Using Gould's Index to win *Rail Baron*

Kyle Suelflow and Owen Suelflow

May 4, 2023

Abstract

Rail Baron is a popular board game in the United States. The object of the game is to buy railroads that allow you to move between different cities in the United States. Building a good network of railroads is imperative to your chances of winning. In this paper, we will discuss using an adjacency matrix to model the graph of railroads. Then, we will use the dominant eigenvector to find the Gould's Index of each railroad, thus giving us a ranking of railroads. Using aspects of *Rail Baron*, we weighted each Gould's index by a variety of factors, which will be discussed below, in order to find the optimal railroads.

1 Introduction

In our family, board games are a staple of our free time. We love the strategy behind them, always allowing for improvement. *Rail Baron* is just one of many games that we love, but it is special to us for a couple reasons. First, we always play with our Dad, and he grew up playing *Rail Baron* too, with his friends. We have been playing *Rail Baron* since we were little, probably 8 or 9 when we first started. This longevity gave us the inspiration and excitement for this analysis. Have we been playing the game wrong all along? Or will our suspicions be confirmed?

1.1 Initial Setup

Our first problem was figuring out how to setup the adjacency matrix. There were a few options. We could have each city as a node, and whether or not they were connected would determine the matrix. The problem with this method is twofold: for starters, it is unclear how a city is connected to another. Do cities have to be on the same railroad? Are San Diego and Chicago considered "connected" because they share the same railroad, even though they are significantly far apart? The "connectedness" of any two cities became increasingly subjective, as there wasn't a good method that worked for all scenarios. The other problem with this method was that having cities as nodes doesn't really do us any good. It might answer the question, "What city is the most important", but we already have a suitable answer to that question: Each city has its own probability of being rolled as a destination, and therefore higher probability cities are more important. Thus, we decided not to use cities as nodes, rather using railroads as nodes. This solved the two problems of the first method by using a simple and concrete process of determining "connectedness", as well as having a significant purpose: Railroads are what a player can purchase during the course of the game; this is where the strategy comes in. Now, we have an adjacency matrix, which we can use to explore our research question: Which railroads are the most valuable?

2 Gould's Index

2.1 Adjacency Matrix

Now that we have established the structure of our graph, we can create our adjacency matrix. Since each node in our graph is a railroad, the edges between nodes can be represented by a connection between two railroads. Our matrix $A = (A_{ij})$, where:

$$A_{ij} = \left\{ \begin{array}{ll} 1, & \text{if railroad i is connected to railroad j} \\ 0, & \text{if railroad i is not connected to railroad j} \end{array} \right\} \tag{1}$$

For example, there is a 1 in the $A_{5,6}$ entry because the fifth railroad as read from the left side of the board to the right, the UP, is connected to the sixth railroad, the CRI & P.

The following descriptions in this subsection and further subsections can be attributed to Straffin and his work in [1]. Now, we are going to modify A to be $B = A + I$, where I is the identity matrix. In this problem, we want to characterize each railroad in terms of how accessible it is. One way to do this is by matrix multiplication of B. By including 1's along the diagonal of B, the ijth entry in B^k tells us the number of paths of length k connecting railroads i and j, including stopovers at railroads along the way. Without the diagonal entries, we would not be including any stopovers.

2.2 Finding Gould's Index

Now, we have established what our matrix looks like and what we want to accomplish. So, how can we use B^k to find an index of accessibility? First, some technical details need to be defined. Our matrix B is non-negative and primitive; the proof can be found in [1]. Because of this, we can apply the following:

Perron-Frobenius Theorem 1 (Perron) *If M is an $n \times n$ nonnegative primitive matrix, then there is an eigenvalue λ_1 such that*

1. *λ_1 is real and positive, and is a simple root of the characteristic equation,*
2. *$\lambda_1 > |\lambda|$ for any eigenvalue $\lambda \neq \lambda_1$,*
3. *λ_1 has a unique (up to constant multiples) eigenvector v_1, which may be taken to have all positive entries.*

Using the above, we can find the eigenvector v_1 associated with the dominant eigenvalue, λ_1. Further, we can express any vector x that is not orthogonal to v_1 as:

