2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

Chelsea Won, and You Bought a T-shirt: Characterizing the Interplay Between Twitter and E-Commerce

Haipeng Zhang*[1], Nish Parikh[2], Gyanit Singh[2], Neel Sundaresan[2]
[1]School of Informatics & Computing, Indiana University, Bloomington, IN
[2]eBay Research Labs, San Jose, CA, USA
Email: zhanhaip@indiana.edu, {nparikh,gysingh,nsundaresan}@ebay.com

***Abstract*—The popularity of social media sites like Twitter and Facebook opens up interesting research opportunities for understanding the interplay of social media and e-commerce. Most research on online behavior, up until recently, has focused mostly on social media behaviors and e-commerce behaviors independently. In our study we choose a particular global e-commerce platform (eBay) and a particular global social media platform (Twitter). We quantify the characteristics of the two individual trends as well as the correlations between them. We provide evidences that about 5% of general eBay query streams show strong positive correlations with the corresponding Twitter mention streams, while the percentage jumps to around 25% for trending eBay query streams. Some categories of eBay queries, such as 'Video Games' and 'Sports', are more likely to have strong correlations. We also discover that eBay trend lags Twitter for correlated pairs and the lag differs across categories. We show evidences that celebrities' popularities on Twitter correlate well with their relevant search and sales on eBay. The correlations and lags provide predictive insights for future applications that might lead to instant merchandising opportunities for both sellers and e-commerce platforms.**

I. Introduction

The explosion of social media has enabled the spread of public opinion, intentions, and observations, sometimes, even at a faster pace than traditional news media [1]–[3]. This has enabled many applications that mine the (social) web to generate political, scientific and real-world values, including predicting election outcomes [4], [5], monitoring the spread of flu [6], modeling ecological phenomena such as snow [7] and earthquakes [3].

There is also extensive work in understanding user sessions, behaviors, and activities on search and e-commerce sites. Such analysis can unlock commercial value through understanding user intentions. Recent work on this includes quantifying the influence of users' domain knowledge on their search behaviors [8], comparing users' search behaviors across different devices [9], understanding user behaviors when their queries return no result [10] and rewriting these queries to increase recall [11]. User's bidding behaviors have also been studied [12]–[15].

We believe that the understanding of behavioral characteristics of both social media and e-commerce domains and their impact on each other would create opportunities of both social and economic value. A step is taken in this direction by analyzing the internal social network of a Chinese e-commerce website, to model how much a buyer will pay for a transaction with a trusted seller [16]. The interplay between Facebook mouth-to-mouth recommendations and a special form of e-commerce, daily deal, have been studied [17]. Though these two papers reveal interesting aspects of the interaction between social network and e-commerce, the correlations between the trends from both domains which reflect the interest and demand of the general public still remain unexplored.

In this paper, we study the interplay of user behavior on a social media site (we choose Twitter) and an e-commerce site (we choose eBay) at a large scale and on a fine grain by analyzing the timestamped Tweets, e-commerce search logs and transaction data. A hypothesized and simplified scenario would be: on Twitter, you see people talking about a victory of your favorite soccer club, Chelsea, after which you go to eBay, search for and eventually purchase a Chelsea T-shirt. We answer the questions such as: What is the correlation between the sudden surge in popularity on social media and e-commerce behaviors? does one platform lead or trail in the burst in activity when such events occur? Our findings suggest that about 5% of general eBay query streams have strong positive correlations with the corresponding Twitter mention streams and for trending queries [18], the percentage goes to around 25%. Queries from certain eBay categories such as 'Sports' have better correlations. We also discover evidences that eBay lags Twitter for correlated pairs of streams. By monitoring the popularities of a list of celebrities, we find that their Twitter popularities correlate with the related eBay search and sales. The correlations and lags can be useful for predictive tasks. For example, by monitoring social media, it can be possible to instantly decide whether a burst is going to lead a sales drive on e-commerce platforms, as a result, relevant sellers and potential buyers will be notified in advance to stimulate the transactions.

In the remainder of this paper, followed by a survey of related work, we describe the methods to quantify the correlation between two streams as well as the lags between the two and apply them to the time series extracted from hundreds of millions of eBay search logs and Tweets. We conduct a case study after which we conclude our work.

*This work was done while this author was an intern at eBay Research Labs.

ASONAM'13, August 25-29, 2013, Niagara, Ontario, CAN

II. Related Work

A lot of recent research has focused on mining the (social) web to create practical, scientific and socio-economic values.

Event detection and trend analysis Traditional news media corpora has been analyzed to explore temporal characteristics. Kotov *et al* [19] extract named entities with correlated temporal bursts from multilingual web news streams. Cook *et al* [20] measure people's periods of fame in a news paper corpus that spans a century. Radinsky *et al* [21] represent word semantics with temporal distributions of related terms in a news archive. There has also been studies on the patterns of the temporal distributions of user-generated web data including search logs [18], [22]–[24], tags [25], [26] and Tweets [27], [28]. As Tweets are short text composed by users expressing their thoughts, opinions and interests, they differ from other user-generated web data or news articles in nature and thus should be handled differently. Much work has been done in event detection on Twitter. Weng *et al* [29], Becker *et al* [30] and Sakaki *et al* [3] examine the text stream in Twitter to detect real-life events. Ritter *et al* [31] not only detect events on Twitter, but also classify them. Marcus *et al* [32] build a system to visualize and summarize events on Twitter. Different from the previous work, our paper is focused on comparing the Twitter stream with a corresponding stream from a global e-commerce platform to study the correlations.

