Wealth in the Form of Getting the "Benefit of the Doubt" and Preferential Attachment Models Supporting It

Proposal for Further Research

Beatrix Carroll
Columbia University
bac2108@columbia.edu | beatrixcarroll@gmail.com

ABSTRACT

In this paper, I argue that there is a power law distribution as to whom receives empathy and privilege in society, and how much they receive.

It has been asserted that in the context of preferential attachment in social networks, there is a Nash-equilibrium game underpinning the evolutionary network formation, and that this equilibrium is responsible for a preferential attachment strategy across nodes.[1]

I would like to propose that this theory applies to a social network game where edges represent the impartial/stereotypical views held between nodes, the wealth metric is how much "benefit of the doubt" society allows a given node (and accordingly, how little), and the aim of each node is to maximize wealth through the adoption of either acceptance of stereotypes as beliefs about other node(s) or the allowance for individuals to determine judgment.

I suggest that we experiment with trying to find empirical ways to measure this "wealth in reputation," and explore some features of nodes/graphs to incorporate into potential wealth metrics so that later it can be determined whether or not they could be involved in the underlying heuristic for the game.

I will also look at other models involving preferential attachment, fairness interventions, optimization, and more, and think about how the empirical data regarding wealth distribution (with respect to my my new form of social capital) should look in the above cases.

CCS CONCEPTS

• **Theory of computation → Social networks, network formation**; ***come back to

KEYWORDS

preferential attachment; network formation games; social networks; ***come back to

1 INTRODUCTION

The power law's emergence has been observed in many instances in the natural, man made world. The study of social networks can show the power law's "rich get richer, poor get poorer" phenomenon occurring in wealth of reputation that one "inherits" from his/her demographic: the small fraction of people with the highest concentration of negative stereotypes concentrated upon them receive most of the social disadvantages, stigma, and hate, and have a more difficult time improving the negative reputation or stereotype on their group(s)'s behalf, and the people who receive the lion's share of social advantages, empathetic trust, and "benefit of the doubt" allowances represent a small percentage of the population, at the very top of the spectrum of positive associations (and have an easier time further improving their reputations).

1.1 The Matthew Effect

As the rich get richer, the rich in reputation get richer in repution.

This phenomenon can be seen on an individual basis. For example, people are more likely to credit a famous scientist with an idea than a lesser known one. This phenomenon has been observed and explored by scientists, and sociologist Robert K. Merton termed it "the Matthew effect".[2] How this could happen has been written about in many scholarly papers that discuss how the power law underpins the aphorism "the rich get richer."[3]

1.2 Is the Matthew Effect "Scale-Free"?

Scale-free characteristics have been shown to emerge in social networks.[4] Perhaps fittingly, the Matthew effect can be observed on a larger scale, where the reputation in question is not that of the individual but that of the group to which they they belong. Many women in the STEM fields can attest to the experience of presenting an insight/breakthrough, and then having questions or comments about their work addressed to a man who is somehow socially associated with them. Obviously this is anecdotal, but it is an experience readers may recognize.

1.2 Measuring Where One Falls on the Empathy-to-Bias Spectrum as "Trust Privelege Wealth"

Throughout the rest of the paper I discuss a form of social capital that I suspect exists but that I want to pin down the mechanics of, and I am calling it "trust privilege wealth." I believe that there is a certain metric for wealth that I call "trust privilege" in society that is linked to the ability of person A to be a "blank slate" in the eyes

of person B. Furthermore, I think that this ability is linked to the propensity of others to feel empathy for you. I think that this wealth is about how many people in the network can willing relate to you as a character, and how much.

