A Social Network Analysis of NBA All-Stars

Biancheng Wang[1]

Abstract

In this paper, we focus on the relationship of public figures using statistical network analysis methods. Specifically, we analyze the network of Instagram following relationship of NBA All-Stars in the most recent 3 seasons. We find that there might be an International "gang" within this network by Latent Space Model but it is not so clear. Triad Census Model implies that this network is more likely to be transitive. With the help of ERGM, we analyze this network from 4 aspects. We find the strong Nationality and Team homophily effects. As for Business Factors, it does show some Brand homophily effect. Also, players who have more NBA experience, more honors and are more active in the social media are more likely to form a tie with others.

Keywords: Network Data, Centrality, Latent Space Model, ERGM

1. Introduction

Many previous researches have implemented analysis on the friendship network, e.g., adolescent friendship using Add Health data and relationship between monks using Sampson's monastery data. However, those relationships are not so complex when comparing with the network of public figures. Adolescents are simpler and more naïve while the groups of monks are quite clear. One might be curious about what kind of factors may influence the relationship network of public figures.

In this paper, we choose professional basketball players, specifically, NBA All-Stars from the most recent 3 years, as our research object. As the best athletes all over the world in this area, it is obvious that they know each other. We can always hear some news about some of NBA players working together during the off-season, which might imply that there exist small groups within these talented. Since it is hard to figure out the true friendship between each other, so we focus on their relationship presented on the social media. Who do they follow on Instagram?

Since this is not a real friendship network, more aspects should be considered in this case. We think the relationship on the social media might be related to the following four aspects: Demographical Characteristics, Social Media Characteristics, Basketball Factors and Business Factors. We mainly use ERGM (Exponential random graph models), a probability model for adjacency matrices, to answer how they influence the formation of ties in a network between public figures.

In the next section, we describe the dataset we have and give visual and numeric overview of our network. Section 3 is dedicated to the social network analysis and we try to figure out some

[1] Department of Statistics, UCLA (email: wangbcbill@ucla.edu)

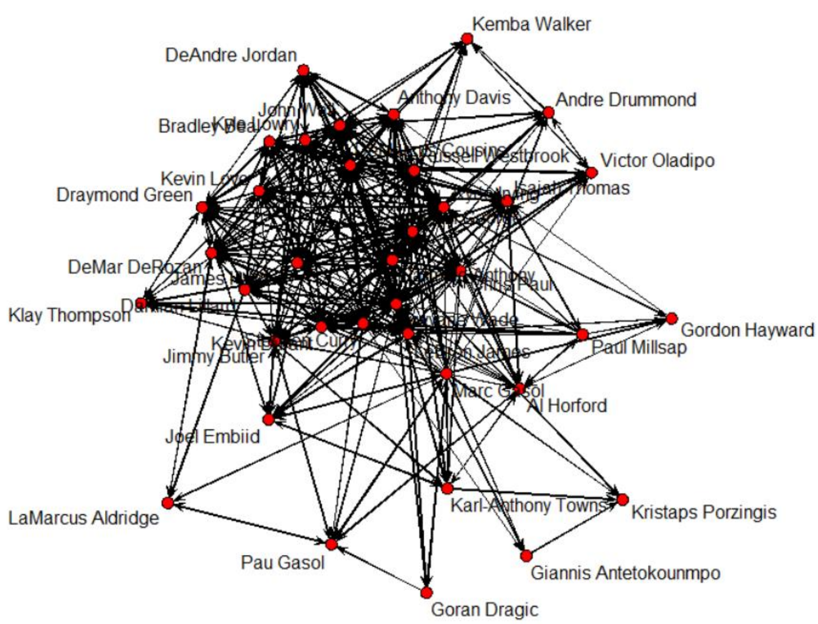
pattern within this network and what characteristics may impact the relationship. Finally, we conclude and give more discussion in Section 4.

2. Data

The dataset we work on in this paper is mostly collected from Wikipedia and Instagram on December 7th, 2018. There are 38 NBA All-Stars from 2015-2016, 2016-2017 and 2017-2018 seasons. However, Kawhi Leonard hardly use social media and does not have a verified Instagram account. Therefore, we only collect data of 37 NBA All-Stars in our network.

It contains two networks. The first one is a network describing the following relationship between those players, which is shown as follow in Figure 1. This is a directed network with 37 vertices and 489 edges. Since there are 1332 possible edges, the density of ties in this network is 0.367. There is only one component which means there is no one in this network who does not follow anyone else and is not followed by others. There is a large cloud with dense ties in the center of the Figure 1 while some players seem to have less ties. In the next following sections, we will give more specific analysis regarding this network.

Figure 1: Instagram Following Network of NBA All-Stars



The second one is a network of teammates relationship. Not only this season but also previous experience of being teammates in NBA are included in this network. This is an undirected network with 37 vertices and 50 edges. Since there are 666 possible edges for this undirected network, the density of ties is 0.075 which is much sparser than the first network. From Figure 2, we can find that 2 large components and 5 small components including 4 isolates. Kemba Walker, Marc Gasol, Giannis Antetokounmpo, and Andre Drummond do not work with any other players in this network. The average of geodesic distances is 1.700. This is a measure of the average length of the shortest path between two players. Although we mainly focus on the first network, we still think plotting Teammates network might be helpful to give us an additional information and better understanding about the relationship among those players.

