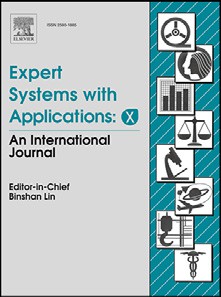
[Expert Systems with Applications: X 3 (2019) 100007](https://doi.org/10.1016/j.eswax.2019.100007)

Contents lists available at [ScienceDirect](http://www.ScienceDirect.com/)

Expert Systems with Applications: X

journal homepage: [www.elsevier.com/locate/eswax](http://www.elsevier.com/locate/eswax)

Mining Twitter data for causal links between tweets and real-world outcomes

Sunghoon Lim[a,](#_bookmark0) Conrad S. Tucker [b,](#_bookmark1)[c,](#_bookmark2)[∗](#_bookmark3)

a *School of Management Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea*

b *Department of Mechanical Engineering, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh, Pennsylvania 15213-3890*

c *Department of Machine Learning, Carnegie Mellon University 5000, Forbes Avenue Pittsburgh, Pennsylvania 15213-3890*

a r t i c l e i n f o a b s t r a c t

*Article history:*

Received 22 January 2019

Revised 11 June 2019

Accepted 2 July 2019

Available online 5 July 2019

*Keywords:*

Expert and intelligent system Social media

Enterprise outcome

Co-occurrence network Time series analysis Granger causality analysis

The authors present an expert and intelligent system that (1) identifies influential term groups having causal relationships with real-world enterprise outcomes from Twitter data and (2) quantifies the appro- priate time lags between identified influential term groups and enterprise outcomes. Existing expert and intelligent systems, which are defined as computer systems that imitate the ability of human decision making, could enable computers to identify the spread of Twitter users’ enterprise-related feedback au- tomatically. However, existing expert and intelligent systems have limitations on automatically identifying the causal effects on enterprise outcomes. Identifying the causal effects on enterprise outcomes is impor- tant, because Twitter users’ feedback toward enterprise decisions may have real-world implications. The proposed expert and intelligent system can support decision makers’ decisions considering the real-world effects of identified Twitter users’ feedback on enterprise outcomes. In particular, (1) a co-occurrence net- work analysis model is exploited to discover term candidates for generating influential term groups that are combinations of enterprise-related terms, which potentially influence enterprise outcomes. (2) Time series models and (3) a Granger causality analysis model are then employed to identify influential term groups having causal relationships with enterprise outcomes with the appropriate time lags. Case studies involving a real-world internet video streaming and disc rental provider as well as an airline company are used to test the validity of the proposed expert and intelligent system for both predicting enterprise outcomes in a long period and predicting the effects of specific events on enterprise outcomes in a short period.

© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## Introduction

Recently, enterprises have successfully used expert and intelli- gent systems (i.e., computer systems that imitate the ability of hu- man decision making [Jackson, 1998](#_bookmark64)) in order to enable computers to extract social media users’ feedback from large-scale and pub- licly available social media data (e.g., Twitter, Facebook, Instagram) [automatically (Holzinger, Krüpl, & Herzog, 2006; Meire, Ballings, & Van den Poel, 2017; Mostafa, 2013). Social media are considered](#_bookmark58) useful not only for enterprises, due to the ease of acquiring users’

∗ Corresponding author at: Industrial Engineering, Engineering Design, Computer Science and Engineering, 213 Hammond Building, State College, PA 16802, USA.

*E-mail addresses:* [sunghoonlim@unist.ac.kr](mailto:sunghoonlim@unist.ac.kr) (S. Lim), [ctucker4@psu.edu](mailto:ctucker4@psu.edu), [conradt@andrew.cmu.edu](mailto:conradt@andrew.cmu.edu) (C.S. Tucker).

feedback, but also for social media users, who can easily post opin- ions related to a wide range of topics ([Tuarob & Tucker, 2015](#_bookmark103)). In particular, due to its popularity and scalability, Twitter has been widely used as a suitable social media platform for expert and in- telligent systems to discover users’ feedback regarding enterprise [decisions (Culnan, McHugh, & Zubillaga, 2010; Daniel, Neves, & Horta, 2017; Ghiassi, Skinner, & Zimbra, 2013; Greer & Ferguson, 2011; Ikeda, Hattori, Ono, Asoh, & Higashino, 2013; Oliveira, Cortez, & Areal, 2017).](#_bookmark78)

While many existing expert and intelligent systems for Twit- ter user feedback discovery could enable computers to analyze the correlation between Twitter users’ feedback (e.g., Twitter users’ [sentiment) and real-world enterprise outcomes (Asur & Huber- man, 2010; Mao, Wei, Wang, & Liu, 2012) automatically, limited](#_bookmark53) contributions have been made to analyze causal effects of Twit-

<https://doi.org/10.1016/j.eswax.2019.100007>

2590-1885/© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

ter users’ feedback on real-world enterprise outcomes. The term “*Twitter users’ feedback*” means information coming directly from Twitter users about what they feel with a company, a product, or a service. The term “*enterprise outcome*” is defined as the metric that is typically used to evaluate the success or failure of an enterprise. These include an enterprise’s stock price, product sales, number of customers, market valuation, etc. The term “*influential term group*” is defined as a combination of enterprise-related terms on Twitter (e.g., “passenger”, “ceo”) that potentially influence enterprise out- comes. Should enterprise decision makers (e.g., CEOs) take Twitter users’ feedback seriously when searching for emerging topics that the future market desires? Have enterprises reversed decisions as a result of Twitter users’ feedback? Causality is differentiated from mere correlation, since a correlation between two events does not imply that one event (e.g., Twitter user feedback) causes the other (e.g., an enterprise outcome) ([Aldrich, 1995](#_bookmark51)).

The main contributions of this research are to propose an ex- pert and intelligent system that enables computers to analyze causal relationships between natural language data (i.e., Twitter data) and real-world events (i.e., real-world enterprise outcomes) automatically. In particular, the proposed expert and intelligent system (1) identifies influential term groups having causal effects on enterprise outcomes from Twitter data and (2) discovers the ap- propriate time lags between identified influential term groups and enterprise outcomes. While existing expert and intelligent systems could enable computers to identify the spread of Twitter users’ enterprise-related feedback automatically, identifying the causal ef- fects on real-world enterprise outcomes is limited. The proposed expert and intelligent system discovers term candidates for influ- ential term groups through (1) a co-occurrence network analysis model. (2) Time series models and (3) a Granger causality anal- ysis model then identify influential term groups that have causal effects on real-world enterprise outcomes. The appropriate time lags between identified influential term groups and enterprise out- comes that have casual relationships are also discovered in order to measure how soon identified influential term groups cause en- terprise outcomes. Solving these problems is challenging due to several reasons:

* There are over 500 million tweets (i.e., Twitter messages) gen- erated each day ([Bodnar, Dering, Tucker, & Hopkinson, 2016](#_bookmark62)), many of which do not have causal relationships with real- world outcomes. As a result, an expert and intelligent system is needed to determine which Twitter users’ feedback causes real-world enterprise outcomes.
* There is no fixed relationship between positive/negative Twit- ter messages and real-world outcomes. I.e., it is possible for negative Twitter users’ sentiment (e.g., the announcement by iHOP that it was changing its name to iHOB) to have a posi- tive impact on real-world outcomes (i.e., the name change in- spired 36 billion social media (e.g., Twitter) users’ impressions and resulted in a 0.7% increase in sales ([Taylor, 2018](#_bookmark98))). Con- versely, negative Twitter users’ sentiment (e.g., backlash over United Airlines’ passenger ejection decision) can result in neg- ative real-world outcomes (i.e., United Airlines’ decreased stock price) ([Ohlheiser, 2017](#_bookmark99)). The challenge is determining the im- pact that Twitter users’ feedback will have on real-world out- comes.
* Given the discovery of Twitter users’ feedback (i.e., influential

term groups) that has been determined to have a causal rela- tionship with real-world outcomes (e.g., stock price), the chal- lenge is determining how long it will take for the discovered Twitter users’ feedback to cause the real-world outcome. An ex- pert and intelligent system is needed that provides enterprise decision makers with a timeline as to how long they have af- ter Twitter users’ feedback before the real-world impacts of this

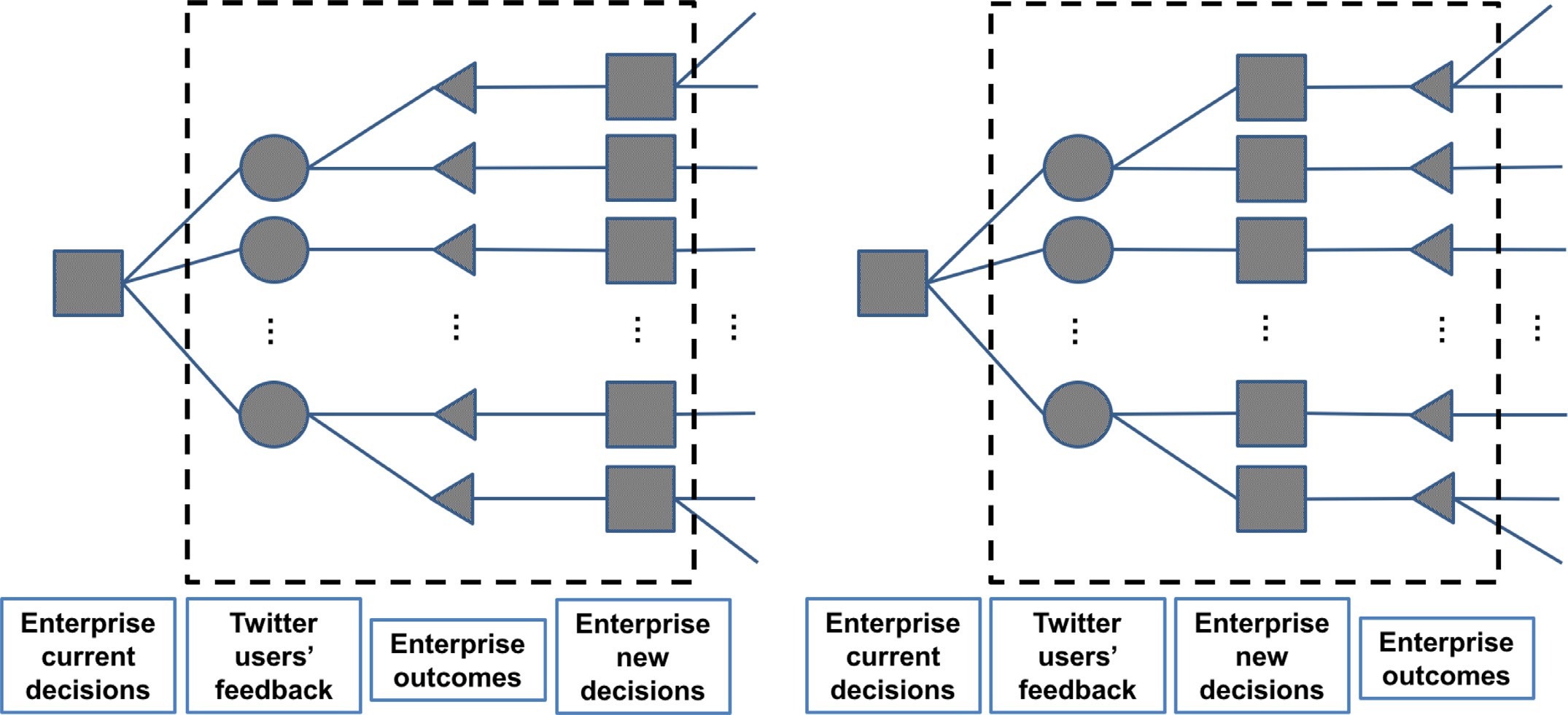
feedback occur. This would enable enterprise decision makers to intervene and explore solutions that may mitigate or reverse the potential negative real-world outcomes (e.g., via an apol- ogy) before they occur.

On the left-hand side of [Fig. 1](#_bookmark4), a current conceptual decision tree describes how Twitter users’ feedback on enterprise decisions could potentially affect (1) enterprise outcomes, such as market sales or stock prices, and (2) enterprise decision makers’ future decisions. Enterprise decisions have causal effects on customer re- actions displayed on social media networks, including Twitter. For instance, *Microsoft*’s digital rights management (DRM) policy an- nouncement – that the *Xbox One* console requires an online check- in at least every 24 hours to validate the ownership of game soft- ware – caused strong negative reactions on Twitter ([Strickl, 2013](#_bookmark114)). Previous research has explored how enterprise decisions cause cer- [tain social media feedback, like Twitter users’ feedback (Bruhn, Schoenmueller, & Schäfer, 2012; Kaplan & Haenlein, 2010). How-](#_bookmark72) ever, little work has been done on exploring which Twitter user feedback (i.e., influential term group in this work) causes certain enterprise outcomes. Considering causal effects of Twitter users’ feedback is important, because negative Twitter users’ feedback can negatively affect other users on Twitter and enterprises’ fu- ture revenue. For instance, *United Airlines’* stock fell 4% and wiped out $800 million in just three days after Twitter raged at *United Airlines* due to the CEO’s email to his employees that defended those who forcibly dragged a passenger off an overbooked flight in 2017 ([Imbert & Thomas, 2017](#_bookmark63)). The proposed expert and intelli- gent system focuses on identifying overall Twitter users’ feedback, instead of only monitoring a few influential users’ Twitter mes- sages that potentially cause enterprise outcomes (e.g., Kylie Jen- ner’s one tweet that wiped out $1.3 billion of Snapchat’s market value [Vasquez, 2018](#_bookmark106)), because only monitoring a few influential users’ tweets is relatively easier than identifying feedback from a large number of tweets written by overall Twitter users (i.e., both influential and ordinary Twitter users) ([Bosch et al., 2013](#_bookmark68)).

This research discovers Twitter users’ feedback (i.e., influential

term groups) that has causal effects on real-world enterprise out- comes. The proposed expert and intelligent system enables en- terprise decision makers to change their current decisions before negative Twitter users’ feedback about those current decisions de- creases future enterprise outcomes. If positive Twitter users’ feed- back about the current decisions is discovered, enterprise deci- sion makers can make new decisions that intensify positive factors identified from Twitter users’ feedback in order to increase future enterprise outcomes. The proposed expert and intelligent system can therefore help enterprise decision makers identify significant Twitter users’ feedback in real-time and at a low-cost. In addition, the proposed expert and intelligent system can be used to assist decision makers’ decisions that will improve future enterprise out- comes, such as market sales or stock prices. The right-hand side of [Fig. 1](#_bookmark4) illustrates the enterprise decision tree based on the proposed expert and intelligent system.

The remainder of the paper is organized as follows. This section provides an introduction and motivation for this work. [Section 2](#_bookmark5) provides the background of related works. [Section 3](#_bookmark9) pro- poses the expert and intelligent system to identify Twitter users’ feedback (i.e., influential term groups) that causes enterprise out- comes. [Section 4](#_bookmark27) introduces case studies involving a real-world in- ternet video streaming and disc rental provider (i.e., *Netflix*) and an airline company (i.e., *United Airlines*) in order to validate the proposed expert and intelligent system not only for predicting en- terprise outcomes in a long period but also for predicting the ef- fects of specific events on enterprise outcomes in a short period. [Section 5](#_bookmark32) presents the experimental results and discussion, and [Section 6](#_bookmark41) concludes the paper.



**Fig. 1.** Enterprise decision trees: without the proposed expert and intelligent system (left-hand side) and with the proposed expert and intelligent system (right-hand side).

## Literature review

The literature review section contains literature related to so- cial media user feedback extraction ([Section 2.1](#_bookmark6)) and using social media data to predict real-world events ([Section 2.2](#_bookmark7)).