For instance, From my experience as a stand up comic, I have often witnessed women enjoying the comedy of men without even thinking about the gender of the comic, yet to a male audience member a female comic is always "a female comic." (I have never heard someone be described as a male comic). Then I have also found that my being a woman causes audiences to project biases onto me at a higher rate than if I were a male, and because of this they have less ability to see me as a tabula rasa. To continue with the comedy example, Jerry Seinfeld is a tabula rasa to his audiences. Compare his tendency to use observational humor (“what's the deal with airplane food”) in his comedy, to that of Joan Rivers, who would instead allocate her stage time to talking about herself, often in terms of gender, sex, and power, in order to address the elephant in the room: the biases that audiences unknowingly placed upon her as a female comic.

I think that there is immense power in the ability to be received as a tabula rasa because it means that in the absence of stereotype, there is room to see someone for who they actually are. And seeing someone as an individual is what leads to feeling empathy for them. And with empathy comes “benefit of the doubt.” Thus in a way, it all comes down to trust.

So throughout the paper, I refer to this social wealth metric that I am trying to explore as “trust privilege wealth,” but I also speak of it in terms of:

1. “benefit of the doubt”
2. empathy,
3. “tabula rasa”-ness,
4. stereotype
5. bias

(which is the order that I see the spectrum as being in.) Also, I believe this type of wealth follows a power law distribution, and I want to do research towards proving that.

Again, so far all I have are my unfounded hunches, that’s why I propose to do the research I outline. Even if the power law distribution of the type of wealth I am describing is verifiable, there could be other explanations as to why some people at the top get all the “benefit of the doubt” that has nothing to do with the empathy and bias dynamics I tried to describe with comedy. For instance, perhaps from a psychological perspective, since the people above you in terms of societal advantage (money, political power, whatever it may be) are in a better position in life, it's appealing to relate to them more for vanity reasons. Still, I would like to use these theories as starting points for rigorous studies the proposal be received with interest.

Furthermore, I think that trust privilege can be viewed on a pairwise basis (where it is used as an edge weight between two nodes, and trust privilege of person A with respect to person B varies depending on some individual trust privilege metric of both members in the edge) or an individual basis (where it is some cumulative function of the pair wise trust privelege edges between the node in question and all other nodes in the network).

In addition to creating a formalized set of empirical data, I seek to examine several possible underpinning models to see how they would explain this phenomenon, which I begin to do in the next section.

2 MATHEMATICALLY MODELING TRUST PRIVILEGE WEALTH AS A POWER LAW DISTRIBUTION

2.1 Distribution of “Benefit of the Doubt” as a Result of a Universal Unique Nash Equilibrium of the Game

In their paper “Preferential Attachment as a Unique Equilibrium,” Avin et al. demonstrate that a simple wealth-based recommendation game can be used to model the evolutionary formation of a (social) network, and that the Preferential Attachment rule occurs as the strategy because it is the unique Nash equilibrium of that wealth-based game.[5]

In their game, whose stopping time τ is unknown, each node’s aim is to maximize its degree in the future. They show how this simple game is equivalent to the game underpinning the evolution of a society, where the nodes are people whose goal is to maximize social capital.[6] In the game, nodes are introduced one at a time, and the new node n must choose at random a node u already in the graph, and decide whether or not to form an edge with it. If not, the node n begins a random walk by choosing at random one of u’s neighbors, deciding whether or not to attach to the neighbor or continue its random walk, and so on. To model the Preferential Attachment rule in the game, any given node in the network with degree k will receive a new edge with probability k/Z, where Z is a normalizer.

They show that what they call the Preferential Attachment strategy profile is the unique Nash equilibrium of the game, and that this equilibrium is universal.[7] They prove that the Preferential Attachment Strategy profile is the only viable choice for establishing connections when joining the network. (Explicit definitions are given in *Appendix 1*.) Furthermore, there are implications for Preferential Attachment in connection with random walks and Young’s Lattice that follow from their results.

From this game theoretical analysis, we know that Preferential Attachment must necessarily emerge as the most rational strategy

for all players in this wealth-based strategy game, where social capital is wealth. And since it is already established that Preferential Attachment leads to the appearance of meaningful characteristics in networks (including a power law degree distribution[8][9], similarity of degree among neighbors[10], and smaller network diameter[11]), the question that begs to be answered is: what exactly is the formula for the social capital that is being used as the metric here? Social capital in this game is linked to node degree, but what exactly are the edges and weights that comprise these wealth values?