Figure 2: Teammates Network of NBA All-Stars

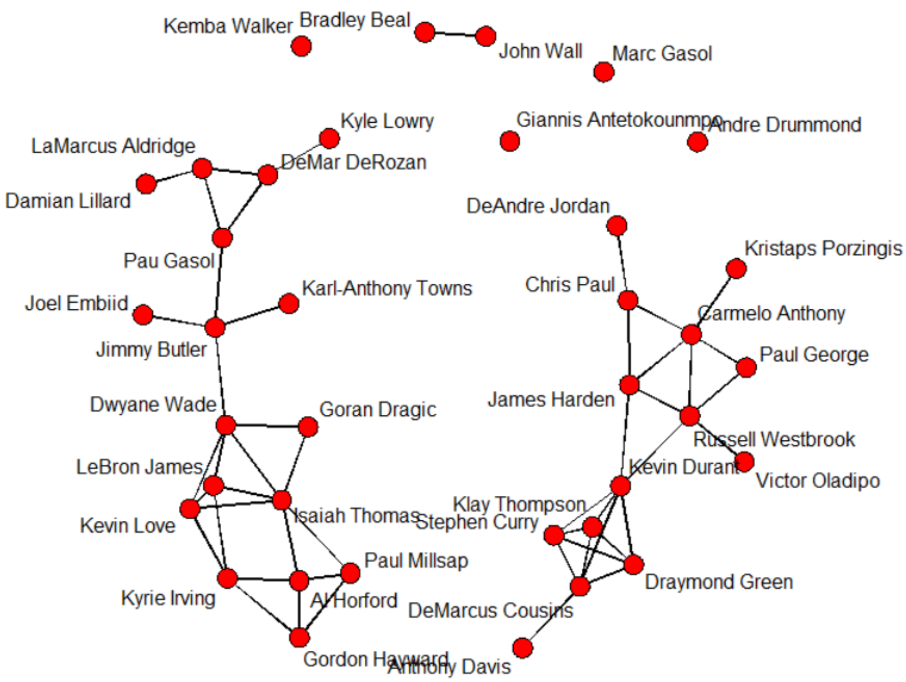


Table 1 gives us all 22 original variables we have, which are regarded as node attributes in network analysis. We can divide them into 4 categories. Within the Basketball Factor, there still exists different types of covariates. University, years in NBA and the number of team changed can stand for his overall basketball experience while the number of draft pick, MVP, All-Star and Champion can be a set of representative variables related to their capacity. We may also define some new variables based on these original variables in the next section.

Table 1: Node Attributes of Network

Demographical Characteristics	Race	Basketball Factor	University
	International		Experience_NBA[2]
	Nationality		Num_of_Team_Change
	State		Draft[3]
	Age		Num_of_MVP
Social Media Characteristics	Ins_Post		Num_of_All_Star
	Ins_Follower		Num_of_Champion
	Ins_Follows		Team

[2] Years in NBA.
[3] The Number of Draft Pick which means the player's Draft Order.

Business Factor	Agent[4]		Division
	Agency[5]		Conference
	Brand		Position_detailed[6]

3. Analysis and Results

3.1 Degree and Centrality Analysis

To capture the sociality of the network, we calculate the indegree and outdegree for each player in 2nd and 3rd columns of Table A1 in Section 5. In our network, the indegree of a player stands for how many players in this network follows him on Instagram while the outdegree stands for how many players he follows. Therefore, we think indegree can tell us how popular this player is or at least how popular he seems to be on the social media.

We can find that the Top 5 players with largest indegree are Kyrie Irving, Carmelo Anthony, Paul George, Dwyane Wade and Lebron James while Goran Dragic, Paul Millsap, Giannis Antetokounmpo, LaMarcus Aldridge and Kristaps Porzingis has the least number of indegree. It is nature to think about what kind of factors will influence the “popularity” of a player. Carmelo Anthony, Dwyane Wade and Lebron James are from 2003 Draft Class which might be the best class in the NBA history and most people believe that they will be a part of NBA Hall of Fame without any doubt. However, those players seem to be not so popular have relatively less NBA All-Star experience. Besides, three of them are also international players. Also, those isolates in Figure 2, tend to have less indegree when comparing with others. This might imply that teammates are more likely to follow each other, which is very understandable.

To better understand the relationship between variables and “popularity”, a correlation plot is shown as below in the 1st row of Figure 3. We define some new variables here. Top_30, Top_14 and Top_3 respectively stand for whether a player is picked at the first-round pick, the lottery pick and the first 3 picks. Generally, gifted player will be picked early during NBA Draft. Change is whether the player changed team during his career. MVP and Champion are whether he has got those titles. We set the median value 4 as a threshold to identify if the player is frequently selected to attend NBA All-Star Game. Since most of players sign their shoes contract with Nike, we add Nike variable which is whether his sponsor is Nike.

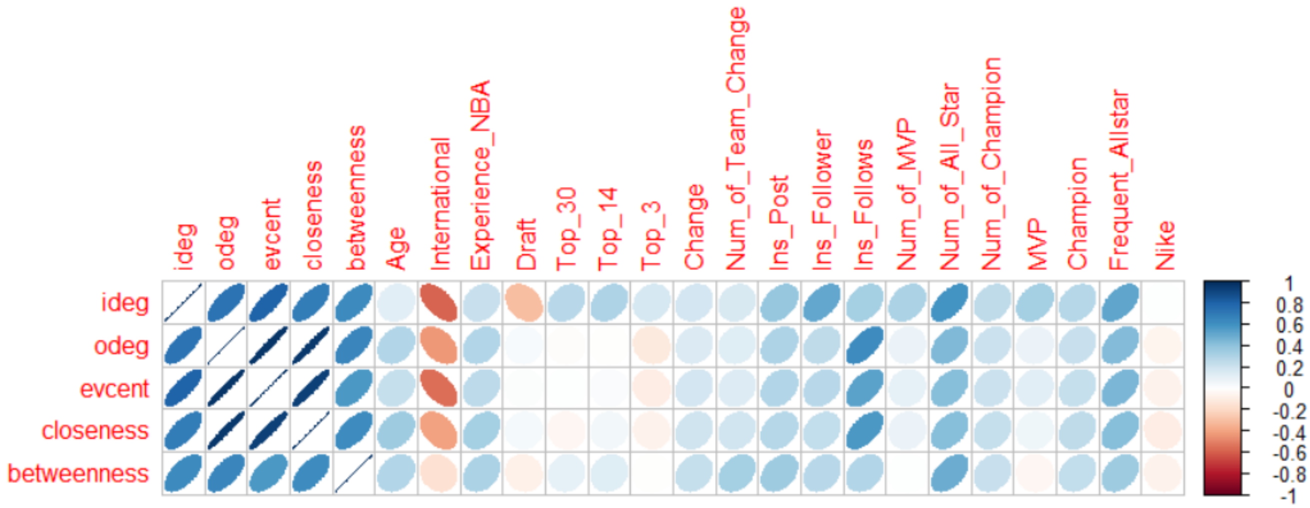
Figure 3: Correlation Plot

---

[4] Data Source: https://hoopshype.com/reps/

[5] Data Source: https://hoopshype.com/reps/

[6] There are 5 types of position in basketball, including Point Guard(PG), Shooting Guard(SG), Small Forward(SF), Power Forward(PF), and Center(C).