* 1. *Social media user feedback extraction*

Identifying customer feedback could enable enterprise deci- sion makers to understand feedback that leads to positive or [negative consumer experiences (Chang, Chou, Wu, & Wu, 2017; Lim & Tucker, 2017). Customer feedback extraction methods from](#_bookmark75) [textual data are an emerging field. Wang, Youn, Azarm, and Kannan (2011) propose a systematic method that extracts cus-](#_bookmark109) tomer data for product design selection based on web-based user- generated content. [Yan, Xing, Zhang, and Ma (2015)](#_bookmark112) develop a novel method that ranks candidate terms in online customer re- views in order to automatically elicit enterprise attributes that in- terest customers. Recently, a sentiment analysis and a case ana- logical reasoning model have been proposed for extracting enter- prise attributes and latent customer needs from online consumer reviews ([Zhou, Jiao, & Linsey, 2015](#_bookmark115)).

Over the past years, social media platforms, including Twitter, [have been widely used for eliciting customer feedback (Li, Chen, Liou, & Lin, 2014). For instance,](#_bookmark73) [Ru](#_bookmark107)[i, Liu, and Whinston (2013) pro-](#_bookmark73) pose a dynamic panel data model in order to investigate how Twit- ter word-of-mouth influences movie sales. A novel decision sup- port system is proposed for filtering valuable intelligence frozen in [the large number of postings (Abrahams, Jiao, Fan, Wang, & Zhang, 2013) as well as finding, categorizing, and prioritizing automotive](#_bookmark47) [defects discussed in social media (Abrahams, Jiao, Wang, & Fan, 2012).](#_bookmark48) [Bao](#_bookmark60) [and Chang (2014) investigate the relationship between](#_bookmark48) product reports written by journalists, social media feedback writ- ten by customers, and product sales (i.e., Amazon book sales) using the New York Times Best Seller List and Amazon user reviews. Us- ing a cluster analysis, [Zhang (2015)](#_bookmark114) investigates the factors affect- ing both a company’s overall information disclosure and its finan- cial information disclosure in social media to discover that compa- nies with high adoption levels attract more interested customers using social media than do companies with low levels. [Tuarob and](#_bookmark103)

[Tucker (2015)](#_bookmark103) propose a mathematical model that discovers latent customer feedback using ground truth data extracted from social media networks. Since ground truth data acquisition from social [media data is expensive or unavailable in some cases, Lim and Tucker (2016) propose a Bayesian-based statistical sampling algo-](#_bookmark76) rithm that extracts customer feedback from social media networks without utilizing ground truth data.

Nevertheless, considering social media user feedback that has causal effects on real-world enterprise outcomes is still limited. This consideration has recently become significant in enterprise decision making. Enterprise outcomes substantially affect business revenue, and real-time social media user feedback is useful for an enterprise’s quick response to the market. In this work, the proposed system identifies social media user feedback that causes real-world outcomes for enterprise decision making.

* 1. *Using social media data to predict real-world events*

Predicting future events is significant for real-world businesses, because the ability to predict future events affects enterprise de- [cision making and relates to market success (Fourt & Wood- lock, 1960). Recently, several researchers have proposed methods](#_bookmark87) that use social media data, including Twitter data, to predict real- world events ([Liu, Wu, Li, & Li, 2015](#_bookmark81)). For example, in order to predict population health indices, [Nguyen et al. (2017)](#_bookmark94) propose a mathematical model based on the distributions of textual fea- tures over Twitter data. [Gerber (2014)](#_bookmark92) presents a crime predic- tion model, based on linguistic analysis and statistical topic mod- eling, that uses spatiotemporally tagged tweets across Chicago, Illi- nois. Researchers show that their prediction methods that use so- cial media data outperform existing predictors that forecast real- [world events, such as Oscar award winners (Bothos, Apostolou, &](#_bookmark69) [Mentzas,](#_bookmark62) [2010), energy utilization patterns (Bodnar, Dering, Tucker,](#_bookmark69) [&](#_bookmark53) [Hopkinson, 2016), and box-oﬃce revenues (Asur & Huberman,](#_bookmark62) [2010; Ding, Cheng, Duan, & Jin, 2017).](#_bookmark53)

In particular, sentiment expressed on social media networks has [been widely utilized to predict enterprise outcomes (Khadjeh Nas-](#_bookmark70) [sirt](#_bookmark102)[oussi, Aghabozorgi, Ying Wah, & Ngo, 2014). Tuarob and](#_bookmark70) [Tucker (2013) propose a mathematical algorithm that quanti-](#_bookmark102) fies the correlation between sentiment on social media networks

**Table 1**

Summary of previous studies and this research on predicting real-world enterprise outcomes.

Reference Correlation between social media users’ sentiment and enterprise outcomes

Correlation between social media user feedback and enterprise outcomes

Causation between social media users’ sentiment and enterprise outcomes

Causation between social media user feedback and enterprise outcomes

Sentiment metric

The unit of analysis

Domain Enterprise Outcome

[Tuarob and Tucker (2015)](#_bookmark103)

√ 3 classes

(negative, neutral, and positive)

Products Product markets Product sales

[Bogle and Potter (2015)](#_bookmark65)

√ 3 classes

(negative, neutral, and positive)

Jamaica Stock Stock markets Stock prices Exchange

[Ding et al. (2017)](#_bookmark83) √ 2 classes

(neutral and positive)

[Ho et al. (2017)](#_bookmark55) √ A range of −2

to 2

Movies Box oﬃces Box oﬃce sales

Companies Stock markets Stock returns

[Asur and Huberman (2010)](#_bookmark53)

√ − Movies Box oﬃces Box oﬃce revenues

[Bollen et al. (2011)](#_bookmark66) √ 6 classes (calm,

alert, sure, vital, kind, and happy)

Dow Jones Industrial Average (DJIA)

Stock markets Stock prices

[Bae and Lee (2012)](#_bookmark59) √ 3 classes

(negative, neutral, and positive)

Audiences of popular users

Social media and real-world

The sentiment of audiences and real-world phenomena

[Smailovic´,](#_bookmark113)

[Grcˇar, Lavracˇ, and Žnidaršicˇ (2013)](#_bookmark113)

[Ranco, Aleksovski, Caldarelli, Grcˇar, and Mozeticˇ (2015)](#_bookmark105)

[Checkley, Higón, and Alles (2017)](#_bookmark77)

√ The ratio of

positive messages

√ 3 classes

(negative, neutral, and positive)

√ A range of 0 to

4 for bullish and bearlish sentiment

Companies Stock markets Stock market

movements Companies Stock markets Stock price returns

Companies Stock markets Financial indicator

(return, trading volume, volatility)

Ours √ √ A range of −5

to 5

Companies Stock markets Stock prices

and product market adoption in order to predict product sales. [Nguyen, Shirai and Velcin (2015)](#_bookmark93) analyze sentiment of the top- ics, which are extracted from social media networks using latent [Dirichlet allocation (LDA), for stock market prediction. Ho, Damien, Gu, and Konana (2017) investigate a potential dynamic relationship](#_bookmark55) between sentiment expressed on social media networks and future stock returns.

[Table 1](#_bookmark8) shows a summary of existing studies and the proposed research on predicting real-world enterprise outcomes. While pre- vious expert and intelligent systems have been widely applied to predict real-world enterprise outcomes, identifying social me- dia user feedback that causes future enterprise outcomes, which can support enterprise decision makers’ future decisions, remains ambiguous. Identifying social media user feedback helps enter- prise decision makers understand how social media user feedback causes future enterprise outcomes. This work’s novelty is that the expert and intelligent system presented identifies not only Twit- ter users’ sentiment but also Twitter user feedback that has causal [effects on real-world enterprise outcomes, while Bollen, Mao, and Zeng (2011) only consider Twitter users’ overall sentiment that af-](#_bookmark66) fects enterprise outcomes through Granger causality analysis. This work also provides the appropriate time lags between identified Twitter user feedback and real-world enterprise outcomes.

## Method

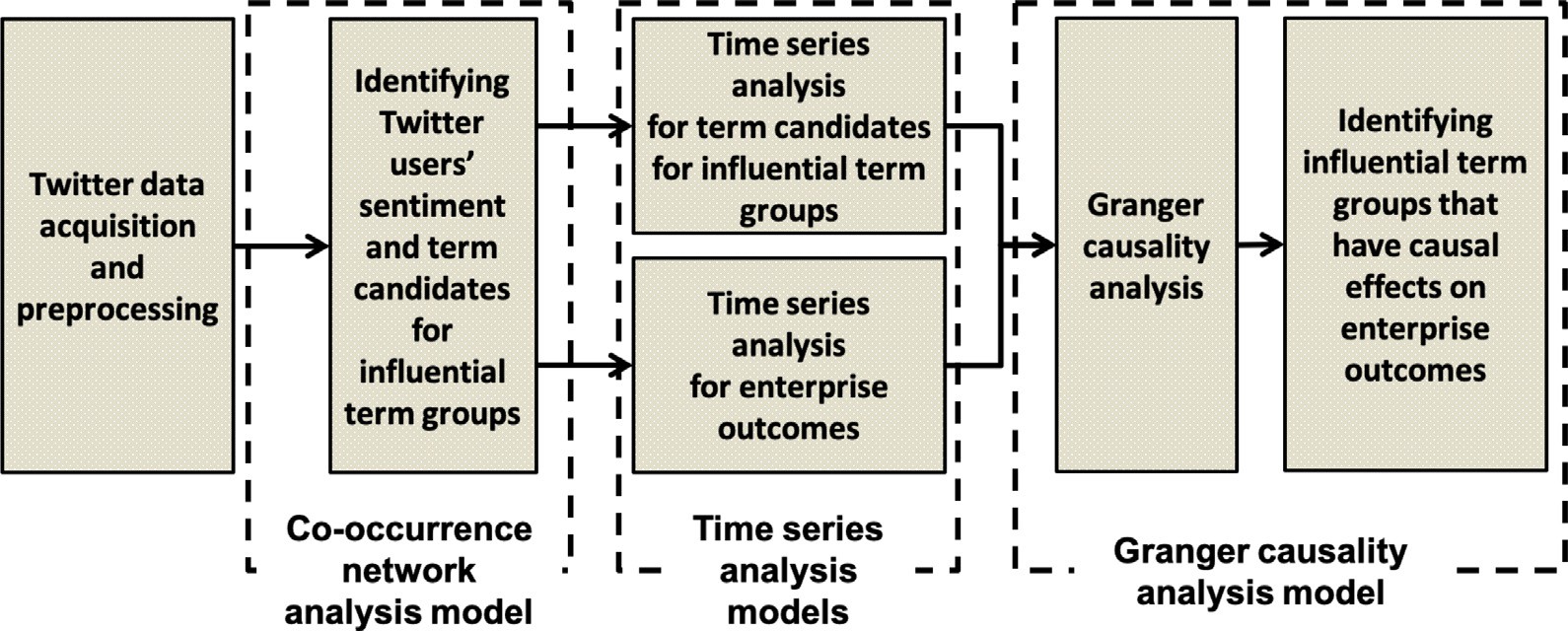
[Fig. 2](#_bookmark10) outlines the proposed expert and intelligent system with the steps involved in identifying Twitter users’ feedback that

causes future enterprise outcomes, along with the appropriate time lags. First, textual and temporal data are extracted from Twitter data, and data preprocessing is employed. Term candidates for in- fluential term groups, which are considered able to potentially af- fect enterprise outcomes, are identified through the co-occurrence network analysis model. Time series analysis models then analyze the patterns of Twitter users’ feedback (i.e., Twitter users’ senti- ment and influential term groups) on enterprise decisions as well as the trend of real-world enterprise outcomes. Finally, the Granger causality analysis model identifies influential term groups, along with Twitter users’ sentiment, that affect enterprise outcomes. The appropriate time lags between identified influential term groups and Twitter users’ sentiment and real-world enterprise outcomes are also discovered through the Granger causality analysis model.

* 1. *Twitter data acquisition and preprocessing*

Tweets containing the name of the company (e.g., “netflix”), along with temporal information, are extracted, because only enterprise-related tweets are necessary for this research. If the name of the enterprise is not unigram (e.g., *United Airlines*), sim- ilar expressions, including the abbreviation of the enterprise, (e.g., “united airlines”, “united airline”, “unitedairlines”, “unitedairline”, “ua”) are considered. Twitter application program interface (API) is used for Twitter data acquisition.

Tweets are filled with noise that can induce unexpected re- sults ([Russell, 2013](#_bookmark110)). Therefore, data preprocessing is necessary for removing noise and enhancing the quality of the experimen-



**Fig. 2.** Overview of the proposed expert and intelligent system.

tal results ([Symeonidis, Effrosynidis, & Arampatzis, 2018](#_bookmark96)). Out-of- vocabulary (OOV) words, such as typos (e.g., “dilivery” instead of “delivery”), single-word abbreviations (e.g., “luv” instead of “love”), and phonetic substitutions (e.g., “2day” instead of “to- day”) are transformed to in-vocabulary (IV) words using exist- [ing OOV word databases (e.g., Apache Lucene (Apache Lucene, 2010) and Spell Checker Oriented Word Lists (SCOWL) and Friends](#_bookmark54) ([Atkinson, 2017](#_bookmark56))), because tweets contain a high rate of OOV words ([Nikfarjam, Sarker, O’Connor, Ginn, & Gonzalez, 2015](#_bookmark97)). In this work, SCOWL and Friends, which is an English word database that con- tains 657,798 words useful for generating high quality word lists suitable for use in spell checkers of most English dialects, is used. A web application checks if a word is in SCOWL and assigns a score that indicates if a word should or should not be included based on its frequency in Google Book’s corpus. Stemming is implemented using the Porter stemming algorithm ([Porter, 2006](#_bookmark104)) in order to im- prove result accuracy. For example, an original tweet “C U 2mor- row” is converted to “see you tomorrow” after preprocessing.

* 1. *Analyzing twitter users’ sentiment and discovering term candidates for influential term groups*

Analyzing Twitter users’ sentiment and discovering term candi- dates for influential term groups are used to identify Twitter users’ feedback that can affect future enterprise outcomes. In this work, time, which has a continuous nature, is approximated as a discrete time window. *t* is defined based on a unit of time (e.g., one day) for the discrete time window. It is assumed that the data belong- ing to the same time frame (i.e., the same sub-interval) are station- ary ([Kaz´mierski & Morawiec, 2011](#_bookmark67)), and this assumption is used in Granger causality analysis in further steps.

* + 1. *Sentiment analysis using Twitter data*

*SentiStrength*, which is a trained sentiment classifier developed by [Thelwall, Buckley, Paltoglou, Cai, & Kappas (2010)](#_bookmark100), is used for sentiment analysis in this research. Each tweet is used as an in- put, and the output has a sentiment score that ranges from −5 to

5. Positive and negative numbers indicate positive sentiment and negative sentiment, respectively, and 0 is neutral. Let *K*(*t*) be the average of sentiment scores for all tweets written during the time period *t* (*t* ∈ {1*, . . . , T* }).

* + 1. *Discovering term candidates for influential term groups*

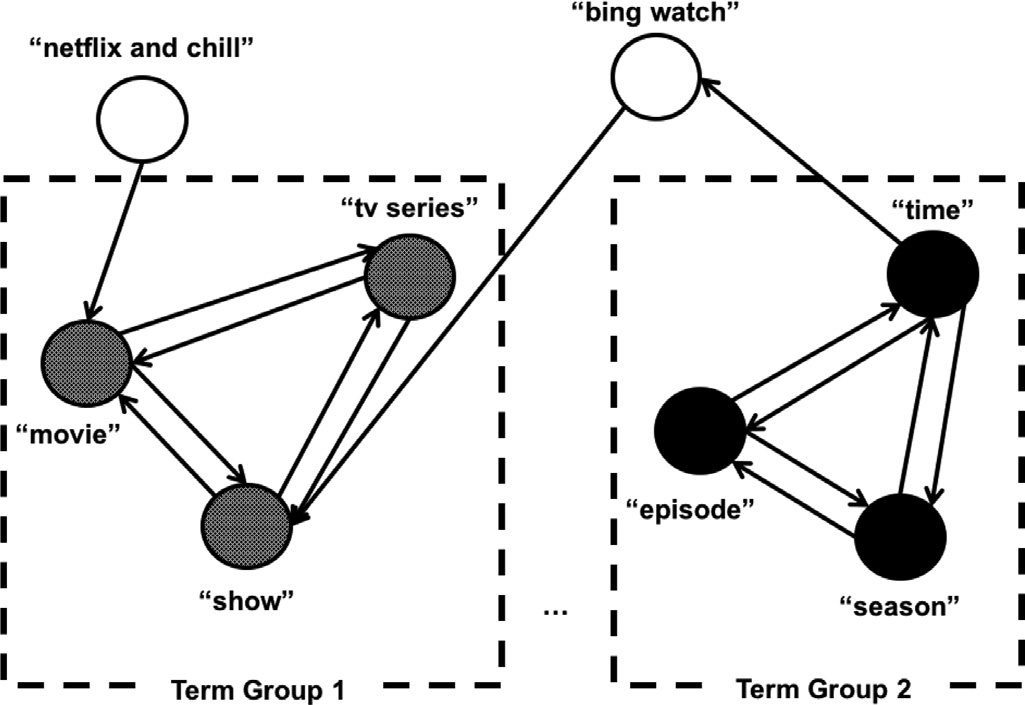
In this research, term candidates for influential term groups are composed of enterprise-related terms by time period. Significant enterprise-related terms can be determined differently by differ- ent applications, because important factors for an insurance com- pany’s outcomes and important factors for an online retailer’s out-

comes can differ. For instance, “shipping” may be one of the most significant enterprise-related terms for online retailers (e.g., *Ama- zon.com*), but it may not be an important factor for insurance com- panies (e.g., *GEICO*). It is therefore necessary to propose a general- ized model for discovering Twitter users’ feedback in order to ap- ply the model to different applications.

