Article

Monitoring Events of Market Competitors: A Text Mining Method for Analyzing Massive Firm-Generated Social Media

Hua Yuan [1], Wenjun Deng [1], Baojun Ma [2,*] and Yu Qian [1]

1   School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 611731, China
2   Key Laboratory of Brain-Machine Intelligence for Information Behavior (Ministry of Education and Shanghai), School of Business and Management, Shanghai International Studies University, Shanghai 201620, China
*   Correspondence: mabaojun@shisu.edu.cn

Abstract: Understanding how competitors act in a market is a critical component of strategic decision-making. In this paper, we propose a method to extract firm events from the textual content generated by firms in the market and explore the competitive relationships among firms based on the spatiotemporal homogeneity of events of different firms. To this end, we first introduce experts to define a series of business events based on the content of corporate-generated texts; then, we propose algorithms to extract and enrich the feature words (triggers) of these business events to form better event classifiers. We subsequently use these classifiers to identify the business events recorded in all online texts published by companies. Finally, based on these results, we can obtain a sequence of activities/events for each firm in the market, which can be used to identify the evolutionary patterns of firms' behavior in the market, as well as their potential competitive relationships. Considering that competition between companies in the market appears to be continuous at the strategic level, but the implementation of competitive behavior is expressed through their "events" in the market, identifying whether companies are "competing" in the market requires timely observation of the information about "events" in the market. However, obtaining accurate market information is complex and costly. Therefore, this study provides a way to bridge the gap between social media data and market competition "events".

Keywords: social media; firm-generated content; text mining; business event; firm–event matrix

1. Introduction

Understanding market competition is important for firms, consumers of products, and investors [1]. In the literature on management, firms have adopted two main streams of approaches to competitor identification, i.e., the supply-based [2,3] and the demand-based approaches [2]. The former is based on the firms' attributes to divide the competitor, while the latter is based on the customers' attributes. The growing complexity of organization and market structures in modern industries brings new issues for managers in identifying competition; for example, it has been found that not all firms will take the same competitive action simultaneously [4]. Thus, another new research stream has been developed to identify competitive relationships between firms by judging the similarity of their activities in the market [2,5]. In this approach, the term competitors refers to a set of firms in a market that share similar behaviors/activities [4,6]. However, previous methods rely on the data collected from traditional sources, such as market research [7], which may be outdated and time-consuming to analyze manually [2,8]. Particularly, such data may lead to "managerial myopia" in identifying competitive relationships [8]. Therefore, a critical step in analyzing market competition is to acquire and analyze relevant data that can effectively facilitate scanning the activities of firms. In this study, we obtain data publicly posted by companies on social media and introduce text analysis techniques to

extract events from them, thus providing timely observation data for identifying potential competition between companies.

With the rapid development of Internet technology, social media has been widely adopted in business environments [9]. For example, firms can use social media to deliver information [10], communicate with followers [11], and build relationships with consumers [12] and other organizations [13]. All the messages posted by firms on their official social media pages are firm-generated content [14]. As a result, microblog platforms, such as Twitter.com and Weibo.com, have provided abundant and timely firm-generated content. These contents are usually related to firms' business events [15], recording a series of events that happened in firms, and showing great value in the form of business strategies [16]. Therefore, it is a feasible way to collect data from social media and extract further information about business events in a market [17].

However, extracting business event information from massive microblog data presents several challenges [18]. First, we know that retrieving business events from massive online texts is a computer-aided computing task; however, there is no formal or precise definition of business events in the field of text information retrieval [19]. Second, the textual data crawled from websites are very noisy, and terms may be varied and ambiguous [20]. These problems make it necessary for us to present a well-designed approach to explore event information from natural language [21]. Along this line, the following questions thus should be carefully addressed:

RQ1: How can we identify a business event from massive online text data?

RQ2: How can we explore the event revolution of competitors in a market?