From Figure 3, we can find that International players tend to be less popular in the player's network. From the aspect of social media, those with more Instagram followers, which means they are popular all over the world, are also popular among other players. Besides, players with more honors and more gift has more player followers. Those findings can help us to fit ERGM for further analysis.

We are also curious about how their positions in the network vary, in other words, the extent to which they are "central" in the network. There are 4 ways to measure the centrality, including degree centrality, eigenvector centrality, closeness centrality and betweenness centrality. The things they are measuring are not the same.

Degree centrality measures how connected an entity is by counting the number of directed links it has to others in the network. The indegree of a node is the node's number of incoming ties while the out degree of a node is the node's number of outgoing ties. Eigenvector centrality measures how connected an entity is and how much direct influence it might have over other connected entities. Closeness centrality measures the proximity of an entity to others. Betweenness centrality counts the number of paths passing through each entity in the network.

Table 2 gives a summary of those different measures. We also include the overall centrality of the network in Table 2.

Table 2: Statistical Summary of Degree and Centrality

	Min	Median	Mean	Max	Overall
Indegree	2	15	13.22	25	0.3901
Outdegree	1	14	13.22	30	
Eigenvector centrality	0.00078	0.17105	0.14519	0.28406	0.1427
Closeness centrality	0.2927	0.6102	0.5956	0.8571	0.2762
Betweenness centrality	0.1769	13.37	26.8919	134.431	0.0877

Besides, more details are given in Table A1 in Section 5. It has shown the indegree, outdegree and different centrality measures for each player. Also, we plot the correlation between centrality measures and other variables in Figure 3. We can find those different measures are highly correlated with each other. Compared with "popularity", a quite similar conclusion can be made for Centrality. American players with more honors and more Instagram activities, including the number of Instagram Posts and Instagram Follows, are more centered in this network.

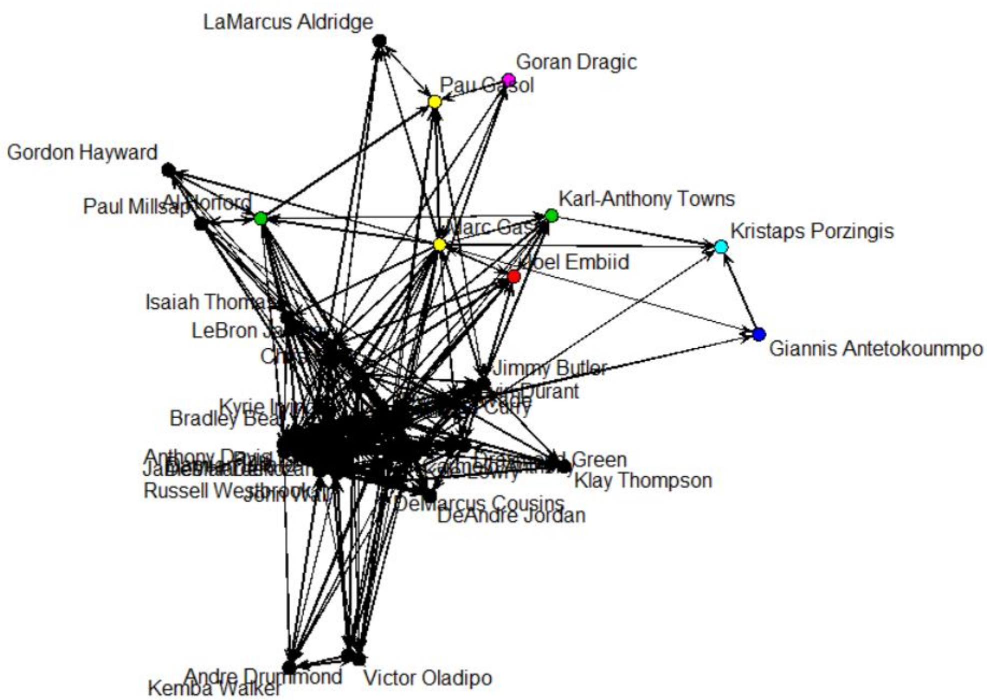
3.2 Latent Space Model

For this section, we want to find out whether there exist some sub-groups within this network. For example, is there any "International gang" or "OG gang" among those talented athletes? A latent space model is the best way to capture the cluster of nodes. We can use it to represent network in a lower dimension and preserve their relative distance at the same time.

Therefore, we fit a two-dimensional latent space model. Since we are more interested the International players group versus American players group and veteran groups versus rookie groups, we plot the results of model with International players marked in Figure 4 and with frequent NBA All-Star Game attendants marked in Figure 5.

Figure 4: Instagram Following Relationship of NBA All-Stars

(International Players marked)

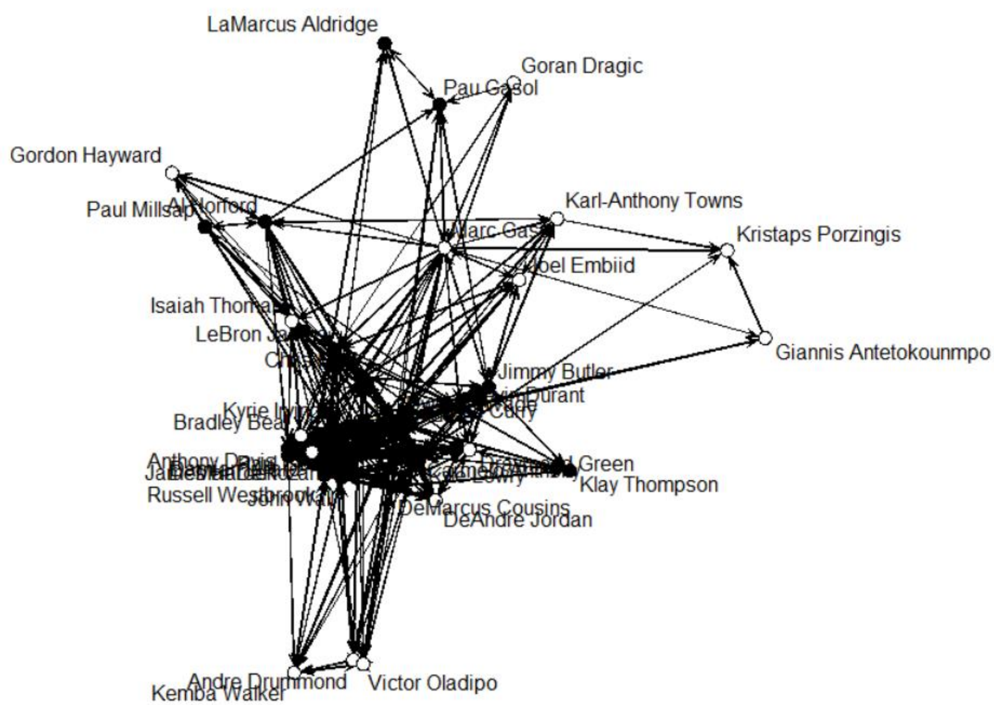


From the plot above, it is hard to find clear groups in this network. However, we can still find some patterns with the help of marks. Many players are clustered at the middle of the plot. Those isolates in Figure 2 are far from the center.

