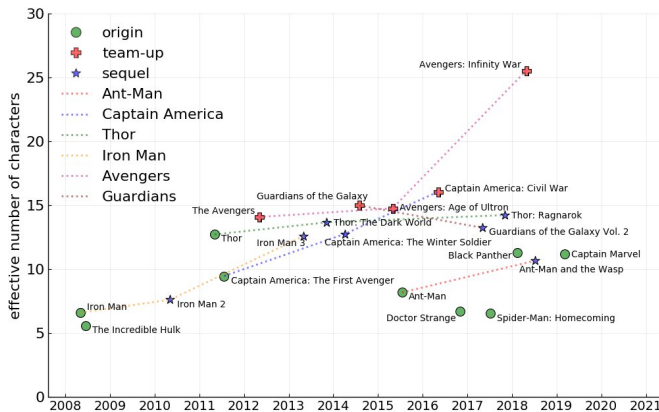


# How the Avengers assemble: Ecology-based metrics model effective cast sizes for Marvel movies

5 July 2019, by Ingrid Fadelli



Effective cast size of each movie in the MCU showing type of movies by shape, and sub-sequences connected by dashed lines. The x-axis is the theatrical release date. Credit: Roughan, Mitchell & South.

In a recent study, researchers at the ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS) in Adelaide have tried to use ecology-related concepts to model effective cast sizes for movies, focusing on characters from the Marvel Cinematic Universe (MCU). Their research, outlined in a paper pre-published on arXiv, gathered interesting findings that could shed light on some factors associated with the success of Marvel movies.

"We're huge fans of the recent suite of Marvel [movies](#), the Marvel Cinematic Universe, as it's called," Matthew Roughan, one of the researchers who carried out the study, told TechXplore. "We feel that the producers, directors, actors and the rest of the large production family are doing something unique in the history of movie making, so we set out to quantify that."

The study carried out by Roughan and his

colleagues Lewis Mitchell and Tobin South lies at the intersection between statistics, [computer science](#) and data science. The researchers collaborated with the data science team at the University of Adelaide, who has been working on a wide range of projects studying the internet, media, and biology.

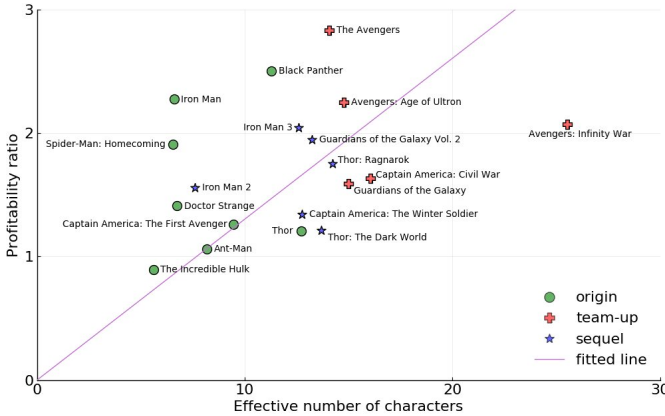
"In biology there is a need to measure the [ecological diversity](#) of a habitat to understand its health and resilience," Roughan explained. "Biologists use a type of measurement that we thought could apply to movies. The hope is that this measurement might be as valuable in analyzing movies as it is in the study of [biodiversity](#)."

Roughan and his colleagues applied a metric commonly used in ecology research to movie casts. This metric is based on the notion of Shannon-entropy, which describes the inherent uncertainty of the distribution of species in a given region, with a higher uncertainty suggesting that there is greater diversity.

"Simply put, if it is harder to guess what species you are observing (assuming you know nothing about taxonomy) at any given moment, there must be more possibilities out there," Roughan said. "An analogy could be a multiple choice question- if it is harder to guess the answer, then there is more entropy. Think of it as measuring how many effective answers there are to the question. Some answers may be obviously wrong, so you don't count them seriously."

In their study, the researchers showed that an entropy-based metric can be generalized using a statistical method called Jensen-Shannon divergence, ultimately offering a measure of the similarity of characters appearing in different movies. This could be particularly useful in

recommender systems for media streaming services, such as Netflix, Amazon Prime Video, etc.



Profitability as a function of effective cast size. Credit: Roughan, Mitchell & South.

"The size of a cast of characters is surprisingly hard to define," Roughan explained. "There are so many small parts that are never-the-less important. Some are credited and many are not, but even the standard for what warrants credit is surprisingly variable. In ecology, they have a similar problem. It's hard to count all the species in a region. However, they have been using a metric based on Shannon entropy to get a grip on that problem."

The application of entropy in contexts other than biology has already been achieved in previous works, for instance [to quantify the size of a person's vocabulary](#).