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n (\alpha_1 \neq 0). \tag{2}$$

We can then apply B^k to both sides, and, using the definition of eigenvalues, as $k \to \infty$,

$$B^k x = \lambda_1^k \alpha_1 v_1 + \lambda_2^k \alpha_2 v_2 + \cdots + \lambda_n^k \alpha_n v_n \tag{3}$$

$B^k x$ will point in the same direction as v_1, as we know that λ_1 is the dominant eigenvalue. Applying this approach back to our adjacency matrix B, we can simply compute the eigenvalues and take the eigenvector v_1 corresponding to the dominant eigenvalue λ_1. *Gould's Index* takes this eigenvector and normalizes it, usually by dividing each component by the sum of all components:

$$v_i = \frac{v_i}{\sum_{j=1}^{28} v_j} \tag{4}$$

Now that we have our *Gould's Index*, we can compare the components to come up with a ranking of the importance/accessibility of each railroad.

2.3 Analysis of Gould's Index

In the context of this problem, Gould's index measures the accessibility, or importance, of a railroad. Specifically, if a railroad is connected to a high number of railroads, its Gould's index will be higher. Even better, if the railroads it is connected to are *also* connected to a high number of railroads, that will increase the importance of the railroad. The results of this analysis are shown in table 1. An interesting pattern has developed. Railroads that are in the central/eastern US, or connected to cities like Chicago, are clustered at the top. The GM & O, for example, does not connect to many cities, but is located in the north south corridor of the east-central US, and because of this, is very accessible. Does the importance of the GM & O change as we begin to consider other factors?

3 Weighting

3.1 Cost of Railroad

The resulting Gould's Index values are not without flaws. One major issue is that these values do not take into account the cost of the railroad. For example, the RF & P has a much lower score than the AT & SF, but of course it does, because the AT & SF costs ten times as much money! Thus, we need to take this variable into account. We decided to divide each Gould's index by it's cost, creating an updated index that is per 1000 dollars. Shown below is the equation we used. X_i is our updated Gould's index, G_i is our original Gould's index, and C_i is the cost of the *ith* railroad. This, of course, changed our results drastically. Before, the indices were essentially just reflections of the cost of the railroad. Now, some smaller railroads rise to the top, as shown in Table 2.

$$X_i = \frac{G_i}{C_i} \tag{5}$$

Railroad	Gould's Index
GM & O	.055
CRI & P	.055
IC	.054
PA	.054
AT & SF	.053
B & O	.053
CB & Q	.052
NYC	.050
C & O	.050
MP	.049

Table 1: The railroads with the highest Gould's Index. The top 10 railroads are shown. Having a high Gould's index means that the railroad is accessible, that it is important. It also means that the railroad connects to other important railroads.

Railroad	Gould's CostIndex
GM & O	.0046
RF & P	.0044
IC	.0039
D & RGW	.0035
T & P	.0031
C & NW	.0031
CB & Q	.0026
C & O	.0025
SLSF	.0024
CMSTP & P	.0024

Table 2: The railroads with the highest Gould's Index, weighted by the cost of the railroad. This means that per 1000 dollars, these railroads have the highest accessibility, or importance, out of the 28 railroads in the game.

Railroad	Gould's City
PA	.0049
AT & SF	.0040
C & NW	.0039
NYC	.0031
B & O	.0026
SP	.0025
UP	.0023
CRI & P	.0018
C & O	.0018
L & N	.0015

Table 3: The railroads with the highest Gould's Index, weighted by their the cities' probabilities that they connect to. Railroads with high Gould's index have are very accessible, and have high importance. In this table, we also weighted by the cities they can reach. Railroads that connect to important cities, such as Chicago, New York, or Philadelphia are ranked highly.

The RF & P shoots up to second place in Table 2, because, while it only connects 5 railroads, it does so at such a cheap price that it might be worth the small investment, according to this analysis. An interesting observation about this top ten is that most of the railroads are located in the east-central US. In fact, a lot of these railroads connect with each other. Next, we will look at which cities railroads are connected too. Does the east-central US region continue to have the strongest railroads?