International players which has colorful marks in Figure 4 are located at the upper right of the plot. But we can find LaMarcus Aldridge is closer to those International players. That might be caused by the culture of his team San Antonio Spurs where there are many International players in this decade. Also, Al Horford, a Dominican professional basketball player seems to be close to the American group. That is also reasonable since he moved to the United States with his family when he was in the middle school and played his high school and university games in this country. There is a common characteristic among International players who are very far from the American group, e.g., Goran Dragic, Pau Gasol, Kristaps Porzingis and Giannis Antetokounmpo. They did not have either high school or university level basketball experience in the United before entering NBA. Thus, although the group is not so clear, we can find some homophily effect based on the plots.

Figure 5: Instagram Following Relationship of NBA All-Stars

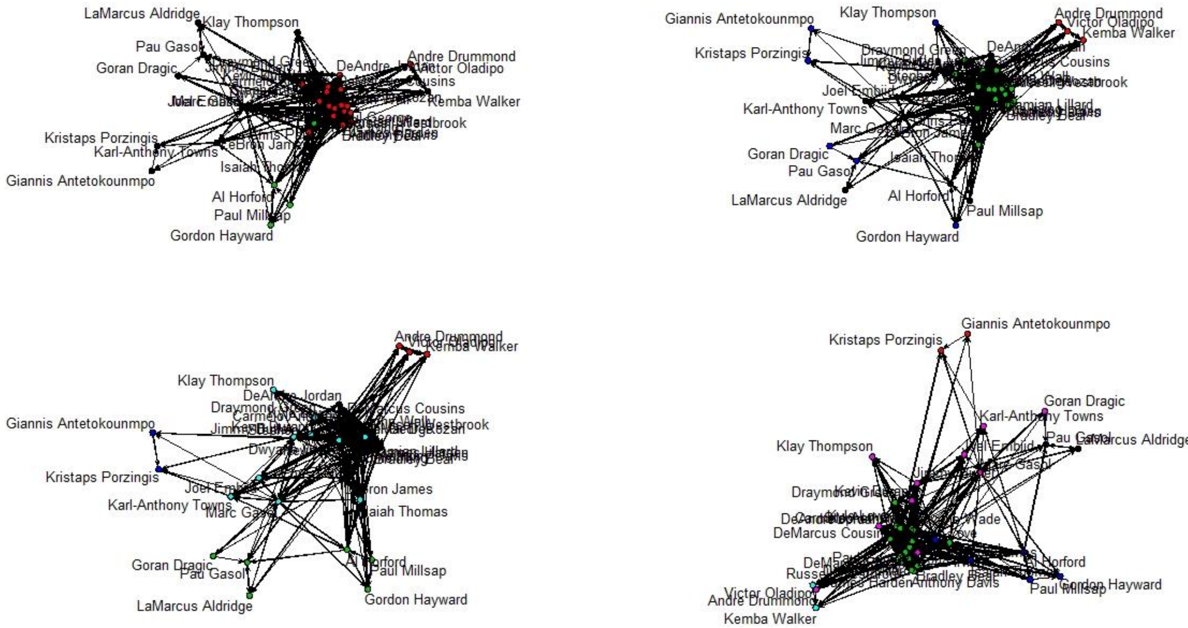
(Frequent NBA All-Star Game Attendants marked)



From Figure 5, the clustering is not as clear as what is shown in Figure 4. There are still some patterns can be found as follows. Most of players in the middle have more NBA All-Stars experience. The lower side of players are all Eastern Conference players who just have only 1 or 2 NBA All-Stars experience and has no superstar teammates. The upper right are all International rookies. Besides, an interesting fact is that the lower right group of players are all from Golden State Warriors.

Figure 6: Instagram Following Relationship of NBA All-Stars

(3,4,5 and 6 sub-groups' Bayesian posterior clusters marked)



For further analysis, we try to fit two-dimensional latent cluster model with different number of sub-groups from 3 to 6 and plot them in Figure 6. But the result is not so convincing especially when looking at the posterior probability. Some of players, e.g., Jimmy Butler and Isaiah Thomas, are quite ambiguous to be classified into any group using all these models.

Therefore, we think the network of public figures showed on the social media is too complex and we cannot simply divide them into clear sub-groups. But the findings we get from the visualization of network by latent space model do help.

3.3 Triad Census Model

Transitivity has been found in many friendship networks. How about this relationship among public figures? Since it is just a kind of relationship shown on the social media, which might be related to business and other factors instead of purely true friendship, the results may be different. In this section, we want to answer this question regarding social balance and transitivity in this network using ERGM. Table 3 shows the number of all 15 types of triad appearing in this network. It is obvious that this network is not balanced since a graph is balanced if all subsets of 3 nodes are balanced.

Table 3: Triad Census

Triad	003	012	102	021D	021U	021C	111D	111U
Count	1734	975	1974	174	70	52	426	421
Triad	030T	030C	201	120D	120U	120C	210	300
Count	59	0	777	115	63	40	438	452

To measure the statistical degree of balance in a network, we include statistics in the model that is count of the number of transitive triads and the count of the number of cyclic triads.

Table 4: Social Balance and Transitivity Analysis with transitive and cyclic triads

	(1)	(2)
mutual	1.477^{***}	1.363^{***}
	(0.165)	(0.164)
nodematch.Team		2.865^{***}
		(0.714)
ttriple	0.025^{*}	0.023^{*}
	(0.010)	(0.009)
ctriple	-0.194^{***}	-0.191^{***}
	(0.034)	(0.032)
AIC	1737	1699
BIC	1753	1720

Note: * p < 0.05; ** p < 0.01; *** p < 0.001. (The significance codes used in this paper are consistent.)