Modeling and prediction Various web corpus have been mined for modeling and prediction of other domains to create socio-economic values. Bollen *et al* [33] estimate public mood on Twitter to predict stock market indices, but does not explore other possibilities where social media and commerce have direct interplays. Jin *et al* [4] estimate the adoption rates of products by monitoring Flickr photos. Due to the photo sharing nature of Flickr and its relatively smaller users basis comparing with Twitter, the estimations are done on coarse temporal grains for only phenomenal products such as iPod. Choi *et al* [34] and Goel *et al* [35] use search engine query logs to predict facts such as sales of video games and motor vehicles, movie box-office revenue and unemployment rate. Though the vast amount of queries reflects the intentions of the general public, they are not doing a side by side and instant comparison of user behaviors on social media where users explicitly express their ideas and that on a global e-commerce platform where large volumes of transactions can be monitored instantly, as the facts such as box-office revenue are usually released with delays. Results from these studies shed light on our attempt to understand a different kind of connection, which is between a popular social media website and a global e-commerce marketplace, at a large scale and on a fine grain.

User behaviors on e-commerce websites User behaviors on e-commerce websites have been analyzed in order to build better services for users and increase the profits for the websites. Lots of the work have been based on understanding the search query logs. Parikh and Sundaresan [18] develop a near real-time burst detection system to suggest trending queries for the buyers and sellers. Singh *et al* [10], [11] study user behaviors in a situation where their initial queries return no matches and build a system to rewrite these queries to increase recall. User behaviors with regard to online auctions have been another theme of research, including cross-bidding [14], last-minute bidding [12], shilling [13] and effect of seller reputation on price [15]. Instead of only focusing on the e-commerce domain itself, we also examine the social media domain to quantify the correlations and explore what drives the sales on e-commerce websites.

Connection between social media and e-commerce The studies at the intersection of social media and e-commerce bring interesting findings. Guo *et al* [16] analyze the internal social network of buyers and sellers on a Chinese e-commerce website, to model how much a buyer will pay for a transaction with a trusted seller. The focus is on the structure of the internal social network and the price of trust, rather than analyzing trends and user behaviors on an open social media platform and a global e-commerce website as we do in this paper. Byers *et al* [17] study the daily deal websites including Groupon and their ties with social media websites such as Facebook and Yelp. On the non-social side of their research, they analyze the incentives for users to purchase and use the parameters of the deal to predict the sales. On the social side, they examine how daily deals affects merchants' reputation on Yelp and they suggest that word-of-mouth recommendations on Facebook benefit daily deal sites. The deals that they study are a special form of e-commerce offering the consumers with localized discounted merchandise in a relatively short period of time while we conduct a study of more general user behaviors in both domains at a large scale and in a long run. Moreover, instead of studying information diffusion in social media, we focus on analyzing the temporal attributes of the information, which reflect users' instant interest and demand.

III. Quantifying Correlations

We start with two datasets - Twitter Tweets and eBay search query logs. The Twitter dataset contains a sample of Tweets, each of which has a short text of the Tweet itself, a user id of the the person composing the Tweet and a timestamp indicating when the Tweet was posted. The eBay dataset contains a sample of eBay search query logs, each of which has a text query, a user id of the person issuing the query, the number of resulting items the user clicks on, the number of bids the resulting items receive from the user, the total number of price-fixed 'Buy It Now' (BIN) item the user purchases and the total prices the user pays on these BIN items.

We focus on keyword phrases appearing in Tweets and eBay queries and for a keyword phrase, we extract two time series of its mentions as a delegate of its popularity over time on Twitter and eBay, respectively. For a keyword phrase such as 'Barack Obama' during a certain period of time, in each time unit, we count the number of unique Twitter users with Tweets containing this keyword phrase as well as the number of unique eBay users ('SEARCH') with queries containing this keyword phrase. For eBay dataset, besides the number of unique users in each time unit, we can also compute the total number of clicks ('VIEW') and bids ('BID') on resulting items, the total number of purchased BIN ('BIN_count') items of the resulting item, the total BIN ('BIN_total') and average BIN prices ('BIN_avg') of these items.

As suggested in [36], we calculate the Pearson Correlation coefficients between the normalized pairs of time series as well as the 2-tailed p-value for testing the null hypothesis that the true correlation coefficient is equal to 0. In the following subsections, we will describe the time series extraction and correlation calculation in detail.

A. Extraction of Time Series

We extract time series which represent the keyword trends from the two datasets.

A simplified scenario of a buyer using eBay is: the user issues a query and clicks on several resulting items, then the user has the option to bid on some of them and she can also choose to purchase some price-fixed items right away using 'Buy It Now' (BIN). For the experiment, we define the set of all user actions $A = \{a_1, a_2, ..., a_q\}$ as a set of tuples of the form $a_i = (u_i, q_i, t_i, v_i, bid_i, bin_i, p_i)$, where u_i is a user, q_i is a query, t_i is a timestamp. v_i is the number of unique items the user clicks on among the resulting items after issuing the query, bid_i is the number of bids the user places on the resulting items, bin_i is the number of BIN items the user purchases among the resulting items and p_i is the total amount of money the user pays for all the BIN items that she purchases.

In order to get the time series of search mentions for a keyword phrase k from time t_s to t_e on eBay, we divide this period of time into m coarse temporal bins. For each bin, we count the number of unique users who issued queries containing the keyword phrase. Let $quant(t_i, t_s, t_e, o)$ be a quantization function that maps a timestamp t_i into one of m temporal bins with the starting time t_s, ending time t_e and temporal bin size o, returning a bin index in the range $[1, m]$. Let $ngram(q_i)$ be a function that replaces the non-numeric and non-alphabetic characters in a text string q_i with spaces and returns a set of all possible $n$grams in the new string. For any keyword k, we then build an m-dimensional vector, counting the number of unique users who issued search queries whose n-grams contain k in each temporal period b from t_s to t_e.

$$U_E(b, k, t_s, t_e) = ||\{u_i|(u_i, q_i, t_i, v_i, bid_i, bin_i, p_i) \in A,\\ k \in ngram(q_i), b = quant(t_i, t_s, t_e, o)\}||,$$

where A includes all a dated from t_s to t_e. Using this approach, we can also extract time series for 'VIEW', 'BID', 'BIN_count', 'BIN_total' and 'BIN_avg' as mentioned above.

