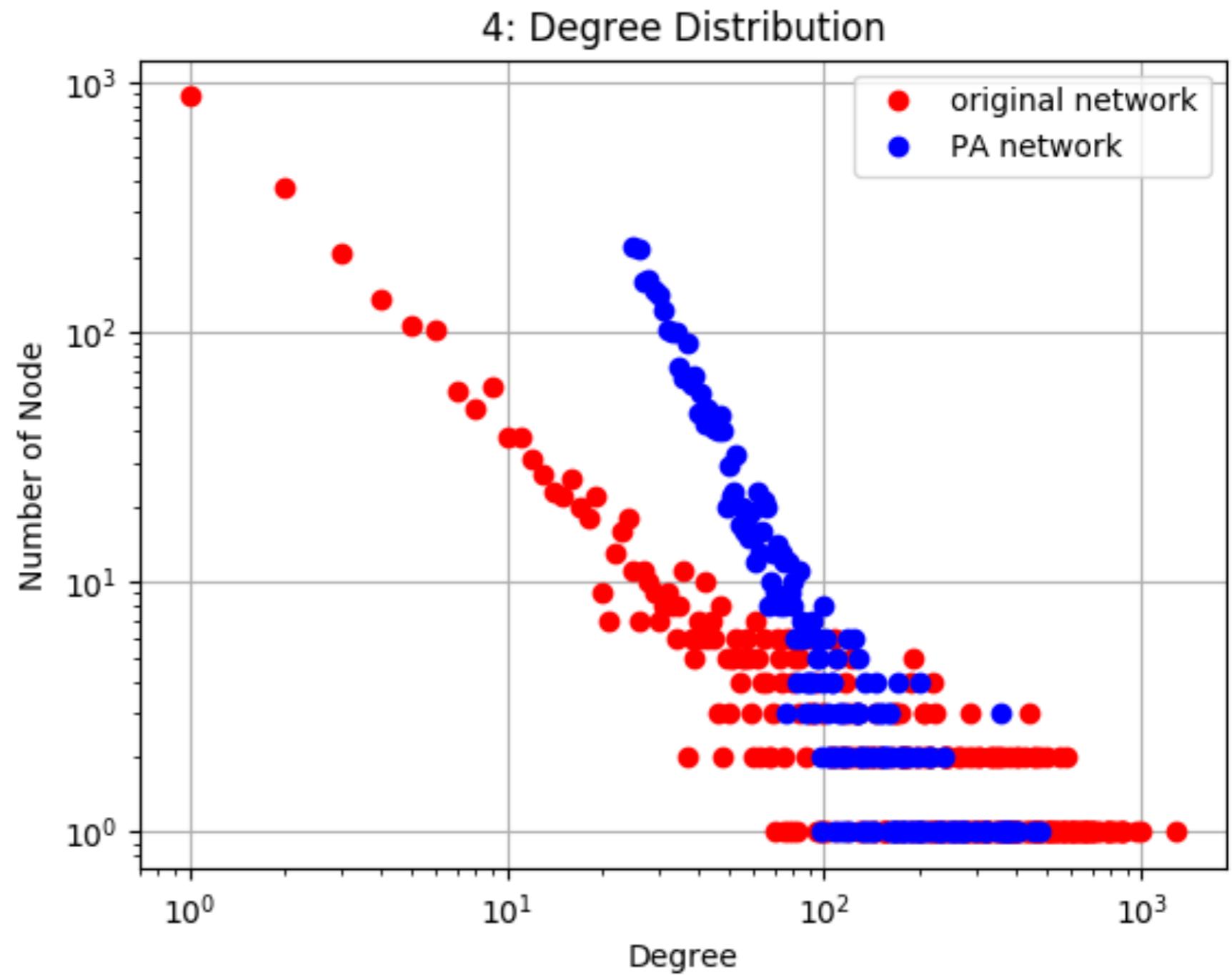


# Scale Free $\alpha$ Comparison

	preferential network	my network
--	----------------------	------------

alpha	3.71	2.11
kmax	480	1290
kmin	25	1
N	3035	3035

$$k_{max} = k_{min} N^{\frac{1}{\alpha - 1}}$$



---

# Conclusion

---

- ❖ Simple graphical analysis on network is useful for tennis player evaluation
- ❖ Both Eigen and Katz centrality gives a good result for player evaluation
- ❖ Player network exhibits scale-free model with  $\alpha \sim 2.1$

---

# Future Work

---

- ❖ New rankings can be explored through different match statistics
- ❖ Constructing & Comparing network for every decade

---

# Source Code & Raw Data

---

- ❖ Raw Data: [https://datahub.io/sports-data/atp-world-tour-tennis-data#resource-match\\_scores\\_1991-2016\\_unindexed](https://datahub.io/sports-data/atp-world-tour-tennis-data#resource-match_scores_1991-2016_unindexed)
- ❖ My Code: [https://github.com/kuangyu0801/WS19\\_ComplexNetworkSystem](https://github.com/kuangyu0801/WS19_ComplexNetworkSystem)

# Appendix & Reference