

# Discovering Opinion Spammer Groups by Network Footprints (Supplementary Material)

Junting Ye      Leman Akoglu

Department of Computer Science, Stony Brook University  
{juyye, leman}@cs.stonybrook.edu

In this supplementary material, we present the evidence we observe in side information (including review time stamps, star ratings, and review text) to verify the suspiciousness of detected groups. We also show that our method is able to detect spammer groups with different behavioral patterns.

*Example Groups:* After using NFS and *GraphStrainer*, we detect a total 16 spammer groups from two real-world datasets, **iTunes** and **Amazon**. We verify each of these groups based on side information and observe collective behaviors of high suspiciousness. For brevity, we discuss further details of two groups, in particular, **iTunes#2** (i.e., group (2) in iTunes) and **Amazon#1** (i.e., group (1) in Amazon) (see Table 1).

**Table 1.** Example Groups from **iTunes** and **Amazon**: #P (#U): number of products (users), t: time stamps, ★: star-ratings. (for distributions): c: concentrated, s: scattered.

Groups	#P	#U	t, ★	#Duplicates/#Reviews	Developer/Owner
<b>iTunes#2</b>	8	38	c,s	29/202	2 same
<b>Amazon#1</b>	10	20	c,c	90/138	all same

*Evidence & Analysis:* To verify the effectiveness of our method, we present the evidence from three perspectives: temporal, rating, and text based evidence (shown in Figure 1). (1) **Temporal evidence:** All the reviews in both **iTunes#2** and **Amazon#1** are concentrated in time (i.e., both collective behaviors), although in different ways. In **iTunes#2** all users write reviews to a given product on the same day, while in **Amazon#1** a given user writes reviews to all products on the same day. (2) **Rating evidence:** In **iTunes#2**, most review ratings (about 80%) are either 4- or 5-star (lower ratings are possibly for camouflage given they have highly suspicious behaviors in other aspects). Ratings in **Amazon#1** are much more extreme. Users only give either 5-star (76%) or 1-star (24%) ratings to all of their target products. (3) **Text evidence:** As shown in the right-most of Figure 1, we find three different patterns when duplicate reviews are observed: (i) different users write duplicates to the same product, (ii) different users write to different products, and (iii) the same user writes to different products. Writing duplicate reviews is a highly suspicious behavior. Interestingly, however, these two groups exhibit different patterns in creating duplicates: **iTunes#2** contains type (i) and (ii) duplicates, while **Amazon#1** only has type (iii) duplicates. Based on this observation, we find that spammer groups behave in different ways. Our method is capable of finding them, since we do not assume/use any particular behavioral patterns, but rather identify suspicious network footprints that emerge as a result of the collusion among spammers.



**Fig. 1.** Abundant evidence of suspicious patterns for spammer groups even though their behavioral patterns are very different.