In this section, enterprise-related term candidates are identi- fied for creating influential term groups. Bigram and trigram key- words, as well as unigram keywords, are considered when identify- ing enterprise-related term candidates from tweets, because some enterprise-related expressions are bigrams or trigrams (e.g., “tv se- ries” for *Netflix*) instead of unigrams. ***S*** is defined as a set of all tweets containing the name of the enterprise (e.g., “netflix”) for the whole time period.

Let ***S1* , *S2*** , and ***S3*** be sets of the frequent unigrams, bigrams, and trigrams, respectively, which are identified in ***S***. Only uni- grams, bigrams, and trigrams that appear more than 0.5% of ***S*** are contained in ***S1* , *S2*** , and ***S3*** , respectively, because it is considered that terms that appear with low frequency (i.e., less than 0.5% of ***S***) do not have significant effects on enterprise-related term can- [didate discovery (Davidov, Tsur, & Rappoport, 2010; Stringam & Gerdes, 2010). *Stop words* (e.g., “it”, “to”) are excluded from ***S1***,](#_bookmark82) because language-specific functional terms and frequently occur- ring words in the English dictionary (i.e., stop words) are super- fluous for enterprise-related term discovery. In addition, only bi- grams and trigrams listed in *Urban Dictionary* ([Peckham, 2009](#_bookmark101))

are considered candidates for ***S2*** and ***S3*** (e.g., “tv series”). This is because Urban Dictionary is widely used in social media analyt- ics as a website that provides user-generated content and voting [mechanisms for defining colloquial terms (Marwick & boyd, 2011; Paul, Agrawal, Liao, & Choudhary, 2016; Peleja, Santos, & Magal- hães, 2014; Thompson, Rivara, & Whitehill, 2015; Wu, Morstatter, & Liu, 2016). Only bigrams and trigrams listed in Urban Dictio-](#_bookmark88) nary are considered term candidates for influential term groups in this research, because most of bigrams and trigrams (e.g., “worth it”, “it free”, “worth it free” in the sentence “Worth it! Free wifi, straight getting paid to Netflix and chill”) are not idiomatic ex- pressions ([Bodnar, Dering, Tucker, & Hopkinson, 2016](#_bookmark62)). Conjunc- tions are not disregarded if they are part of the bigrams or tri- grams. A frequent bigram in ***S2*** or a frequent trigram in ***S3***, which is composed of a frequent unigram(s), is considered a different fre- quent terms with frequent unigrams for the same reason. For in- stance, “chill” and “netflix and chill”, which have different mean- ings, are considered different terms. Four-grams or larger are not considered in this research, because (1) it is assumed that ***S1*, *S2***, or ***S3*** contain the sub-sequences of the appropriate four-gram or larger terms ([Fürnkranz, 1998](#_bookmark90); [Bodnar, Dering, Tucker, & Hopkin-](#_bookmark62)

[son, 2016](#_bookmark62)), and (2) tweets commonly consist of short messages and have a limit of 140 characters ([Lim, Tucker, & Kumara, 2017](#_bookmark80)).

The co-occurrence network analysis model is proposed to clus- ter the co-occurring unigrams, bigrams, and trigrams for creat- ing influential term groups. A co-occurrence is a term intercon- nection based on their paired existence within a specified unit of text ([Lim, Tucker, Jablokow, & Pursel, 2018](#_bookmark79)). For instance, the terms “free” and “netflix and chill” co-occur if they both appear in a par- ticular tweet (e.g., “Worth it! Free wifi, straight getting paid to Net- flix and chill”). Let *wij* be a co-occurrence weight of the term *ki* for

another keyword *kj* (*ki, kj* ∈ {***S***1 ∪***S***2 ∪***S***3 }, *i* /= *j*). In this research, *wij* is

defined in a method similar to the term expansion ranking func- tion, which is introduced by [Ruthven and Lalmas (2003)](#_bookmark111) as [Eq. (1)](#_bookmark15).

*wi* is defined as [Eq. (2)](#_bookmark16).

¡ *nij/*¡*ni* − *nij* ¢

¯ *ij*

*n*

*j*

*ij*

*i*

*n j* − *nij* ¯

*i*

*wij* = *log*

*j*

*ij*

*i*

*n* − *n* ¢*/*¡*n* − *n* − *n*

+ *n* ¢ · ¯ *n*

— *n* − *n*

¯ (1)

**Fig. 3.** An example of a partial co-occurrence graph *G* for *Netflix.*

*w**i* = *wi*1 + *wi*2 + ··· + *wii*−1 + *wii*+1 + ··· + *wiI ,* ∀*i, j* = 1*, . . . , I*

# (2)

where:

*n*: the total number of tweets in ***S*** (i.e., a set of all tweets con- taining the enterprise name for the whole time period)

*ni* : the number of tweets containing the enterprise name and a term *ki*

*nij* : the number of tweets containing the enterprise name as well as terms *ki* and *kj* (*i*/=*j*)

*I* = |***S***1 ∪ ***S***2 ∪ ***S***3|

A weighted adjacency matrix *A*, which is a co-occurrence ma- trix among the terms from ***S1* , *S2*** , and ***S3*** , is then generated based on the co-occurrence weights as [Eq. (3)](#_bookmark18). A matrix *A* is not a trian-

*wij wji*

gular symmetric matrix, since *wi* /= *wj* .

*k*1 *k*2 *k*3 ··· *kI*

*w*1

··· ⎥

considered an independent term group if the number of tweets containing both the enterprise name (e.g., the term “netflix”) and the term *ki* is greater than or equal to the minimum value of the number of tweets containing the enterprise name and at least one term belonging to the same term group identified from the graph *G* (e.g., “episode”, “season”, or “time”). In this work, *Xp*(*t*) is defined as [Eq. (4)](#_bookmark17) for period *t* based on its definition.

*X (t )* = *n( p, t) ,* ∀ *p* = 1*,* ··· *, P,* ∀*t* = 1*,* ··· *, T* (4)

*p n(t)*

where:

*n*(*t*): the total number of tweets containing the enterprise name for period *t*.

*n*(*p, t*): the number of tweets containing the name of the en-

terprise and at least one term belonging to the *p*th term group for

*k*1 ⎡ − *w*12

*w*1

*k*2 ⎢

*w*13 *w*1

··· *w*1*I* ⎤

period *t*, where the number of identified term groups is *P*.

*w*21 *w*2

—

⎢ *w*31

*w*23 *w*2

*w*32

*w*2*I w*2

*w*3*I* ⎥

⎥

*A* = *k*3 ⎢ *w*3

*w*3 − ··· *w*3

(3)

. ⎢⎣

.

*wI*1

.

*wI*2

. .

*wI*3

. ⎥⎦

* 1. *Time series analysis models*

Time series models analyze a sequence of data points over

*kI wI wI*

*wI* ··· −

time in order to extract the given data’s statistical characteris-

An edge-weighted graph *G* of co-occurrence can be generated based on a matrix *A*. A graph *G* is expressed as a directed graph, since a matrix *A* is not a triangular symmetric matrix and the di- rection is important in this case. Only components of the matrix *A* that are above or equal to the average weights for each row

tics and predict future values based on previously observed data ([Hamilton, 1994](#_bookmark57)). Time series models are often used to predict [the change of enterprise outcomes (Lee, Cho, Kwon, & Sohn, 2019; Luo, Zeng, & Duan, 2016; Rosas-Romero, Díaz-Torres, & Etchev- erry, 2016; Weng, Lu, Wang, Megahed, & Martinez, 2018). In this](#_bookmark71)

(i.e., wij

wi

that is greater than or equal to  1 ) are used to con-

*I*−1

research, two kinds of time series analysis models are proposed

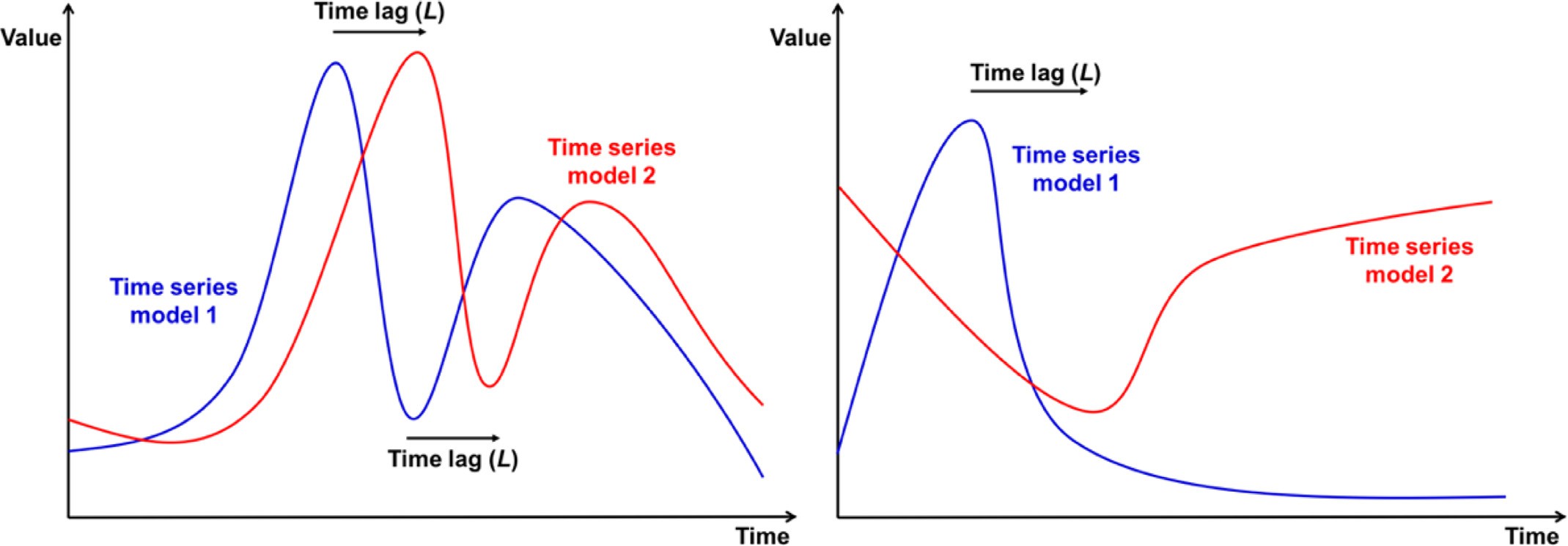
struct the co-occurrence graph *G*, because the components hav- ing below the average weights are not considered to have a sig- nificant co-occurrence compared to other terms in the document (i.e., all tweets). In addition, each term group is defined as strongly connected components in the graph *G*, in a method similar to [Cataldi, Di Caro, and Schifanella (2010)](#_bookmark74). [Fig. 3](#_bookmark14) shows an example of a (partial) co-occurrence graph for *Netflix*, where ***S1***={“movie”, “show”, “episode”, “season”, “time”}, ***S2***={“bing watch”, “tv series”}, and ***S3***={“netflix and chill”}. In this example, two term groups that contain multiple terms are identified: one contains the keywords “movie”, “show”, and “tv series”, and the other contains the key- words “episode”, “season”, and “time”. Some single components are not disregarded, because some frequent terms can be the term groups themselves, even if they do not have strongly connected neighbors (e.g., the term “netflix and chill” does not have strongly connected neighbors in [Fig. 3](#_bookmark14).). In this work, a single term *ki* is

for identifying Twitter user’s feedback that has causal effects on

enterprise outcomes: (1) a time series analysis model that ana- lyzes Twitter users’ sentiment and influential term groups and (2) a times series analysis model that analyzes real-world enterprise outcomes.

Time series analysis model used for analyzing Twitter users’ feedback quantifies the trend of Twitter users’ sentiment and in- fluential term groups. *K*(*t*) is a time series analysis model for Twit- ter users’ sentiment (see [Section 3.2.1](#_bookmark12)), and *Xp*(*t*) (*p* 1,…,*P*) is a time series analysis model for the *p*th term groups for period *t* (see [Section 3.2.2](#_bookmark13)). Time series analysis models used for analyzing enterprise outcomes investigate the trend of real-world enterprise outcomes over time. Let *Yj* (*t*) (*j* = 1,…,*J*) be a time series analysis model used for analyzing the *j*th enterprise outcome (e.g., market sales, stock prices) for period *t*, where the number of the enter- prise outcomes is *J*.

=



**Fig. 4.** Illustrations of two different time series models having positive causal relationships (left-hand side) and negative causal relationships (right-hand side).

* 1. *Granger causality analysis model for twitter users’ feedback and real-world enterprise outcomes*

*X* is said to cause *Y* if it can be shown that *X* provides statis- tically significant information about the future values of *Y*. Causal- ity is differentiated from mere correlation, because a correlation between two events does not imply that one event (e.g., Twit- ter users’ feedback) causes the other (e.g., an enterprise outcome) ([Aldrich, 1995](#_bookmark51)). While traditional regressions consider correlations, causality analysis can test a predictive causal relationship between two time series models ([Diebold, 2001](#_bookmark84)). First introduced by C. W. Granger, Granger causality analysis identifies whether one station- ary time series model can be used to predict the future values of the other stationary time series model. A Granger causal relation-

[Eq. (8)](#_bookmark24) is the same time series analysis model with [Eq. (7)](#_bookmark23) ex- cept one additional regressor (i.e., *Xp(t l )*). [Eq. (9)](#_bookmark25) is a time se- ries analysis model for the *jth* enterprise outcomes that does not reflect causal effects of any Twitter users’ sentiment or influential term group (i.e., the reduced model):

**Ho** *( p, j).* Neither *K(t)* or *Xp(t )* causes *Yj (t ), (B(l)*

—

= 0 *or Cp(l )* = 0*,* ∀*l* = ··· *, M)* (5)

**Ha** *( p, j).* Either *K(t)* or *Xp(t )* causes *Yj (t ), B(l)*

/= 0 *or Cp(l )* /= 0*,* ∃*l* = ··· *, M)* (6)

ship is defined based on two principles ([Granger, 1969](#_bookmark52)). The first *M M*

principle is *temporal precedence of causes*, which means that the ef- *Yj (t )* = Σ *A j (m)* · *Yj (t* − *m)* + Σ *B(i)* · *K(t* − *i)*

fect does not happen prior to its cause. This first principle is com- [monly accepted by existing probabilistic causation theories (Good,](#_bookmark95)

*m*=1

*M*

*i*=*l*+1

[1961a,b; Suppes, 1970).](#_bookmark95)

The second principle is predictability and can be validated [through the following mathematical approach (Eichler, 2012; Liu & Bahadori, 2012). On the left-hand side of](#_bookmark86) [Fi](#_bookmark19)[g. 4 illustrates how one](#_bookmark86) time series model (i.e., the blue line) has positive causal effects on

+ Σ *Cp(i)* · *Xp(t* − *i)* + *ε j,l*+1*,*

*i*=*l*+1

∀ *j* = 1*,* ··· *, J,* ∀*t* = *M,* ··· *, T* (7)

the other time series model (i.e., the red line) with a time lag *L*. On *M M*

the right-hand side of [Fig. 4](#_bookmark19) shows how one time series model (i.e., *Yj (t )* = Σ *A j (m)* · *Yj (t* − *m)* + Σ *B(i)* · *K(t* − *i)*

the blue line) has negative causal relationships with the other time series model (i.e., the red line) with a time lag *L*. Granger causality

*m*=1

*M*

*i*=*l*+1

analysis is applicable to this work, since time series analysis mod- els (i.e., *K*(*t*), *X**p*(*t*) (*p* = 1,…,*P*), and *Yj* (*t*) (*j* = 1,…,*J*)) are divided into

+ Σ *Cp(i)* · *Xp(t* − *i)* + *ε j,l ,*

*i*=*l*+1

sub-intervals based on a unit of time *t* and it is assumed that the

time series are stationary in each sub-interval in this work (see [Section 3.2](#_bookmark11)) ([Liu & Bahadori, 2012](#_bookmark85)). A time lag *L* is not set to zero, because (1) it is known that a time delay in enterprise outcome re- [sponses (e.g., stock price responses) to news or events exists (Hou & Moskowitz, 2005) and (2) this research is used for predicting fu-](#_bookmark61)

*Yj (t )* =

∀ *j* = 1*,* ··· *, J,* ∀*t* = *M,* ··· *, T* (8)

Σ *A j (m)* · *Yj (t* − *m)* + *ε j, m*=1

*M*

ture outcomes that lead to future market success rather than dis- covering the correlation of current enterprise outcomes with Twit- ter users’ feedback. In a further step, the appropriate values of a time lag *L* are also discovered.

