Analyzing Rumor Propagation through Graph Theory

Tanmay Chopra

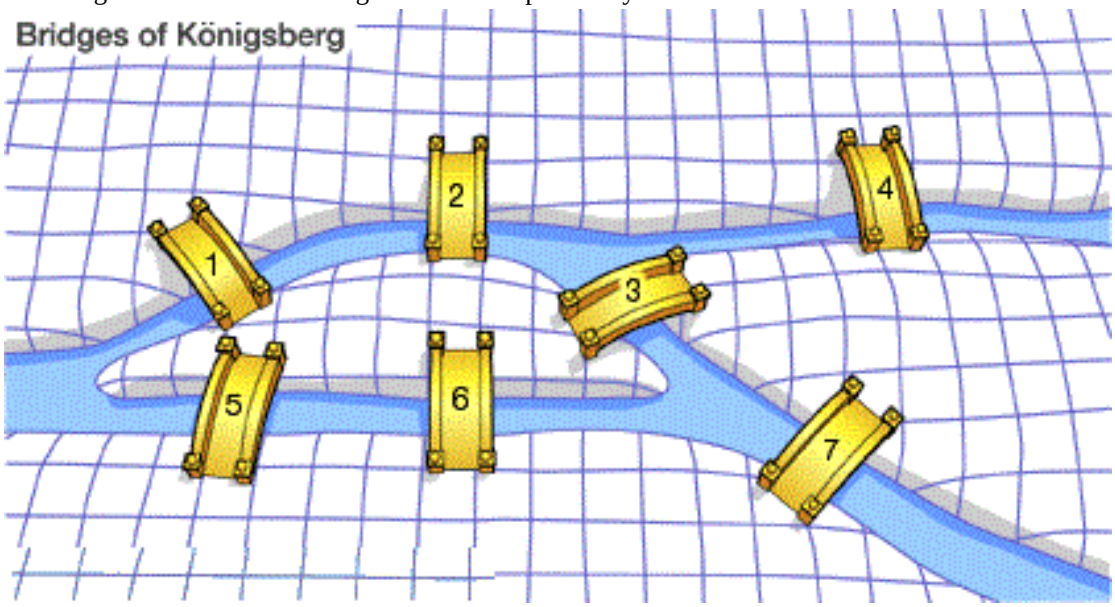
Abstract:

This Research report aims to analyze how graph theory can be applied to rumor propagation in limited social settings with a special emphasis on schools. In this paper, a modified Nearest Neighbour Algorithm has been used to reach a probabilistic weight for the edges of an Eulerian Graph and an equation has been proposed for the same, which is intended to complement the SIR model used to normally study rumor propagation by focusing on the propagated content, instead of the propagators.

Introduction:

The city of Königsberg in Prussia (now Kaliningrad, Russia) was set on both sides of the Pregel River, and included two large islands which were connected to each other and the mainland by seven bridges. The problem was to devise a walk through the city that would cross each bridge once and only once, with the provisos that: the islands could only be reached by the bridges and every bridge once accessed must be crossed to its other end. The starting and ending points of the walk need not to be the same.[1] No one could assert without a doubt that no solution existed, but no one was able to find a way either.