In this paper, we focus on competition at the level of the market behavior of firms, i.e., firms take similar market actions in a defined market, targeting the same users at the same time, in what is called competition (events). To this end, we present a framework to extract business events from massive microblog texts, and then, based on the extracted information, we propose a novel method to explore the event information among competitors in a market. Our contribution to the existing literature is threefold. First, we propose a method for extracting business events from massive firm-generated text. The method consists of three tasks: expert annotation of events in the seed text, feature word extraction, and feature word enrichment, so that we can build a better event classifier to recognize the events recorded in the text. Second, we propose a method to follow the evolution of various events within a particular company, which can be used to observe the behavior of companies in the market from the perspective of time. Third, by defining competitors in a market as a group of firms with similar actions/activities [4], we propose a simple and effective method to explore the competition between companies at the level of market event.

The paper is organized as follows. Section 2 gives an overview of the literature on business event detection and market intelligence with social media. Section 3 details the methods for collecting data, extracting business events, and exploring market competition. Section 4 summarizes the results. Section 5 identifies the managerial implications and the limitations. Section 6 summarizes this work.

2. Literature Review

2.1. Social-Media-Based Market Intelligence

One early work in market intelligence (MI) is the BrandPlus platform [22], which explores consumers' buzzwords about brands and companies. An automatic method by Netzer et al. [23] identified which brands are discussed in consumer forums for the markets of sedan cars and diabetes drugs. Xu et al. [24] focused on extracting comparative relationships from Amazon customer reviews. Wu et al. [25] developed a recommendation system based on the historical weblog posts of users. Recently, many scholars have conducted studies of information extraction from social media for business reasons, and most studies have been focused on the commercial value of enterprise microblogs, such as brand analysis [26], marketing promotion [27], and customer engagement [28]. For example, Onishi and Manchanda [29] specified a log-linear system for market outcomes

(sales) and the volume of blogs, and their results suggested that new and traditional media act synergistically. The analysis results in Colicev et al. [30] show that user-generated content has a stronger relationship with awareness and satisfaction, while firm-generated content is more effective for consideration and purchase intent.

Although monitoring business processes has become a major practical concern [31], scant attention has been paid to firms' market behavior from the perspective of firm-generated content. Unlike previous work, this study focuses on firms' business events in the market, which differs greatly from unorganized personal or social events in terms of their aims, contents, and objectives [32]. Business events may constitute some observable actions or circumstances (e.g., a new competitor's product) [33]. In particular, due to the nature of our particular context, these social news-based topics frequently change according to real-world events [21]. However, the literature does not provide a method to represent business events in an easy-to-calculate way [33], which motivates us to fill this gap with text mining technology.

2.2. Market Competition and Firm Action in a Market

In both academic and commercial literature, the main tool for explaining rivals' behavior is game theory models [34], in which it is presumed that all players use the same basic principles to take strategic actions in a market [35]. However, in the real world, game theory models become unwieldy when a competitor has many options for management actions or when there are multiple competitors, each of whom might react differently [36]. Therefore, it is equally important for market managers to know the actions of competitors in a timely manner so as to know their strategies.

In the literature on market competition, one consensus is to use market segmentation and category management strategies to differentiate market competition [37,38].

However, it has been shown that a firm may not recognize a potential competitor even if its action appears obvious in the market [39]. This is primarily because most companies rely on incomplete data, such as market research [7], to evaluate changes in a market. Particularly, such data may lead to a "managerial myopia" [40] in identifying competitive relationships [8].

There is a wealth of research in the literature on analyzing competition in the marketplace, and studies related to "observable events" in the marketplace can be briefly categorized as competitor identification, competitive positioning, dynamic competitive monitoring based on resources and capabilities, and product competitiveness analysis. Some examples of typical studies are listed in Table 1. We can briefly summarize the following characteristics of the work on market competition. Firstly, managerial myopia in identifying competitive threats is a well-recognized phenomenon. Secondly, broad competitor identification is an increasingly important task for managers, but there is no particularly efficient solution. Finally, online data have attracted the attention of managers, but the literature on its application to market competition is still scarce.