In this research, the null hypothesis (i.e., *H0*(*p, j*)) and the alter- native hypothesis (i.e., *Ha*(*p, j*)) are defined as [Eqs. (5)](#_bookmark20) and [(6)](#_bookmark21), re- spectively. *Yj* (*t*) can be expressed as [Eq. (7)](#_bookmark23), which is a time series analysis model for the *jth* enterprise outcomes that reflect causal effects of (1) Twitter users’ sentiment (*K*(*t*), set *p* = 0) or [(2)](#_bookmark16) the *pth* influential term group (*Xp*(*t*) (*p* = 1,…,*P*)) (i.e., the full model).

∀ *j* = 1*,* ··· *, J,* ∀*t* = *M,* ··· *, T* (9)

where:

*X1*(*t*): the value of the average sentiment scores at time *t*

*K*(*t*): the average of sentiment scores for all tweets written dur- ing the time period *t*

*Xp*(*t*): the proportion of tweets which mention a term from the *p*th influential term group, out of all tweets containing the enter- prise name at time *t*, ∀ *p* = 1*,* ··· *, P*

*Yj* (*t*): the value of the *j*th enterprise outcome at time *t,* ∀ *j* =

1*,* ··· *, J*

*Aj*(*m*): Coeﬃcients for indicating the influence of the previous

value of *Yj*(*t*), *j* = 1*,* ··· *, J*

∀

*B*(*i*): Coeﬃcients for indicating the influence of the previous value of *K*(*t*)

*Cp*(*i*): Coeﬃcients for indicating the influence of the previous value of *Xp*(*t*), *p* = 1*,* ··· *, P* + 1

∀

*P*: the number of term groups extracted from Twitter users’ feedback

*J*: the number of enterprise outcomes used for Granger causality analysis

*M*: the number of previous values used for predicting *Yj*(*t*) (i.e., the maximum possible value of time lags)

ɛ*j, l*, ɛ*j*: the prediction error*, j* = 1*,* ··· *, J, l* = 0*,* ··· *, M* − 1

∀ ∀

*M* (i.e., the number of previous values used for predicting *Yj*(*t*)) is set to 8 as the default value in this research. Commonly used lag parameter values of Granger causality for future market anal-

ysis is not greater than 8, and 8 is commonly used as the num- [ber of previous values for predicting *Yj*(*t*) (Bollen et al., 2011; Liew, 2004; Thornton & Batten, 1985). *H0*(*p, j*) is not rejected if and](#_bookmark66) only if *Cp(*1*)* = *Cp(*2*)* = *. . .* = *Cp(M)* = 0 in [Eq. (7)](#_bookmark23) (or [Eq. (8)](#_bookmark24)). Par- tial *F*-tests are used to test the hypothesis. If the *p*-value is less

than *α* (i.e., the significance level, implying that it is permissible to have an *α* probability of incorrectly rejecting the null hypoth- esis), the null hypothesis (i.e., *H0*(*p, j*)) is rejected. It is then con- cluded that causal effects of (1) Twitter users’ sentiment (set *p* 0) or (2) the *pth* term group on the *jth* enterprise outcome exist. The proposed Granger causality analysis model identifies Twitter users’ feedback that affects enterprise outcomes over the whole time pe- riod *(t* ∈ 1*, . . . , T )*. Decision makers can set the period for train- ing the systems (i.e., *T*: the number of previous time units used for discovering Twitter users’ feedback), while it is known that *T* ≥ 14 [is appropriate for the default value (Bollen et al., 2011; Makrehchi, Shah & Liao, 2013; Si et al., 2013).](#_bookmark66)

=

{ }

The next step is identifying the appropriate time lag for each

**Algorithm 1** Identification of the appropriate value of a time lag *L*.

**STEP 1** Set *l* = *M*−1.

**STEP 2** If *C* is less than or equal to the critical value *χ* 2 1 *α* , go to STEP 3.

*N*

*j*

1*,* −

Otherwise, go to STEP 4.

**STEP 3** Set *l* = *l*−1 and go to STEP 2.

**STEP 4** Stop. Set *L* = *l* + 1 and return *L* (i.e., the appropriate value of a time lag).

Twitter users’ feedback (i.e., influential term groups) helps enter- prise decision makers change their decisions (or make new deci- sions) in real-time in order to predict and improve future enter- prise outcomes. For instance, suppose that the term group “price” and “discount”, is identified as an influential term group that has causal effects on future enterprise outcomes, and the appropriate time lag is four days. It is expected that real-time Twitter users’ feedback regarding a price discount affects future enterprise out- comes after four days. The proposed expert and intelligent sys- tem then helps decision makers analyze Twitter users’ feedback regarding a price discount in order to improve future enterprise outcomes.

## Applications

This section introduces case studies involving an internet video streaming and disc rental provider (i.e., *Netflix*) and an airline com- pany (i.e., *United Airlines*). These case studies are used to verify the proposed expert and intelligent system for a relatively long period (i.e., one year) and for a specific event (i.e., the *United Express Flight 3411* incident [Victor & Stevens, 2017](#_bookmark108)), respectively. Experiments are conducted on a 2.5 GHz Intel Core i7 with 16GB RAM using *Python*

2.7.14 and *R* 3.4.3. Two Twitter datasets, which were random sub- samples using Twitter API, along with temporal information, are exploited for the case studies of *Netflix* and *United Airlines*, respec-

Twitter users’ sentiment or term group. If ˆ 2

*σ*

*j,l*

(i.e., the estimated

tively. Tweets not written in English are disregarded in this case

study. Apache Lucene API as well as Spell Checker Oriented Word

variance of ɛ*j, l* in [Eq. (8)](#_bookmark24)) is statistically significantly smaller than

*σ*ˆ 2

*j,l*+1

(i.e., the estimated variance of *ε*

*j,l*+1

in [Eq. (7)](#_bookmark23)), it is con-

Lists are used to transform OOV words to IV words. The Fox stop

list ([Fox, 1989](#_bookmark89)) is used to remove stop words. Default values are

cluded that *Xp(t* − *l)* has a significant causal effect on predict-

ing *Yj*(*t*). The measure of causality at time lag *l* is defined as [Eq. (10)](#_bookmark28). The measure of causality can be tested using the like- [lihood ratio (LR) test statistic (](#_bookmark49)[Eq. (11)](#_bookmark29)[) (Gourieroux & Monfort, 1997;](#_bookmark49) [Gelper](#_bookmark91)[, Lemmens, & Croux, 2007):](#_bookmark49)

*σ*ˆ 2

used to set *α* (i.e., 0.05) and *M* (i.e., 8).

The results of the proposed expert and intelligent system, which considers keyword frequencies and co-occurrences for dis- covering Twitter users’ feedback (see [Section 3.2.2](#_bookmark13)), are compared with the results of a random-keyword-sampling method in or- der to validate the need to consider keyword frequencies and co-

*C* = ln  *j,l*+1

*σ*ˆ 2

*j*

*j,l*

# (10)

occurrences. A random-keyword-sampling method is defined as a method that, instead of considering keyword frequencies and co-

*LR* = 2*(*log *L*¡*θj*¢ − log *L*¡*θ*

*j*+1

¢*)* (11)

occurrences, randomly samples terms from all tweets as term can- didates for influential term groups.

where:

log *L(θj*+1*)*: the likelihood at the time series model that does not contain the additional regressor *Xk(t* − *l)* (i.e., [Eq. (7)](#_bookmark23)), with *θ j* the parameter vector collecting the estimate of all *Aj*(*m*), *B*(*i*), all *Cj*(*i*), and the variance of the error term

log *L*(*θj*): the likelihood at the time series model that contains

the additional regressor *Xk(t* − *l)* (i.e., [Eq. (8)](#_bookmark24)), with *θ j* the parame- ter vector collecting the estimate of all *Aj*(*m*), *B*(*i*), all *Cj*(*i*), and the variance of the error terms

[Gourieroux and Jasiak (2001)](#_bookmark50) show that *LR* = *N* × *Cj* ∼ *χ* 2,

1

* 1. *Case study for* Netflix

On the one hand, a case study for *Netflix* validates whether or not the proposed expert and intelligent system is applicable for identifying Twitter users’ feedback that has causal effects on enter- prise outcomes for a relatively long period (i.e., one year). Twitter data over a time period from January 4, 2016 to December 30, 2016 in the United States is used in the case study for *Netflix*. Among the whole dataset, 17,170,347 tweets containing the term “netflix” are

where *N* is the total number of observations. *χ* 2

1*,*1−*α*

is the *α*-upper

extracted for this case study.

quantile of the chi-squared distribution, where the degree of free-

dom is 2. [Algorithm 1](#_bookmark26) summarizes the steps to identify the appro- priate value of a time lag *L*. These steps are similar to the proce- dure proposed by [Gourieroux & Monfort, 1997](#_bookmark49):

The proposed expert and intelligent system supports enterprise decision makers’ future decisions based on identified Twitter users’ feedback (see the right-hand side of [Fig. 1](#_bookmark4)). In particular, identified

*Netflix*’s *Nasdaq* index is selected as the *Netflix*-related outcome

for this case study, because it is the only *Netflix*-related daily outcome. Most other *Netflix*-related outcomes (e.g., market sales, the number of subscribers) are quarterly outcomes, which are too sparse for detailed analysis. *J* is 1, because only one enterprise out- come (i.e., the *Netflix*’s *Nasdaq* index) is used in this case study. Let *t* be the *tth* day on weekdays from January 4, 2016 to December

**Table 2**

*Netflix*-related attributes identified from the co-occurrence graph *G* and by the random-keyword-sampling method, their *p*-values, and the appropriate time lags (*L*).

***S1*** = {“watch”, “show”, “series”, “movie”, “season”, “chill”, “need”, “day”, “tv”, “time” “love”, “youtube”, “episode”, “night”, “start”, “bing”, “hulu”, …, “trailer”, …} (the total number of unigrams: 33)

***S2*** = {“bing watch”, “tv series”, “tv show”} (the total number of bigrams: 3)

***S3*** = {“netflix and chill”} (the total number of trigrams:1)

The number of n-grams in the group *p*-value for enterprise outcome (stock price) *L*

Unigrams Bigrams Trigrams

Twitter users’ sentiment − − − 0.31 −

*Netflix*-related term group identified from the graph G (by the proposed model)

“watch”, “bing”, “hulu”, “bing watch”

**“show”, “series”, “movie”, “tv”, “tv show”, “tv series”**

3 1 0 0.49 −

**4 2 0 0.04 4**

**“season”, “time”, “episode” 3 0 0 0.03 3**

“need”, “day”, “love” 3 0 0 0.17 −

“chill”, “night”, “netflix and chill” 2 0 1 0.21 −

**“youtube”, “start”, “trailer” 3 0 0 0.03 3**

*Netflix*-related term group consisting of randomly sampled keywords (by the

random-keyword-sampling method)

“kid”, “girl”, “origine”, “stranger things”

“essay”, “pandora”, “documentary”, “thank”, “luke cage”, “fuller house”

3 1 0 0.77 −

4 2 0 0.55 −

“walk”, “life”, “day” 3 0 0 0.86 −

“eat”, “good”, “cute” 3 0 0 0.79 −

“random”, “website”, “haters back off”

2 0 1 0.96 −

“best”, “daredevil”, “suggestion” 3 0 0 0.55 −

30, 2016 (i.e., 252 weekdays), and *T* is set to 252. *Y1*(*t*) is defined as the *Netflix*’s *Nasdaq* index.

* 1. *A case study for* United Airlines

On the other hand, a case study for *United Airlines* validates whether or not the proposed expert and intelligent system is ap- plicable to identify Twitter users’ feedback having causal effects on enterprise outcomes for the specific event (i.e., the *United Express Flight 3411* incident) and a relatively short period (i.e., 15 week- days). Twitter data in a time period from April 3, 2017 to April 24, 2017 (i.e., from five weekdays before the incident on April 9 to ten weekdays after the incident) in the United States is used in the case study for *United Airlines*. Among the whole dataset, 140,286 tweets containing the term “united airlines”, “united air- line”, “unitedairlines”, “unitedairline”, or “ua” (i.e., the abbreviation for “united airlines”) are extracted for this case study.

Like the case of *Netflix*, the *United Airlines*’ *Nasdaq* index is used as the *United Airlines*-related outcome for this case study as well, because it is the only *United Airlines*-related daily outcome. Other *United Airlines*-related outcomes, such as market sales, are quar- terly outcomes, which are too sparse for detailed analysis. *J* is 1, because only one enterprise outcome (i.e., the *United Airlines*’ *Nas- daq* index) is selected in this study. Let *t* be the *tth* day on week- days from April 3, 2017 to April 24, 2017 (i.e., 15 weekdays) and *T* is set to 15. *Y1*(*t*) is defined as the *United Airlines*’ *Nasdaq* index on the *tth* day. In addition, sensitivity analysis is implemented to discover the effects of changing time windows and window sizes.

