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

R. Teal Witter & Alex Lyford

Middlebury College

January 16, 2020
Ticket to Ride (USA)
Overview

Routes
▶ Long routes are overvalued ... and can be used to easily win.
▶ We can find a better route scoring scheme with indicator random variables.

Destination Tickets
▶ Players with some Destination Tickets perform better than players with others ... why?
▶ We use regression to identify the best Destination Tickets.
## Current Route Values

<table>
<thead>
<tr>
<th>Route Length</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points Scored</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Points per Train</td>
<td>1.00</td>
<td>1.00</td>
<td>1.33</td>
<td>1.75</td>
<td>2.00</td>
<td>2.50</td>
</tr>
</tbody>
</table>
# Current Route Values

<table>
<thead>
<tr>
<th>Route Length</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points Scored</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Points per Train</td>
<td>1.00</td>
<td>1.00</td>
<td>1.33</td>
<td>1.75</td>
<td>2.00</td>
<td>2.50</td>
</tr>
</tbody>
</table>

**Arguments for**
- Collecting many trains of the same color is hard

**Arguments against**
- Is it really *that* hard?
- Only one route can be claimed per turn
- Collecting multiple colors simultaneously helps
Games with routes of length at most \( k \)

For all games, all 45 trains will be collected over 23 turns.\(^2\)

<table>
<thead>
<tr>
<th>( k )</th>
<th>Composition</th>
<th>Points</th>
<th>Turns</th>
<th>Points per Turn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 \times 45</td>
<td>45</td>
<td>23 + 45</td>
<td>0.66</td>
</tr>
<tr>
<td>2</td>
<td>2 \times 22, 1 \times 1</td>
<td>45</td>
<td>23 + 23</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>3 \times 15</td>
<td>60</td>
<td>23 + 15</td>
<td>1.58</td>
</tr>
<tr>
<td>4</td>
<td>4 \times 11, 1 \times 1</td>
<td>78</td>
<td>23 + 12</td>
<td>2.23</td>
</tr>
<tr>
<td>5</td>
<td>5 \times 9</td>
<td>90</td>
<td>23 + 9</td>
<td>2.81</td>
</tr>
<tr>
<td>6</td>
<td>6 \times 7, 3 \times 1</td>
<td>109</td>
<td>23 + 8</td>
<td>3.52</td>
</tr>
</tbody>
</table>

\(^2\)We ignore locomotives collected from the five face up cards.
Win Rate in Simulated Games

Wins by Strategy in 20,000 Games

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Win Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungry</td>
<td>0.28</td>
</tr>
<tr>
<td>Path</td>
<td>0.07</td>
</tr>
<tr>
<td>One Step</td>
<td>0.29</td>
</tr>
<tr>
<td>Long Route</td>
<td>0.35</td>
</tr>
</tbody>
</table>
How *should* routes be valued?

Idea: value = expected time to collect \(^3\)

\(^3\)measured in number of cards rather than turns
Expected number to find $k$ blue cards

Without loss of generality, our goal is to calculate the expected numbers to find $k$ blue cards.
Expected number to find $k$ blue cards

$N_k :=$ number of cards until $k$ blue cards are found

$C :=$ all cards

$B :=$ blue cards

$x \in C \setminus B$

$I_{x,k} := \begin{cases} 
1 & \text{if } x \text{ appears before the } k^{th} \text{ blue card} \\
0 & \text{otherwise} 
\end{cases}$

Then

$$N_k = k + \sum_{x \in C \setminus B} I_{x,k}$$
Expected number to find $k$ blue cards

$$N_k = k + \sum_{x \in C \setminus B} I_{x,k}$$

Taking expectation,

$$E[N_k] = k + (|C \setminus B|) \times E[I_{x,k}]$$

Recall $E[I_{x,k}] = 1 \times P(I_{x,k} = 1) + 0 = P(I_{x,k} = 1)$.  

