Uses and misuses of mathematics in analysing social networks

Peter Hegarty

Department of Mathematics, CTH/GU

11 June, 2014

- First part:

- First part:

  The friend of my friend is my friend

- First part:

  The friend of my friend is my friend

- Second part:

- First part:

  The friend of my friend is my friend

- Second part:

  The enemy of my enemy is my friend

- First part:

  The friend of my friend is my friend

- Second part:

  The enemy of my enemy is my friend
  The enemy of my friend is my enemy

- First part:

  The friend of my friend is my friend

- Second part:

  The enemy of my enemy is my friend
  The enemy of my friend is my enemy
  The friend of my enemy is my enemy

- First part:

  The friend of my friend is my friend

- Second part:

  The enemy of my enemy is my friend
  The enemy of my friend is my enemy
  The friend of my enemy is my enemy

- The second part is more controversial as it equates absence of friendship with emnity

- First part:

  The friend of my friend is my friend

- Second part:

  The enemy of my enemy is my friend
  The enemy of my friend is my enemy
  The friend of my enemy is my enemy

- The second part is more controversial as it equates absence of friendship with emnity
- Nevertheless, "balance" captures the idea that connections are formed in a network primarily on the basis of affinity, rather than other factors

Definition

A **triad** in a graph G is an induced subgraph on three vertices.

Definition

A **triad** in a graph G is an induced subgraph on three vertices.

- Hence in a graph with n vertices, there are $\binom{n}{3}$ triads.

Definition

A **triad** in a graph G is an induced subgraph on three vertices.

- Hence in a graph with n vertices, there are $\binom{n}{3}$ triads.
- A triad is **balanced** if it contains either one or three edges.

Definition

A **triad** in a graph G is an induced subgraph on three vertices.

- Hence in a graph with n vertices, there are $\binom{n}{3}$ triads.
- A triad is **balanced** if it contains either one or three edges.
- Hence, in a network where relationships are formed because of affinity, there should be a greater proportion of balanced triads than in a random network of the same edge density.

Definition

A **triad** in a graph G is an induced subgraph on three vertices.

- Hence in a graph with n vertices, there are $\binom{n}{3}$ triads.
- A triad is **balanced** if it contains either one or three edges.
- Hence, in a network where relationships are formed because of affinity, there should be a greater proportion of balanced triads than in a random network of the same edge density.
- If this is not the case, then it indicates that there is a different sociological dynamic in the network.

- Let G be a graph with n nodes and edge density p, i.e.:

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3,$$

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3, \quad \mathcal{E}_1 = \binom{n}{3}3p(1-p)^2,$$

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3, \quad \mathcal{E}_1 = \binom{n}{3}3p(1-p)^2,$$
$$\mathcal{E}_2 = \binom{n}{3}3p^2(1-p),$$

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3, \quad \mathcal{E}_1 = \binom{n}{3}3p(1-p)^2,$$
$$\mathcal{E}_2 = \binom{n}{3}3p^2(1-p), \quad \mathcal{E}_3 = \binom{n}{3}p^3.$$

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3, \quad \mathcal{E}_1 = \binom{n}{3}3p(1-p)^2,$$

$$\mathcal{E}_2 = \binom{n}{3}3p^2(1-p), \quad \mathcal{E}_3 = \binom{n}{3}p^3.$$

- Now take an actual network and let $\mathcal{N}_i$, $i = 0,1,2,3$, be the actual numbers of i-edge triads.

- Let G be a graph with n nodes and edge density p, i.e.: $p = [\#\text{ edges in } G]/\binom{n}{2}$.
- If the edges were placed at random (E-R model), then the expected numbers of i-edge triads, for $i = 0, 1, 2, 3$, would be

$$\mathcal{E}_0 = \binom{n}{3}(1-p)^3, \quad \mathcal{E}_1 = \binom{n}{3}3p(1-p)^2,$$
$$\mathcal{E}_2 = \binom{n}{3}3p^2(1-p), \quad \mathcal{E}_3 = \binom{n}{3}p^3.$$

- Now take an actual network and let $\mathcal{N}_i$, $i = 0,1,2,3$, be the actual numbers of i-edge triads. In a "balanced network" we expect to find

$$\mathcal{N}_0 < \mathcal{E}_0, \quad \mathcal{N}_1 > \mathcal{E}_1, \quad \mathcal{N}_2 < \mathcal{E}_2, \quad \mathcal{N}_3 > \mathcal{E}_3.$$

Kadushin's two key assertions are the following:

Kadushin's two key assertions are the following:

Assertion 1: "There are 1,575 symmetric dyads in the network (triad type 3-102 in chapter 2, figure 2) ... The number of dyads was much greater than would have been found by chance".