## Results and discussion

* 1. *Results for* Netflix *case study*

[Table 2](#_bookmark30) shows the partial results of ***S1*, *S2***, and ***S3***, which are *Netflix*-related unigrams, bigrams, and trigrams that appear more than 0.5% of ***S*** (see [Section 3.2.2](#_bookmark13)), respectively (i.e., components of the co-occurrence graph *G*). The full results can be found in the

[Appendix A](#_bookmark44) ([Table A1](#_bookmark45)). [Table 2](#_bookmark30) illustrates six *Netflix*-related term groups that consist of unigrams, bigrams, and trigrams, along with their numbers, identified from the co-occurrence graph *G*. It also provides another six *Netflix*-related term groups, which consist of the same numbers of unigrams, bigrams, and trigrams identified by the random-keyword-sampling method. [Table 2](#_bookmark30) shows the *p*-values of *H0*(*p*, 1) and the appropriate time lags (i.e., *L*) identified through [Algorithm 1](#_bookmark26). Only *Netflix*-related term groups that have *p*-values less than 0.05 provide the appropriate time lags (see [Section 3.4](#_bookmark22)). [Fig. 5](#_bookmark34) illustrates the time series of *Netflix*’s *Nasdaq* index as well as identified influential term groups, which has *p*-values less than

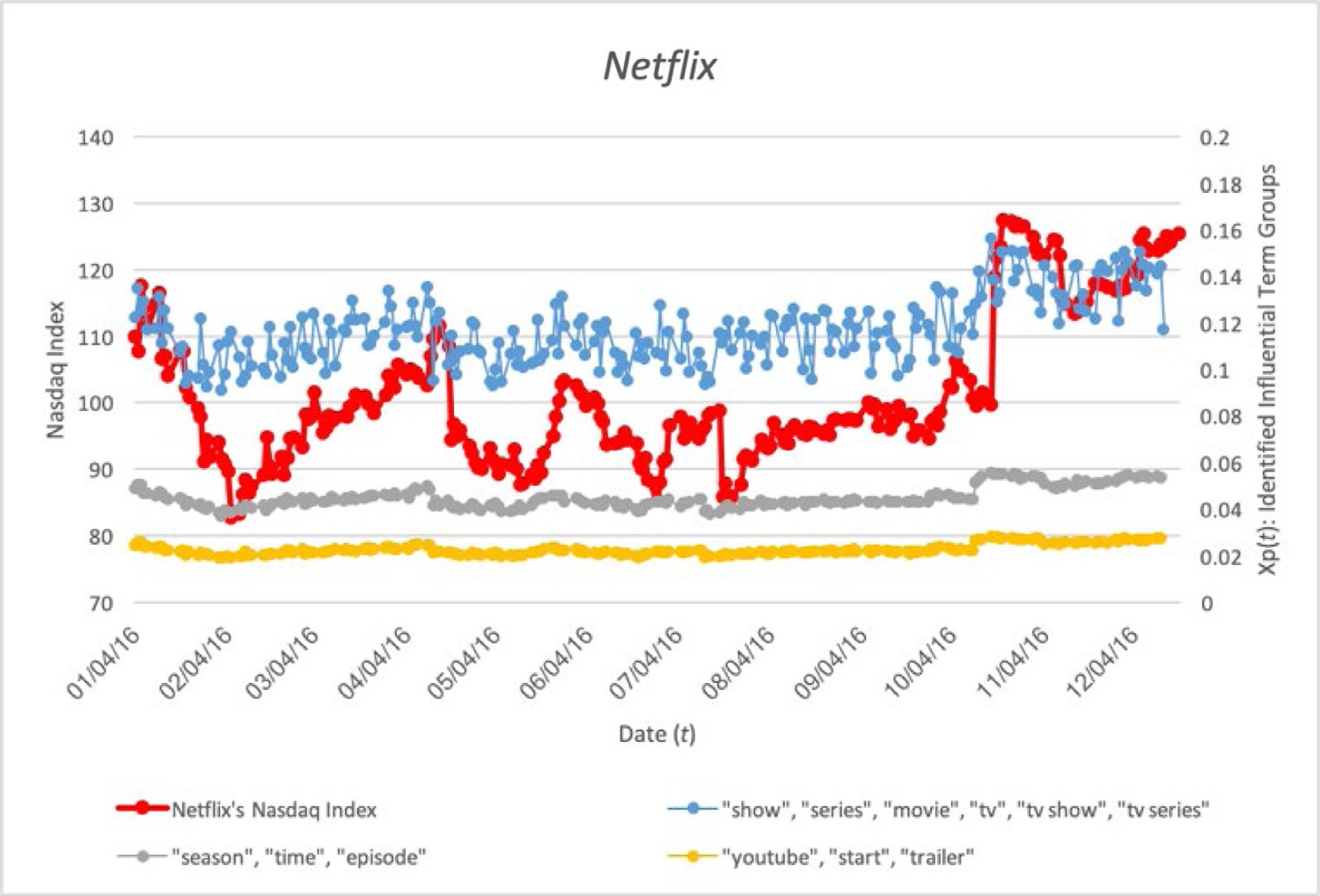
0.05 in [Table 2](#_bookmark30). [Figs. 6](#_bookmark35) and [7](#_bookmark36) show the partial graphs of [Fig. 5](#_bookmark34) in order to highlight the causal relationships between identified in- fluential term groups and (1) an increase in *Netflix*’s *Nasdaq* index and (2) a decrease in *Netflix*’s *Nasdaq* index, respectively.

* 1. *Results for* United Airlines *case study*

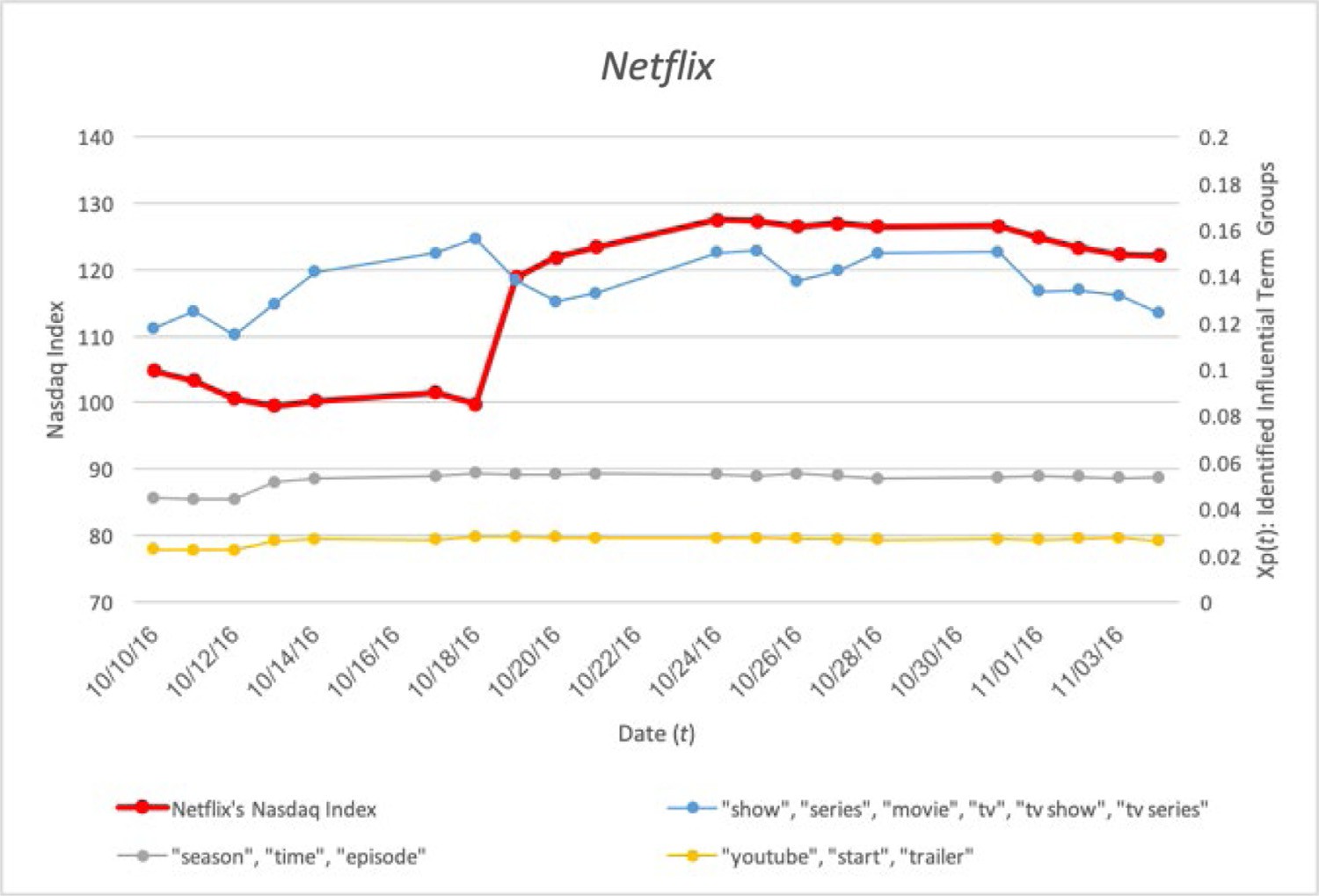
[Table 3](#_bookmark37) illustrates the partial results of ***S1*, *S2***, and ***S3***, which are *United Airlines*-related unigrams, bigrams, and trigrams that appear more than 0.5% of ***S*** (see [Section 3.2.2](#_bookmark13)), respectively (i.e., compo- nents of the co-occurrence graph *G*). The full results can be found in the [Appendix A](#_bookmark44) ([Table A2](#_bookmark46)). It shows five *United Airlines*-related term groups that consist of unigrams, bigrams, and trigrams, along with their numbers, identified from the co-occurrence graph *G*. [Table 3](#_bookmark37) also shows another five *United Airlines*-related term groups, which consist of the same numbers of unigrams, bigrams, and tri- grams, identified by the random-keyword-sampling method. It in- dicates the *p*-values of *H0*(*p*, 1) and the appropriate time lags (i.e., *L*) are identified through [Algorithm 1](#_bookmark26) as well. Only *United Airlines*- related term groups, which have *p*-values less than 0.05, provide the appropriate time lags (see [Section 3.4](#_bookmark22)). [Fig. 8](#_bookmark38) illustrates the time series of *United Airlines*’ *Nasdaq* index as well as identified in- fluential term groups, which has *p*-values less than 0.05 in [Table 3](#_bookmark37).

* + 1. *Sensitivity analysis for United Airlines case study*

While *Netflix* case study uses a relatively long period (i.e., one year), *United Airlines* case study uses a short period (i.e., 15 week-



**Fig. 5.** Time series of *Netflix*’ *Nasdaq* index as well as identified influential term groups.

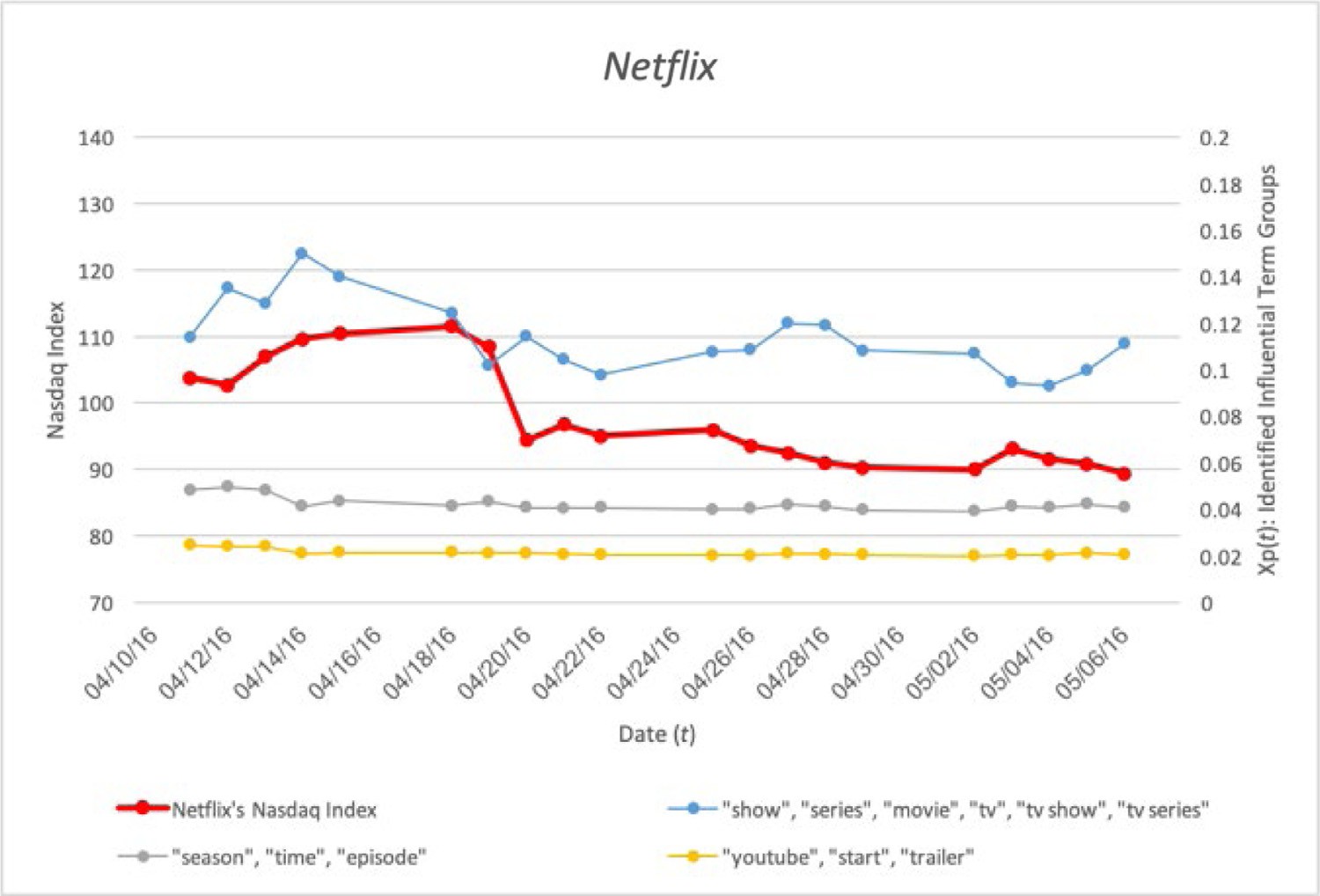


**Fig. 6.** A partial graph of [Fig. 5](#_bookmark34) (October 10, 2016–November 5, 2016).

days), so selecting time windows and window sizes can be sensi- tive to the results of *United Airlines*’ case study. Sensitivity analy- sis is therefore implemented in order to discover whether or not different time windows (i.e., different first day) and different win- dow sizes (i.e., different *T*) affect the causal relationships between Twitter users’ feedback and *United Airlines*’ *Nasdaq* index. [Table 4](#_bookmark39) shows the results of sensitivity analysis using three different first days (i.e., April 3, 2017, March 30, 2017, and April 5, 2017) and three different *T* (i.e., 15, 10, and 20) for the same identified in- fluential term groups. It illustrates the influential term groups, the *p*-values of *H*0(*p*, 1), and the appropriate time lags (i.e., *L*) for each time window and window size.

* 1. *Overall discussion*

According to [Sections 5.1](#_bookmark33) and [5.2](#_bookmark31), the proposed expert and in- telligent system can be used for identifying Twitter users’ feedback not only to predict future enterprise outcomes over a long term ([Section 5.1](#_bookmark33)) but also to identify the effects of specific events on fu- ture enterprise outcomes in a short period ([Section 5.2](#_bookmark31)). Identified Twitter users’ feedback can be used by enterprise decision mak- ers to change their current decisions before Twitter users’ feedback about current decisions or events affecting future enterprise out- comes (i.e., time lags), as shown on the right-hand side of [Fig. 1](#_bookmark4). The results of [Tables 2](#_bookmark30) and [3](#_bookmark37) (and [Tables A1](#_bookmark45) and [A2](#_bookmark46)) show that



**Fig. 7.** A partial graph of [Fig. 5](#_bookmark34) (April 11, 2016–May 7, 2016).

**Table 3**

*United Airlines*-related attributes identified from the co-occurrence graph *G* and by the random-keyword-sampling method, their *p*-values, and the appropriate time lags (*L*).