Table 1. Research literature related to business market.

Literature	Research Goals	Methods	Contributions
[2,40–42]	Develop models for competitor identification	Hypotheses test	Propose a cognitive framework for the identification of competitors
[5–10,43–48]	Competitive positioning; testing market structure	Empirical	Managers must be mindful to incorporate new information proactively from many sources and to actively discard old, automatic maps in order to develop reliable maps for changing environments
[11–13,49–51]	Dynamic competition identification; resource-based competition	Concept model	Develop a market-based and resource-based framework for broad competitor identification
[14–16,52–54]	Product competitive advantage analysis	Data mining	Identify competitors from UGC

In addition to analyzing competitors from the perspective of market strategy, in this study, we believe that a critical step in analyzing market competition is to acquire and analyze relevant data that can effectively facilitate scanning the activities of firms. Now that firm-generated content on social media platforms provides us with rich data to observe the behavior of the enterprises [8], the tasks in analyzing market competition, therefore, are to explore the events of competitive firms from these data and to analyze the market competition.

3. Research Methodology

3.1. Research Problem Statement

It is generally believed that competition among firms is characterized by negative interdependence [55]; that is, one side attaining its goal will decrease the probability of the other side successfully attaining it [56]. Accordingly, in this study, the competitors in a market are referred to as a set of firms that share similar actions/activities [4,6]. Therefore, the basic problem we want to address is to identify the behavioral events of firms from textual data and measure the similarity (or relevance) of their behavior.

Suppose that there are L firms $\mathbb{F} = \{F_1, F_2, \ldots, F_l, \ldots, F_L\}$ in a competitive market, and, at the observation time $t \in [1, T]$, firm $F \in \mathbb{F}$ published a piece of media b_t^F (such as online news, a blog, a report, etc.) about a business event via social media. Here, b_t^F can be deemed as a combination of a series of terms:

$$b_t^F = \{w_1, w_2, \ldots, w_s \ldots, \} \tag{1}$$

where w_s represents a semantic term in b_t^F. Furthermore, we assume that there are K types of business events, $\mathbb{E} = \{E_1, E_2, \ldots, E_k, \ldots, E_K\}$, and each event $E \in \mathbb{E}$ can be triggered by a set of featured terms:

$$E = \{v_1, v_2, \ldots, v_j \ldots, v_m\} \tag{2}$$

Given two enterprises F_i and F_j in the market $\mathbb{F}$, if they simultaneously conduct the same business event of $E \in \mathbb{E}$, then we believe that they have a competitive relationship regarding business event E. This competitive relationship often occurs when the two companies' target markets are the same [57]. In this way, the task of exploring market competition is then transformed to determine whether $b_t^{F_i}$ and $b_t^{F_j}$ contain the same business event, E. To this end, the research question can be specified as the following two tasks based on business event analysis from massive firm-generated content:

- Obtaining the complete representation of $\mathbb{E}$; that is, obtaining the total triggers (features) for each event $E \in \mathbb{E}$;
- Determining the semantic similarity between $b_t^{F_i}$ and E, where $i \in [1, L]$, $t \in [1, T]$ and $E \in \mathbb{E}$.

In this work, we propose a feature annotation and enhancement (FAE) method to address the first task and then deal with the second task by determining the relevance of the reported business events between any two companies in the market.

3.2. FAE Method: Extracting Business Event

Figure 1 presents an overview of the FAE system for exploring business events from massive firm-generated content in social media.