Similarly, we extract the time series of mentions for a keyword phrase k on Twitter. We build an m-dimensional vector $U_T(b, k, t_s, t_e)$, counting the number of unique users who posted Tweets of which the n-grams contain k in each temporal period b from t_s to t_e.

For all the vectors extracted, we perform $L1$-norm.

B. Pearson's Correlation Co-efficient and a t-test

We compute the Pearson's correlation co-efficient r between the eBay vector E and the corresponding Twitter vector T for each keyword phrase. It measures the linear dependence between two datasets and gives a co-efficient in $[-1, 1]$ with +1 indicating exact positive linear relationship, -1 indicating exact negative linear relationship and 0 indicating no correlation. It is computed as the covariance of E and T divided by the product of their standard deviations:

$$r = P(E, T) = \frac{cov(E, T)}{\sigma_E \sigma_T}$$

We perform Student's t-test on Pearson's r of the pair of vectors and the null hypothesis is that they are uncorrelated (the actual r is 0). If the underlying variables follow a bivariate normal distribution, the sampling distribution of Pearson's r would have a Student's t-distribution with degrees of freedom $n = 2$. According to [37], this holds approximately for sample sizes that are not small, even when the observed values are not normal. Therefore, even though E and T might not have a bivariate normal distribution, we can still use the two-tailed p-value from the t-test as an approximation of confidence of correlation.

IV. Computing Lag

We quantify the evidence that one stream lags its counterpart in another domain by applying a moving window method. It finds a shift that maximizes the Pearson's r between the pair of streams. Though the evidence from one pair of streams is not strong, when we compute the lags for a massive amount of keyword phrases, there would be enough signals providing the insights. We choose this method instead of the popular Granger causality test [38] as suggested in [33], because fixed temporal bins are not suitable to capture and quantify the instant sudden changes in popularity at a fine grain. For example, with the bin size fixed to one day, it might not be capable of capturing the changes happen in one day, if one stream lags another by minutes or hours; if the bin size is set to a couple of minutes or hours, there might not be enough counts in each bin; the lag might also vary for different pairs of streams and different types of events which might only be captured by windows moving at finer grains.

For each pair of streams, we shift one stream by a period of time Δt to compute the vector and calculate the Pearson correlation between the shifted vector and the other vector which is not shifted. We define the shifted period of time Δt which maximizes the Pearson's r to be the lag. We give the formal definitions here. Given k, b, t_s and t_e, the Pearson correlation co-efficient between eBay vector and Twitter vector is calculated as:

$$f(\Delta t) = \begin{cases} P(U_E(t_s + \Delta t, t_e), U_T(t_s, t_e - \Delta t)), \Delta t >= 0 \\ P(U_E(t_s, t_e - |\Delta t|), U_T(t_s + |\Delta t|, t_e)), \Delta t < 0 \end{cases}$$

where one of the streams is shifted by Δt. k and b are omitted as they are fixed across all the streams in comparison. If $\Delta t >= 0$, the starting point of the eBay stream is shifted to a later time stamp and meanwhile the end point of the Twitter stream is shifted to an earlier time stamp to ensure that the two resulting vectors have the same dimension. If $\Delta t < 0$, the starting point of the Twitter stream is shifted to a later time stamp and the end point of the eBay stream is shifted to an earlier time stamp. The lag l is computed as:

$$l = \underset{\Delta t \in [-\mathcal{T}, \mathcal{T}], \Delta t \in \mathbb{Z}}{\operatorname{argmax}} f(\Delta t),$$

where positive l suggests eBay lags Twitter by l while negative l suggests Twitter lags eBay by $|l|$.

V. Experiments and Results

A. General Correlations

We select the top queries from the long tail distribution of eBay queries as general keyword phrases that we monitor. For the top 150,000 queries on eBay (on July 11, 2012), we monitor the period of time from January 1 2012 to March 31 2012 on eBay and Twitter. In this experiment, the eBay dataset contains a large representative randomized sample of the queries logs and the Twitter dataset includes 1% of the Tweets posted on Twitter. We set thresholds on both daily query counts and daily Twitter mentions to ensure the density of the vectors that we extract. For average eBay daily query counts, we require the numbers to be no less than 20 and for average Twitter mentions, we set the threshold to 5 per day. This gives us a list of 16,099 keyword phrases that can represent general queries on eBay. For the convenience of later references, we name it as GeneralQueryList. For each phrase in this list, we get two time series of daily mentions from both eBay and Twitter. For each pair of time series, we compute the Pearson correlation coefficient r and the p-value of the t-test with the null hypothesis that the two time series are generated from two uncorrelated systems.

The statistics in Table I suggest that around 5.25% of these queries can be considered as having significant positive correlations with their counterparts on Twitter while only around 0.65% show significant negative correlations at $p = 0.0005$. Here we consider several confidence levels from 0.0001 to 0.01 and calculate the fractions of positively and negatively correlated keyword phrase pairs. Notice that when we increase the confidence level, the ratio between positively correlated pairs and negatively correlated pairs increases, suggesting the possible noises are removed.