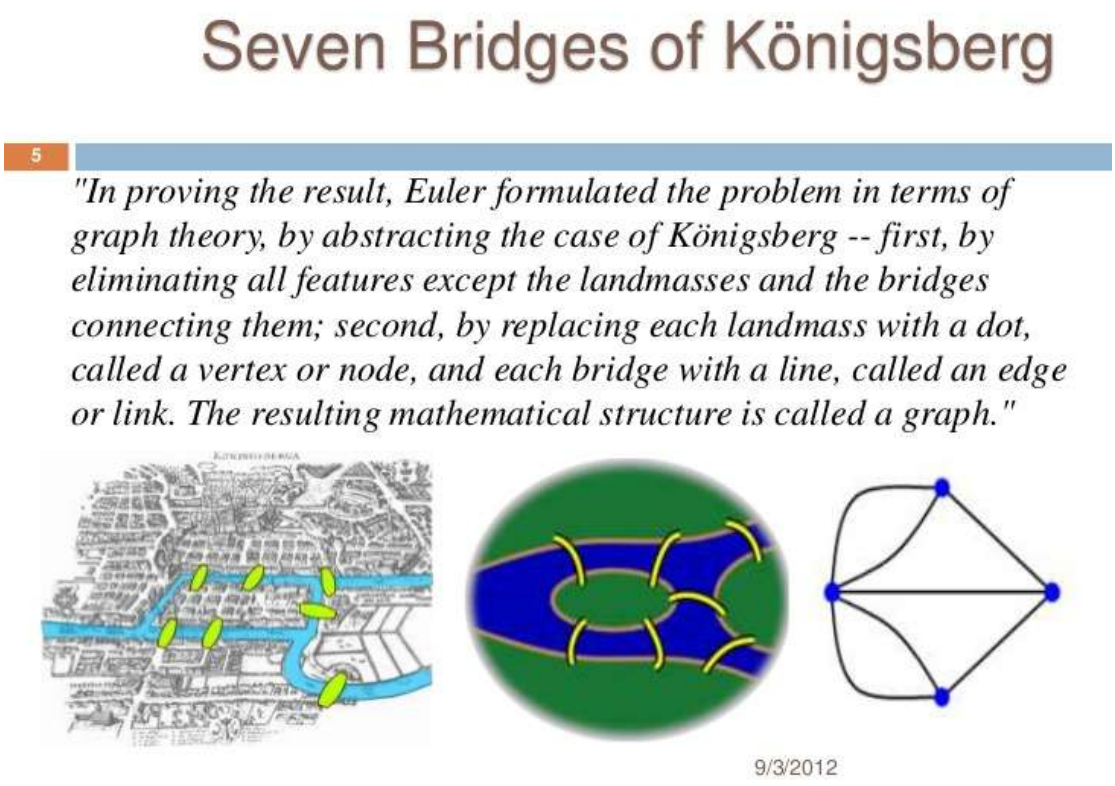
This problem was then handed to Leonard Euler, a mathematician, to solve. Euler intuitively soon realized it was not possible, but he didn't have any scientific proof to back his claims and set about searching for the same. And this gave rise to Graph Theory.



Representative Image of the Problem [1]

Euler realized that the only things that mattered were the number of landmasses and the number of bridges connected to each of them.

He represented the problem as a graph, with the vertices as the nodes. This was the inception of graph theory. This was a format created by Euler himself, in order to represent this problem in a way that could lead to a solution. The landmasses were objects, while the bridges represented the various relations between these objects.



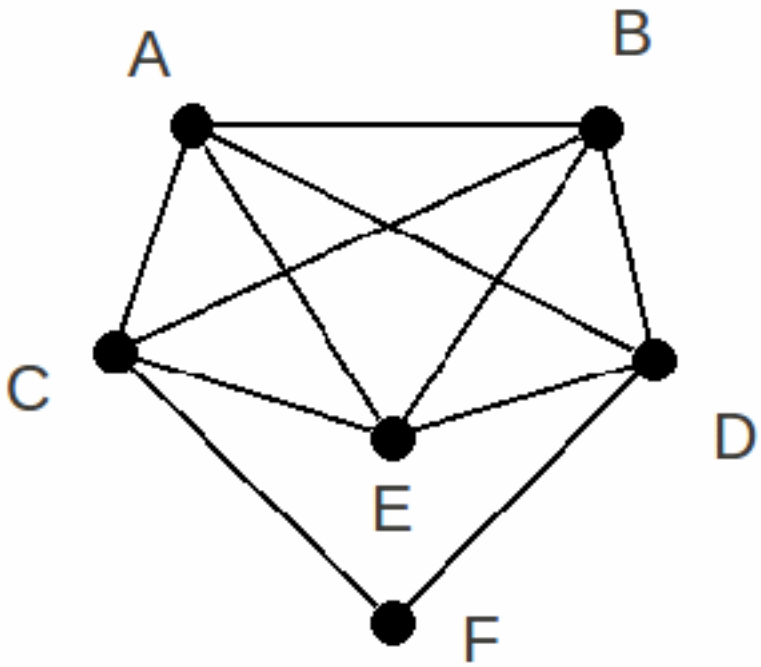
Outline of his Method [2]