3.2 Probabilities of Cities

Another aspect that our original Gould's Index does not consider is which cities are railroads connected to. To tackle this question, we determined which cities were on a railroad's route, and added up the probability of going to that city. After completing a route, a player will roll some dice to determine their next destination. First, they will roll for the region they will be going to, and second, the city in that region. Each city has a certain combination(s) of dice rolls, and so each city has a probability of going there. We added up each cities' probability in a railroads' network, and multiplied it's Gould's Index by that number. Shown below is the equation we used. Y_i is the updated Gould's index, G_i is the original Gould's index, and c_j is the probability that the *jth* connected city is rolled.

$$Y_i = G_i \sum_{j=1}^{n} c_j \tag{6}$$

To give some context, a city with a high probability is Chicago, and a low probability city is Norfolk. The resulting table is shown in Table 3. This table makes sense; The railroads that cost the most are generally at the top. There is a reason they cost a lot, after all; they are longer, and can connect to more cities that way. One surprise is the C & NW. It only costs 14 thousand, the cheapest in Table 3, and yet it ranks number 3. We believe this is because it is an important link between the western railroads and Chicago. Chicago is a high probability city and an important railroad hub, as many railroads connect to Chicago.

3.3 Cost & City Probabilities

As outlined above, two important characteristics of a railroad are its cost, and its city probability. It is useful to consider these separately, but we wanted to see what would happen if we combined them. Which railroad gives us the perfect balance?

$$Y_i = \frac{G_i \sum_{j=1}^{n} c_j}{C_i} \tag{7}$$

Looking at Table 4, we see that the PA continues to lead the way, while the northeastern and east-central railroads are heavily featured. Interestingly, the GM & 0 and the IC have very similar scores, and this is reflected on the map. They both cover virtually the same portion of the US. What is tipping the GM & O above the IC?

Railroad	GouldsCost & City(10^4)
PA	5.02
C & NW	4.59
GM & O	4.21
IC	3.82
B & O	3.72
C & O	3.62
NYC	3.55
CB & Q	3.20
AT & SF	2.99
CMSTP & P	2.76

Table 4: The railroads with the highest Gould's Index, weighed by both their cost and the cities with which it a railroads connects to. A Railroad with a high Gould's index is an important railroad, that it connects to other railroads, and that it is accessible. We also factored in the cost of the railroad, and the cities that it connects to. A high value means that the railroad is important. NOTE: All values in the right hand column are multiplied by 10^4, to make the table easier to understand.

Railroad	GouldsCost & City & NumRail(10^4)
C & NW	2.75
PA	1.64
NYC	1.11
B & O	1.10
AT & SF	.99
C & O	.89
L & N	.85
CMSTP & P	.70
GM & O	.66
SAL	.66

Table 5: These are our final rankings. The Gould's index, which represents a railroads accessibility, is weighted by the cost of the railroad, the amount of cities it connects too, and the number of railroads that connect to those cities. The C & NW blew out the competition, and that really surprised us. NOTE: values on the right hand side are multiplied by 10^4 to make the table easier to understand.

3.4 Cost, City Probabilities, & Number of Railroads per City

One final integral aspect of strategy in *Rail Baron* is the interaction between players and their railroads. If player 1 had to take player 2's railroad, player 1 must then pay player 2 10,000 dollars each time they use player 2's railroad. This makes "owning" cities a viable strategy, meaning that a player owns all railroads that connect to the city. To incorporate this component into our analysis, we took each cities' probability, and divided it by the number of railroads that connect to it. For example, a city like Chicago, with numerous railroads connecting to it, goes down in importance while Las Vegas, with only 1 railroad connecting it, increases its importance. Luck plays a big factor in whether this strategy pays off, but if someone ends up going to a city you own, you'll be in for a hefty payday. Shown below is the equation we used. The introduction of the new variable is represented by N_j, where N is the number of railroads connected to the *jth* city.

$$Y_i = \frac{G_i \sum_{j=1}^{n} \frac{c_j}{N_j}}{C_i} \tag{8}$$

Interestingly, the C & NW crushes the competition, overtaking the PA for first place. Then follows the 3 northeastern stalwarts, and a hodgepodge of other railroads. The results are shown in table 5. A similar table to the last, but a few rails moved up or down a significant amount.

