Applications of Graph Theory and Probability in the Board Game *Ticket to Ride*

R. Teal Witter
rtealwitter@gmail.com
Middlebury College
Middlebury, Vermont

Alex Lyford
alyford@middlebury.edu
Middlebury College
Middlebury, Vermont

ABSTRACT

In the board game *Ticket to Ride*, players race to claim routes and connect cities on a map of the U.S. In this work, we identify winning strategies for and potential improvements to *Ticket to Ride* by applying probabilistic and graph-theoretic concepts. We find that longer routes are overvalued, presenting a simple winning strategy for opportunistic players. The scoring scheme we propose—based on indicator random variables—prevents exploitation from this strategy and improves the competitive nature of the game. Using a variety of game data visualizations, we also investigate why players who connect particular pairs of cities perform better than others. In addition, we build a statistical model from the effective resistance of the game's underlying graph structure to suggest how to choose the best pairs of cities.

ACM Reference Format:
R. Teal Witter and Alex Lyford. 2020. Applications of Graph Theory and Probability in the Board Game *Ticket to Ride*. In *International Conference on the Foundations of Digital Games (FDG '20), September 15–18, 2020, Bugibba, Malta.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3402942.3402963

1 INTRODUCTION

1.1 Previous Work

Board games have proved to be a productive area of mathematical intrigue [1, 3, 10]. One game with particularly elegant mathematical underpinnings is *Ticket to Ride.* [1] In this game, players race to connect cities and build railroads on a map of the U.S. Mathematically, the board can be thought of as a graph where cities represent nodes and the routes between them represent edges. The educational implications of *Ticket to Ride* have been extensively analyzed [2, 7, 9]. Previous work has also focused on improving player strategies in *Ticket to Ride*. Silva et al. used simulations to compare different heuristic strategies [5]. In related papers, the same authors explore the game space to find trends in the way *Ticket to Ride* is played and experiment with new maps and decks [4, 6].

[1] A. Moon. Ticket to Ride. [Board game]. Days of Wonder: Los Altos, CA, 2004.

FDG '20, September 15–18, 2020, Bugibba, Malta

ACM ISBN 978-1-4503-8807-8/20/09...$15.00
https://doi.org/10.1145/3402942.3402963

1.2 *Ticket to Ride* Gameplay

We give a cursory overview of *Ticket to Ride* and invite readers who have not previously played the game to familiarize themselves with the rules available online (linked here).

In *Ticket to Ride*, players race to claim routes and connect cities on a map of the U.S. shown in Figure 1. Each player starts with 45 physical trains and may use special colored cards to place them on the board. Players earn points for the length of the routes they claim according to Table 1 and whether they connect particular pairs of cities specified on Destination Tickets. The game ends when a player runs out of physical trains.

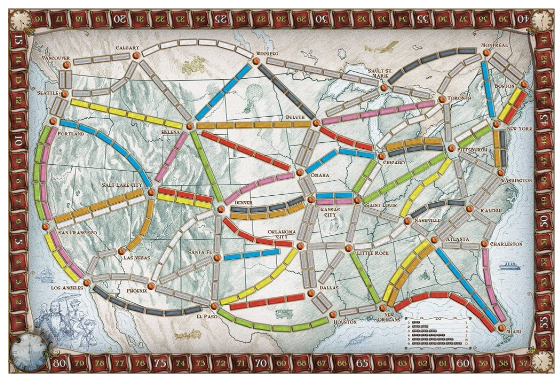

Figure 1: The *Ticket to Ride* U.S. board.

1.3 Our Results

We extend the mathematical interpretations of *Ticket to Ride* to devise player strategies and an improved scoring scheme that more fairly assigns points for building routes. Our approach calculates the expected time until certain moves can be made in the board game. We then argue that this expected time should be proportional to the reward for these moves.

To investigate the difficulty of connecting pairs of cities, we use both effective resistance and linear regression. Effective resistance is a measure of the difficulty an electrical current undergoes to travel between one node and another. We implement two existing algorithms to calculate the effective resistances between nodes and apply them to the *Ticket to Ride* board. By comparing the effective resistance, the reward of collecting a pair of cities, and the most successful pairs of cities in simulated games, we suggest strategies for players to pick the best cities to connect.

