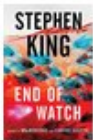


Analyzing book reading behavior on Goodreads to predict Amazon Bestsellers

September 26 2018, by Ingrid Fadelli

Status Updates From End of Watch (Bill Hodges T...



End of Watch (Bill Hodges Trilogy, #3)











by Stephen King (Goodreads Author)

[all editions](#) | [this edition](#)

[show all updates](#) | [only show updates with text](#)

Status Updates (showing 91-120 of 16,316)

order by

-  Caleb is starting
– Jun 07, 2016 09:27AM [Add a comment](#) 
-  Caleb is 8% done
– Jun 07, 2016 09:27AM [Add a comment](#) 
-  Kristina is 50% done
– Jun 07, 2016 09:40AM [Add a comment](#) 
-  Edward Lorn is 5% done
King-verse Alert: Inside View
– Jun 07, 2016 09:40AM [Add a comment](#) 
-  Kevin is starting
– Jun 07, 2016 09:50AM [Add a comment](#) 

Status page for “End of Watch (Bill Hodges Trilogy #3) by Stephen King”. Credit: Maity, Panigrahi & Mukherjee.

Researchers at Northwestern University, Microsoft Research India, and

the Indian Institute of Technology Kharagpur have recently developed a model to predict whether a book will become a bestseller on Amazon within 15 days of its publication. Their model, outlined in a study pre-published on arXiv, works by analyzing reading behavior on the online platform Goodreads.

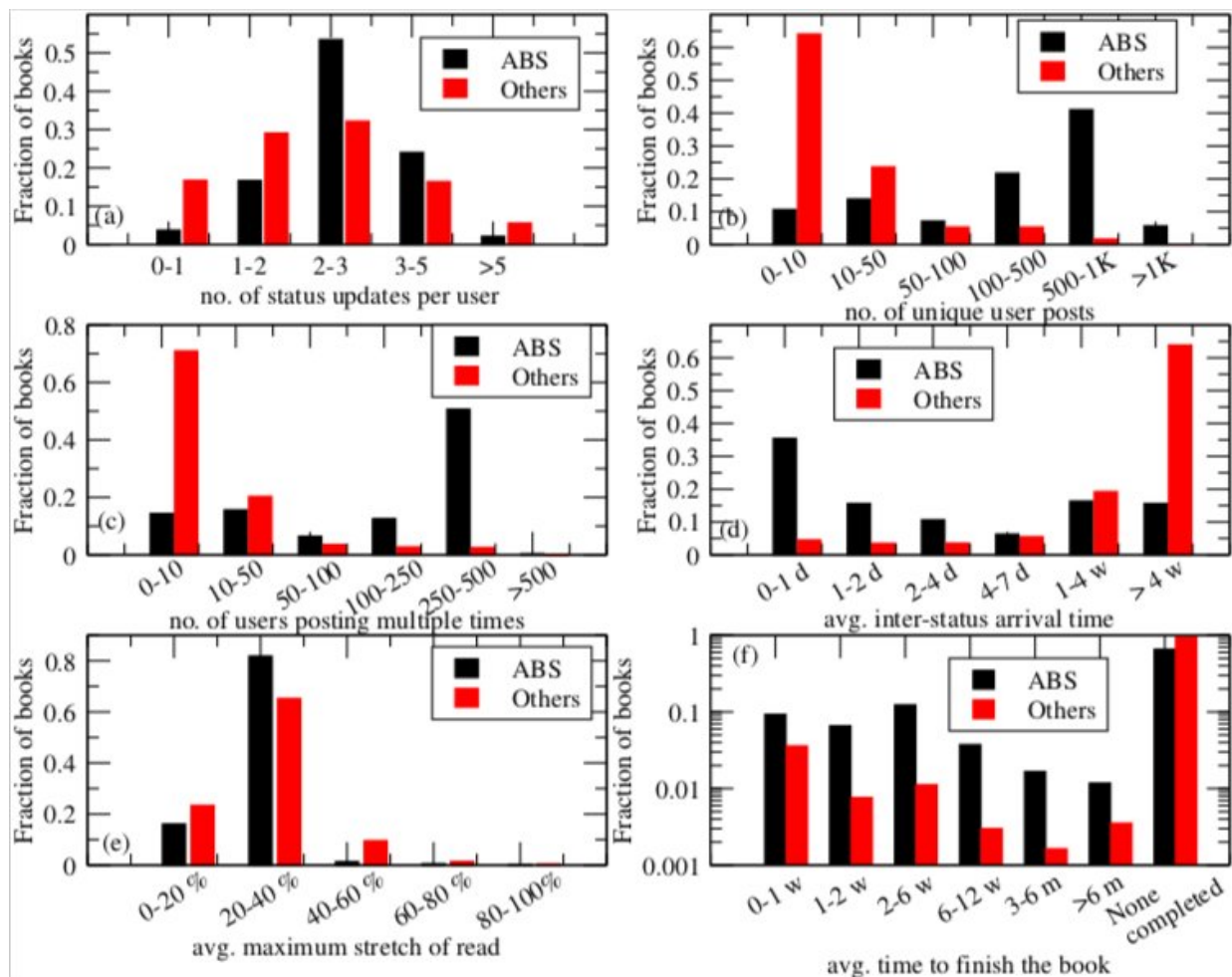
"We have been working on analyzing the [popularity](#) dynamics of various social media entities, such as hashtags in Twitter, topics in Quora etc." Animesh Mukherjee, one of the researchers who carried out the study told TechXplore. "We felt that a similar approach could be taken to analyze the popularity of books and we found Goodreads to be ideal for this investigation."

