



Temporal Opinion Spam Detection by Multivariate Indicative Signals



Thanks

Stony Brook University



Junting Ye
Stony Brook University



Santhosh Kumar
Stony Brook University



Leman Akoglu
Stony Brook University

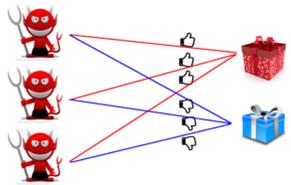


Problem

Background

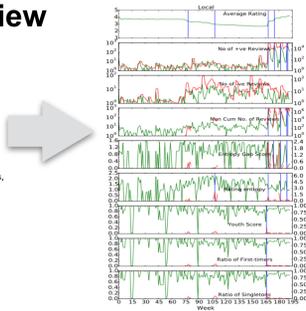
Important to businesses' revenue!

Pay spammers to write **fake reviews**



Problem Statement

Input: products' review streams



Output: targeted products at time t .

Approach

Indicative Signals

Name	Range	Suspicious if
Average Rating	[1, 5]	Change
Number of (+/-) Reviews	[0, ∞]	Increase
Rating Entropy	[0, log ₂ 5]	Decrease
Ratio of Singletons	[0, 1]	Increase
Ratio of First-timers	[0, 1]	Increase
Youth Score	[0, 1]	Increase
Temporal Gap Entropy	[0, max ϵ]	Decrease

†With windows size ΔT and logarithmic binning, number of bins is $\lceil \log_2 \Delta T \rceil + 1$ and $max\epsilon = \log_2(\lceil \log_2 \Delta T \rceil + 1)$.

Overview

1. Temporal Signal Extraction;

2. Anomaly Detection in **Lead Signal**;

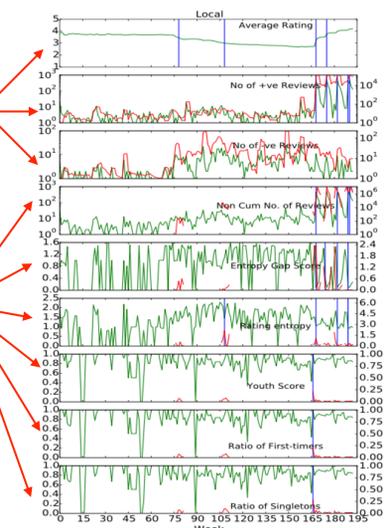
- CUSUM for average rating;
- Autoregressive model (AR) for others;

3. Anomaly Detection in **Supporting Signals**;

- Analyze **local values** only when "alarms" triggered by lead signal;
- Use **AR** to detect anomalies;

4. Suspiciousness Quantification;

- 4 features** to characterize anomalies;
- Integrate** features into single value;



Contributions

- Problem formulation:** Descriptive and general to detect spam reviews by monitoring indicative signals;
- New methodology:** Online and efficient algorithm;
- Validated** the method on real-world datasets.

Code and Data available:
<http://www3.cs.stonybrook.edu/~juyye/>

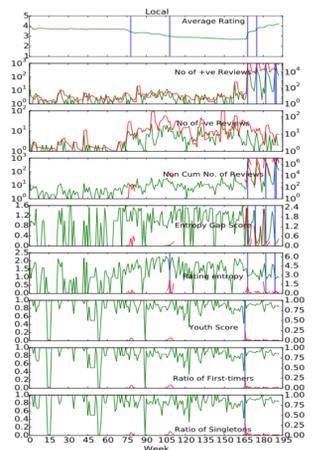
Experiments

Data sets

Dataset	Reviewer#	Product#	Review#	Start Date	End Date	Week#
iTunes	966K	15K	1.1M	Jul. 2008	Apr. 2012	198
Flipcart	1.1M	550K	3.3M	Aug. 2011	Jan. 2015	180

Case 1

- Product from iTunes
- Burst in # of positive reviews: **every 7 weeks**;
- Duplicate review texts, e.g.
 - "Great app for gamers"
 - "Great App For Gaming news"
 - "Must have app for gamers"
 - "One of my favorite apps"
- Synchronized** extreme ratings (5-star)



Case 2

- Product from Flipcart
- Week 35** is detected suspicious;
- Spammer group** is detected;

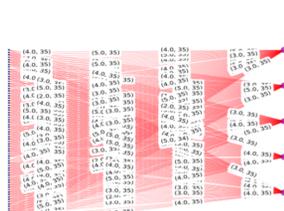


Figure 12: Reviewers (left) concurrently spammed multiple hair products (right) from FLIPKART, including those in Figure 10. Edge labels: (rating, campaign week)

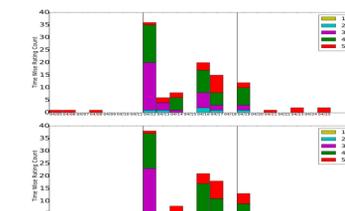


Figure 11: Daily review counts before, during, and after the spam campaign week 35 for two FLIPKART products in Figure 10.

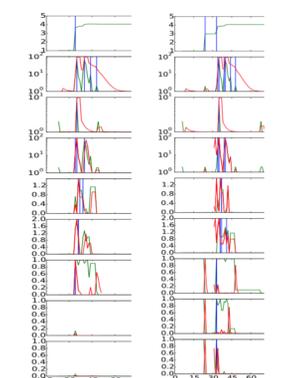


Figure 10: Partial time series for 9 indicative signals for two different products from FLIPKART that were spammed by same reviewers during same time periods (week 35 and 40).

Case 3

- Product from Flipcart
- 125 5-star reviews** in 2 days, while less than **10 reviews** during other weeks;
- These reviewers also reviewed another book, **same author, same order** during the **same days**!

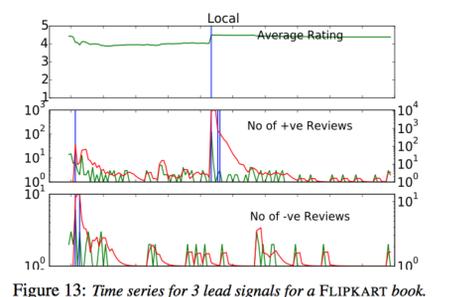


Figure 13: Time series for 3 lead signals for a FLIPKART book.

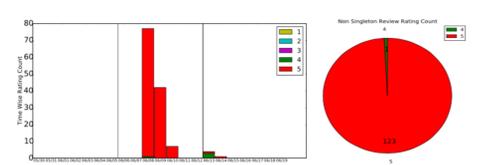


Figure 14: (left) Daily review counts, and (right) Rating distribution by non-singletons in week 35 for product in Figure 13.