The overarching question of this work is as follows: What mathematical structure in *Ticket to Ride* can be exploited to optimize player strategies? Our answer builds on existing results and our own applications of mathematical concepts to enhance player strategies and propose a better scoring scheme.

1.4 Simulations

We used Silva et al.'s code (github.com/fernandomsilva/Ticket-to-Ride-Engine) to simulate *Ticket to Ride* games. We run our simulations on the USA map. However, our techniques can apply to other maps. For simulated four-player games, we include each of Silva et al.'s four agent types: The Destination Hungry Agent chooses Destination Tickets at the beginning of the game and tries to connect them for the rest of the game. The Route Focused Agent works toward a longer Destination Ticket, then claims longer routes. The One Step Agent operates by choosing the most advantageous move at each step without a long-term strategy. The Long Route Agent claims longer routes (defined as four trains or more) with a preference for routes that help connect the player's initial Destination Tickets. For two-player games, we simulate an equal number of all six pairings of the four agents.

2 VALUE OF ROUTES

2.1 Longer Routes are Overvalued

The reward for owning routes substantially increases for longer routes as observed in Table 1. The ostensible reason for the increasing value per train is that it takes much more time to collect the cards needed to claim longer routes than shorter routes. However, in Section 2.2 we show that the expected number of turns to collect k cards of a specific color is linear in relation to k (rather than polynomial as the current scoring method suggests).

Route Length	1	2	3	4	5	6
Points Scored	1	2	4	7	10	15
Points per Train	1	1	$1.\overline{3}$	1.75	2	2.5

Table 1: The points scored for building routes by length of route and per the number of trains.

There are several reasons that longer routes are overwhelmingly more valuable in the current scoring scheme. First, the reward per train for the longest route is *2.5* times that of the shortest route. Assuming a player has six train cards (of the same color), she gains 15 points from claiming the longest possible route but only 6 points for claiming six of the shortest routes. Second, players may claim only one route per turn. In the same scenario, a player must spend six turns claiming six of the shortest routes but only one turn claiming the longest route. Third, while it is very unlikely to initially pick six of the same color of trains, collecting four or so colors at a time almost guarantees that a player can accumulate sets of six trains of the same color.

k	Optimal Games	Points	Turns	Points per Turn
1	1 x 45	45	23 + 45	0.66
2	2 x 22, 1 x 1	45	23 + 23	0.98
3	3 x 15	60	23 + 15	1.58
4	4 x 11, 1 x 1	78	23 + 12	2.23
5	5 x 9	90	23 + 9	2.81
6	6 x 7, 3 x 1	109	23 + 8	3.52

Table 2: Optimal games that can be played with routes of at most length *k*. At least 23 turns must be spent collecting all 45 train car cards.

We illustrate optimal games in Table 2. As a strategy, claiming only routes of length six is incredibly advantageous: the number of points is staggering and can be quickly claimed in only a few turns. A potential flaw is that the two initial Destination Tickets may subtract from the total score unless they are easily connected with routes of length six.

2.2 Expected Cards to Build a Route

In order to simplify our analysis, we make two assumptions. First, we ignore the possibility of a player collecting a wild card. We justify this decision with the observation that taking a wild card from the face up pile counts as both cards a player may collect in her turn. The loss of an extra card is almost never worth it (given that a player is collecting multiple colors at once) except in end games where there is pressure to quickly claim a route. Second, we assume that other players draw cards uniformly so the probability that a particular color of card appears stays constant throughout the game. While in some games a player may specifically target a color (perhaps defensively), the relative scarcity of routes with the targeted color returns the proportion of each color to equilibrium.

Recall that claiming a route of length k requires k cards of one particular color. Our goal is to find the random variable N_k that represents the number of cards we need to draw until we find k of a particular color. We contend that $E[N_k]$ is a more equitable reward for routes of length k than the current scoring scheme that, as discussed in Section 2.1, overvalues longer routes.

Let C be the set of all cards and k be a fixed integer between 1 and 6, inclusive. Without loss of generality, say we are looking for k blue cards and call the set of blue cards B. In order to find the number of cards it takes to find k blue cards, think of our well-shuffled deck as blue cards separated by non-blue cards. For example, our deck may have the ordering $xxxbxbxxxx...xxbxxx$ where x is a non-blue card in $C \setminus B$ and b is a blue card in B.

Our strategy is to write N_k in terms of indicator random variables. Let $I_{k,x}$ be the indicator that takes value 1 if non-blue card x appears before the k^{th} blue card and 0 otherwise.