One idea is to think of the network as three sets of edges (or networks) that overlay the same nodes: The first network (denoted as E_a) is made up of unsigned directed edges that denote familiarity, and the second is a complete graph made up of positively-weighted, directed edges, and the third is a complete graph made up of negatively-weighted, directed edges.

In the first graph, edges represent whether or not node n_1 "knows" n_2, personally or by name and individual reputation. Familiarity can be one way, in the case of celebrities.

In the second overlaying network (denoted as E_+), we imagine that all nodes do not know each other individually at all, and construct positively-weighted, directed edges representing the empathy/"benefit of the doubt" node n_1 would have for n_2 based solely on the demographics that n_2 belongs to, disregarding possible experience with the individual if it exists. An edge exists between every node and every other node in this graph regardless of whether or not the individual nodes are connected by an acquaintance edge and thus this set of edges is complete. If n_1 does not have a net positive perception of n_2, the edge between them in E_+ has a weight of 0.

In the third network (denoted as E_-), negatively-weighted, directed edges follow the same logic for existence and direction as the set of positive edges just described, but here their negative sign represents the acceptance of negative reputations/stereotypes/biases. If n_1 harbors no negative views for n_2 the would-be negative edge between them is just weighted as 0.

The social network formation game here is still one in which new nodes are introduced over time until an unknown stopping time τ. The goal of each node is still to maximize its expected wealth metric.

Maybe the wealth heuristic used in determining whether or not to adopt new connections is some function of signed directed edges, where positive directed edges represent empathy/benefit of the doubt one node would give to another, and negative directed edges represent the acceptance of negative reputations/ stereotypes/ biases, based solely on demographic information. This question is something that I would like to research.

No matter the exact nature of the true metric at play in a game, the individual's goal is to maximize one's wealth; however, society would be be much improved for the vast majority of people (the people in the long tail) if everyone's goal was to have a wealth as close to 0 as possible.

I tried to make a toy model of how demographic-based biases can compound (a phenomenon known today as "intersectionality") to see if there is evidence of the power law there, and make any other observations.

In this toy model there are 2 binary categories of people, one is "letter" and the other is "number". Each node in my network must belong to one letter and one number. The 2 categories overlap, just as there is intersection between demographic groups in real life.

In Appendix A.2, the calculations for toy example's trial metrics are given in detail. Charts for some of the metrics experimented with are included here as well in Fig. 1.

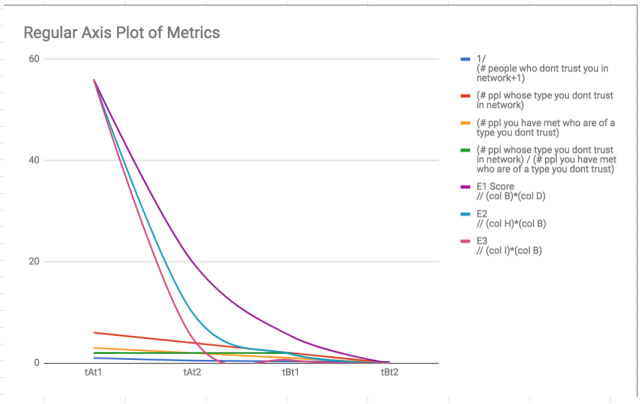

Figure 1: looking at the metrics, some of them seem to exhibit a power law.