In this research report, the endeavor has been to implement the concepts of Graph Theory to Rumor Propagation, a everyday activity that plays a role in all our lives in some capacity-. A rumor is defined as 'information or a story that is passed from person to person but has not been proven to be true' [3].
Every relation known to man can be represented as a graph, such as the Internet, where every device connected is a node and the connections between the devices are the edges[4]. Even social networks are commonly represented as graphs, with each user representing a node and the connections between users being depicted by edges[5].

Till now, various models and algorithms have been proposed and used for information passing and routing like the Nearest Neighbour Algorithm [6], the Djikstra's Algorithm[7] etc. The Nearest Neighbour Algorithm is the algorithm where information is transmitted to the closest node, but the downside to this is that it only looks at the distance between the nodes, or for specifically a single variable being the edge variable and information is passed to only one node, whereas in rumor passing, one person may transfer information to multiple nodes.

The Djikstra's Algorithm looks at the shortest path between any two nodes, but the downside to that is that there is no definite ending node to a rumor, making this algorithm tough to implement here.
In this study, we will see the propagation of rumors in any social setting as a weighted graph using the concepts of graph theory The aim of this paper is to model the spreading of rumors physically in everyday school life and try to control this spread or alter it, something every teenager wants to know about, to use to his/her advantage!
Here, the intention is to use a scaled version of the Nearest Neighbor algorithm, but enable multiple information transfers using probability. Furthermore, the edge value will no be the distance, but the probability, which depends on multiple variables.
To start off, existing data about rumor propagation was analyzed. Most data was regarding viral rumors on social media, which varied slightly in terms of setting as compared to a school rumor in terms of

number of possible nodes and the edges between each node. In social media, a post reaches every friend, while in school, every friend needs to be told personally by another friend.
The next step was to interview a set of 100 teenagers across various age groups was interviewed and asked what factors would make them pass on a rumor and what wouldn't, after which a few harmless rumors were spread around to test these factors.



Here, A, B, C, D, E, F represent different individuals in schools like teachers, students etc. [9]

Material and Method- Proposed Hypothesis

Analogous to Graph Theory, every individual in the school will be considered a node on the graph and the edges of the graph represent the transfer of the rumor from one person to the other.
The weights of the graphs represent the probability of the rumor being passed on from one person to the next. The person starting the rumor will be considered the source node.
The aspects that rumor propagation depends on are[15]:

1) Believability of the rumor [8]
2) Sociability of person or persons starting and or spreading the rumor[13]
3) Likeability of person about whom the rumor is. (Factor added from student poll)
4) Interestingness of rumor (accounting for maliciousness which impacts equation in a manner inversely proportional to likeability)[12]
5) Convenience of propagating the rumor (Factor added from student poll)
6) Interaction Time (Factor added from student poll)
7) Emotional or psychological or material reward by spreading the rumor versus Risk involved in spreading of the rumor[9] (Risk-Reward factor)

The believability of the rumor is directly proportional to the weight, the less believable a rumor is, the less likely it is that it will be propagated.
The sociability of the person starting and/or spreading the rumors decides the number of edges connected to the node. A more social person will have more edges connected to the node, meaning more possible nodes that the information can be propagated to, spreading and speeding up the propagation.

The likeability of a person/persons about whom the rumor is a very vital factor because a rumor about a disliked person will spread much faster, especially if negative.
Interestingness of the rumor is directly related to the rate of spread of the rumor. An interesting rumor will spread incredibly fast compared to a boring one.
Convenience of spreading the rumor is a factor that will apply to all other nodes except the source node, because the source node has some external motivation to spread the rumor for sure, while others may or may not.
As interaction time between two nodes increases, they are more likely to discuss a host of matters. Hence, as interaction time increases, probability of spreading rumors increases as they discuss more matters and have lesser subjects competing for more time. If interaction time between nodes is short, they are less likely to spend that time discussing a rumor.
Furthermore, the reward plays a vital factor. The research on procrastination, which would apply in terms of the activity of spreading this rumor, factors this in as well [11] At the same time, risk factor applies in every situation. Humans are programmed to evaluate risk vs. reward in each activity, making it a vital part of this study.