In Figure 1, the left path is the learning process, which is used to obtain the formation of a collection of business events. The middle path is the word-embedding process. It is used to represent the words in the corpus in an easy-to-calculate form, for example, word vectors [58]. The right path is the inference process. Its main function is to retrieve information on business events in firm-generated content, based on which we can further explore the competition in a market.
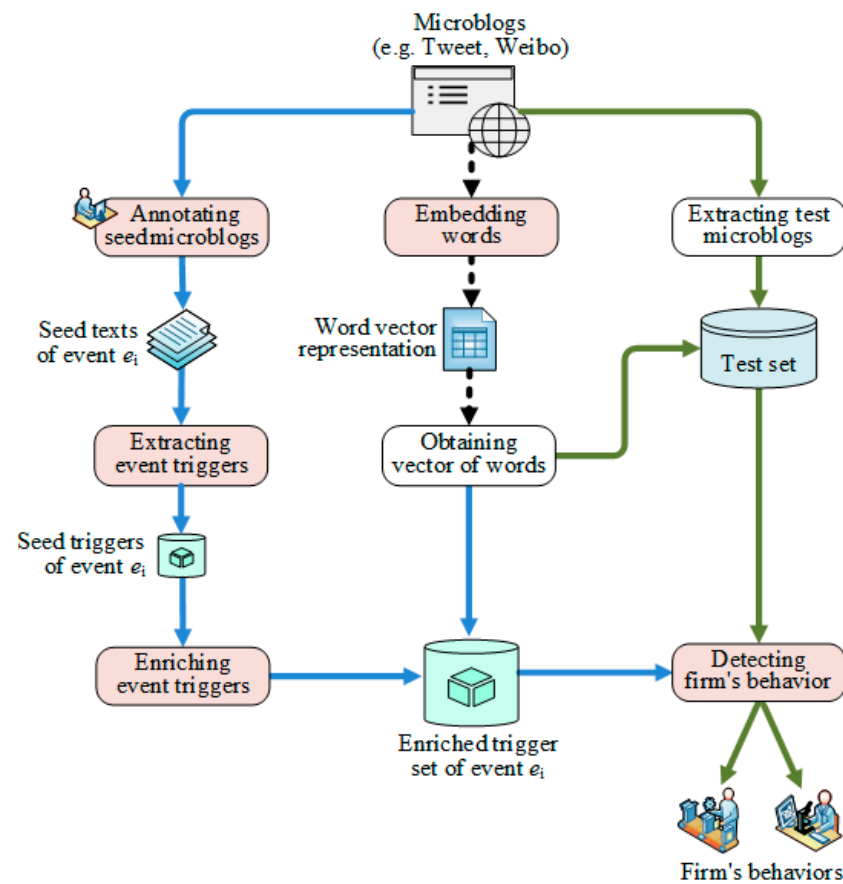
Figure 1. The research framework.

3.2.1. Word-Embedding Process

The word-embedding training method in NLP is used to find word representations that are useful for predicting the surrounding words in a sentence or document. Mathematically, word embedding is a process to obtain a vector representation of $vec(w_i)$ for word w_i in the corpus:

$$w_i \rightarrow vec(w_i) = (w_{i1}, \ldots, w_{ij} \ldots, w_{iD}) \tag{3}$$

where w_{ij} is the j-th dimension value of the vector for a word w_i, and D denotes the dimension of the trained vector. The benefit of word vector representation is that when we calculate the semantic relationship between any two words w_i and w_j in the corpus, it can be equivalently performed on the corresponding $vec(w_i)$ and $vec(w_j)$.

3.2.2. Learning Process

In this subsection, we propose a semi-supervised approach to learn the formation (i.e., trigger set) of each business event $E \in \mathbb{E}$, which consists of three parts: annotating seed microblogs, extracting seed triggers, and enriching event triggers.

In the process of annotating seed microblogs, we invited some experts to identify K types of business events in the market and annotate a fixed number of microblog posts as the seed texts for each event $E_k, k \in [1, K]$. Such a set of microblog text is denoted by $SeedBlog(E_k)$.