In order to examine the correlations in behaviors across the two domains for different classes of queries, we break these queries into the 36 eBay meta categories[1] and some categories have higher portions of strongly correlated queries at the confidence level of 0.0005. As we may expect, the categories such as ‘Video Game’ and ‘Sports Mem, Cards & Fan Shop’ which are more likely to have sales driving news events such as a championship match and releasing of a video game appear in the top 5 categories are shown in Table II, as apposed to categories like ‘Home & Garden’ and ‘Pottery & Glass’. Among all the general queries, ‘Sports Mem, Cards & Fan Shop’ takes up 7.34% which is a relatively large portion.

We then select the trending queries detected by the system described in [18] from the GeneralQueryList mentioned above, resulting in 730 queries which we name as TrendQueryList. As shown Table III, 24.93% of the trending keyword phrase pairs have very strong correlations, compared to 5.25% for the general pairs. At 0.01, a relaxed level of confidence, almost 30% of them are correlated. At the same level, the ratio between positively correlated pairs and negatively correlated pairs is much higher than that in Table I. When we break these trending queries into the eBay meta categories, we find that some categories demonstrate much more correlations. 69.23% of the queries in ‘Sports Mem, Cards & Fan Shop’ are strongly correlated at the confidence level of 0.0005 as shown in Table IV. Even for the ‘Music’ category which ranks at 5th, its percentage is 15% higher than that of the top category in general queries.

TABLE I. Fractions of correlated pairs of keyword phrases at different confidence levels. The ratio between positively correlated pairs and negatively correlated pairs increases as the confidence level increases, suggesting the possible noises are removed.

p-value	pos_corr	neg_corr	pos/neg
0.01	8.52%	2.49%	3.42
0.005	7.34%	1.77%	4.14
0.001	5.75%	0.80%	7.18
0.0005	5.25%	0.65%	8.07
0.0001	4.35%	0.32%	13.59

TABLE II. Top 5 eBay meta categories ranked by portions of queries with strong correlations at the confidence level of 0.0005.

Category	pos_corr
Video Games	21.28%
DVDs & Movies	14.20%
Entertainment Memorabilia	13.47%
Sports Mem, Cards & Fan Shop	13.45%
Tickets	13.04%

B. Lag between Two Streams

From the GeneralQueryList, we get a list of pairs which are strongly correlated using a threshold of Pearson’s r at 0.4 and that results in 690 queries which we name as CorrelatedQueryList. For these queries, we calculate the lags in the shift range of $[-5000, 5000]$ minutes with positive values indicating eBay lags Twitter and vice versa.

In order to reduce noise, we require the increase in Pearson’s r to be at least 5% to be considered that there is a lag. We then plot the histogram of the lags as shown in Figure 1. The average lag value is 290 minutes (4.83 hours) and 61.30% of the 690 queries have positive lag values as opposed to 50% if the lags are normally distributed around a mean value of 0, suggesting that for keyword pairs correlated on a daily basis, eBay is lagging Twitter. For the pairs where eBay lags Twitter,

TABLE III. For trending keyword pairs, fractions of correlated pairs of keywords at different confidence levels. At the same level, the ratio between positively correlated pairs and negtively correlated pairs is much higher than that in Table I.

p-value	pos_corr	neg_corr	pos/neg
0.01	29.58%	1.64%	18.03
0.005	28.21%	1.09%	25.88
0.001	25.61%	0.13%	197
0.0005	24.93%	0.13%	191.76
0.0001	22.46%	0%	INF

TABLE IV. For trending keyword pairs, top 5 eBay meta categories ranked by portions of queries with strong correlations at the confidence level of 0.0005.

Category	pos_corr
Sports Mem, Cards & Fan Shop	69.23%
Video Games & Movies	64.28%
Cell Phones & PDAs	52.94%
Entertainment Memorabilia	37.50%
Music	36.66%

[1]There is 36 meta categories at the top level of the eBay merchandise ontology composed by domain experts. The list is available on www.ebay.com. A query is assigned to a category according to what the majority of users issuing the same query clicked on and purchased historically.

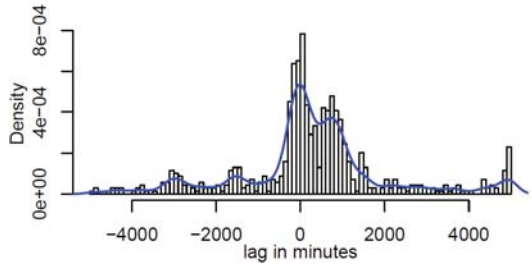

Fig. 1. Histogram of the lags with the density curve for CorrelatedQueryList. The average lag time of eBay from Twitter is 4.83 hours and 61.30% of the 690 queries have positive lag values. Most of the lags are distributed on the positive half of the x-axis.

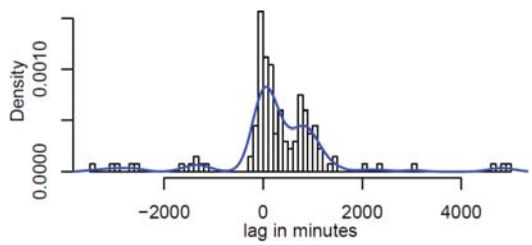

Fig. 2. Histogram of the lags in Sports from CorrelatedQueryList with the density curve. 74% have positive lag values and as a result, the majority of lags is distributed on the positive side of the x-axis.

the average lag is 1214 minutes (20.2 hours) and the third quantile is 1441 minutes (24.0 hours).

Among the 36 eBay meta categories, 'Clothing, Shoes & Accessories' ('Clothing') and 'Sports Mem, Cards & Fan Shop' ('Sports') have more than 100 keywords out of the 690 keywords and we choose to observe their lag patterns as shown in Figure 2 and Figure 3. For 'Clothing', only 45.6% have positive lag values, while for 'Sports' the percentage is 70.14%. This suggests that for certain categories, the signal is stronger that Twitter is leading eBay.