Then

$$E[N_k] = k + (|C \setminus B|) \times P(I_{x,k} = 1)$$
Expected number to find $k$ blue cards

Think of the deck as non-blue cards separated by blue cards into $|B| + 1$ piles (possibly of size 0):

$$xxxbxbbxxxbxbb \ldots xbxxxxb$$

Then $P(I_{x,k} = 1)$ is $k/(|B| + 1)$
Expected number to find $k$ blue cards

$$
E[N_k] = k + (|C \setminus B|) \times \frac{k}{|B| + 1}
$$

$$
= \left(1 + \frac{110 - 12}{12 + 1}\right) k
$$

$$
= \frac{111}{13} k
$$

Thus our scoring should be linear!!
Choosing a scalar

Perhaps somewhere between 3.5 and 5?
Best and Worst

Best: Montreal/Vancouver, New York/Seattle

Worst: Boston/Miami
Effective Resistance?

A measure of connectivity between two nodes on a graph: electric flow from one node to another

$$\text{resistance} = \min_{\text{flows}} \sum_{\text{routes}} (\text{flow on route})^2$$

More, shorter paths $\rightarrow$ lower resistance
Effective Resistance!

Destination Tickets by Difficulty and Reward
Colored by Difference from Expected Proportion of Wins

1 Denver/El Paso
2 Houston/Kansas City
3 Atlanta/New York
4 Calgary/Salt Lake City
5 Chicago/New Orleans
6 Duluth/Houston
7 Helena/Los Angeles
8 Nashville/Sault St. Marie
9 Chicago/Santa Fe
10 Atlanta/Montreal
11 Oklahoma City/Sault St. Marie
12 Los Angeles/Seattle
13 Miami/Toronto
14 Duluth/El Paso
15 Denver/Pittsburgh
16 Phoenix/Portland
17 Dallas/New York
18 Little Rock/Winnipeg
19 Boston/Miami
20 Houston/Winnipeg
21 Santa Fe/Vancouver
22 Montreal/New Orleans
23 Calgary/Phoenix
24 Chicago/Los Angeles
25 Atlanta/San Francisco
26 Nashville/Portland
27 Montreal/Vancouver
28 Los Angeles/Miami
29 Los Angeles/New York
30 New York/Seattle
Rankings by Minimum Path Length and Residual

- New York/Seattle
- Los Angeles/New York
- Los Angeles/Miami
- Montreal/Vancouver
- Chicago/Los Angeles
- Nashville/Portland
- Atlanta/San Francisco
- Calgary/Phoenix
- Denver/Pittsburgh
- Montreal/New Orleans
- Houston/Winnipeg
- Little Rock/Winnipeg
- Duluth/El Paso
- Phoenix/Portland
- Santa Fe/Vancouver
- Oklahoma City/Sault St. Marie
- Chicago/Santa Fe
- Miami/Toronto
- Helena/Los Angeles
- Los Angeles/Seattle
- Dallas/New York
- Nashville/Sault St. Marie
- Duluth/Houston
- Calgary/Salt Lake City
- Atlanta/Montreal
- Chicago/New Orleans
- Denver/El Paso
- Boston/Miami
- Houston/Kansas City
- Atlanta/New York
Summary

Routes
▶ Long routes are overvalued ... and can be used to easily win.
▶ We can find a better route scoring scheme with indicator random variables.

Destination Tickets
▶ Players with some Destination Tickets perform better than players with others ... why?
▶ We use regression to identify the best Destination Tickets.
Thank you!

Questions?
Correlations with overall wins

- **Destination Tickets by Path Length and Wins**
  - Pearson coefficient: 0.64
  - p-value: 0.00014

- **Destination Tickets by Resistance and Wins**
  - Pearson coefficient: -0.204
  - p-value: 0.281

- **Destination Tickets by Residual and Wins**
  - Pearson coefficient: 0.766
  - p-value: 8.25e-07