4 Takeaways

4.1 Building a Winning Network

We began our project by taking a look at *Rail Baron* and thinking about how we could try to find the optimal strategy using Gould's Index. After carefully weighting each railroad to reflect the rules and constraints of the game, we arrived at a ranking of the best railroads using all three of our weights in subsection 3.4. Now, we want to take our rankings and develop a network of railroads such that we can reach as many cities as possible. We will assume that it is a four player game, and that we will be able to comfortably buy 6 railroads. Which 6 should we buy? Taking a look at the rankings, we can see that the four main railroads in the Northeast are extremely important. Failure to control at least one of the four makes winning the game very difficult, as the Northeast is the most likely region to go to. We will go for the highest ranked Northeast railroad, the **PA**. The NYC or B & O would work well too.

The next obvious choice for a railroad is our top-ranked road, the **C & NW**. This road is important for a couple of reasons. It connects to Chicago and Minneapolis/St. Paul, which are two important hubs where many roads connect to. It also connects to Rapid City and Casper, two places who only have one or two railroads connected, which increases their importance. Our next road is the **AT & SF**. This railroad is the second-most expensive in the game, tied with the UP and only behind the SP. This will be our main outlet into the Southwest and South Central regions of the game. While expensive, it is worth it as it covers a large swath of land and connects us to many cities, including Chicago, which acts as a hub for our network.

Next, we need something to get us to the Northwest. Looking at our rankings, the **CMSt.P & P** is a logical choice. It gets us to the Northwest, while also connecting to Chicago. Interestingly, when combined with the C & NW, we own the connection between Chicago and MSP. This is important, as other players who own railroads in the Northwest such as the GN or NP will have to find other ways of getting to the North Central and Northeast, as those roads end in MSP and do not get to Chicago. The next railroad to buy that gets us to new cities would be the **L & N**. This road gets us further into the South Central and the Southeast. One interesting thing that helps rank it highly is the amount of cities that are connected to it where each city only has a couple railroads connected. The L&N gets us to a lot of new cities, and makes it harder for other players to navigate around this area.

Finally, we chose our 10th ranked railroad, the **SAL**. This road connects to the L & N and covers almost all of the Southeast. The main reason it is ranked highly is because it gets to a lot of cities that only have it and the ACL connected to them. This usually results in one or two players not being able to access these cities, which generates money for the owners of the ACL and SAL. Overall, these six railroads, the **PA**, **C & NW**, **AT & SF**, **CMSt.p & P**, **L & N**, and **SAL** make up our optimal railroad network. After analyzing the network, we determined that 71.1% of the time we will roll a city that we are able to get to solely on our railroads, meaning we do not have to pay another player. For only six railroads, that is pretty good. Even if we roll a city we cannot get to with just our railroads, it won't be far off track. In fact, the city farthest from our railroads is Salt Lake City, which is only 7 moves away. Our network is shown in Figure 1.

4.2 Limitations & Further Inquiry

In the end, our method determined the optimal railroads to buy in a way that does make sense with our experience playing the game. There were certainly a few railroads that we did not expect to be near the top, such as the C & NW, but for the most part it worked as predicted. The PA rose near the top as expected, as did the other main railroads in the Northeast.

One limitation in our model was that our choices for how we weighted the Gould's index were arbitrary. We could have chosen different weights, or we could have placed more emphasis on certain weights than others by weighting our weights. In that case, we might have looked to place more of an emphasis on the probability that a given railroad can reach the destination city, as in the end that is the most important aspect of the game - can you reach your destination city without using another player's railroad. We also did not factor in one other popular strategy when playing the game: The Superchief. This strategy involves not buying any railroads until the player has the $40,000 required for the Superchief, which allows the player to roll 3 dice when moving their pawn instead of two. While our goal with this project does not concern this strategy, it would have been interesting to consider it in our model. It would have been difficult to quantify the strategy, but certainly would've been intriguing.

For further inquiry, we might look to analyze other games with a similar approach. *Ticket to Ride*, another popular railroad-based game, would be a great game to use. It is a bit simpler, but with different rules that would present their own challenges an opportunities. A classic example of Gould's index is the game of *Risk*, which would be fun to work with. Overall, any game that can be represented as a graph could be a candidate.