***S1*** = {“flight”, “passenger”, “drag”, “overbook”, “remove”, “man”, “video”, “plane”, “ceo”, “customer”, “forcible”, “seat”, “incident”, …, “news”, “apology”, …, “time”, …, “munoz”, “oscar”, …, “youtube”, …, “social”, …, “media”, “twitter”, …} (the total number of unigrams: 114)

***S2*** = {“oscar munoz”, “social media”} (the total number of bigrams: 2)

***S3*** = *ø* (the total number of trigrams: 0)

The number of n-grams in the group p-value for enterprise outcome (stock price) *L*

Unigrams Bigrams Trigrams

Twitter users’ sentiment − − − 0.24 −

*United Airlines*-related term group identified from the graph G (by the proposed model)

**“passenger”, “drag”, “remove”, “plane”, “seat”**

**5 0 0 0.03 2**

“flight”, “overbook” 2 0 0 0.15 −

*United Airlines*-related term group consisting of randomly sampled keywords (by the

random-keyword-sampling method)

“man”, “video”, “customer”, “youtube”

**“ceo”, “apology”, “munoz”,**

**“oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media”**

“incident”, “forcible”, “news”, “time”

“fatigue”, “number”, “europe”, “exit”, “demo”

4 0 0 0.29 −

**7 2 0 0.04 2**

4 0 0 0.22 −

5 0 0 0.64 −

“lesson”, “die” 2 0 0 0.79 −

“trash”, “gate”, “collect”, “vote” 4 0 0 0.41 −

“damn”, “mention”, “agent”,

“cover”, “leader”, “risk”, “san francisco”, “jet lag”

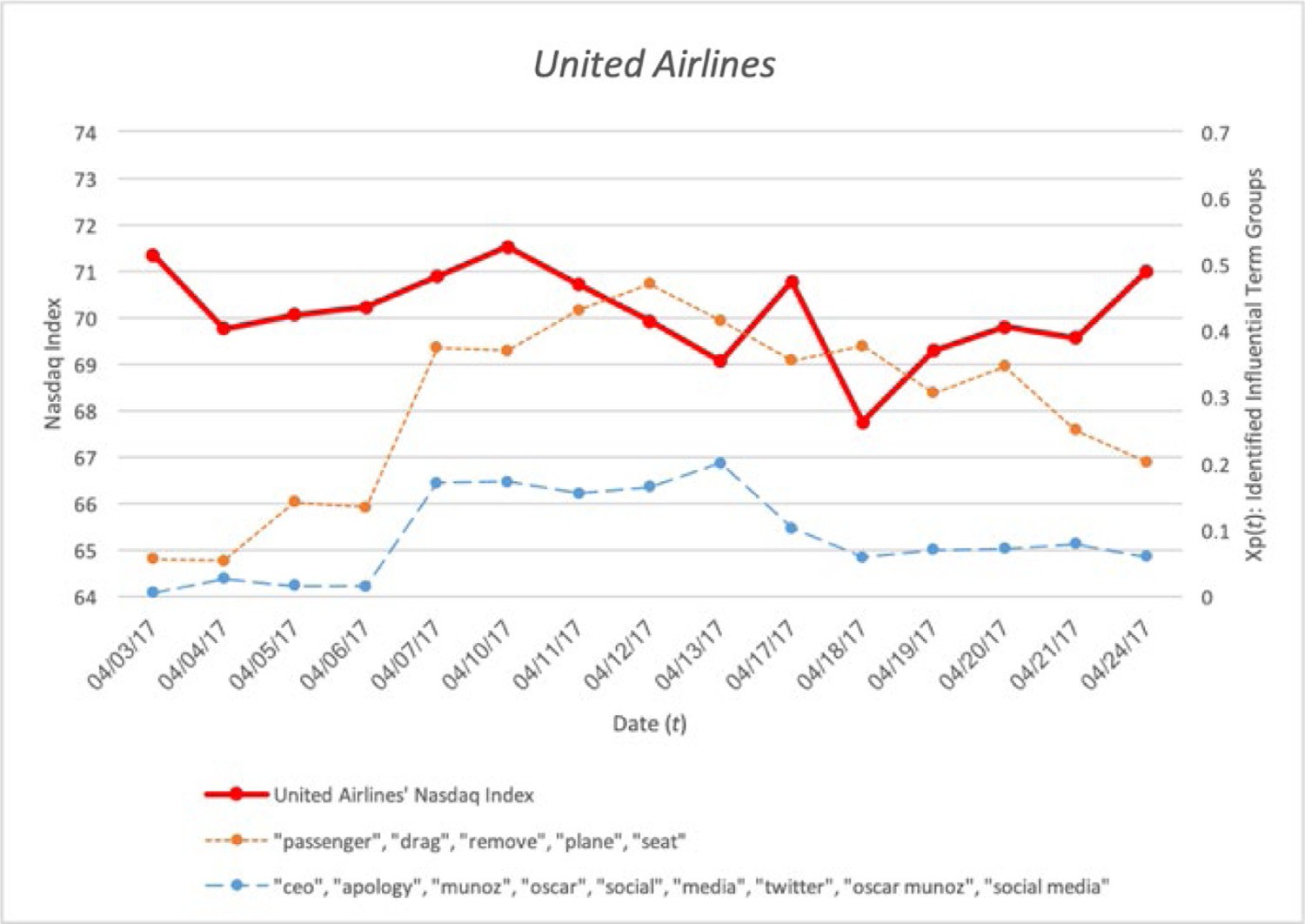
“photo”, “bathroom”, “refund”, “woman”

7 2 0 0.64 −

4 0 0 0.59 −

frequent enterprise-related terms are different for different appli- cations. They also indicate that most frequent *Netflix*-related terms and frequent *United Airlines*-related terms are unigrams (i.e., 33 out of 37 and 114 out of 116, respectively). It is postulated that tweets commonly consist of short messages ([Lim & Tucker, 2016](#_bookmark76)). The number of selected frequent *United Airlines*-related terms (i.e., 116) is more than three times greater than the number of selected frequent *Netflix*-related terms (i.e., 37). It is postulated that spe- cific events (e.g., the *United Express Flight 3411* incident in this

case study) lead to the use of certain terms (e.g., “drag”, “remove”, “forcible” in [Tables 2](#_bookmark30) and [A1](#_bookmark45)), but further analysis in necessary in the future. [Tables 2](#_bookmark30) and [3](#_bookmark37) illustrate that frequent bigrams and tri- grams and their subsequences are contained in the same group (e.g., “watch”, “bing”, and “bing watch” in [Table 2](#_bookmark30) and “munoz”, “oscar”, “social”, “media”, “oscar munoz”, and “social media” in [Table 3](#_bookmark37)). These results demonstrate the reason why only bigrams and trigrams listed in *Urban Dictionary* are considered and four- grams or larger sizes are disregarded in this research.



**Fig. 8.** Time series of *United Airlines*’ *Nasdaq* index as well as identified influential term groups.

**Table 4**

The results of sentiment analysis for *United Airlines* case study.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Case | First day | *T* | Influential term group | p-value for enterprise outcome (stock price) | *L* |
| 1 | April 3, 2017 | 15 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.03 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.04 | 2 |
| 2 | April 3, 2017 | 10 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.02 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.03 | 2 |
| 3 | April 3, 2017 | 20 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.03 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.04 | 2 |
| 4 | March 30, 2017 | 15 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.04 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.04 | 2 |
| 5 | March 30, 2017 | 10 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.04 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.03 | 2 |
| 6 | March 30, 2017 | 20 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.03 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.04 | 2 |
| 7 | April 5, 2017 | 15 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.02 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.03 | 2 |
| 8 | April 5, 2017 | 10 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.03 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.03 | 2 |
| 9 | April 5, 2017 | 20 | “passenger”, “drag”, “remove”, “plane”, “seat” | 0.04 | 2 |
|  |  |  | “ceo”, “apology”, “munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “social media” | 0.04 | 2 |

According to [Tables 2](#_bookmark30) and [3](#_bookmark37), the proposed expert and intelli- gent system, which considers term frequencies and co-occurrences, identifies three influential term groups that affect *Netflix*’s *Nas- daq* index over a long term and two influential term groups that affect *United Airlines’ Nasdaq* index in a short period. (The pro- posed expert and intelligent system identifies three influential term groups that affect *United Airlines*’ *Nasdaq* index when *T* = 10.) By contrast, term combinations selected by the random-keyword- sampling method (i.e., randomly sampled terms) are not causally associated with stock prices of both *Netflix* and *United Airlines*. It is concluded that considering term frequencies and co-occurrences is necessary to identify Twitter users’ feedback that has causal ef- fects on enterprise outcomes. [Tables 2](#_bookmark30) and [3](#_bookmark37) also show that only using overall Twitter users’ sentiment is not appropriate for pre- dicting future enterprise outcomes in this case study, because their *p*-values are greater than 0.05 (i.e., 0.31 in [Table 2](#_bookmark30) and 0.24 in [Table 3](#_bookmark37), respectively). However, further research should investigate

these phenomena in detail. It is postulated that keywords related to launching new series or episodes (e.g., “show”, “series”, “movie”, “season”, “episode”, “start”, “trailer” in [Table 2](#_bookmark30)) can be used to pre- dict *Netflix*’s future outcomes (i.e., stock prices in this case study). It is also postulated that terms related to the *United Express Flight 3411* incident (e.g., “passenger”, “drag”, “remove”, “plane”, “seat” in [Table 3](#_bookmark37)) and terms related to the *United Airlines* CEO’s follow- up action (e.g., “ceo”, “apology”, “oscar munoz”, “social media” in [Table 3](#_bookmark37)) had casual effects on *United Airlines*’ stock prices. On the other hand, randomly sampled terms (e.g., “kid”, “girl”, “ori- gin”, “stranger things” in [Table 2](#_bookmark30) and “fatigue”, “number”, “europe”, “exit”, “demo” in [Table 3](#_bookmark37)) are not appropriate for identifying Twit- ter users’ feedback that causes both *Netflix*’s and *United Airlines*’ outcomes (i.e., stock prices). For example, [Table 5](#_bookmark40) illustrates a ran- dom subsample of (1) tweets containing terms relating launching new series/episodes (i.e., “episode”, “trailer”, and “series”) and (2) tweets not containing those terms in order to identify whether or

**Table 5**

A random subsample of (1) tweets containing terms relating launching new series/episodes and (2) tweets not containing those terms.

Tweet Terms relating launching new

series/episodes

Whether or not the tweet is actually related to launching new series/episodes

New **episode** of orphan black out on **Netflix** and I couldn’t be happier “episode” Yes

**Netflix** announces ’The Little Prince’ release with beautifully moving **trailer** “trailer” Yes

Drew Barrymore to Star In Upcoming **Netflix** Comedy **Series**, Santa Clarita Diet, and I’m Glad “series” Yes

**Netflix** and Chill Type of Night – No

@dyxxxxxx Need to get **Netflix** – No

not terms in the influential term groups are indeed used in tweets relating to launching new series/episodes on *Netflix*.

On the one hand, [Fig. 5](#_bookmark34) shows that all three identified influ- ential term groups (i.e., Twitter users’ feedback) for *Netflix* posi- tively affect *Netflix*’s *Nasdaq* index (see [Table 2](#_bookmark30)). On the other hand, [Fig. 6](#_bookmark35) indicates that all two identified influential term groups for *United Airlines* negatively affect *United Airlines’ Nasdaq* index (see [Table 3](#_bookmark37)). [Tables 2](#_bookmark30) and [Fig. 5](#_bookmark34) illustrate that time lags between the time series of the identified influential term groups and the time series of *Netflix*’s outcomes have a value of 3 or 4 in this case study, which are also known as common values of the appropriate time lags of existing Granger causality models for future market anal- ysis ([Bollen et al., 2011; Thornton & Batten, 1985](#_bookmark66)). On the other hand, [Table 3](#_bookmark37) and [Fig. 6](#_bookmark35) show that time lags between the time se- ries of all two identified influential term groups and the time se- ries of *United Airlines*’ outcomes have a value of 2. It is postulated that time lags for a specific event(s) (e.g., the *United Express Flight 3411* incident in this case study) are relatively shorter than com- mon values of the appropriate time lags (i.e., 3 or 4), but further investigation is necessary.

[Table 4](#_bookmark39) illustrates that the first two influential term groups (i.e., “passenger”, “drag”, “remove”, “plane”, “seat” and “ceo”, “apology”,

“munoz”, “oscar”, “social”, “media”, “twitter”, “oscar munoz”, “so- cial media”) have causal relationships with *United Airlines*’ *Nasdaq* index and their appropriate time lags are 2 for all nine cases. These results indicate that different time windows and different window sizes do not significantly affect the causal relationship between the identified influential term groups and enterprise outcomes (i.e., *United Airlines*’ *Nasdaq* index).

## Conclusions and future work

The novel contribution of this research is to propose the expert and intelligent system that enables computers to identify Twitter users’ feedback having causal effects on real-world enterprise out- comes automatically. While existing expert and intelligent systems can be used to identify the spread of Twitter users’ enterprise- related feedback ([Asur & Huberman, 2010; Mao et al., 2012](#_bookmark53)), iden- tifying the causal relationships between Twitter users’ feedback and enterprise outcomes is limited. Twitter users’ feedback hav- ing causal effects, which is identified using the proposed expert and intelligent system, supports decision makers to improve future enterprise outcomes (e.g., future stock price values, market sales).

This paper is comprised of four main steps. First, textual and temporal information are extracted from Twitter data, and data preprocessing is exploited. Term groups that potentially affects en- terprise outcomes are identified through the co-occurrence net- work analysis model. Time series analysis models are implemented for analyzing (1) the trend of identified Twitter users’ feedback (i.e., Twitter users’ sentiment and term groups) and (2) the trend of real-world enterprise outcomes. Finally, the Granger causal anal- ysis model identifies Twitter users’ feedback that has causal ef- fects on enterprise outcomes. The Granger causality analysis model identifies the appropriate time lags between identified Twitter

users’ feedback (i.e., Twitter users’ sentiment or influential term groups) and enterprise outcomes as well.

Case studies involving Twitter data related to (1) a real-world internet video streaming and disc rental provider (i.e., *Netflix*) and

(2) an airline company (i.e., *United Airlines*) are used to test the va- lidity of this research. Experimental results present several insight- ful implications. While the proposed expert and intelligent system considers not only frequent unigrams but also frequent bigrams and trigrams for identifying influential term groups, experimen- tal results of two case studies illustrate that most frequent terms of influential term groups are unigrams, because Twitter messages typically consist of short messages. Experimental results of *United Airlines* case study illustrate that specific events (e.g., the *United Ex- press Flight 3411* incident) induce Twitter users to use certain terms (e.g., “drag”, “remove”, “forcible”). Experimental results also indi- cate that identified influential term groups can be used to predict enterprise future outcomes, since it is discovered that identified influential term groups (i.e., (1) terms related to launching new series or episodes of *Netflix* case study and (2) terms related to the *United Express Flight 3411* incident and the *United Airlines* CEO’s follow-up action of *United Airlines* case study) had causal effects on enterprises’ stock prices.

It is therefore concluded that the expert and intelligent system presented in this work, which considers term frequencies and co- occurrences, can be useful for identifying Twitter users’ feedback that causes future market success and for helping enterprise deci- sion makers improve future enterprise outcomes. It is postulated that the proposed expert and intelligent system can be applied to predict long-term future enterprise outcomes and to identify the effects of a specific event(s) on short-term future enterprise out- comes.

Although this research proposes the expert and intelligent sys- tem that enables computers to identify Twitter users’ feedback having causal relationships with enterprise outcomes automati- cally, there are still several ways to improve its performance as below:

* This research considers only overall Twitter users’ sentiment re- garding the company, but future work will also consider Twit- ter users’ sentiment for each influential term group separately, because Twitter users’ sentiment can be different for the differ- ent events regarding the same company (e.g., the *United Express Flight 3411* incident and the *United Airlines* CEO’s follow-up ac- tion).
* Future research will consider not only Twitter users’ sentiment and enterprise-related terms but also Twitter user information (e.g., gender, age, and posting frequency) in order to improve the performance of the proposed expert and intelligent system.
* Synonyms, which are not considered in this research, will be considered in future research, because there exist different terms with a similar meaning (e.g., “see” and “watch”), which can be included in the same influential term group.
* Topic models (e.g., latent Dirichlet allocation) can be applied for generating a co-occurrence graph *G* in order to improve the performance of the co-occurrence network analysis model.
* The authors will consider low-cost and real-time information sources for the proposed expert and intelligent system (e.g., Facebook, online forums, online news articles, online surveys), other than Twitter used in this research, because Twitter users’ feedback only represents some part of the opinions of people regarding the company.

## Conflict of interest

None.

## CRediT authorship contribution statement

**Sunghoon Lim:** Conceptualization, Methodology, Formal analy- sis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Conrad S. Tucker:** Con- ceptualization, Methodology, Validation, Writing - review & edit- ing, Supervision, Project administration, Funding acquisition.