Table 4 above gives us the results of fitting ERGM including triad statistics. The model in Column (2) is also conditional on the Team homophily effects. The second model has a better fit based on AIC and BIC. The significant positive coefficient implies that players in the same team are more likely to follow each other. Considering the momentum of the team, players, especially the superstar, need to tell the public that they are a consolidated team even though their relationships are not such close. Following other superstars in the same team is just a way to show their solid and unified relationships.

In both two models, since the coefficient of the transitive triad statistics is significantly positive while the coefficient of the cyclic triads is significantly negative, it appears to be a general preference for balanced friendship ties.

To add more evidence, we also use Triad Census model shown in Table 4. Since we find a strong Team homophily effect, we still fit the model conditional on such effect. Tapered ERGM is used here for a better fit.

Table 4: Social Balance and Transitivity Analysis using Triad Census Model

	(1)	(2)
edges	-0.418^{***}	-0.523^{***}
	(0.066)	(0.073)
nodematch.Team	2.467^{***}	2.221^{***}
	(0.634)	(0.628)
triadcensus.012	-0.071^{*}	
	(0.031)	
triadcensus.102	-0.094^{***}	
	(0.020)	
triadcensus.021D	-0.014	
	(0.050)	

triadcensus.021U	0.015	
	(0.107)	
triadcensus.021C	-0.509***	-0.640***
	(0.139)	(0.137)
triadcensus.111D	-0.185***	-0.103**
	(0.043)	(0.039)
triadcensus.111U	-0.188***	-0.104**
	(0.040)	(0.040)
triadcensus.201		-0.092***
		(0.027)
triadcensus.300		0.216***
		(0.039)
AIC	1647	1614
BIC	1694	1651

First, we fit ERGM with 7 types of triad census which contributes to the large proportion of triangles in the network. According to the results in Column (1) of Table 4, Team homophily effect is still significantly positive. It seems that there appear to be a general preference for transitive friendship ties since the coefficients of the number of intransitive triads, e.g., 021C, 111D and 111U, are significantly negative.

Next, we include all triad census into the model and then select the most significant ones to improve the fitting of models. The AIC and BIC of the model in Column (2) of Table 4 are better than the previous one. From the goodness-of-fit diagnostics in Figure A1 and A2 in section 5, we can find that the second model does perform better in edge-wise shared partners and minimum geodesic distance. MCMC diagnostics[7] also look pretty good and they are converged.

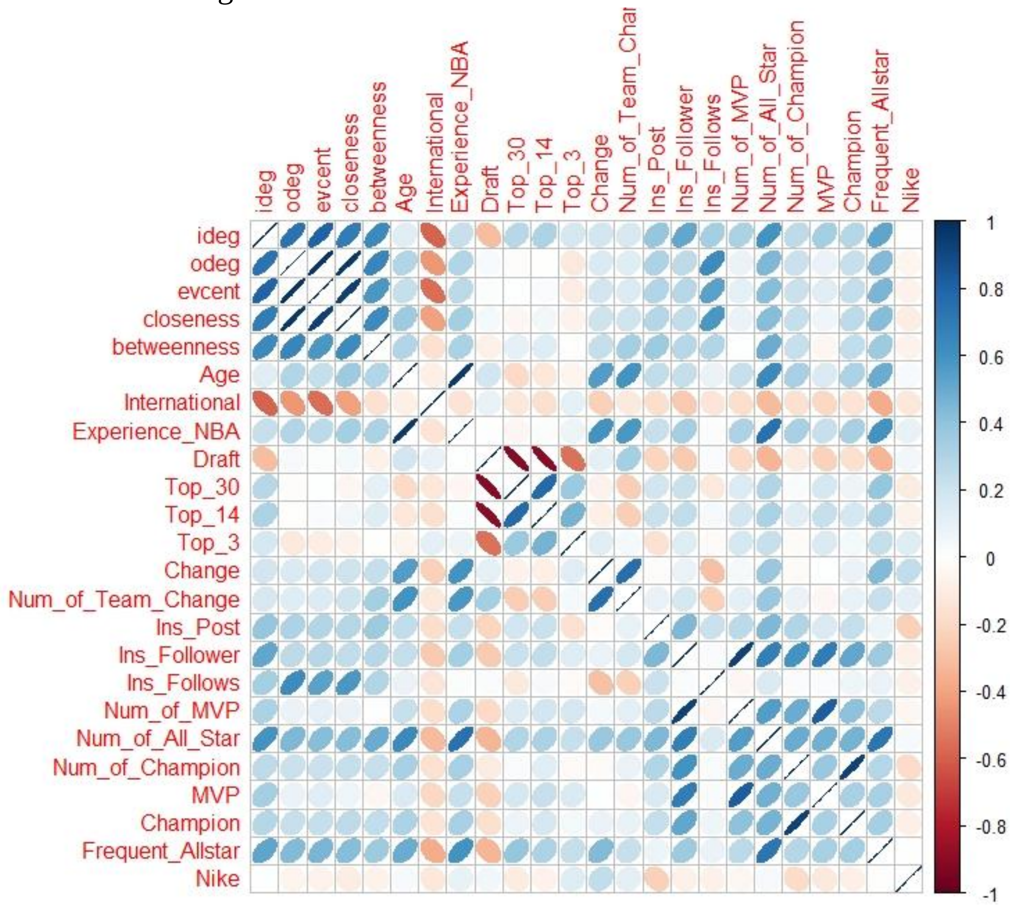
Besides what we found in the previous model, the number of transitive triad 300 has a significantly positive coefficient while the number of another intransitive triad 201 has a significantly negative coefficient, implying that this network is more likely to be transitive. This means that there is a tendency that a friend of your friend is also your friend.

3.4 Further Analysis by ERGM

In this section, we will find the effects of those characteristics by ERGM. We have 4 categories of variables and we plot the correlation among those numerical variables in Figure 7. From the figure below, some highly correlated variables, especially within the same category, can be easily found. Also, some variables cross the category can also show high correlation, for example, Age and some Basketball Factor standing for a player’s honor.