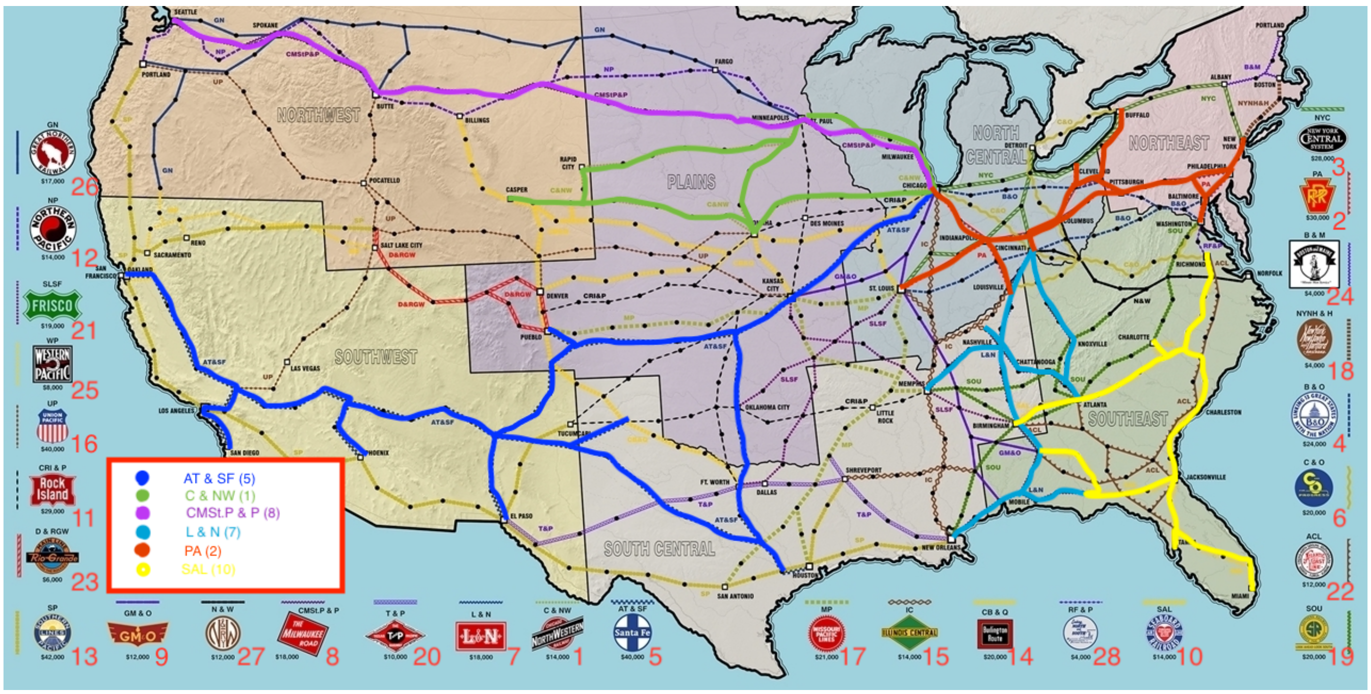

Figure 1: Shown above is a modified game board of Rail Baron. Each railroad is listed on the edge of the board, and the number next to each railroad shows it's ranking based on our analysis in subsection 3.4, with 1 being the best railroad and 28 being the worst. The highlighted railroads on the board show the optimal network of railroads, as chosen by us. The legend in the bottom left corner displays a color-coded guide for each railroad highlighted. We wanted to display a network of railroads because one railroad on its own isn't good enough; To win the game, one must concoct a network of railroads such that they can get to as many cities as possible.

References

[1] Straffin, Philip D. “Linear Algebra in Geography: Eigenvectors of Networks.” *Mathematics Magazine*, vol. 53, no. 5, 1980, pp. 269–76.

[2] Erickson, R.S., Erickson, T.F. Jr., ”Rail Baron”, *Avalon Hill*, 1977.

[3] Ewert Technologies, ”Rail Baron Tools”, *IOS App Store*, 2021.