Again we pick the trending queries from the Correlated-QueryList and it gives us 164 keyword phrases. The average lag value is 662 minutes (11.03 hours) and 76.82% of them have positive lag values compared to 290 minutes and 61.30% for general queries. The histogram is shown in Figure 4. Noticing that 10 keyword phrases have lags of over 4000 minutes, we plot two of them here with shifted eBay curves in red marked with asterisks: one is a general keyword phrase 'air conditioner' with a lag of 4950 mins (3.43 days) and the other is a more specific product keyword phrase 'droid 4' with a lag of 4981 mins (3.45 days) as shown in Figure 5 and Figure 6. After shifting, the Pearson's r between Twitter and eBay for 'air conditioner' goes from 0.42 to 0.48 and for 'droid 4', the value goes from 0.45 to 0.57 which suggests a margin of over three days for eBay to respond to Twitter for these two keyword phrases. 'droid 4' is a Motorola cellphone released on Feb 10, 2012 but 3 days before that, there were Tweets and news articles about Verizon confirming the releasing date which contributed to the Twitter burst on Feb 7.

In histograms in Figure 1, 2 and 4, the gaps between neighboring local maximas are usually around a day (1440 minutes), which might be related with the daily repetitions of users' activity pattern.

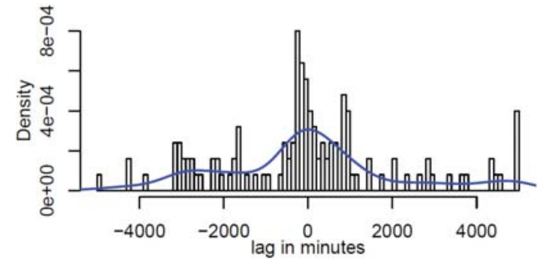

Fig. 3. Histogram of the lags in Clothing from CorrelatedQueryList with the density curve. 45.6% have positive lag values and the lags are distributed more evenly on both sides of 0.

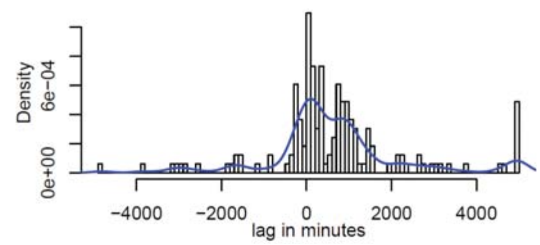

Fig. 4. Histogram of the lags for trending keywords with the density curve. 76.82% of them have positive lag values, the majority of lags is also distributed on the positive side of the x-axis.

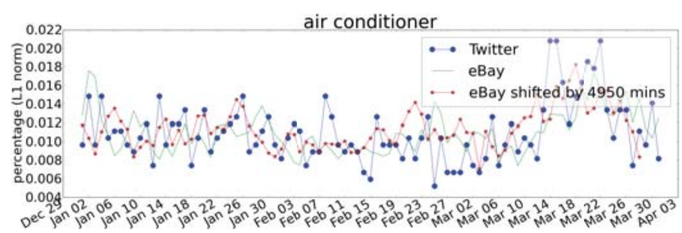

Fig. 5. Twitter trend, eBay trend and shifted eBay trend for 'air conditioner'. The lag (shift) is 3.43 days.

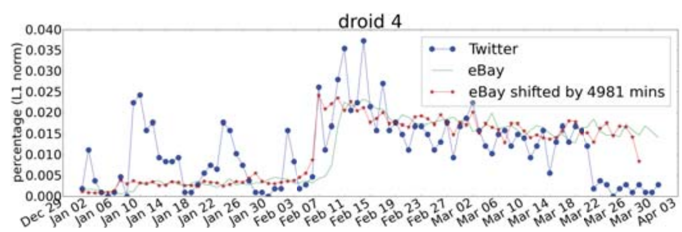

Fig. 6. Twitter trend, eBay trend and shifted eBay trend for 'droid 4'. The lag (shift) is 3.45 days.

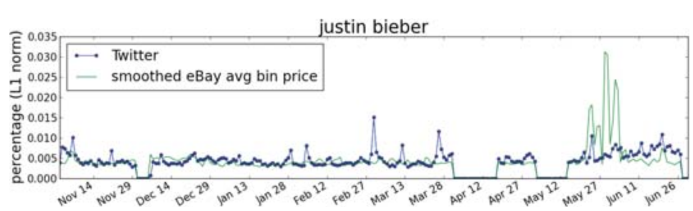

Fig. 7. Twitter trend and smoothed eBay average BIN price for 'justin bieber', with a smoothing window size of 2. Comparing with no smoothing, the Pearson's r goes from 0.183 to 0.233 with a p-value of 0.0007. (Zero values are due to missing data.)

TABLE V. For 100 celebrity keywords, fractions of correlated pairs of keywords at different confidence levels. For SEARCH and VIEW, the pairs are best correlated among the 6 statistics. BID, BIN_count and BIN_total demonstrate moderate correlations.

	SEARCH		VIEW		BID		BIN_count		BIN_total		BIN_avg	
p-value	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr
0.01	46%	0%	44%	0%	14%	0%	16%	2%	12%	0%	5%	0%
0.005	45%	0%	43%	0%	11%	0%	14%	1%	11%	0%	3%	0%
0.001	39%	0%	38%	0%	10%	0%	11%	1%	9%	0%	1%	0%
0.0005	37%	0%	30%	0%	8%	0%	7%	1%	7%	0%	0%	0%
0.0001	37%	0%	30%	0%	8%	0%	7%	0%	7%	0%	0%	0%

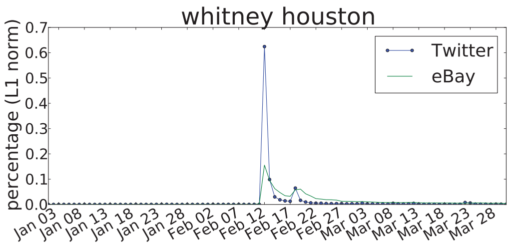

Fig. 8. Twitter trend and eBay trend for ‘whitney houston’. The bursts on both streams correlate well on the day she passed away and the day for her funeral. Users’ interest drops slower on eBay which suggests a different pattern of attention from Twitter.