By imagining more specific, mathematical definitions of trust privilege wealth in coming research than I have been able to do for this proposal, I seek to expand upon the work done by Avin et al. Later, in the course of the research that this proposal is suggesting, testing these formulations for wealth metrics within the provided game theory model would give rise to various portraits of what networks would look like corresponding to the various wealth metrics. If a wealth metric formulation is correct (or close to correct), the corresponding network portrait would be reflected in empirical data, and we will have backed into an understanding of the underlying, rational-from-a-game-theory-perspective type of thinking that accounts for and perpetuates the power law distribution of empathy/trust.Such an understanding can certainly provide insight into how to remedy the situation.

2.3 Reputation Wealth As a Growing Investment, And Stereotype Threat As an Accruing Debt

"Stereotype Threat" refers to the psychological effects (and their practical repercussions) on members of groups who are stereotyped negatively when those members are trying to perform those tasks themselves.[12] The case of women in STEM fields provides a clear and researched example of how stereotype threat leads to “the poor in repute stay poor in repute” even just internally, looking only at the immediate impact on self-view.[13]

Perhaps demographic reputation capital has some sort of compounding interest that has been attempted to be explained by stereotype threat. There could be an actual mathematical formula and economic theory to elucidate the effects of stereotype threat and help quantify them, and I think this area deserves more research.

2.4 “Benefit of the Doubt” Wealth as the Result of an Optimization Problem

Next, I want to explore the hypothesis that there is an optimization model (resulting in a tempered preferential attachment rule) underpinning the adoption of reputations/biases, using the work of D’Souza et al. as a foundation.[14] In their paper, “Emergence of Tempered Preferential Attachment from Optimization,” they provide proof of Mandelbrot’s proposal that underlying optimization problems can give rise to power law distributions.[15]

Therefore just as I want to look into ways that trust privilege wealth could be a function of a preferential attachment inducing game equilibrium (as explained inin the previous section), I want to look into ways that this wealth could be a function of an underlying optimization problem.

Also, I propose to try to prove one way or another which case it is (Preferential Attachment as a result of the game equilibrium described in the previous section, or Preferential Attachment as a result of an optimization problem) by looking for consistency in the data with a power law (equilibrium), or with a power law multiplying an exponential decay (optimization).

3.3 A Short-term Intervention for Long Term Empathy Wealth Redistribution

In their paper, “A Short Term Intervention for Long Term Market Fairness in the Labor Market,” Hu et al. show that there is a “reinforcing nature of asymmetric outcomes resulting from groups’ divergent accesses to resources and as a result, investment choices.” They show that fairness interventions that look at the bigger picture, going beyond case by case checks on procedures, will be increasingly needed to avoid falling into an enduring inequitable equilibria in the labor force.[16]

Because of the power law being at play in our game of maximizing reputation / empathy wealth, the presence of this reinforcing nature and inequitable equilibria seen in their game seems to suggest that the their intervention research could inform interventions for our reputation wealth problem.

4 CONCLUSION

I hope that by considering the new form of social wealth I have described, and considering that it could be power law distribution, we will be able to identify the process whereby people who are the most burdened by not being seen as individuals become more burdened (poor get poorer).

In the future, I hope to explore how one could address and stop this process by possibly engineering "tipping points" of reputation that re-assign "wealth of repute" and thus node edge weight values, wherein suddenly poor nodes are the new rich nodes and rich nodes are the new poor. Perhaps we are already seeing that tipping point get closer. Maybe one day, I will hear someone say “I just saw the best male comic!”

A APPENDICES

A.1 Game and Strategy

A.1.1 Definition of Preferential Attachment Strategy

The preferential attachment strategy (PA) is the strategy by which, given a degree sequence $D = (d_1, \ldots d_n)$ and a node of degree k $\in$ D is contacted with probability

$$PA(k, D) = PA(k) = \frac{k}{2(t-1)}$$

A.1.2 Definition of Preferential Attachment Strategy Profile

The preferential attachment profile, (Πpa), is the strategy profile where players vt play arbitrarily for t $\leq$ 4, and players vt play according to pa for t $\geq$ 5.[17]