The number of cards until the k^{th} blue card is the number of blue cards k plus the number of non-blue cards before the k^{th} blue. Written as an equation, $N_k = k + \sum_{x \in C \setminus B} I_{k,x}$. We are interested in the long-run average number of cards drawn until the k^{th} blue card so we take the expectation of N_k and distribute over addition via linearity of expectation. Then $E[N_k] = E[k] + (|C| - |B|) \times E[I_{k,x}]$. Since k is fixed, $E[k]$ is simply k. Since B is a subset of C and both sets are fixed, $E[|C \setminus B|]$ is $|C| - |B|$. Thus we need only find $E[I_{k,x}]$.

To calculate $E[I_{k,x}]$, think of the deck as $|B| + 1$ sequences of non-blue cards separated by $|B|$ blue cards. (It is possible for a sequence to be of length zero in the case that two blue cards are adjacent to each other or in the case that the first or last card is blue.) Since we assume the deck is well-shuffled, the non-blue cards are uniformly distributed across the $|B| + 1$ sequences: it is as likely for card $x \in C \setminus B$ to be in any one sequence as any other.

When looking for only one blue card, we see that card x appears before the first blue if and only if x is in the first sequence. Since x is uniformly distributed across all $|B| + 1$ sequences, the probability that x appears before the first blue $P(I_{1,x})$ is $1/(|B| + 1)$. Similarly, the probability that x appears before the second blue $P(I_{2,x})$ is

$2/(|B|+1)$. By induction, the probability that x appears before the k$^{\text{th}}$ blue $\mathrm{P}(I_{k,x})$ is $k/(|B|+1)$.

Recall that $I_{k,x}$ takes value 1 if x appears before the k$^{\text{th}}$ blue and 0 otherwise. Then, conditioning on $I_{k,x}$, we write its expectation

$$\mathrm{E}[I_{k,x}] = 1 \times \mathrm{P}(I_{k,x}) + 0 \times (1 - \mathrm{P}(I_{k,x})) = \mathrm{P}(I_{k,x}) = \frac{k}{|B|+1}.$$

Through substitution, we have

$$\mathrm{E}[N_k] = k + (|C| - |B|) \times \left(\frac{k}{|B|+1}\right) = \left(1 + \frac{|C| - |B|}{|B|+1}\right) k.$$

and, by plugging in the total number of cards $|C|$ and total number of blue cards $|B|$,

$$\mathrm{E}[N_k] = \left(1 + \frac{110 - 12}{12 + 1}\right) k = \frac{111}{13} k.$$

While $111/13$ has no obvious units or interpretation, we have learned that $\mathrm{E}[N_k]$ is proportional to k. That is, the expected number of cards needed to purchase a route is linearly related to the length of the route.

2.3 Optimal Route Value via Simulations

Using the number of cards needed to purchase a route as a proxy for effort, it is clear from Section 2.2 that the reward for a route should be proportional to its length. Let α be the points per train so that $\mathrm{E}[N_k] = \alpha k$. In order to choose a good α, we simulate *Ticket to Ride* games and track how various strategies perform. For each α between 1 and 7 in .2 increments, we simulate 1000 games where the reward for a route of length k is αk. The results appear in Figure 2. (In the 20,000 games we simulated with the actual, non-linear scoring scheme, the Hungry strategy won 28% of games, the Path strategy won 8%, the One Step strategy won 29%, and the Long Route strategy won 37%. Note that the percentages add to 102% due to rounding and some tied games.)

It is evident that the Long Route agent wins more as α increases. The Hungry, Path, and One Step agents seem to distribute the games they otherwise would have won equally. The question of how many points per train to use in the scoring scheme is more a question of game design. However, the game would be more exciting if the three best strategies–Hungry, One Step, and Long Route–are closely tied. Thus we recommend an α value somewhere between 3.5 and 5.5.

3 DESTINATION TICKETS

3.1 Proportions of Wins

We now investigate the effect of collecting Destination Tickets on winning the game. We simulate 10,000 two-player and 10,000 four-player games. For each Destination Ticket, we calculate the proportion of games that the player holding the Destination Ticket won. Our results appear in Figure 3. Note that the vertical red lines at .25 and .5 represent the expected proportion of games players would win if Destination Tickets had no effect.