In online documents, an event trigger is a key term that most clearly expresses the exact meaning of a business event [59], and most of these key terms are verbs or nominalizations [60]. Accordingly, in the process of extracting seed triggers, we segmented all of the microblogs in $SeedBlog(E_k)$ into terms and then focused on the verbal terms in $SeedBlog(E_k)$ [14]. Along this line, we sorted the terms in $SeedBlog(E_k)$, and asked experts

to evaluate the verbs based on their contributions to event E_k. As such, we can sift out a set of initial seed triggers from $SeedBlog(E_k)$ for E_k.

Based on the annotation results (i.e., the event E_k and its initial seed triggers), we can further implement the part of enriching event triggers. To this end, we first obtained the word vector for each initial trigger $v \in E_k$ and then used the following method to represent the vector of E_k:

$$vec(E_k) = \frac{1}{|E_k|} \sum_{v_j \in E_k} tf\left(v_j^k\right) vec\left(v_j^k\right) \qquad (4)$$

where $tf\left(v_j^k\right)$ and $vec\left(v_j^k\right)$ denote the term frequency and the vector representation of v_j^k, respectively.

Next, to diminish the biased nature of the experts' annotations, we also considered the selection of new words from outside the experts' annotations to enhance the features in the trigger. For this purpose, we considered that the words that can help characterize the event E_k can be those that have a strong semantic similarity with E_k (measured by *COS*) or those that interact strongly with the seed words in E_k (measured by *PMI*). Given a candidate word w, we can calculate the similarity between $w \in \mathbb{B}(w \notin E_k)$ and E_k based on the following semantic-contextual similarity:

$$SIM(w, E_k) = COS(w, E_k) + \lambda * PMI(w, E_k) \qquad (5)$$

where $COS(w, E_k) = \frac{vec(w) \cdot vec(E_k)}{\|vec(w)\| \cdot \|vec(E_k)\|}$ represents the semantic similarity between w and E_k, in which $vec(w)$ and $vec(E_k)$ denote the vector representation of term w and behavior E_k, respectively. $PMI(w, E_k) = \sum_{v \in E_k \setminus \{w\}} log\left(\frac{p(w,v) + \epsilon}{p(w) \cdot p(v)}\right)$ measures the co-occurrence relationship (contextual similarity) between in w and existing triggers in E_k, in which ϵ is a smoothing coefficient. According to (5), if the value of $SIM(w, E_k)$ is very high, then w is likely to be a new trigger for E_k. In this way, the trigger set of E_k is enriched to $E_k \bigcup \{w\}$.

3.2.3. Inference Process

Given a microblog text b_t, its vector representation can be specified as:

$$vec(b_t) = \frac{1}{|b_t|} \sum_{w_j \in b_t} tf\left(w_j\right) vec\left(w_j\right) \qquad (6)$$

where $tf\left(w_j\right)$ denotes the term frequency of w_j.

Accordingly, we can introduce a function $f(b_t, E_k)$ to predict whether an event E_k was reported in b_t or not by calculating the similarity between $vec(b_t)$ and E_k. Therefore, b_t can be considered to be a piece of host text for an event E_{k*} if the following relationship holds:

$$k^* = \begin{cases} \operatorname{argmax}_k f(b_t, E_k), & \text{if } \max_k f(b_t, E_k) \geq \mu; \\ NULL, & \text{otherwise.} \end{cases} \qquad (7)$$

Here, the parameter μ is a predefined threshold. If the value of $f(b_t, E_k)$ is lower than the threshold of μ, we label b_t as "*NULL*", which means there is no event contained in b_t; otherwise, b_t will be labeled with the k^*-th event in accordance with the maximum value of $f(b_t, E_k)$.