A.1.3 Additional Game Information

Universal Nash Equilibrium in the game

Evolution of the society/network’s wealth is measured by the sequence $\overline{\alpha}$, where

$$\overline{\alpha} = (\alpha_t),\ t \geq 1,$$

$\alpha_t \in [0,1]$ is a parameter that measures the wealth of the the society at time t,

τ = *stopping time* where the utilities are evaluated, and

$\bar{\alpha}$, τ *are unknown to the players*

The Nash Equilibrium that the PA strategy profile gives rise to is universal, meaning that it holds for all $\bar{\alpha}$ and all τ [18]

A.2 Toy Example Trial Metrics in Detail

Fig 0 shows my example of a simple network whose nodes follow these types:

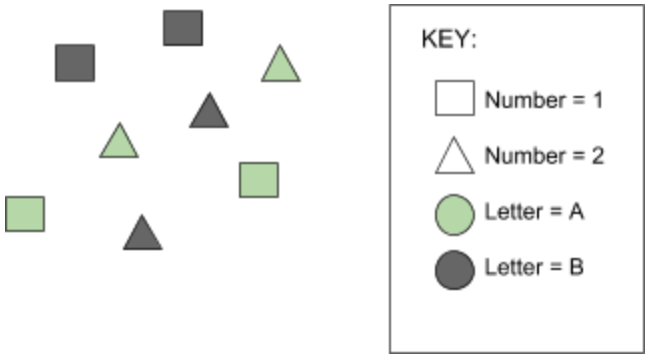

Figure A. 0: A possible set of nodes for toy example

For each category, the binary types are attributed different amounts pre-existing empathy/reputation wealth. (Fig. 1.)

Trust Privelege Values inherent to each type within each category: (higher the better, 0 is best possible)				
Category	Type within Category	Privelege value	Type within Category	Privelege value
Letter	t1	0	t2	-2
Number	tA	0	tB	-5

Figure A.1: Pre-existing trust privelege by type within individual category

Compounded Trust Privelege Values by Compunded Type				
Compunded Type	Letter Type	Number Type	empathy wealth of demographic	count of nodes in toy network
tAt1	tA	t1	0	2
tAt2	tA	t2	-2	2
tBt1	tB	t1	-5	2
tBt2	tB	t2	-7	2

Figure A.2: Compounded Trust Privelege Values by Compounded Type

Edge Weights: Trust Privelege ([Col]) - Trust Privelege ([Row]):				
	tAt1	tAt2	tBt1	tBt2
tAt1	0	2	5	7
tAt2	-2	0	3	5
tBt1	-5	-3	0	2
tBt2	-7	-5	-2	0

Figure A.3: One type of metric for edge weights to try: the trust privelege of one node minus the trust privelege of the other node.

Various Metrics by Compounded Type										
	1/ (# people who dont trust you in network+1)	sum of trust for you in network	sum of trust for you in your netork, normalized	(# ppl whose type you dont trust in network)	(# ppl you have met who are of a type you dont trust)	(# ppl whose type you dont trust in network) / (# ppl you have met who are of a type you dont trust)	E1 Score // (col B)*(col D)	E2 // (col H)*(col B)	E3 // (col I)*(col B)	E4 // (col D)*(col E)
tAt1	1	28	56	6	3	2	56	56	56	336
tAt2	0.5	12	40	4	2	2	20	10	5	160
tBt1	0.333333333	-12	16	2	1	2	5.333333333	1.777777778	0.5925925926	32
tBt2	0.25	-28	0	0	0	#DIV/0!	0	0	0	0

Figure A.4: Various Metrics to look into for relevence in empirical data and game theory model

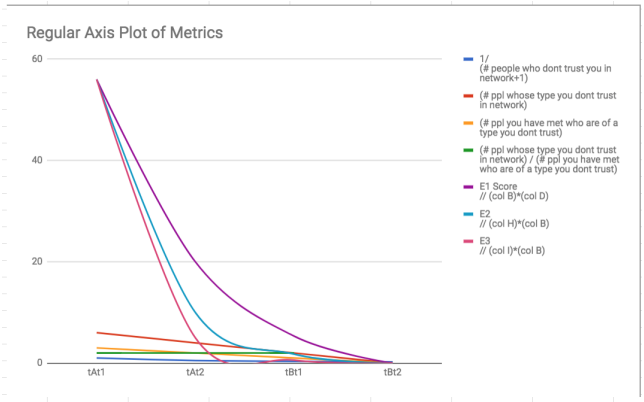

Figure A.5: looking at the metrics, some of them seem to exhibit a power law.