While most win less often than expected (we say that a Destination Ticket "wins" if the player who possesses it wins the game), roughly 10 of the 30 Destination Tickets win more often than expected. Of these, the ones that win the most are Montreal to Vancouver, New York to Seattle, and Chicago to Los Angeles. An

Figure 2: The proportion of wins of each of the four strategies by the number of points earned per train in 1000 games.

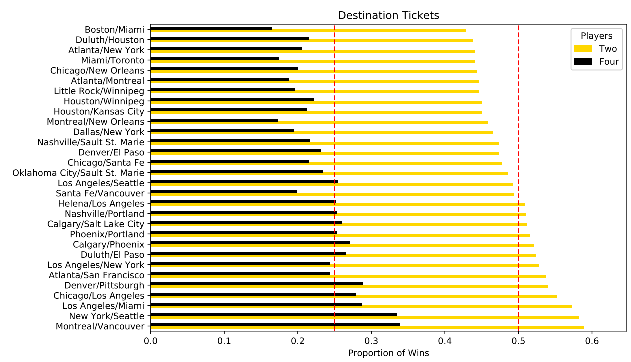

Figure 3: For each Destination Ticket, the proportion of two-player and four-player games that the player with the Destination Ticket won.

inspection of Figure 1 shows that the paths between the cities in these three Destination Tickets include the longest routes on the board. However, to gain a better understanding of what makes some Destination Tickets more advantageous than others, we investigate the characteristics that most closely correlate with winning.

3.2 Effective Resistance

The *Ticket to Ride* board can be mathematically interpreted as a graph where cities represent nodes and routes represent edges. One of the most powerful and natural measures of connectivity between two nodes on a graph is effective resistance [8]. Imagine that the *Ticket to Ride* board were a large electrical circuit with a unit current entering at city a and leaving at city b. Effective resistance indicates how much work a current needs to exert to get from a to b. In general, the number of paths and length of routes determine the effective resistance: cities with fewer paths and longer routes have a higher effective resistance between them than cities with many paths and shorter routes. We calculate the effective resistance for each Destination Ticket in order to gain insight into what makes some more advantageous than others.

We use two separate algorithms to find the effective resistance of Destination Tickets [8, 11]. Both algorithms calculate the effective resistance by finding the eigenvalues of the Laplacian matrix for a given graph. We let the weight between two cities be the length of the route connecting them.

A natural question is how well Destination Tickets' effective resistance correlates with wins. The relationship between resistance and the difference from expected proportion of wins has an R^2 value of .042 and a p-value of .281. Thus we would conclude that resistance is not related to winning. However, resistance is correlated with completing a route. The relationship between resistance and rate of Destination Ticket completions has an R^2 value of .295 and a p-value of .00192. Clearly, resistance is a good measure of the difficulty of connecting a pair of cites: the harder a pairs of cities is to connect, the lower the proportion of completions.

The weak relationship between resistance and wins and the strong relationship between resistance and completions motivates our analysis in Section 3.3.

3.3 Reward per Difficulty

We plot Destination Tickets in Figure 4 by the difficulty of connecting their cities (effective resistance) against their reward. The least squares line of best fit gives the predicted reward as a function of difficulty. We would predict, for example, that the reward for connecting the cities in a Destination Ticket with effective resistance .4 would be 12 points. However, completing the 2$^{\text{nd}}$ Destination Ticket Houston/Kansas City earns 5 points whereas completing the 30$^{\text{th}}$ Destination Ticket New York/Seattle earns 22 points even though both have resistance approximately .4. We conclude that Houston/Kansas City is very undesirable and New York/Seattle is very desirable. Coloring by the difference from the expected proportion of wins confirms our conclusion: Houston/Kansas City has value -.08 (winning 42% of two-player and 17% of four-player games) while New York/Seattle has value .09 (winning 59% of two-player and 34% of four-player games).

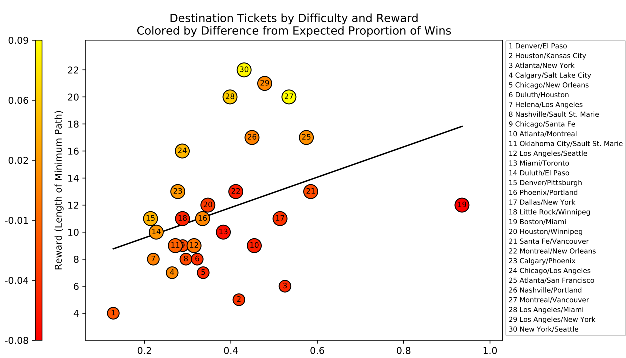

Figure 4: Destination Tickets plotted by the difficulty of connecting their cities and the reward.