A book's popularity depends on a multitude of factors and can be measured using several parameters. In their study, the researchers focused on how book reading characteristics influence its popularity. They performed a cross-platform analysis of Goodreads entities and tried to link these with the volume of sales for books on Amazon.

"We followed the intuition that the popularity of books is mostly driven by its readers, hence the motivation to extract book reading behavior to understand the future popularity of books," Mukherjee said. "One of the best ways to quantify the popularity of books is to look at its sale record. Thus, we tried to quantify the notion of popularity in terms of Amazon bestsellers."

To begin with, the researchers analyzed the collective reading behavior of users on Goodreads. They then quantified different characteristic features of Goodreads entities, which could be used to identify differences between Amazon bestsellers and other lesser-selling books. Finally, they developed a machine learning-based model that uses these characteristic features to predict whether a book will become a bestseller 15 days after its publication.

"We used state-of-the-art machine learning models to perform our predictions," Mukherjee explained. "We observed that the ratings and reviews received by a book on Goodreads are not as effective in predicting the bestsellers as the reading status post patterns of the users. For example, in Goodreads, a reader can post how much of the book has been read, which page he/she is on, can comment about the book etc. We find these features to be very effective in predicting whether the book is going to be a bestseller in the future."



Characteristic properties of Goodreads users' status posts: distribution of a) number of status updates per user b) number of unique users updating status c) number of users updating multiple times d) inter-status arrival time e) average

maximum stretch of reading f) average time to finish reading for ABS vs other books. Credit: Maity, Panigrahi & Mukherjee.

Their model achieved a very promising average accuracy of 88.72 percent in predicting books that would become Amazon bestsellers a few weeks after their publication. Their method, which was based on features derived from user posts and genre-related properties, attained an improvement of 16.4 percent compared to baseline methods that only use traditional popularity factors, such as book ratings or reviews.

"One of the most important insights that we obtain from this study is that the Amazon bestseller books might not necessarily be qualified by high-quality review text of the readers or a high volume of ratings," Mukherjee said. "In contrast, a large majority of them have reader status post patterns that strongly distinguish them from the rest of the books."

The researchers also evaluated how well their method could predict two further types of books: highly rated ones that receive a large number of reviews but are not bestsellers (HRHR), and Goodreads Choice Award-nominated (GCAN) books that are not bestsellers. They achieved a high average accuracy of 87.1 percent for GCAN and of 86.22 percent for HRHR books.

"We believe that this work is an important contribution to the current literature as it not only unfolds the collective reading behavior of a social book-reading platform through a rigorous measurement study but also establishes a strong link between two orthogonal channels – Goodreads and Amazon," Mukherjee said.

The model developed by Mukherjee and his colleagues could foster the development of tools that bridge Amazon and Goodreads via new cross-

platform policy designs. They believe that such interactions might be one of the reasons behind Amazon's acquisition of Goodreads in March 2013. The researchers are now looking to expand their study by further analyzing reading behaviors of users.

"There are several directions that we plan to explore in future," Mukherjee said. "One is to investigate the popularity of different genres of books—for instance, what are the status post patterns across different genres of [books](#)? Another is to study the inter-dynamics of genre and reader demographics. For example, how do reading behaviors of males differ from females, or how do they differ across various continents?"

More information: Analyzing social book reading behavior on Goodreads and how it predicts Amazon Best Sellers.
arXiv:1809.07354v1 [cs.SI]. arxiv.org/abs/1809.07354

© 2018 Tech Xplore

Citation: Analyzing book reading behavior on Goodreads to predict Amazon Bestsellers (2018, September 26) retrieved 20 April 2024 from <https://techxplore.com/news/2018-09-behavior-goodreads-amazon-bestsellers.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.