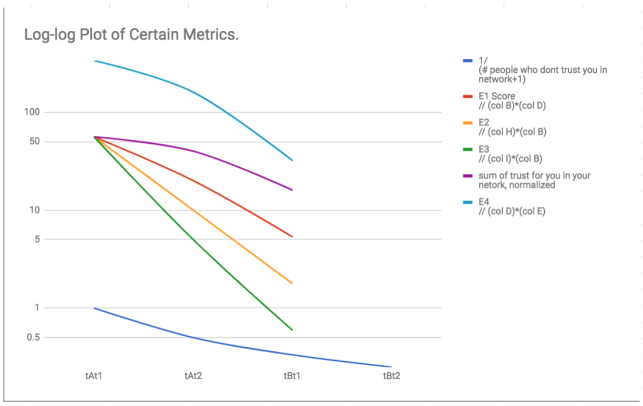

Figure A.6: looking at the metrics again, on a log-log graph

ACKNOWLEDGMENTS

This work is supported by Social Networks COMS 6998 at Columbia University in the City of New York, Spring 2018

REFERENCES ****

[1] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (pp. 559–568).

[2] "The role of the Matthew effect in science - ScienceDirect." https://www.sciencedirect.com/science/article/pii/S0039368106000252. Accessed 7 May. 2018.

[3] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (pp. 559–568).

[4] "Emergence of scale-free characteristics in socio-ecological systems ...." 11 Jun. 2015, https://www.nature.com/articles/srep10448. Accessed 7 May. 2018.

[5] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (pp. 559–568).

[6] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (p. 559).

[7] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (pp. 559–568).

[8] Béla Bollobás, Oliver Riordan, Joel Spencer, Gábor Tusnády, et al. The degree sequence of a scale-free random graph process. Random Structures & Algorithms, 18(3):279–290, 2001.

[9] Remco Van Der Hofstad. Random graphs and complex networks, 2016.

[10] Paul L Krapivsky and Sidney Redner. Organization of growing random networks. Physical Review E, 63(6):066123, 2001.

[11] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. Combinatorica, 24(1):5–34, 2004.

[12] "Stereotype Threat and College Academic Performance: A Latent ...." https://www.ncbi.nlm.nih.gov/pubmed/23950616. Accessed 7 May. 2018.

[13] "Why Do Women Underperform Under Stereotype Threat? - CiteSeerX." http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.901.1226&rep=rep1&type=pdf. Accessed 7 May. 2018.

[14] "Emergence of tempered preferential attachment from optimization ...." http://www.pnas.org/content/104/15/6112.short. Accessed 7 May. 2018.

[15] "Emergence of tempered preferential attachment from optimization ...." http://www.pnas.org/content/104/15/6112.short. Accessed 7 May. 2018.

[16] "A Short-term Intervention for Long-term Fairness in the Labor Market." 30 Nov. 2017, https://arxiv.org/abs/1712.00064. Accessed 7 May. 2018. p1389-90.

[17] "A Short-term Intervention for Long-term Fairness in the Labor Market." 30 Nov. 2017, https://arxiv.org/abs/1712.00064. Accessed 7 May. 2018. p1389-90.

[18] Avin, C., Cohen, A., Fraigniaud, P., Lotker, Z., & Peleg, D. (2018). Preferential Attachment as a Unique Equilibrium (p. 561).