Define a Destination Ticket's residual as the difference between its predicted reward and actual reward. For the same difficulty, a Destination Ticket with a high residual has a greater reward than one with a low residual. Clearly, the residuals of their Destination Tickets should affect whether players win: greater point values increase the number of points relative to the expended resources. We compare how closely residual and path length correlate with wins: the relationship between path length and wins has an R^2 value of .410 and a p-value of .00014 while the relationship between residual and wins has an R^2 value of .587 and a p-value of 8.25×10^{-7}. Both variables have low p-values and the residual has a sufficiently larger R^2 value. We therefore conclude that residual is a better metric of finding advantageous Destination Tickets than simply the minimum length path between cities.

4 LIMITATIONS & FUTURE WORK

We rely heavily on *Ticket to Ride* simulations but are limited to four fairly simple heuristic strategies. Future work could explore additional *Ticket to Ride* strategies in relation to the four we used as well as match players with the same strategies against each other. Our analysis of collecting Train Car cards ignored wild cards while our analysis of resistance ignored the routes that can be collected with any color. Future work could incorporate both nuances.

Our route value recommendations and Destination Ticket rankings are specific to the USA map. Future work may explore how our analysis applies to the more than twenty other *Ticket to Ride* maps. (Silva et al. indicate this is a productive avenue because their four heuristic strategies perform radically differently on other maps [5].) While our work can be directly extended to other *Ticket to Ride* maps, we do not know of any additional games with graph-theoretic structure that our techniques could apply to. Finally, we only explore changing the reward for routes. Future work could analyze changing the reward for Destination Tickets in conjunction with the reward for routes.

4.1 Acknowledgements

The authors would like to thank Bill Peterson for the idea behind the expected time calculation in Section 2.2.

REFERENCES

[1] Stephen D Abbott and Matt Richey. 1997. Take a Walk on the Boardwalk. *The College Mathematics Journal* 28, 3 (1997), 162–171.
[2] Wen-Chih Chang, Yan-Da Chiu, and Mao-Fan Li. 2008. Learning Kruskal's Algorithm, Prim's Algorithm and Dijkstra's Algorithm by Board Game. In *International Conference on Web-based Learning*. Springer, 275–284.
[3] Michael Cornelius. 1986. An Historical Background to Some Mathematical Games. *Mathematics in School* 15, 1 (1986), 47–49.
[4] Fernando de Mesentier Silva, Scott Lee, Julian Togelius, and Andy Nealen. 2017. AI as Evaluator: Search Driven Playtesting of Modern Board Games.. In *AAAI Workshops*.
[5] Fernando de Mesentier Silva, Scott Lee, Julian Togelius, and Andy Nealen. 2017. AI-based Playtesting of Contemporary Board Games. In *Proceedings of the 12th International Conference on the Foundations of Digital Games*. ACM, 13.
[6] Fernando de Mesentier Silva, Scott Lee, Julian Togelius, and Andy Nealen. 2018. Evolving Maps and Decks for Ticket to Ride. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*. ACM, 48.
[7] Peter Drake and Kelvin Sung. 2011. Teaching Introductory Programming with Popular Board Games. In *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education*. ACM, 619–624.
[8] Wendy Ellens, FM Spieksma, P Van Mieghem, A Jamakovic, and RE Kooij. 2011. Effective Graph Resistance. *Linear Algebra Appl.* 435, 10 (2011), 2491–2506.
[9] Darren Lim. 2007. Taking Students out for a Ride: Using a Board Game to Teach Graph Theory. *ACM SIGCSE Bulletin* 39, 1 (2007), 367–371.
[10] Alexander Lyford, Thomas Rahr, Tina Chen, and Benjamin Kovach. 2019. Using Camels to Teach Probability and Expected Value. *Teaching Statistics* 41, 1 (2019), 18–24.
[11] Fa-Yueh Wu. 2004. Theory of Resistor Networks: the Two-point Resistance. *Journal of Physics A: Mathematical and General* 37, 26 (2004), 6653.