## Acknowledgments

[This work is in part funded by the Air Force Oﬃce of Scien- tific Research (AFOSR) grant](https://doi.org/10.13039/100000181) [F](#_bookmark42)[A9550-18-1-0108 and the](https://doi.org/10.13039/100000181) [NSF](https://doi.org/10.13039/100000001) [I/UCRC](https://doi.org/10.13039/100000181) Center for Healthcare Organization Transformation (CHOT) award #[1624727](#_bookmark43). Any opinions, findings, or conclusions found in this pa- per are those of the authors and do not necessarily reflect the views of the sponsors. The authors of this work are grateful for help with programming and experimentation from Chonghan Lee. The authors would like to acknowledge Lawrence Lee and Punnavit Akkarapitakchai for data acquisition as well.

## Appendix A

[Tables A1](#_bookmark45) and [A2](#_bookmark46)

**Table A1**

The full results of 33 frequent unigrams (S1), 3 bigrams (S2), and 1 trigram (S3) from *Netflix*-related tweets (the total number of tweets containing the term “netflix”: 17,170,347 (i.e., *n*)).

Frequent

*Netflix*-related terms (*ki* )

Total number of tweets containing *ki* (*ni* )

*ni (*%*)*

*n*

Frequent

*Netflix*-related terms (*ki* )

Total number of tweets containing *ki* (*ni* )

*ni (*%*)*

*n*

Frequent

*Netflix*-related terms (*ki* )

Total number of tweets containing *ki* (*ni* )

*ni (*%*)*

*n*

Frequent

*Netflix*-related terms (*ki* )

Total number of tweets containing *ki* (*ni* )

“watch” “show” “series” “movie” “season” “chill” “need” “day” “tv” “time” “love”

2660,969 829,666 789,664 744,068 661,513 645,678 463,376 402,788 367,761 340,095 317,565

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 15.50 | 4.83 | 4.60 | 4.33 | 3.85 | 3.76 | 2.70 | 2.35 | 2.14 | 1.98 | 1.85 |
| “youtube” | “episode” | “night” | “start” | “bing” | “hulu” | “final” | “stream” | “origin” | “video” | “bed” |

255,250 237,726 226,130 217,020 209,527 205,471 197,702 190,729 187,719 179,639 178,132

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1.49 | 1.38 | 1.32 | 1.26 | 1.22 | 1.20 | 1.15 | 1.11 | 1.09 | 1.05 | 1.04 |
| “trailer” | “account” | “home” | “give” | “sleep” | “internet” | “shit” | “spotifi” | “recommend” | “call” | “week” |

154,320 142,683 135,612 130,248 121,486 120,615 116,930 106,839 106,322 100,286 96,194

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.90 | 0.83 | 0.79 | 0.76 | 0.71 | 0.70 | 0.68 | 0.62 | 0.62 | 0.58 | 0.56 |
| “bing watch” | “tv series” | “tv show” |  |  |  |  |  |  |  |  |

123,551 94,093 90,448

*ni (*%*)* 0.72 0.55 0.53

*n*

Frequent

*Netflix*-related terms (*ki* )

Total number of tweets containing *ki* (*ni* )

“netflix and chill”

394,129

*ni (*%*)* 2.30

*n*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Frequent *United* | “flight” | “passenger” | “drag” | “overbook” | “remove” | “man” | “video” | “plane” | “ceo” | “customer” | “forcible” |
| *Airlines*-related terms (*ki* ) | | | | | | | | | | | |
| Total number of tweets 18,124 | | 17,591 | 12,980 | 10,830 | 8508 | 7818 | 7783 | 7543 | 6797 | 6135 | 5177 |
| containing *ki* (*ni* )  *ni (*%*)* 12.92 | | 12.54 | 9.25 | 7.72 | 6.06 | 5.57 | 5.55 | 5.38 | 4.85 | 4.37 | 3.69 |
| Frequent *United* “seat” | | “incident” | “people” | “doctor” | “go” | “news” | “apology” | “say” | “beat” | “time” | “pepsi” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 4438 | | 4016 | 3869 | 3718 | 2994 | 2991 | 2898 | 2890 | 2887 | 2850 | 2832 |
| containing *ki* (*ni* )  *ni (*%*)* 3.16 | | 2.86 | 2.76 | 2.65 | 2.13 | 4.85 | 4.37 | 3.69 | 3.16 | 2.86 | 2.76 |
| Frequent *United* “need” | | “munoz” | “oscar” | “shame” | “pay” | “boycott” | “trump” | “show” | “think” | “way” | “youtube” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 2705 | | 2662 | 2493 | 2431 | 2417 | 2408 | 2260 | 2178 | 2177 | 2096 | 2033 |
| containing *ki* (*ni* )  *ni (*%*)* 2.65 | | 2.13 | 1.78 | 1.73 | 1.72 | 1.72 | 1.61 | 1.55 | 1.55 | 1.49 | 1.45 |
| Frequent *United* “disgust” | | “happen” | “treat” | “hope” | “stock” | “use” | “give” | “see” | “police” | “bad” | “travel” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 2021 | | 2001 | 1992 | 1982 | 1975 | 1894 | 1882 | 1880 | 1866 | 1856 | 1769 |
| containing *ki* (*ni* )  *ni (*%*)* 1.44 | | 1.43 | 1.42 | 1.41 | 1.41 | 1.35 | 1.34 | 1.34 | 1.33 | 1.32 | 1.26 |
| Frequent *United* “book” | | “look” | “today” | “know” | “thing” | “watch” | “want” | “sky” | “day” | “social” | “assault” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 1684 | | 1677 | 1677 | 1653 | 1619 | 1606 | 1534 | 1520 | 1494 | 1479 | 1477 |
| containing *ki* (*ni* )  *ni (*%*)* 1.20 | | 1.20 | 1.20 | 1.18 | 1.15 | 1.14 | 1.09 | 1.08 | 1.06 | 1.05 | 1.05 |
| Frequent *United* “media” | | “twitter” | “friend” | “guy” | “ticket” | “dao” | “come” | “call” | “resign” | “crisis” | “kick” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 1439 | | 1421 | 1439 | 1403 | 1377 | 1373 | 1367 | 1319 | 1278 | 1262 | 1257 |
| containing *ki* (*ni* )  *ni (*%*)* 1.03 | | 1.01 | 1.03 | 1.00 | 0.98 | 0.98 | 0.97 | 0.94 | 0.91 | 0.90 | 0.90 |
| Frequent *United* “market” | | “shit” | “week” | “offer” | “year” | “fire” | “staff” | “chicago” | “board” | “boss” | “sue” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 1215 | | 1179 | 1158 | 1142 | 1138 | 1131 | 1120 | 1108 | 1088 | 1086 | 1074 |
| containing *ki* (*ni* )  *ni (*%*)* 0.87 | | 0.84 | 0.83 | 0.81 | 0.81 | 0.81 | 0.80 | 0.79 | 0.78 | 0.77 | 0.77 |
| Frequent *United* “david” | | “asian” | “fuck” | “ask” | “face” | “class” | “thank” | “work” | “fight” | “drop” | “troll” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 1073 | | 1054 | 1050 | 1045 | 1034 | 1032 | 1007 | 990 | 982 | 963 | 944 |
| containing *ki* (*ni* )  *ni (*%*)* 0.76 | | 0.75 | 0.75 | 0.74 | 0.74 | 0.74 | 0.72 | 0.71 | 0.70 | 0.69 | 0.67 |
| Frequent *United* “train” | | “problem” | “help” | “club” | “world” | “slogan” | “stop” | “poor” | “sure” | “reason” | “flgiht3411” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 934 | | 907 | 895 | 893 | 884 | 868 | 867 | 852 | 852 | 847 | 841 |
| containing *ki* (*ni* )  *ni (*%*)* 0.67 | | 0.65 | 0.64 | 0.64 | 0.63 | 0.62 | 0.62 | 0.61 | 0.61 | 0.60 | 0.60 |
| Frequent *United* “public” | | “avoid” | “air” | “leg” | “hear” | “rule” | “love” | “may” | “hospital” | “delta” | “post” |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 813 | | 791 | 785 | 782 | 777 | 765 | 764 | 752 | 744 | 742 | 738 |
| containing *ki* (*ni* )  *ni (*%*)* 0.58 | | 0.56 | 0.56 | 0.56 | 0.55 | 0.55 | 0.54 | 0.54 | 0.53 | 0.53 | 0.53 |
| Frequent *United* “free” | | “tweet” | “internet” | “china” |  |  |  |  |  |  |  |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 731 | | 726 | 715 | 705 |  |  |  |  |  |  |  |
| containing *ki* (*ni* )  *ni (*%*)* 0.52 | | 0.52 | 0.51 | 0.50 |  |  |  |  |  |  |  |
| Frequent *United* “oscar munoz” | | “social media” |  |  |  |  |  |  |  |  |  |
| *Airlines*-related terms | |  |  |  |  |  |  |  |  |  |  |
| (*ki* )  Total number of tweets 2303 | | 1008 |  |  |  |  |  |  |  |  |  |
| containing *ki* (*ni* )  *ni (*%*)* 1.64 | | 0.72 |  |  |  |  |  |  |  |  |  |

**Table A2**

The full results of frequent unigrams (S1), bigrams (S2), and trigrams (S3) from *United Airlines*-related tweets (the total number of tweets containing the term “united airlines”, “united airline”, “unitedairlines”, “unitedairline”, or “ua”: 140,286 (i.e., *n*)).

*n*

*n*

*n*

*n*

*n*

*n*

*n*

*n*

*n*

*n*

*n*

*n*

## References

Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A., & Zhang, Z. (2013). What’s buzzing in the blizzard of buzz? Automotive component isolation in social media post- [ings. *Decision Support Systems, 55*(4), 871–882. https://doi.org/10.1016/j.dss.2012. 12.023.](https://doi.org/10.1016/j.dss.2012.12.023)

Abrahams, A. S., Jiao, J., Wang, G. A., & Fan, W. (2012). Vehicle defect discovery from [social media. *Decision Support Systems, 54*(1), 87–97. https://doi.org/10.1016/j. dss.2012.04.005.](https://doi.org/10.1016/j.dss.2012.04.005)

[Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical* *Science, 10*, 364–376.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0003)

Apache Lucene. (2010). Retrieved from <http://lucene.apache.org/>.

Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. In *Web* *intelligence and intelligent agent technology (WI-IAT), 2010 IEEE/WIC/ACM interna-* [*tional conference on, 1,* Retrieved from (pp. 492–499). http://ieeexplore.ieee.org/ abstract/document/5616710/.](http://ieeexplore.ieee.org/abstract/document/5616710/)

[Atkinson, K. (2017). *Spell checker oriented word lists,* Retrieved from http://wordlist. aspell.net.](http://wordlist.aspell.net/)

Bae, Y., & Lee, H. (2012). Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers. *Journal of the American So-* [*ciety for Information Science and Technology, 63*(12), 2521–2535. https://doi.org/ 10.1002/asi.22768.](https://doi.org/10.1002/asi.22768)

Bao, T., & Chang, T. S. (2014). Why amazon uses both the New York Times Best Seller List and customer reviews: An empirical study of multiplier effects on product sales from multiple earned media. *Decision Support Systems, 67*, 1–8. <https://doi.org/10.1016/j.dss.2014.07.004>.

Bodnar, T., Dering, M. L., Tucker, C., & Hopkinson, K. M. (2016). Using large-scale so- cial media networks as a scalable sensing system for modeling real-time energy utilization patterns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 1–14. <https://doi.org/10.1109/TSMC.2016.2618860>.

Bogle, S. A., & Potter, W. D. (2015). SentAMaL-A sentiment analysis machine learning stock predictive model. In *Proceedings on the international conference on artificial* [*intelligence (ICAI),* Retrieved from (p. 610). http://search.proquest.com/openview/ 966827dbd4e94e543dbdfe642aed4536/1?pq-origsite=gscholar.](http://search.proquest.com/openview/966827dbd4e94e543dbdfe642aed4536/1?pq-origsite=gscholar)

[Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0010)

[Bosch, H., Thom, D., Heimerl, F., Püttmann, E., Koch, S., Krüger, R., et al. (2013). Scat- terblogs2: Real-time monitoring of microblog messages through user-guided filtering. *IEEE Transactions on Visualization and Computer Graphics, 19*(12), 2022–2031.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0011)

[Bothos, E., Apostolou, D., & Mentzas, G. (2010). Using social media to predict future events with agent-based markets. *IEEE Intelligent Systems, 25*(6), 50–58.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0012)

Bruhn, M., Schoenmueller, V., & Schäfer, D. B. (2012). Are social media replacing tra- ditional media in terms of brand equity creation. *Management Research Review, 35*(9), 770–790. <https://doi.org/10.1108/01409171211255948>.

Cataldi, M., Di Caro, L., & Schifanella, C. (2010). Emerging topic detection on twitter based on temporal and social terms evaluation. In *Proceedings of the tenth inter-* [*national workshop on multimedia data mining: 4* Retrieved from http://dl.acm. org/citation.cfm?id=1814249.](http://dl.acm.org/citation.cfm?id=1814249)

Chang, H.-L., Chou, Y.-C., Wu, D.-Y., & Wu, S.-C. (2017). Will firm’s marketing efforts on owned social media payoff? A quasi-experimental analysis of tourism prod- ucts. *Decision Support Systems*. <https://doi.org/10.1016/j.dss.2017.12.011>.

Checkley, M. S., Higón, D. A., & Alles, H. (2017). The hasty wisdom of the mob: How market sentiment predicts stock market behavior. *Expert Systems with Applica- tions, 77*, 256–263. <https://doi.org/10.1016/j.eswa.2017.01.029>.

[Culnan, M. J., McHugh, P. J., & Zubillaga, J. I. (2010). How large US companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive, 9*(4).](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0017)

Daniel, M., Neves, R. F., & Horta, N. (2017). Company event popularity for financial markets using Twitter and sentiment analysis. *Expert Systems with Applications, 71*, 111–124. <https://doi.org/10.1016/j.eswa.2016.11.022>.

[Davidov, D., Tsur, O., & Rappoport, A. (2010). Enhanced sentiment learning using twitter hashtags and smileys. *Proceedings of the 23rd international conference on*](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0019)[*computational linguistics: Posters*. Association for Computational Linguistics. (pp. 241–249).](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0019)

[Diebold, F. X.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0020) [(2001). *Elements of forecasting South-Western*. Cincinnati: Ohio](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0020).

Ding, C., Cheng, H. K., Duan, Y., & Jin, Y. (2017). The power of the “like” button: The impact of social media on box oﬃce. *Decision Support Systems, 94*, 77–84. <https://doi.org/10.1016/j.dss.2016.11.002>.

[Eichler, M. (2012). Causal inference in time series analysis. *Causality: Statistical Per- spectives and Applications*, 327–354.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0022)

[Fourt, L. A., & Woodlock, J. W. (1960). Early prediction of market success for new grocery products. *The Journal of Marketing*, 31–38.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0023)

Fox, C. (1989). A stop list for general text. *ACM SIGIR Forum, 24*, 19–21. Retrieved from <http://dl.acm.org/citation.cfm?id=378888>.

[Fürnkranz, J. (1998). A study using n-gram features for text categorization. *Austrian Research Institute for Artificial Intelligence, 3*(1998), 1–10.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0025)

Gelper, S., Lemmens, A., & Croux, C. (2007). Consumer sentiment and consumer spending: Decomposing the Granger causal relationship in the time domain. *Ap- plied Economics, 39*(1), 1–11. <https://doi.org/10.1080/00036840500427791>.

Gerber, M. S. (2014). Predicting crime using twitter and kernel density estimation.

*Decision Support Systems, 61*, 115–125. <https://doi.org/10.1016/j.dss.2014.02.003>.

Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Ex-* [*pert Systems with Applications, 40*(16), 6266–6282. https://doi.org/10.1016/j.eswa. 2013.05.057.](https://doi.org/10.1016/j.eswa.2013.05.057)

[Good,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0029) [I. J. (1961a). A causal calculus (I). *The British Journal for the Philosophy of Sci- ence, 11*(44), 305–318.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0029)

[Good,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0030) [I. J. (1961b). A causal calculus (II). *The British Journal for the Philosophy of Science, 11*(45), 43–51.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0030)

[Gourieroux, C., & Jasiak, J. (2001). *Financial econometrics*. Oxford: Princeton Univer- sity Press.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0031)

[Gourieroux, C., & Monfort, A. (1997). *Time series and dynamic models (Vol. 3)*. Cam- bridge: Cambridge University Press.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0031a)

[Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424–438.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0032)

Greer, C. F., & Ferguson, D. A. (2011). Using twitter for promotion and branding: A content analysis of local television twitter sites. *Journal of Broadcasting & Elec- tronic Media, 55*(2), 198–214. <https://doi.org/10.1080/08838151.2011.570824>.

