**Visualizing Trading Card Game Metadata**

Jericho McLeod

[Jmcleod3@gmu.edu](mailto:Jmcleod3@gmu.edu)

**Background and Motivation**

The trading card game, Magic: The Gathering™ (MTG), produced by Wizards of the Coast, LLC, has become one of the largest such games in the world since its inception in 1993. Over that time, it has grown to have 35,000,000 active players (Webb, 2018) who build and compete with 60 cards decks chosen from a set of 19,989 cards (Gamepedia, 2019). The quantity of cards and deck size mean there are 1.226859x10^176 possible combinations in total, not accounting for decks that are non-functional, and not considering sideboards, a second set of cards from which a player can reconstruct a deck mid-tournament. Given the complexity of the game, the rate at which new cards are added, and the element of randomness inherent in games involving shuffled decks of cards, it is unlikely to ever be solved.

With the rise of the internet, however, players still managed to converge on the most competitive combinations of cards that had been discovered. Over time, the increased demand for specific cards has led to a rich secondary market for cards, where the cards making up more competitive decks has been driven up. This makes the cost of changing competitive decks high in some formats.

The prize pools of tournaments are high enough to encourage such changes in competitive decks despite the costs. Players choose decks based on cost, performance, the difficulty of play, and preference. The costs of decks, and of swapping between two decks at a particular time, is deterministic given the public market for cards. Performance can be measured to some degree through tournament results, though the random aspect of the game confounds this to some degree. I am primarily interested in better understanding the behavior of players for the more difficult to analyze facets of MTG deck selection.

**Project Objectives**

Understanding player decisions for deck difficulty, preference, and performance, first requires examination of the discrete data available for costs and observed deck performance. This project aims to create a framework for studying cost and performance by relating the most popular decks in a network encoding these facets as a first step in better understanding player behavior through future research.

An additional benefit of the project is its use to players. Once the framework for reviewing costs is created and automated, the site will remain an up-to-date tool for players to use to receive more transparent and consolidated information regarding the cost and performance of decks.

**Data**

Data for tournament results is published in numerous locations. I will gather tournament results and decklists from the website MTGTop8 using Python.

Additional information regarding cards, and validation of card names, will be conducted by using data from the website MTGJSON, where the full list of cards and feature data is publicly available. This data will also be handled using Python.

Pricing data, if needed, will be obtained from TCGPlayer, a common market website for MTG that publishes market prices for cards based on the history of the cards accepted price among buyers and sellers. TCGPlayer offers a developer API that will be utilized in the final deliverable of this project to serve updated prices at the time the page is served.

**Data Processing**

The data is easily accessible on webpages, through cleaned JSON files, or through APIs, thus will be obtained in a relatively clean format. Validation will still be performed on the individual card lists of tournament entrant decks by comparing them to card names from the JSON list of all cards. The validation script will then return the highest cosine similarity of the encoded characters, showing the most likely correct names, to simplify any manual identification that must take place.

The quantity of decks appears high at first glance, as the target period of January through March of 2019 contains 611 legacy decks in tournament results, but of these, the top 12 most common decks account for 62% of decks played. This means that the most vital component of the framework can be built upon a smaller sample while still containing a large number of deck variants if data processing presents problems.

**Visualization Design Concepts**

**Design 1:**

Data will be visualized as a network graph showing the top played deck designs in the Legacy format as nodes. Node size will encode data metadata for the percentage of tournament entrants using the deck, and node color will encode data for win percentages. Edges will encode data for the quantity of shard cards between decks as thickness. Costs will be encoded either as tints and shades of a single color or displayed as a tooltip when hovering over an edge. When a single deck type is selected, the above graph will be replicated for variants of the chosen deck design. Tools would be implemented for exact values when hovering over elements.

**Design 2:**

Given that the costliest cards in the format are a subset called Lands, decks could be linked showing land costs of switching lands. An interactive tool for inputting lands owned by a player using the site would show the deck designs available to players without additional land investment, and the costs of moving to other decks. The available decks without investment would be encoded as full opacity nodes, while additional investment decks would be displayed with partial opacity. The costs of swapping would be encoded in edge thickness, with a tooltip to display exact values as well as win rates for nodes.

**Design 3:**

Another perspective would be primarily review decks by win percentages, relating decks by average win rates across tournaments to account for variations in metagame analysis. In this case, decks would be linked by coappearance in tournaments, helping to illuminate higher overall performance. Win rate would be a size encoding, and deck cost could potentially be encoded as color data. There are imperfections in this format from the available data, however, as merely appearing in a tournament does not necessarily indicate that two particular decks competed directly.

**Must-Have Features**

* A network visualization MTG Legacy Tournament deck data
* Data encoding two of the factors used by players to make decisions for which decks to play

**Additional Features**

* Additional linked data for player utility, such as deck-lists and pricing data
* A list of the card difference for two selected decks

**Project Schedule**

* **March 26.** Proposal due
* **April 2.** Data gathering completed
* **April 9.** Data manipulation code completed
* **April 16.** (Milestone Demo) Draft site completed, visualization work has begun, placeholders on site for final product
* **April 23.** Visualizations completed
* **April 30.** Dazzlers completed (overflow time if behind)
* **May 7.** Process book completed
* **May 13.** Process book due; complete any remaining work
* **May 14.** Present final product

# Bibliography

Gamepedia. (2019, 2 16). *Magic: The Gathering Statistics and Trivia*. Retrieved from Gamepedia: https://mtg.gamepedia.com/Magic:\_The\_Gathering/statistics\_and\_trivia

MTGJSON. (2019, 2 26). *MTGJSON*. Retrieved from https://mtgjson.com/json/AllCards.json

MTGTop8. (2019, 3 23). *Legacy Events*. Retrieved from MTGTop8: https://www.mtgtop8.com/format?f=LE

TCGPlayer. (2019, 2 26). *TCGPlayer APIs*. Retrieved from TCGPlayer: http://developer.tcgplayer.com/

Webb, K. (2018, 12 18). *With more than 35 million players worldwide, Magic the Gathering is giving back to its community with a brand new game and $10 million in esports prize money*. Retrieved from Business Insider: https://www.businessinsider.com/magic-the-gathering-announces-10-million-esports-program-for-2019-2018-12

Wizards of the Coast, LLC. (2019, 2 16). *Rules | Magic: The Gathering*. Retrieved from Wizards: https://magic.wizards.com/en/game-info/gameplay/rules-and-formats/rules

**Links**

<https://www.businessinsider.com/magic-the-gathering-announces-10-million-esports-program-for-2019-2018-12>

<https://mtg.gamepedia.com/Magic:_The_Gathering/statistics_and_trivia>

<https://en.wikipedia.org/wiki/Magic:_The_Gathering_rules>

<https://www.mtgtop8.com/format?f=LE>

<https://mtgjson.com/json/AllCards.json>

<http://developer.tcgplayer.com/>