[7] Due to the limit of pages, we do not report the results of MCMC diagnostics.

Figure 7: Correlation Plot for all Numerical Variables



3.4.1 Effects by Demographical Characteristics

First, we start with Demographical Characteristics. We also fit ERGM conditional on Team homophily effect as before. For Demographical Characteristics, we think Race, International, Nationality and State[8] are measuring a quite similar latent variable. Since there are too many Nationalities and States, we only use nodematch instead of nodemix which is used for International and Race.

From the results in Table 5, we can find that a strong International, Nationality and State homophily effect from their significant positive coefficients. As the Age of player increases, he tends to have more relationship with other players.

The most interesting finding is about the attributes International and Race. Since the coefficient of mix.International.0.0 is larger than mix.International.0.1 and both of them are significant, it seems that American players are more willing to follow American players. The coefficient of mix.International.1.0 is significantly negative, which implies that International players prefer following other International players. The homophily does exist. Besides, from the results of using differential homophily term, we can find that the homophily effect among American players is

[8] For International players, if he played high school basketball in the United States, we use the state his high school is located. Otherwise, we just use the country they come from.

larger than the one among International players. As for Race, there is a strong homophily effect among Black athletes, but such effect is not statistical significant when applying to White athletes. The difference of homophily effect of International and Race might be the reason why those American and Black players are more clustered than those International and White players in Figure 4.

In all those models presented in Table 5, Model in Column (4) seems to have the best fit based on AIC and BIC. Therefore, we will then fit ERGM conditional on both Team and Nationality homophily effects in the next few sections.

Table 5: Effects by Demographical Characteristics conditional on Team homophily effect

	(1)	(2)	(3)	(4)	(5)	(6)
edges	-3.562***	-0.747**	-1.990***	-1.853***	-0.703***	-1.414**
	(0.703)	(0.286)	(0.144)	(0.130)	(0.061)	(0.132)
nodematch.Team	3.741***	4.115***	4.066***	3.975***	3.809***	3.816***
	(0.725)	(0.749)	(0.744)	(0.740)	(0.725)	(0.730)
nodecov.Age	0.050***					
	(0.012)					
mix.International.0.0		0.658*				
		(0.295)				
mix.International.1.0		-0.967**				
		(0.340)				
mix.International.0.1		-1.592***				
		(0.368)				
mix.International.1.1		NA				
nodematch.International.0			1.901***			
			(0.161)			
nodematch.International.1			1.243***			
			(0.320)			
nodematch.Nationality				1.775***		
				(0.148)		
nodematch.State					1.150***	
					(0.263)	
nodematch.Race.Black						1.047***
						(0.148)
nodematch.Race.White						0.224
						(0.451)
AIC	1659	1499	1500	1501	1658	1622
BIC	1675	1530	1520	1517	1673	1643

3.4.2 Effects by Social Media Characteristics

Social Media Characteristics can tell us some social behavior of the players. Here are some of assumptions we want to test in this part of analysis. The number of followers may indicate whether he is popular on the social media. The player who has more follows might result in a higher probability that he follows other players while the one who has more post might result in a higher probability of being followed. In Figure 7, we can find a high positive correlation between the number of followers and posts. Therefore, it would be a wise decision to include only one of them in the final model we choose.

Since the numbers of Instagram posts, followers and follows are too large, we scale them first in the data preprocessing step. Then we fit ERGM and present the results in Table 6. All variables have significant positive coefficients which correspond well to our assumptions above. With the help of AIC and BIC as well as ANOVA Table, the model in Column (4) has a best fit using Social Media Characteristics conditional on both Team and Nationality homophily effect.

Table 6: Effects by Social Media Characteristics

conditional on both Team and Nationality homophily effect

	(1)	(2)	(3)	(4)
edges	-1.813^{***}	-1.780^{***}	-1.977^{***}	-1.896^{***}
	(0.131)	(0.131)	(0.142)	(0.143)
nodematch.Team	4.043^{***}	4.006^{***}	4.461^{***}	4.554^{***}
	(0.739)	(0.743)	(0.760)	(0.766)
nodematch.Nationality	1.686^{***}	1.646^{*}	1.839^{*}	1.687^{*}
	(0.150)	(0.150)	(0.159)	(0.161)
nodecov.Ins_Post	0.266^{***}			
	(0.045)			
nodecov.Ins_Follower		0.247^{***}		0.266^{***}
		(0.046)		(0.047)
nodecov.Ins_Follows			0.477^{***}	0.490^{***}
			(0.049)	(0.049)
AIC	1467	1473	1401	1370
BIC	1488	1494	1421	1395

3.4.3 Effects by Business Factor

Just like other industry, many Business Factor may influence the decision of insiders. Therefore, we cannot just ignore the impact from the business world. Still we fit ERGM conditional on both Team and Nationality homophily effect.

From Table 7, it seems that the Agent or Agency homophily effect is not so important. It makes sense since it just means they hire the same employee. However, the Brand homophily effect is significant. If players sign the shoes contract with the same company, they might have more chances for them to attend the same business activity together. Thus, it is still hard to determine whether this is related to the business decision or this is just a true friendship.

Table 7: Effects by Business Factor

conditional on both Team and Nationality homophily effect

	(1)	(2)	(3)
edges	-1.961***	-1.852***	-1.898***
	(0.140)	(0.130)	(0.133)
nodematch.Team	3.957***	3.974***	4.000***
	(0.740)	(0.740)	(0.742)
nodematch.Nationality	1.790***	1.776***	1.788***
	(0.148)	(0.148)	(0.148)
nodematch.Brand	0.298**		
	(0.134)		
nodematch.Agent		-0.070	
		(0.363)	
nodematch.Agency			0.449
			(0.231)
AIC	1498	1503	1499
BIC	1519	1524	1520

3.4.4 Effects by Basketball Factor

There are two sub-categories within Basketball Factor. One is more related to the capacity a player showed in basketball as well as his experience. The other one includes the position he plays in the game, which NBA Division and Conference he comes from as well as which University[9] he played for before going to NBA.

All the following analysis are the effects conditional on both Team and Nationality homophily effect. From the Column (1) of Table 8, we can find that Year in NBA has a positive influence on the formation of a tie. This is the same as Age in Table 5 since most players enter NBA with quite similar ages. The Column (2) seems that players with high picks are less likely to follow those with low picks. But the homophily of NBA draft pick is not significant. As shown in Column (3), it looks like whether a player have won MVP is not so important in this network. Players who have won Championship and are frequently selected to be NBA All-Stars are more likely to be followed. Therefore, players with more basketball skills will be a bonus for being followed by other players.