Roughan and his colleagues used it to measure the most effective number of characters for movies, focusing on Marvel movies. Their analyses were mostly based on data from public sources, such as transcribed movie scripts, yet the researchers also created a new dataset specifically for this study.

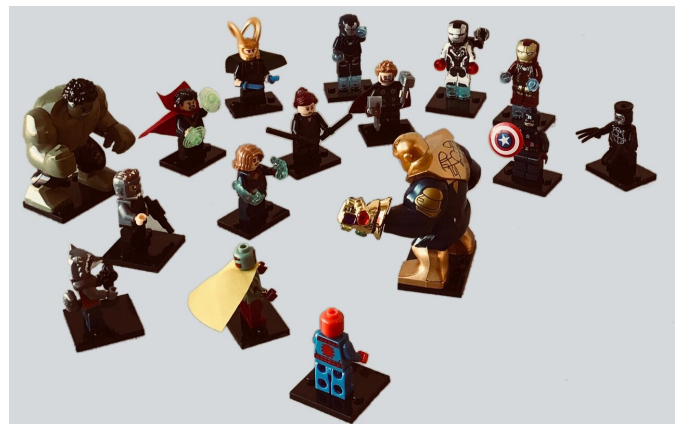
"We watched the entire MCU again (that was the fun bit) and annotated it with information about the conflicts in the movies," Roughan said. "That allowed us to measure how much each [character](#) participated in each movie. From there the entropy

calculation is actually pretty easy mathematics."

Based on the data they gathered, the researchers compared different Marvel movies based on their cast sizes. This allowed them to identify patterns in the data, for instance clustering movies into groups based on particular cast characteristics.

"The biggest surprise for us was that the effective cast size is correlated with the profitability of the movies, with a bigger role-call translating in bigger profits," Roughan said. "However, you have to be very careful about such results. What we observe is only a correlation—we can't get causation from that. We think the true reason for the correlation isn't just that audiences like bigger casts. The real reason is part of the uniqueness of the MCU."

According to Roughan, MCU producers created a series of movies that pave the way towards the assembly of Marvel characters. They first released movies that focused on individual characters, such as Iron Man, Captain America and Thor, then ultimately featured all of these characters as part of the Avengers team. They then repeated this process by releasing origin movies for new characters, leading to increasingly bigger "teams."



Credit: Matthew Roughan.

"That took a special kind of vision, to be willing to develop these characters over multiple movies to build up to an amazing culmination over a period of

years," Roughan added. "It's so different from the typical franchise, which is a series of sequels (and sometimes prequels) with roughly the same set of characters."

Although the results gathered by Roughan and his colleagues do not clarify whether cast sizes have directly influenced profits made from Marvel movies, they offer some interesting insight about the correlation between these two factors. In addition, the researchers showed how metrics used in ecology research can be applied to studies that focus on entirely different topics.

"I think we are just scraping the surface here," Roughan said. "What makes a movie or a franchise work is tremendously complex, and you cannot undervalue the contributions of the brilliant directors, actors and other artists who created these movies. Historically, media analysis has been in the hands of social scientists, who analyze the human pieces of the puzzle, identify tropes, and describe how we feel about movies."

According to Roughan, data science could soon aid our understanding of many different research areas. For example, by analyzing the large amount of data collected over the years, data scientists could better understand factors associated with the success (or failure) of movies, as well as TV-series, books, and so on.

Roughan believes that this shift in the perceived value of [data science](#) resembles what happened a few decades ago, when sports teams started realizing that hard data and statistics could drive them towards victory. In the case of movies, studies such as the one carried out by him and his colleagues could ultimately inform new productions, providing valuable insight into factors that might determine their failure or success.

"At a deeper level, stories are so important for humans," Roughan said. "It is fair to say, I think, that stories are what make us human; what differentiates us from the rest of the natural world. We would really like to make a contribution to understanding how and why that is so."

**More information:** How the Avengers assemble:

ecological modelling of effective cast sizes for movies. arXiv:1906.08403 [cs.SI]. [arxiv.org/abs/1906.08403](https://arxiv.org/abs/1906.08403)

ALEPH-ZERO-HEROES: [aleph-zero-heros.info/](http://aleph-zero-heros.info/)

Philippe Thoiron. Diversity index and entropy as measures of lexical richness, *Computers and the Humanities* (2006). [DOI: 10.1007/BF02404461](https://doi.org/10.1007/BF02404461)

© 2019 Science X Network

APA citation: How the Avengers assemble: Ecology-based metrics model effective cast sizes for Marvel movies (2019, July 5) retrieved 17 September 2021 from <https://techxplore.com/news/2019-07-avengers-ecology-based-metrics-effective-sizes.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*