Kadushin's two key assertions are the following:

Assertion 1: "There are 1,575 symmetric dyads in the network (triad type 3-102 in chapter 2, figure 2) ... The number of dyads was much greater than would have been found by chance".

Assertion 2: "There are 45 (symmetric) triads in the entire network (triad type 16-300 in chapter 2, figure 2), also far more than expected by chance".

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- The actual numbers for this graph are

$$\mathcal{N}_0 = 3971, \quad \mathcal{N}_1 = 1575, \quad \mathcal{N}_2 = 393, \quad \mathcal{N}_3 = 45.$$

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- The actual numbers for this graph are

$$\mathcal{N}_0 = 3971, \quad \mathcal{N}_1 = 1575, \quad \mathcal{N}_2 = 393, \quad \mathcal{N}_3 = 45.$$

- So this graph is quite unbalanced !

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- The actual numbers for this graph are

$$\mathcal{N}_0 = 3971, \quad \mathcal{N}_1 = 1575, \quad \mathcal{N}_2 = 393, \quad \mathcal{N}_3 = 45.$$

- So this graph is quite unbalanced ! Where did Kadushin's first assertion come from ?

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- The actual numbers for this graph are

$$\mathcal{N}_0 = 3971, \quad \mathcal{N}_1 = 1575, \quad \mathcal{N}_2 = 393, \quad \mathcal{N}_3 = 45.$$

- So this graph is quite unbalanced ! Where did Kadushin’s first assertion come from ? After consultation with him, he provided us with his numbers, obtained using `Pajek`:

$$\mathcal{E}_{1,\kappa} \approx 190.68... \quad \mathcal{E}_{3,\kappa} \approx 0.04...$$

- There are 34 nodes and 78 edges in the graph, hence $p = 78/\binom{34}{2} = 78/561 \approx 0.139$.
- This yields

$$\mathcal{E}_0 \approx 3819, \quad \mathcal{E}_1 \approx 1850, \quad \mathcal{E}_2 \approx 299, \quad \mathcal{E}_3 \approx 16.$$

- The actual numbers for this graph are

$$\mathcal{N}_0 = 3971, \quad \mathcal{N}_1 = 1575, \quad \mathcal{N}_2 = 393, \quad \mathcal{N}_3 = 45.$$

- So this graph is quite unbalanced ! Where did Kadushin’s first assertion come from ? After consultation with him, he provided us with his numbers, obtained using `Pajek`:

$$\mathcal{E}_{1,\kappa} \approx 190.68... \quad \mathcal{E}_{3,\kappa} \approx 0.04...$$

- What ??????

- The key to understanding his mistake is found in the sentence "The overall density is 0.139 with 156 connections out of a total of 1,112 possible connections".

- The key to understanding his mistake is found in the sentence "The overall density is 0.139 with 156 connections out of a total of 1,112 possible connections".
- Note $156 = 2 \times 78$. Leaving aside that $2 \times 561 = 1{,}222$, not 1,112, why is Kadushin counting every edge twice ?

- The key to understanding his mistake is found in the sentence "The overall density is 0.139 with 156 connections out of a total of 1,112 possible connections".
- Note $156 = 2 \times 78$. Leaving aside that $2 \times 561 = 1{,}222$, not 1,112, why is Kadushin counting every edge twice ?
- Because he is comparing the network with random **directed** networks !

- The key to understanding his mistake is found in the sentence "The overall density is 0.139 with 156 connections out of a total of 1,112 possible connections".
- Note $156 = 2 \times 78$. Leaving aside that $2 \times 561 = 1{,}222$, not 1,112, why is Kadushin counting every edge twice ?
- Because he is comparing the network with random **directed** networks ! Thus,

$$\mathcal{E}_{1,\kappa} = \binom{n}{3} 3p^2(1-p)^4 = 190.68\ldots \quad \mathcal{E}_{3,\kappa} = \binom{n}{3} p^6 = 0.04\ldots$$

- The key to understanding his mistake is found in the sentence "The overall density is 0.139 with 156 connections out of a total of 1,112 possible connections".
- Note $156 = 2 \times 78$. Leaving aside that $2 \times 561 = 1{,}222$, not 1,112, why is Kadushin counting every edge twice ?
- Because he is comparing the network with random **directed** networks ! Thus,

$$\mathcal{E}_{1,\kappa} = \binom{n}{3} 3p^2(1-p)^4 = 190.68\ldots \quad \mathcal{E}_{3,\kappa} = \binom{n}{3} p^6 = 0.04\ldots$$

- You cannot assume relationships are symmetric when making the graph and then compare it with random graphs in which this assumption is dropped !