Table 8: Effects by Basketball Factor (Capacity and Experience)

conditional on both Team and Nationality homophily effect

	(1)	(2)	(3)	(4)	(5)
edges	-2.778***	-1.787***	-1.963***	-0.999***	-1.234***
	(0.284)	(0.146)	(0.495)	(0.301)	(0.166)
nodematch.Team	3.884***	3.992***	4.001***	3.899***	3.774***
	(0.739)	(0.743)	(0.742)	(0.744)	(0.740)
nodematch.Nationality	1.727***	1.778***	1.750***	1.742***	1.534***
	(0.149)	(0.150)	(0.150)	(0.149)	(0.153)

[9] For some International players who did not have University basketball experience in the United States, we use the country they come from instead. For those American players who did not have University basketball experience, we use the high school he graduated from instead.

Nodecov.Experience_NBA	0.053***				
	(0.014)				
mix.Top_14.0.0		0.007			
		(0.299)			
mix.Top_14.1.0		-0.585***			
		(0.178)			
mix.Top_14.0.1		0.169			
		(0.165)			
mix.Top_14.1.1		NA			
mix.MVP.0.0			0.050		
			(0.479)		
mix.MVP.1.0			0.021		
			(0.504)		
mix.MVP.0.1			0.686		
			(0.504)		
mix.MVP.1.1			NA		
mix.Champion.0.0				-0.966***	
				(0.292)	
mix.Champion.1.0				-0.806**	
				(0.312)	
mix.Champion.0.1				-0.689*	
				(0.312)	
mix.Champion.1.1				NA	
mix.Frequent_All_Star.0.0					-1.158***
					(0.225)
mix.Frequent_All_Star.1.0					-0.699***
					(0.166)
mix.Frequent_All_Star.0.1					-0.533**
					(0.163)
mix.Frequent_All_Star.1.1					NA
AIC	1489	1494	1497	1497	1473
BIC	1510	1531	1534	1533	1509

In Table 9, we test the effect of other Basketball Factors. Most variables do not show the statistical significance. One interesting finding is that players in the same Conference are less likely to form a tie as we can see a significantly negative coefficient in Column (2). Players in the same Conference have more chance to compete both for game and some awards. This might be the reason behind this phenomenon.

Table 9: Effects by Basketball Factor (Others)

conditional on both Team and Nationality homophily effect

	(1)	(2)	(3)	(4)
edges	-1.825***	-1.745***	-1.871***	-1.867***
	(0.132)	(0.140)	(0.130)	(0.134)

nodematch.Team	4.170***	4.111***	3.993***	3.977***
	(0.761)	(0.744)	(0.741)	(0.740)
nodematch.Nationality	1.771***	1.788***	1.776***	1.768***
	(0.148)	(0.148)	(0.148)	(0.149)
nodematch.Division	-0.221			
	(0.198)			
nodematch.Conference		-0.249*		
		(0.126)		
nodematch.University			0.850	
			(0.450)	
nodematch.Position				0.050
				(0.130)
AIC	1502	1499	1500	1503
BIC	1523	1520	1520	1524

4. Discussion

In this research, we analyze a relationship among public figures with many available variables. The part of Degree and Centrality analysis gives us some ideas about what factors may influence this Instagram Following relationship. The Latent Space Model does not tell us clear sub-groups since this relationship is more complex than the friendship among monks. We find strong homophily effects of Team and Nationality and have done a lots of effect analysis conditional on such two homophily effects. Generally, Basketball Factors standing for Capacity and Experience of a player and Social Media Characteristics are important in forming relationship while Business Factors and some other Basketball Factors are not so important.

There are still more things we can try. For example, it is hard to tell the influence of Business Factor apart in our analysis if we want to know whether it is a purely true friendship. In other words, we cannot make such inference based on our results. Maybe we can try to fit a ERGM with more variables included but there might be some other issues, including highly correlated variables and a slow convergence rate. Besides, we can try to include the node indegree covariate (nodeicov in R) or node outdegree covariate (nodeocov in R) which may give us more interpretation about our variables.

5. Appendix

Table A1: Degree and Centrality of NBA All-Stars

	Player name	Indegree	Outdegree	Eigenvector centrality	Closeness centrality	Betweenness centrality
1	LeBron James	22	13	0.155	0.610	26.590
2	DeMarcus Cousins	22	21	0.228	0.706	38.806
3	Anthony Davis	17	15	0.174	0.621	10.475
4	Kevin Durant	16	14	0.171	0.610	10.473

5	Kyrie Irving	25	18	0.192	0.667	94.565
6	LaMarcus Aldridge	4	2	0.013	0.409	1.935
7	Bradley Beal	14	16	0.195	0.643	5.406
8	Goran Dragic	2	4	0.040	0.522	0.360
9	Andre Drummond	7	8	0.094	0.545	2.266
10	Paul George	24	23	0.253	0.735	61.248
11	Kevin Love	15	21	0.247	0.706	17.148
12	Victor Oladipo	8	8	0.091	0.522	5.234
13	Kristaps Porzingis	4	1	0.013	0.400	35.677
14	Kemba Walker	9	4	0.043	0.474	1.654
15	John Wall	21	20	0.218	0.679	38.465
16	Russell Westbrook	21	15	0.164	0.600	23.983
17	Stephen Curry	19	22	0.228	0.706	52.654
18	Giannis Antetokounmpo	3	1	0.001	0.293	1.272
19	DeMar DeRozan	17	16	0.191	0.632	48.275
20	Joel Embiid	9	8	0.090	0.554	11.325
21	James Harden	18	10	0.130	0.545	4.348
22	Jimmy Butler	16	17	0.176	0.643	42.451
23	Draymond Green	16	16	0.185	0.621	13.370
24	Al Horford	6	13	0.123	0.610	22.979
25	Damian Lillard	15	16	0.188	0.632	14.806
26	Kyle Lowry	13	20	0.232	0.692	11.840
27	Klay Thompson	8	8	0.106	0.529	0.908
28	Karl-Anthony Towns	7	4	0.030	0.486	6.654
29	Carmelo Anthony	24	18	0.205	0.655	111.678
30	Isaiah Thomas	15	14	0.173	0.600	46.176
31	Paul Millsap	2	10	0.113	0.563	0.832
32	Marc Gasol	5	24	0.185	0.750	52.428
33	Gordon Hayward	4	4	0.040	0.480	0.177
34	DeAndre Jordan	13	13	0.171	0.590	3.929
35	Pau Gasol	7	3	0.023	0.486	12.487
36	Chris Paul	18	19	0.207	0.667	27.695
37	Dwyane Wade	23	30	0.284	0.857	134.431