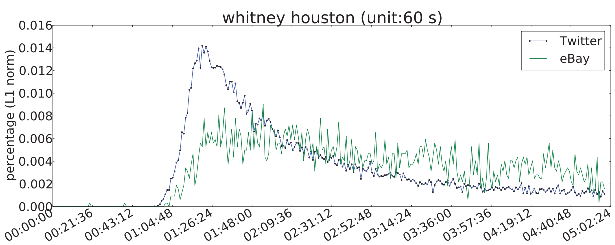

Fig. 9. Twitter trend and eBay trend for ‘whitney houston’ with unit set to 60 seconds. Twitter burst rises faster than eBay and peaks earlier as well. Users’ interest still drops quicker on Twitter after the peak.

C. Celebrity Watching

Celebrity news often trends in Twitter [1] and we want to know when celebrity news - deaths, movies, arrests, games, etc. trends in one network what influence it has on the other network. Some possible scenarios include: the news of a celebrity death triggers a book release, an Oscar nomination is followed by sudden increases in sales of movies in which the celebrity starred, and a celebrity retirement stimulates fans’ passion for memorabilia. Here we explore the relationships between popularity of celebrities on Twitter and their related searches and sales on eBay. We make use of Forbes ‘The Celebrity 100’ list [2] which includes celebrities based on media visibility and entertainment-related earnings.

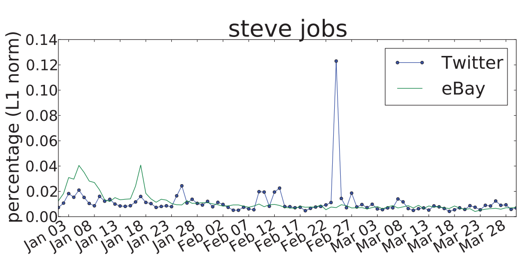

Fig. 10. Twitter trend and eBay trend for ‘steve jobs’. The first two big bursts on eBay are related with the production of Steve Jobs action figures while the huge burst on Twitter is related with his birthday. The relatively weak correlation suggests the interplay between two streams is sometimes profound.

From Nov 1, 2011 to Jun 27, 2012 (excluding 34 missing days), for each celebrity we compute the daily statistics including number of searches (SEARCH), number of views (VIEW), number of bids (BID), number of BINs (BIN_count), total BIN price (BIN_total) and average BIN price (BIN_avg) such that for each celebrity on the list, there will be 6 time series from eBay. For each of the eBay time series, we compute the Pearson’s r between itself and its corresponding Twitter mention time series as well as the p-value of the t-test. We threshold on different confidence levels to examine the percentage of positively correlated pairs and that of negatively correlated pairs and the results are shown in Table V.

It shows that only for BIN_count at lower confidence levels, there are negatively correlated pairs. It also suggests that for SEARCH and VIEW, the pairs are best correlated among the 6 statistics. BID, BIN_count and BIN_total demonstrate moderate correlations and a possible explanation for this is that sometimes users just want to check what is on eBay but will not necessarily make up their minds purchasing.

The average BIN price (BIN_avg) is relatively less correlated only 5% are correlated at 0.01 confidence level. A possible reason for this is that it takes time for the price to reflect the changes in popularity (the prices for BIN items are set when the sellers list the items) and the prices might be insensitive in nature. In order to further understand this, we calculate the BIN_avg on a certain day as the the average BIN price of n days in the future (BIN_total/BIN_count), together with the current day. We show the average Pearson’s r in Table VI where n ranges from 0 to 14 and the portions of correlated pairs at different confidence levels for $n \in [0, 4]$ in Table VII. From these two tables, the average Pearson’s r peaks when $n = 2$ and at the same time, greater portion (9%) of pairs are correlated at the confidence level of 0.01 which suggests the impact of celebrities’ popularity on the average prices of their related items in a 3-day window. Here we show one example for Justin Bieber, with the window size 2 in Figure 7. Comparing with no smoothing, the Pearson’s r goes from 0.183 to 0.233 with a p-value of 0.0007.

D. Peakiness of Two Streams

The peakiness of a usage stream from a website reflects its users’ attention as well as the nature of the platform itself. We measure the general peakiness of the eBay stream and the Twitter stream by computing their average second moment as suggested in [26]. For a vector v that represents a time series, its second moment is computed as:

[2]www.forbes.com/celebrities/list/, May 2012

TABLE VI. AVERAGE PEARSON'S r BETWEEN TWITTER MENTION TIME SERIES AND EBAY AVERAGED BIN PRICE TIME SERIES IN A $(n+1)$-DAY WINDOW. THE AVERAGE PEARSON'S r PEAKS WHEN $n = 2$.

n	0	1	**2**	3	4	5	6	7	8	9	10	11	12	13	14
avg_r	0.018	0.015	**0.020**	0.018	0.012	0.009	0.010	0.013	0.014	0.012	0.012	0.013	0.011	0.007	0.005