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.
- Let's give him the benefit of the doubt and assume the latter.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.
- Let's give him the benefit of the doubt and assume the latter.
- The really salient feature of the graph is the overrepresentation of 2-edge triads: $\mathcal{N}_2 = 393 > 299 = \mathcal{E}_2$.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.
- Let's give him the benefit of the doubt and assume the latter.
- The really salient feature of the graph is the overrepresentation of 2-edge triads: $\mathcal{N}_2 = 393 > 299 = \mathcal{E}_2$.
- This is indicative of **stardom** or **hierarchy**.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.
- Let's give him the benefit of the doubt and assume the latter.
- The really salient feature of the graph is the overrepresentation of 2-edge triads: $\mathcal{N}_2 = 393 > 299 = \mathcal{E}_2$.
- This is indicative of **stardom** or **hierarchy**.
- Ordinary club members were driven, especially after the schism, to form relationships with high-ranking members, rather than with one another.

- One could simply conclude that Zachary was wrong to assume friendships were mutual when making his graph.
- So either he was aware of this problem, but made his graph undirected anyway, or he was unable to pick up such "tensions" between apparent friends in his observations.
- Let's give him the benefit of the doubt and assume the latter.
- The really salient feature of the graph is the overrepresentation of 2-edge triads: $\mathcal{N}_2 = 393 > 299 = \mathcal{E}_2$.
- This is indicative of **stardom** or **hierarchy**.
- Ordinary club members were driven, especially after the schism, to form relationships with high-ranking members, rather than with one another.
- It is in this sense that relationships in the karate club were not based primarily on mutual affinity.

BIG IDEA: To explain the appearance of mature networks in terms of preferential attachment of new nodes to old ones as the network grew.

BIG IDEA: To explain the appearance of mature networks in terms of preferential attachment of new nodes to old ones as the network grew.

PIONEERS: The paper “Emergence of Scaling in Random Networks” by Barbarási and Albert (Science 1999).

BIG IDEA: To explain the appearance of mature networks in terms of preferential attachment of new nodes to old ones as the network grew.

PIONEERS: The paper "Emergence of Scaling in Random Networks" by Barbarási and Albert (Science 1999).

HOT EXAMPLE: World Wide Web (WWW), which people could see growing before their eyes at the time.

BIG IDEA: To explain the appearance of mature networks in terms of preferential attachment of new nodes to old ones as the network grew.

PIONEERS: The paper "Emergence of Scaling in Random Networks" by Barbarási and Albert (Science 1999).

HOT EXAMPLE: World Wide Web (WWW), which people could see growing before their eyes at the time.

FOCUS ON: The so-called "scale-free" property, which is most popularly expressed by the idea that the distribution of vertex degrees follows a power law, i.e.:

Let $P(d)$ be the proportion of vertices that have degree d. Then

$$P(d) \propto d^{-\gamma},$$

for some fixed positive constant γ.

Let $P(d)$ be the proportion of vertices that have degree d. Then

$$P(d) \propto d^{-\gamma},$$

for some fixed positive constant γ.

This is quite different from a Poisson distribution

$$P(d) \approx e^{-\lambda} \frac{\lambda^d}{d!},$$

where $\lambda =$ average degree. For $G(n,p)$, $\lambda = (n-1)p$.

Let $P(d)$ be the proportion of vertices that have degree d. Then

$$P(d) \propto d^{-\gamma},$$

for some fixed positive constant γ.

This is quite different from a Poisson distribution

$$P(d) \approx e^{-\lambda} \frac{\lambda^d}{d!},$$

where $\lambda =$ average degree. For $G(n,p)$, $\lambda = (n-1)p$.

Barbarási-Albert proposed a specific growth model and claimed their simulations indicated convergence to a power law distribution with $\gamma = 2.9 \pm 0.1$.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don’t know about their simulations) and the first step was to clean it up.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don't know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don't know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.
- We define a sequence of graphs G_m^1, G_m^2, ... inductively.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don’t know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.
- We define a sequence of graphs G_m^1, G_m^2, $\dots$ inductively.
- G_1^1 consists of a single vertex with a loop.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don't know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.
- We define a sequence of graphs $G_m^1, G_m^2, \dots$ inductively.
- G_1^1 consists of a single vertex with a loop.
- At step t we add one new vertex v_t and m new incident edges.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don't know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.
- We define a sequence of graphs $G_m^1, G_m^2, \ldots$ inductively.
- G_1^1 consists of a single vertex with a loop.
- At step t we add one new vertex v_t and m new incident edges.
- The edges are added one at a time.

To even define a preferential attachment growth model rigorously is a non-trivial exercise. The text in B-A did not make sense (I don't know about their simulations) and the first step was to clean it up.