Figure A1: Goodness-of-fit Diagnostics for Model in Column (1) of Table 5

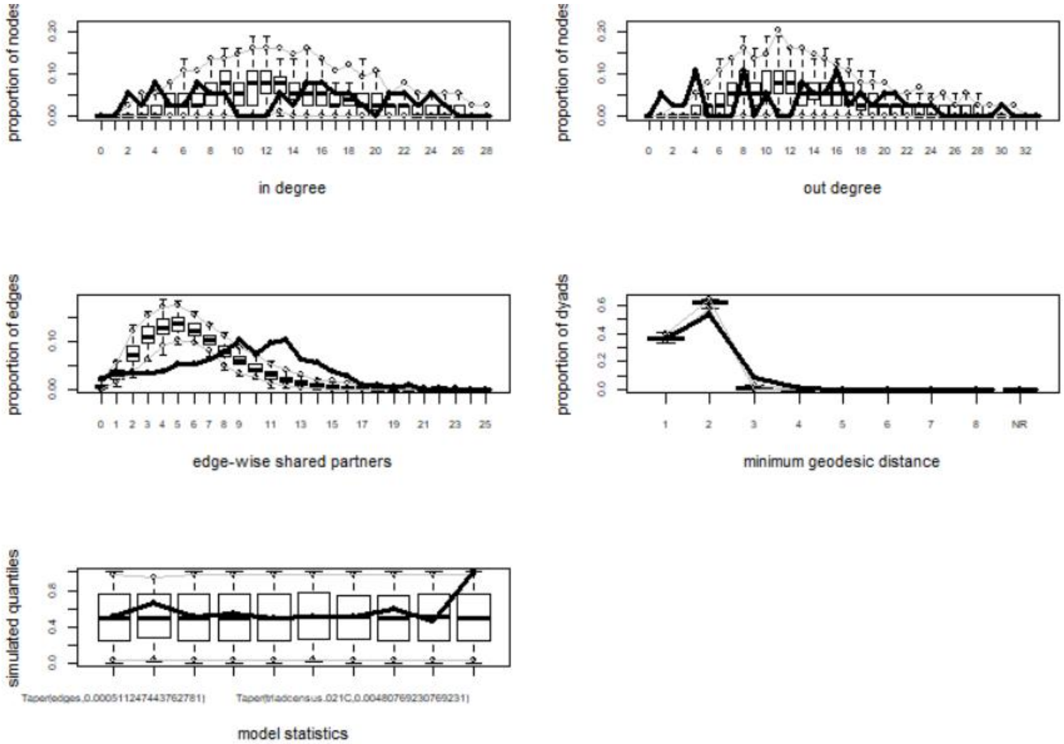
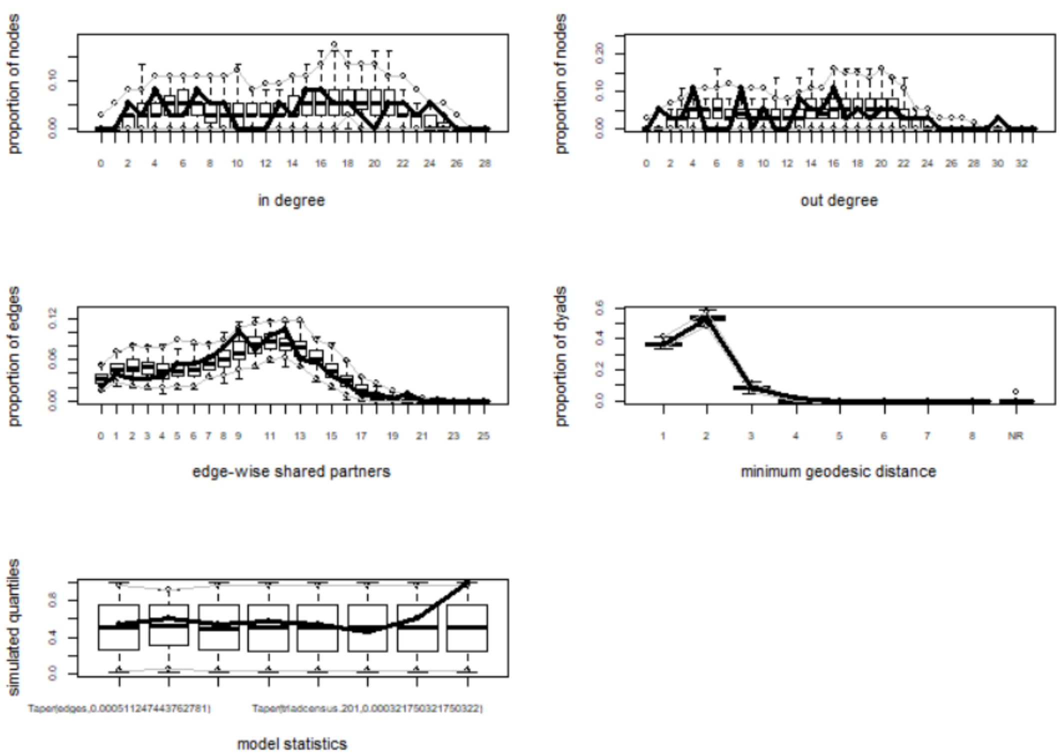


Figure A2: Goodness-of-fit Diagnostics for Model in Column (2) of Table 5



6. Reference

[1] David R Hunter, Steven M Goodreau and Mark S. Handcock. 2008. “Goodness of Fit of Social Network Models”. *Journal of the American Statistical Association*, 103, 1, pp. 248-258.

[2] Peter D. Hoff, Adrian E. Raftery, and Mark S. Handcock. 2002. “Latent Space Approaches to Social Network Analysis”. *Journal of the American Statistical Association*, 97, 1090-1098.