[Hamilton, J. D.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0034) [(1994). *Time series analysis*. Princeton: Princeton university press](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0034).

Ho, C.-S., Damien, P., Gu, B., & Konana, P. (2017). The time-varying nature of social media sentiments in modeling stock returns. *Decision Support Systems, 101*, 69–

81. <https://doi.org/10.1016/j.dss.2017.06.001>.

Holzinger, W., Krüpl, B., & Herzog, M. (2006). Using ontologies for extracting prod- uct features from web pages. In *International semantic web conference,* Retrieved from (pp. 286–299). <http://link.springer.com/chapter/10.1007/11926078_21>.

Hou, K., & Moskowitz, T. J. (2005). Market Frictions, price delay, and the cross- [section of expected returns. *Review of Financial Studies, 18*(3), 981–1020. https:](https://doi.org/10.1093/rfs/hhi023)

[//doi.org/10.1093/rfs/hhi023.](https://doi.org/10.1093/rfs/hhi023)

Ikeda, K., Hattori, G., Ono, C., Asoh, H., & Higashino, T. (2013). Twitter user profil- ing based on text and community mining for market analysis. *Knowledge-Based Systems, 51*, 35–47. <https://doi.org/10.1016/j.knosys.2013.06.020>.

Imbert, F., & Thomas, L. (2017). *United shares slide as PR nightmare catches up* [*with investors* April 11. CNBC Retrieved from http://www.cnbc.com/2017/04/11/ pr-nightmare-catches- up-with-united-investors-as-shares-slide.html.](http://www.cnbc.com/2017/04/11/pr-nightmare-catches-up-with-united-investors-as-shares-slide.html)

[Jackson, P. (1998). *Introduction to expert systems*. Boston: Addison-Wesley Longman Publishing Co., Inc.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0040)

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and [opportunities of social media. *Business Horizons, 53*(1), 59–68. https://doi.org/10. 1016/j.bushor.2009.09.003.](https://doi.org/10.1016/j.bushor.2009.09.003)

[Kaz´mierski, T. J., & Morawiec, A. (2011). System specification and design languages: Selected contributions from FDL 2010 *(Vol. 106)*. New York: Springer Science & Business Media.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0042)

Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applica- tions, 41*(16), 7653–7670. <https://doi.org/10.1016/j.eswa.2014.06.009>.

Lee, T. K., Cho, J. H., Kwon, D. S., & Sohn, S. Y. (2019). Global stock market investment strategies based on financial network indicators using machine learning tech- [niques. *Expert Systems with Applications, 117*, 228–242. https://doi.org/10.1016/j. eswa.2018.09.005.](https://doi.org/10.1016/j.eswa.2018.09.005)

Li, Y.-M., Chen, H.-M., Liou, J.-H., & Lin, L.-F. (2014). Creating social intelligence for [product portfolio design. *Decision Support Systems, 66*, 123–134. https://doi.org/ 10.1016/j.dss.2014.06.013.](https://doi.org/10.1016/j.dss.2014.06.013)

[Liew, V. K.-S.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0046) [(2004). *Which lag length selection criteria should we employ?*](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0046).

[Lim,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0047) [S., & Tucker, C. S. (2016). A bayesian sampling method for product feature extraction from large-scale textual data. *Journal of Mechanical Design, 138*(6), 061403.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0047)

Lim, S., & Tucker, C. S. (2017). Mitigating online product rating biases through the discovery of optimistic, pessimistic, and realistic reviewers. *Journal of Mechanical Design, 139*(11), 111409. <https://doi.org/10.1115/1.4037612>.

[Lim,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0049) [S., Tucker, C. S., Jablokow, K., & Pursel, B. (2018). A semantic network model for measuring engagement and performance in MOOCs. *Computer Applications in Engineering Education, 26* ***(5)***, 1481–1492 https://doi.org/10.1002/cae.22033.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0049)

Lim, S., Tucker, C. S., & Kumara, S. (2017). An unsupervised machine learning model for discovering latent infectious diseases using social media data. *Journal of Biomedical Informatics, 66*, 82–94. <https://doi.org/10.1016/j.jbi.2016.12.007>.

Liu, L., Wu, J., Li, P., & Li, Q. (2015). A social-media-based approach to predicting [stock comovement. *Expert Systems with Applications, 42*(8), 3893–3901. https:](https://doi.org/10.1016/j.eswa.2014.12.049)

[//doi.org/10.1016/j.eswa.2014.12.049.](https://doi.org/10.1016/j.eswa.2014.12.049)

[Liu,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0052) [Y., & Bahadori, M. T. (2012). *A survey on granger causality: A computational view*.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0052)

[University of Southern California.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0052)

Luo, B., Zeng, J., & Duan, J. (2016). Emotion space model for classifying opinions [in stock message board. *Expert Systems with Applications, 44*, 138–146. https:](https://doi.org/10.1016/j.eswa.2015.08.023)

[//doi.org/10.1016/j.eswa.2015.08.023.](https://doi.org/10.1016/j.eswa.2015.08.023)

[Makrehchi, M., Shah, S., & Liao, W. (November, 2013). *Stock prediction using even- t-based sentiment analysis* (pp. 337–342).](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0054)

[Mao,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0055) [Y., Wei, W., Wang, B., & Liu, B. (2012). Correlating S&P 500 stocks with Twitter data. *Proceedings of the first ACM international workshop on hot topics on inter- disciplinary social networks research*. ACM.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0055) [(pp. 69–72).](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0055)

Marwick, A. E., & boyd, d (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society, 13*(1), 114–133. <https://doi.org/10.1177/1461444810365313>.

Meire, M., Ballings, M., & Van den Poel, D. (2017). The added value of social me- dia data in B2B customer acquisition systems: A real-life experiment. *Decision Support Systems, 104*, 26–37. <https://doi.org/10.1016/j.dss.2017.09.010>.

Mostafa, M. M. (2013). More than words: Social networks’ text mining for consumer [brand sentiments. *Expert Systems with Applications, 40*(10), 4241–4251. https:// doi.org/10.1016/j.eswa.2013.01.019.](https://doi.org/10.1016/j.eswa.2013.01.019)

Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications, 42*(24), 9603–9611. <https://doi.org/10.1016/j.eswa.2015.07.052>.

Nguyen, T., Larsen, M. E., O’Dea, B., Nguyen, D. T., Yearwood, J., Phung, D., et al. (2017). Kernel-based features for predicting population health indices from [geocoded social media data. *Decision Support Systems, 102*, 22–31. https://doi. org/10.1016/j.dss.2017.06.010.](https://doi.org/10.1016/j.dss.2017.06.010)

Nikfarjam, A., Sarker, A., O’Connor, K., Ginn, R., & Gonzalez, G. (2015). Pharma- covigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the Ameri-* *can Medical Informatics Association*. <https://doi.org/10.1093/jamia/ocu041>.

Ohlheiser, A. (April 11, 2017). The full timeline of how social media turned United into the biggest story in the country. *The Washington Post*. Retrieved [from https://www.washingtonpost.com/news/the-intersect/wp/2017/04/11/ the-full-timeline-of- how- social-media-turned-united-into- the-biggest-story-in](https://www.washingtonpost.com/news/the-intersect/wp/2017/04/11/the-full-timeline-of-how-social-media-turned-united-into-the-biggest-story-in-the-country/?utm_term=.f76760ebaec9)

[-the- country/?utm\_term=.f76760ebaec9.](https://www.washingtonpost.com/news/the-intersect/wp/2017/04/11/the-full-timeline-of-how-social-media-turned-united-into-the-biggest-story-in-the-country/?utm_term=.f76760ebaec9)

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using twitter to predict returns, volatility, trading vol- ume and survey sentiment indices. *Expert Systems with Applications, 73*, 125–

144. <https://doi.org/10.1016/j.eswa.2016.12.036>.

[Paul, A., Agrawal, A., Liao, W., & Choudhary, A. (2016). AnonyMine: Mining anony- mous social media posts using psycho-lingual and crowd-sourced dictionaries. *22nd ACM SIGKDD workshop on sentiment discovery and opinion mining*.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0064)

[Peckham, A. (2009). *Urban dictionary: Fularious street slang defined*. Andrews McMeel Publishing.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0065)

Peleja, F., Santos, J., & Magalhães, J. (2014). Reputation analysis with a ranked sentiment-lexicon. In *Proceedings of the 37th international ACM SIGIR conference* *on research & development in information retrieval - SIGIR ’14* (pp. 1207–1210). <https://doi.org/10.1145/2600428.2609546>.

Porter, M. F. (2006). An algorithm for suﬃx stripping. *Program, 40*(3), 211–218. <https://doi.org/10.1108/00330330610681286>.

Ranco, G., Aleksovski, D., Caldarelli, G., Grcˇar, M., & Mozeticˇ, I. (2015). The effects [of Twitter sentiment on stock price returns. *PloS One, 10*(9), e0138441. https:](https://doi.org/10.1371/journal.pone.0138441)

[//doi.org/10.1371/journal.pone.0138441.](https://doi.org/10.1371/journal.pone.0138441)

Rosas-Romero, R., Díaz-Torres, A., & Etcheverry, G. (2016). Forecasting of stock re- turn prices with sparse representation of financial time series over redundant [dictionaries. *Expert Systems with Applications, 57*, 37–48. https://doi.org/10.1016/ j.eswa.2016.03.021.](https://doi.org/10.1016/j.eswa.2016.03.021)

Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect [of tweets on movie sales. *Decision Support Systems, 55*(4), 863–870. https://doi. org/10.1016/j.dss.2012.12.022.](https://doi.org/10.1016/j.dss.2012.12.022)

[Russell, M. A. (2013). *Mining the social web: Data mining Facebook, Twitter, LinkedIn, Google*](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0071)+[*, GitHub, and more*. O’Reilly Media, Inc.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0071)

Ruthven, I., & Lalmas, M. (2003). A survey on the use of relevance feedback for information access systems. *The Knowledge Engineering Review, 18*(2), 95–145. <https://doi.org/10.1017/S0269888903000638>.

[Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., & Deng, X. (2013). Exploiting topic based twitter sentiment for stock prediction. In Proceedings of the 51st annual meet- ing of the association for computational linguistics (volume 2: Short papers), 2 (pp. 24–29).](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0073)

Smailovic´, J., Grcˇar, M., Lavracˇ, N., & Žnidaršicˇ, M. (2013). Predictive sentiment anal- ysis of Tweets: A stock market application. In A. Holzinger, & G. Pasi (Eds.). In Human-computer interaction and knowledge discovery in complex, unstruc- tured, big data*: 7947* (pp. 77–88). [https://doi.org/10.1007/978- 3- 642-39146- 0\_8](https://doi.org/10.1007/978-3-642-39146-0_8).

Strickl, D. (May 30, 2013). *Microsoft is paying attention to neg- ative feedback centered on Xbox One’s DRM* Retrieved January [28, 2016, from VR-Zone website Http://vr-zone.com/articles/ microsoft-is-paying-attention-to-negative-feedback-centered- on- xbox- ones- drm/32710.html.](http://Http//vr-zone.com/articles/microsoft-is-paying-attention-to-negative-feedback-centered-on-xbox-ones-drm/32710.html)

Stringam, B. B., & Gerdes, J. (2010). An analysis of Word-of-Mouse ratings and guest comments of online hotel distribution sites. *Journal of Hospitality Marketing & Management, 19*(7), 773–796. <https://doi.org/10.1080/19368623.2010.508009>.

[Suppes,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0077) [P. (1970). *A probabilistic theory of causality*. Amsterdam: North-Holland Pub- lishing Company.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0077)

Symeonidis, S., Effrosynidis, D., & Arampatzis, A. (2018). A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analy- [sis. *Expert Systems with Applications, 110*, 298–310. https://doi.org/10.1016/j.eswa. 2018.06.022.](https://doi.org/10.1016/j.eswa.2018.06.022)

Taylor, K. (August 2, 2018). IHOP’s burger sales quadrupled after its contro- [versial IHOb name change. *Business Insider*. Retrieved from https://www. businessinsider.com/ihop-burger-sales-grew-with-ihob-name-change-2018- 8.](https://www.businessinsider.com/ihop-burger-sales-grew-with-ihob-name-change-2018-8)

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for In-* [*formation Science and Technology, 61*(12), 2544–2558. https://doi.org/10.1002/asi. 21416.](https://doi.org/10.1002/asi.21416)

Thompson, L., Rivara, F. P., & Whitehill, J. M. (2015). Prevalence of Marijuana-related traﬃc on Twitter, 2012–2013: A content analysis. *Cyberpsychology, Behavior, and Social Networking, 18*(6), 311–319. <https://doi.org/10.1089/cyber.2014.0620>.

[Thornton, D. L., & Batten, D. S. (1985). Lag-length selection and tests of Granger causality between money and income. *Journal of Money, Credit and Banking, 17*(2), 164–178.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0082)

Tuarob, S., & Tucker, C. S. (2013). Fad or here to stay: Predicting product market adoption and longevity using large scale, social media data. *ASME 2013 inter- national design engineering technical conferences and computers and information* [*in engineering conference* V02BT02A012–V02BT02A012. Retrieved from http:// proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=1830257.](http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=1830257)

[Tuarob, S., & Tucker, C. S. (2015). Automated discovery of lead users and latent prod- uct features by mining large scale social media networks. *Journal of Mechanical Design, 137*(7), 071402.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0084)

Vasquez, J. (February 22, 2018). *In one tweet, Kylie Jenner wiped out*

*$1.3 billion of snap’s market value* Retrieved from Bloomberg website [https://www.bloomberg.com/news/articles/2018-02- 22/snap- royalty-kylie- jenner-erased-a-billion-dollars-in-one-tweet.](https://www.bloomberg.com/news/articles/2018-02-22/snap-royalty-kylie-jenner-erased-a-billion-dollars-in-one-tweet)

Victor, D., & Stevens, M. (April 10, 2017). *United airlines passenger is dragged from* [*an overbooked flight*. The New York Times Retrieved from https://www.nytimes. com/2017/04/10/business/united- flight-passenger-dragged.html.](https://www.nytimes.com/2017/04/10/business/united-flight-passenger-dragged.html)

Wang, L., Youn, B. D., Azarm, S., & Kannan, P. K. (2011). Customer-driven prod- uct design selection using web based user-generated content. In *ASME 2011 in- ternational design engineering technical conferences and computers and informa-* [*tion in engineering conference,* Retrieved from (pp. 405–419). http://proceedings. asmedigitalcollection.asme.org/proceeding.aspx?articleid=1640611.](http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=1640611)

Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short- term stock prices using ensemble methods and online data sources. *Expert Sys- tems with Applications, 112*, 258–273. <https://doi.org/10.1016/j.eswa.2018.06.016>.

Wu, L., Morstatter, F., & Liu, H. (2016). SlangSD: Building and using a sentiment dictionary of slang words for short-text sentiment classification. Retrieved from <http://arxiv.org/abs/1608.05129>.

Yan, Z., Xing, M., Zhang, D., & Ma, B. (2015). EXPRS: An extended pagerank method for product feature extraction from online consumer reviews. *Information & Management, 52*(7), 850–858. <https://doi.org/10.1016/j.im.2015.02.002>.

Zhang, J. (2015). Voluntary information disclosure on social media. *Decision Support Systems, 73*, 28–36. <https://doi.org/10.1016/j.dss.2015.02.018>.

[Zhou,](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0091) [F., Jiao, R. J., & Linsey, J. S. (2015). Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews. *Journal of Mechanical Design, 137*(7), 071401.](http://refhub.elsevier.com/S2590-1885(19)30007-1/sbref0091)