TABLE VII. FOR $n \in [0, 4]$, PORTIONS OF CORRELATED PAIRS OF KEYWORDS AT DIFFERENT CONFIDENCE LEVELS. WHEN $n = 2$, A GREATER PORTION (9%) OF PAIRS ARE CORRELATED AT THE CONFIDENCE LEVEL OF 0.01.

	n=0		n=1		**n=2**		n=3		n=4	
p-value	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr	p_corr	n_corr
0.01	5%	0	6%	0	**9%**	0	8%	1%	7%	0
0.005	3%	0	4%	0	8%	0	8%	0	7%	0
0.001	1%	0	4%	0	4%	0	4%	0	4%	0
0.0005	0	0	1%	0	3%	0	4%	0	2%	0
0.0001	0	0	1%	0	3%	0	4%	0	2%	0

$$second\_moment(v) = v \cdot v = \sum_{i=1}^{n} v_i^2.$$

For the GeneralQueryList mentioned in Section V-A, we compute the average second moment of their time series of mentions on eBay and Twitter. The value for eBay is 0.011 while the value for Twitter is 0.016, suggesting that Twitter streams are more peaky than eBay streams. The reason might be that as a form of news media, Twitter's user attention spikes and drops more quickly.

VI. CASE STUDIES

We select some cases to demonstrate the characteristics of the two trends that we observe. Two keyword phrases, 'Whitney Houston' and 'Steve Jobs', are selected as they are both popular and easy to be mapped to real world news events.

The 'Whitney Houston' streams shown in Figure 8 appear to have a very strong correlation (Pearson's $r = 0.794$, p-value = 5.44e-21). The bursts on both streams correlate well on the day Whitney Houston passed away and the day for her funeral. But there are some differences in users' pattern of attentions. Twitter is more bursty as a news media but after the burst, users have the information and stop talking about it, while on eBay, users still remain interested in purchasing.

We then examine the period from 00:00 to 05:00, Feb 12, 2012 during which public got to know Whitney Houston passed away. As shown in Figure 9 where the unit is set to 60 seconds, we can see how queries and Tweets arrived at a finer grain and it suggests that Twitter burst rises faster than eBay and peaks earlier as well. At this scale, we are still able to observe that users' interest drops quicker on Twitter after the peak suggesting that users remained interested in purchasing on eBay after they knew the news.

For the 'Steve Jobs' streams shown in Figure 10, the correlation is not strong (Pearson's $r = 0.118$, p-value = 0.264). We observe that there are two big peaks in January for eBay while the corresponding peaks on Twitter are relatively weaker which is not so usual as Twitter is often more bursty than eBay. We manually look at the 240 tweets containing 'Steve Jobs' on January 3 corresponding to the first big burst on eBay. Over a third of them are about a Chinese factory planning to make unauthorized lifelike replica of Steve Jobs[3] and some of the tweets express the willingness to purchase. Meanwhile, we check the eBay queries containing 'Steve Jobs' on the same day, about 33% of them were related to action figures or bobble heads while the ratio for that on Jan 1 is only 5%. For the second burst on eBay, we check the related Tweets and eBay queries on Jan 17. A third of the Tweets were about the Chinese factory canceling the production of the action figures and 50% of the eBay queries were related to the action figures. A news article[4] pointed out that people were still selling and buying the action figures on eBay and the highest price reached \$2500.

For the huge burst on Twitter on February 24, there is no obvious corresponding burst on eBay. But when we check the Tweets on that day, it turns out that it was Steve Jobs birthday and a lot of people were memorializing him by Tweeting. All these indicate that the interplay between user behavior on Twitter and that on eBay sometimes can be profound and this kind of correlations/non-correlations could offer important insight on discovering the factors that drive sales on e-commerce platforms.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed techniques to quantify the correlations and lags between the trend on social media and that on e-commerce. We also examined the individual characteristics of the two streams through a case study and by measuring their peakiness. We discover evidences that:

- About 5% of the eBay query streams have strong positive correlations with their corresponding Twitter streams. For trending queries, the percentage jumps to around 25%.
- Some categories of general queries are more likely to have such correlations. For example, for the 'Video Games' category 21.28% of the query trends are strongly correlated and the percentage is 14.20% for 'DVDs & Movies'. It is also more significant for trending queries. For example, the percentage is about 70% for queries from 'Sports Mem, Cards & Fan Shop'.
- For correlated pairs of streams, eBay stream lags Twitter stream and for trending queries and queries in categories such as 'Sports', the lag is more obvious and potentially useful for predictive tasks.
- Celebrities' popularities on Twitter correlate their search and sale trends on eBay. There is also signals

[3] www.foxnews.com/tech/2012/01/03/steve-jobs-action-figure-planned-for-february, Fox News

[4] www.pcworld.com/article/248238/maker_of_steve_jobs_action_figure_kills_project.html, PCWorld

that they have an impact on the prices of the related merchandise.

- Twitter trend is more peaky than eBay trend.

To summarize, we observe that e-commerce activity correlates well with social media yet lags it especially in certain domains like 'Sports' and 'DVDs & Movies'. This is more prominent when users react to events and happenings. A possible reason which explains the correlation is that the domain of sports and entertainment is eventful and attracts eyeball, making it more likely to generate news of direct commercial values. A reason for the lags might lie in the nature of the two platforms, one is a social media platform which generates and delivers latest news fast while the other one serves the need for purchasing in response to the latest news. We believe that access to online social streams can enable near real-time merchandising of relevant products.

In future work, we plan to predict the sales on e-commerce platforms, according to signals in social media. Ultimately, the system will be able to detect events in social media, classify them into sales-driving and non-sales-driving events and estimate the sales of corresponding products on e-commerce platforms. With such a system, we can recommend relevant items for sellers and buyers well in advance to increase the transactions on e-commerce platforms. This will involve techniques such as event detection, text mining and machine learning. A first step in this direction would be focusing on the sub-domains (e.g. 'Sports' and 'DVDs & Movies') of queries which demonstrate strong correlations and obvious lags, in order to build up a predictive model that can be generalized.