- Fix a positive integer m.
- We define a sequence of graphs $G_m^1, G_m^2, \dots$ inductively.
- G_1^1 consists of a single vertex with a loop.
- At step t we add one new vertex v_t and m new incident edges.
- The edges are added one at a time. Each time an edge is added at step t, its other endpoint v_i is chosen randomly from among $v_1, \dots, v_t$ according to the rule

$$\mathbb{P}(v_i = v_s) = \begin{cases} \frac{d_{G_{\text{now}}}^{\text{tot}}(v_s)}{2E+1}, & \text{if } s < t, \\ \frac{1}{2E+1}, & \text{if } s = t. \end{cases}$$

$$\mathbb{P}(v_i = v_s) = \begin{cases} \frac{d^{\text{tot}}_{G_{\text{now}}}(v_s)}{2E+1}, & \text{if } s < t, \\ \frac{1}{2E+1}, & \text{if } s = t. \end{cases} \tag{1}$$

$$\mathbb{P}(v_i = v_s) = \begin{cases} \frac{d_{G_{\text{now}}}^{\text{tot}}(v_s)}{2E+1}, & \text{if } s < t, \\ \frac{1}{2E+1}, & \text{if } s = t. \end{cases} \tag{1}$$

- **Obs!** The model allows for loops and multiple edges.

$$\mathbb{P}(v_i = v_s) = \begin{cases} \frac{d^{\text{tot}}_{G_{\text{now}}}(v_s)}{2E+1}, & \text{if } s < t, \\ \frac{1}{2E+1}, & \text{if } s = t. \end{cases} \quad (1)$$

- **Obs!** The model allows for loops and multiple edges.
- d^{tot} denotes total degree, which means loops contribute two, whereas an edge between v_t and v_s, where $t > s$, contributes one to the *indegree* of v_s and one to the *outdegree* of v_t.

$$\mathbb{P}(v_i = v_s) = \begin{cases} \frac{d^{\text{tot}}_{G^{\text{now}}}(v_s)}{2E+1}, & \text{if } s < t, \\ \frac{1}{2E+1}, & \text{if } s = t. \end{cases} \quad (1)$$

- **Obs!** The model allows for loops and multiple edges.
- d^{tot} denotes total degree, which means loops contribute two, whereas an edge between v_t and v_s, where $t > s$, contributes one to the *indegree* of v_s and one to the *outdegree* of v_t.
- E is the current number of edges. Thus (1) says that vertices which currently have higher degree are more likely to get new connections ⇒ **preferential attachment**.

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described.

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described. Let $\mathcal{N}_m^n(d)$ denote the number of vertices of indegree d, hence total degree $m + d$, in G_m^n.

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described. Let $\mathcal{N}_m^n(d)$ denote the number of vertices of indegree d, hence total degree $m + d$, in G_m^n. Let

$$\alpha_{m,d} := \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}.$$

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described. Let $\mathcal{N}_m^n(d)$ denote the number of vertices of indegree d, hence total degree $m + d$, in G_m^n. Let

$$\alpha_{m,\,d} := \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}.$$

Let $\epsilon > 0$.

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described. Let $\mathcal{N}_m^n(d)$ denote the number of vertices of indegree d, hence total degree $m + d$, in G_m^n. Let

$$\alpha_{m,d} := \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}.$$

Let $\epsilon > 0$. Then with probability 1, as $n \to \infty$ we have

Theorem

Let $m \in \mathbb{N}$ and let $(G_m^n)_{n \geq 1}$ be the sequence of random graphs just described. Let $\mathcal{N}_m^n(d)$ denote the number of vertices of indegree d, hence total degree $m + d$, in G_m^n. Let

$$\alpha_{m,\,d} := \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}.$$

Let $\epsilon > 0$. Then with probability 1, as $n \to \infty$ we have

$$(1-\epsilon)\alpha_{m,\,d} \leq \frac{\mathcal{N}_m^n(d)}{n} \leq (1+\epsilon)\alpha_{m,\,d}$$

for every d in the range $0 \leq d \leq n^{1/15}$.

THE GOOD:

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE BAD:

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE BAD:

The theorem predicts a specific power $\gamma = 3$.

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE BAD:

The theorem predicts a specific power $\gamma = 3$. Name your favorite positive real number γ and you can probably find a paper in the literature showing data following a power law $P(d) \propto d^{-\gamma}$.

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE BAD:

The theorem predicts a specific power $\gamma = 3$. Name your favorite positive real number γ and you can probably find a paper in the literature showing data following a power law $P(d) \propto d^{-\gamma}$.

THE UGLY:

THE GOOD:

We have some mathematically rigorous support for the hypothesis that preferential attachment is a crucial mechanism in the growth of real networks.

THE BAD:

The theorem predicts a specific power $\gamma = 3$. Name your favorite positive real number γ and you can probably find a paper in the literature showing data following a power law $P(d) \propto d^{-\gamma}$.

THE UGLY:

The bound of $d \leq n^{1/15}$ in the theorem means that it effectively says nothing whatsoever about any real network !!

Pause while I switch to a different set of slides ...