In light of the above, the following equation is proposed:
Edge Weight = *a*Believability**b*InteractionTime + *c*Sociability/ + *d*Convenience**e*Interestingness/*f*Likeability) + Risk-Reward Factor

Where a,b,c,d,e,f are all constant values for given settings.

Classical SIR equation[14]focuses on people and their actions while equation proposed above compliments it by focusing on content of rumor and social dynamic.

$I_1 + S_2 \xrightarrow{\lambda} S_1 + S_2,$
(when spreader meets with the ignorant, it becomes spreader at rate λ)
$S_1 + R_2 \xrightarrow{\delta} R_1 + R_2,$
(when a spreader contacts with stifler, the spreader becomes a stifler at the rate δ)
$S_1 + S_2 \xrightarrow{\delta} R_1 + S_2,$
(when a spreader contacts with another spreader, initiating spreader becomes a stifler at the rate δ)
$S \xrightarrow{\sigma} R.$
(σ is the rate to stop spreading of a rumor spontaneously.)

RESULT:
After setting factors, school individuals were polled to see how many agreed with which factors came into play.

Serial Number	Factor	Students Agreeing
1	Believability of Rumour	80%
2	Sociability of Spreaders	67%
3	Likeability of Target	73%
4	Interestingness of Rumour	94%
5	Convenience	85%
6	Interaction time	61%
7	Risk versus Reward	89%

The formula works in most school-related settings and office settings for example, when the interaction time and convenience factors are extremely high, even a rumor will low believability, spreads: like one in our school about a boy not bathing for years.

The formula does seem to lack efficiency in settings conducive only to gossip, that is, where groups of people congregate to discuss gossip only, like tea parties since even with low interaction time, rumors with low believability tend to spread extremely fast.

The main lacking in the formula is the value of these constants since different settings have completely varied rumor-propagating rates and most factors are relative to conditions. Furthermore, it is possible that all factors may not have been accounted for as social settings often result due to completely disconnected, unforeseen factors.

CONCLUSION:

This research has aimed to establish basic research factors and an understanding of rumor propagation in social settings. By limiting the target group, the model is more accurate. The research shows that this subject is a lot more complicated than it appears. The intermingling of these factors and their effect on other factors further complicates matters but makes for an even more interesting area of study. More importantly, this research shows how everything can be analyzed using math and how math, and especially graph theory- is an integral part of our everyday lives!

References:

[1] http://topologydrexel.wikispaces.com/Seven+Bridges
[2]http://www.slideshare.net/Tech_MX/graph-theory-1
[3]http://www.merriam-webster.com/dictionary/rumor
[4]https://www.techopedia.com/definition/4993/network-map
[5]http://pc57724.uni-regensburg.de/morgan/teaching/CS104-Social_Networking.pdf
[6]http://www.austincc.edu/powens/+Topics/HTML/06-5/06-5.htm
[7]https://www.cs.auckland.ac.nz/software/AlgAnim/dijkstra.html
[8]http://www.academic-journals.org/ojs2/index.php/IJCSE/article/viewFile/759/18
[9] http://astarmathsandphysics.com/a-level-maths-notes/191-d1/3332-eulerian-graphs.html
[10]http://www.sciencedaily.com/releases/2008/10/081009144325.htm
[11] http://www.telegraph.co.uk/news/science/science-news/3660232/Academics-invent-a-mathematical-equation-for-why-people-procrastinate.html
[12]Olles, Deana B., "Rumor propagation on random and small world networks" (2006). Thesis. Rochester Institute of Technology
[13]http://www.hindawi.com/journals/mpe/2010/631357/
[14] http://arxiv.org/pdf/1208.6063.pdf
[15]Polled aspects were added after an a large group of students independently submitted the factors they thought could be relevant to the study.