VIII. Acknowledgments

We thank Prof. David Crandall for helpful discussions and advice, and Qieyun Dai for proofreading this paper.

References

[1] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in *WWW*, 2010.

[2] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *LREC*, 2010.

[3] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *WWW*, 2010.

[4] X. Jin, A. Gallagher, L. Cao, J. Luo, and J. Han, "The wisdom of social multimedia: Using Flickr for prediction and forecast," in *ACM MM*, 2010.

[5] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welpe, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *ICWSM*, 2010.

[6] J. Ginsberg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," *Nature*, vol. 457, 2009.

[7] H. Zhang, M. Korayem, D. Crandall, and G. LeBuhn, "Mining photo-sharing websites to study ecological phenomena," in *WWW*, 2012.

[8] R. White, S. Dumais, and J. Teevan, "Characterizing the influence of domain expertise on web search behavior," in *WSDM*, 2009.

[9] M. Kamvar, M. Kellar, R. Patel, and Y. Xu, "Computers and iphones and mobile phones, oh my!: a logs-based comparison of search users on different devices," in *WWW*, 2009.

[10] G. Singh, N. Parikh, and N. Sundaresan, "User behavior in zero-recall ecommerce queries," in *SIGIR*, 2011.

[11] ——, "Rewriting null e-commerce queries to recommend products," in *WWW*, 2012.

[12] A. Roth and A. Ockenfels, "Last-minute bidding and the rules for ending second-price auctions: Evidence from ebay and amazon auctions on the internet," *The American Economic Review*, 2002.

[13] H. Shah, N. Joshi, A. Sureka, and P. Wurman, "Mining ebay: Bidding strategies and shill detection," *WEBKDD 2002-Mining Web Data for Discovering Usage Patterns and Profiles*, pp. 17–34, 2003.

[14] S. Anwar, R. McMillan, and M. Zheng, "Bidding behavior in competing auctions: Evidence from ebay," *European Economic Review*, 2006.

[15] D. Houser and J. Wooders, "Reputation in auctions: Theory, and evidence from ebay," *Journal of Economics & Management Strategy*, 2006.

[16] S. Guo, M. Wang, and J. Leskovec, "The role of social networks in online shopping: information passing, price of trust, and consumer choice," in *EC*, 2011.

[17] J. Byers, M. Mitzenmacher, and G. Zervas, "Daily deals: Prediction, social diffusion, and reputational ramifications," in *WSDM*, 2012.

[18] N. Parikh and N. Sundaresan, "Scalable and near real-time burst detection from ecommerce queries," in *KDD*, 2008.

[19] A. Kotov, C. Zhai, and R. Sproat, "Mining named entities with temporally correlated bursts from multilingual web news streams," in *WSDM*, 2011.

[20] J. Cook, A. Das Sarma, A. Fabrikant, and A. Tomkins, "Your two weeks of fame and your grandmother's," in *WWW*, 2012.

[21] K. Radinsky, E. Agichtein, E. Gabrilovich, and S. Markovitch, "A word at a time: computing word relatedness using temporal semantic analysis," in *WWW*, 2011.

[22] M. Vlachos, C. Meek, Z. Vagena, and D. Gunopulos, "Identifying similarities, periodicities and bursts for online search queries," in *SIGMOD*, 2004.

[23] A. Kulkarni, J. Teevan, K. Svore, and S. Dumais, "Understanding temporal query dynamics," in *WSDM*, 2011.

[24] M. Mohebbi, D. Vanderkam, J. Kodysh, R. Schonberger, H. Choi, and S. Kumar, "Google correlate whitepaper," http://www.google.com/trends/correlate/whitepaper.pdf, 2011.

[25] L. Kennedy, M. Naaman, S. Ahern, R. Nair, and T. Rattenbury, "How flickr helps us make sense of the world: context and content in community-contributed media collections," in *ACM MM*, 2007.

[26] H. Zhang, M. Korayem, E. You, and D. Crandall, "Beyond co-occurrence: discovering and visualizing tag relationships from geo-spatial and temporal similarities," in *WSDM*, 2012.

[27] J. Yang and J. Leskovec, "Patterns of temporal variation in online media," in *WSDM*, 2011.

[28] J. Lin and G. Mishne, "A study of "churn" in tweets and real-time search queries," in *ICWSM*, 2012.

[29] J. Weng and B. Lee, "Event detection in twitter," *ICWSM*, 2011.

[30] H. Becker, M. Naaman, and L. Gravano, "Beyond trending topics: Real-world event identification on twitter," 2011.

[31] A. Ritter, Mausam, O. Etzioni, and S. Clark, "Open domain event extraction from twitter," in *KDD*, 2012.

[32] A. Marcus, M. Bernstein, O. Badar, D. Karger, S. Madden, and R. Miller, "Twitinfo: aggregating and visualizing microblogs for event exploration," in *CHI*, 2011.

[33] J. Bollen, H. Mao, and X.-J. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, 2011.

[34] H. Choi and H. Varian, "Predicting the present with google trends," *Economic Record*, 2012.

[35] S. Goel, J. Hofman, S. Lahaie, D. Pennock, and D. Watts, "Predicting consumer behavior with web search," *PNAS*, 2010.

[36] M. Sayal, "Detecting time correlations in time-series data streams," *Hewlett-Packard Company*, 2004.

[37] M. Kendall and A. Stuart, *The Advanced Theory of Statistics, Volume 2: Inference and Relationship*. Hafner Publishing Company, 1961.

[38] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, 1969.