3.3. Method for Exploring Market Competition

To explore competition among companies in the market in the time period of $t \in [1, T]$, we collected all the content generated by $F_j \in \mathbb{F}$, and then introduced the FAE method to identify the event information in each of the content. As a result, the event information

identified from the contents published by all companies in the market can be represented along the timeline *t* as a *firm–event matrix*, as shown below:

$$
\Psi_{E_k} = \begin{pmatrix}
F/time & 1 & \cdots & t & \cdots & T \\
F_1 & \psi_{E_k}(1,1) & \cdots & \psi_{E_k}(1,t) & \cdots & \psi_{E_k}(1,T) \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
F_j & \psi_{E_k}(j,1) & \cdots & \psi_{E_k}(j,t) & \cdots & \psi_{E_k}(j,T) \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
F_L & \psi_{E_k}(L,1) & \cdots & \psi_{E_k}(L,t) & \ddots & \psi_{E_k}(L,T)
\end{pmatrix}
\tag{8}
$$

where the variable $\psi_{E_k}(j,t)$ in Equation (8) denotes the occurrences of a given business event E_k in firm F_j at time *t*.

In the market, the sequence of an event E_k conducted by a firm F_j can be described as a vector of $1 \times T$ size , $\psi_{E_k}^{F_j} = \{\psi_{E_k}(j,1), \psi_{E_k}(j,2), \ldots, \psi_{E_k}(j,T)\}$. Then, the event-based competition between firms F_i and F_j can be measured by their Pearson correlation coefficient:

$$
comp_{E_k}(F_i, F_j) = \frac{cov\left(\psi_{E_k}^{F_i}, \psi_{E_k}^{F_j}\right)}{var\left(\psi_{E_k}^{F_i}\right) \cdot var\left(\psi_{E_k}^{F_j}\right)}
\tag{9}
$$

Similarly, at time *t*, all the events conducted by firm F_j can be described as a random vector of $1 \times K$ size, $\psi_t^{F_j} = \{\psi_{E_1}(j,t), \psi_{E_2}(j,t), \ldots, \psi_{E_K}(j,t)\}$. Then the time-based competition between F_i and F_j can be specified as:

$$
comp_t(F_i, F_j) = \frac{cov\left(\psi_t^{F_i}, \psi_t^{F_j}\right)}{var\left(\psi_t^{F_i}\right) \cdot var\left(\psi_t^{F_j}\right)}
\tag{10}
$$

Obviously, method $comp_{E_k}(F_i, F_j)$ measures the competition between two firms regarding a *business event* E_k, and method $comp_t(F_i, F_j)$ measures the competition between two firms at time *t*.

Further, along the timeline, we concatenate all the event sequences of firm F_j as follows: $\psi^{F_j} = \left\{\psi_{E_1}^{F_j}, \psi_{E_2}^{F_j}, \ldots, \psi_{E_k}^{F_j}\right\}$. In this way, we can propose a method to characterize the overall competition situation of F_i and F_j in the market as follows:

$$
comp(F_i, F_j) = \frac{cov\left(\psi^{F_i}, \psi^{F_j}\right)}{var(\psi^{F_i}) \cdot var\left(\psi^{F_j}\right)}
\tag{11}
$$

In a competitive market, competition among enterprises is characterized by negative interdependence [55]; that is, one side attaining its goal will decrease the probability of the other side successfully attaining their goal [56]. Accordingly, Equation (11) can be used as a feasible metric to monitor the event-based market competition.

4. Experiment Results

4.1. Data Collection

In this study, our focal market is the mobile phone production industry in China. We collected data from a leading online social media site, Weibo.com, to test the proposed method. Weibo.com is one of the most influential providers of microblogging services in China, and it is similar to Twitter, providing user services to create content and manage their accounts with access available from information devices (e.g., PCs and mobile phones). The

dominance of Weibo in China's social media websites makes it a good choice for retrieving late-breaking news.

In the industry channel of Weibo.com, there are 121 mobile phone manufacturers in China that have established Weibo accounts. We crawled all of the microblog documents (denoted as "Weibo") posted by these manufacturers. Each crawled Weibo mainly contains the following information: account ID, authentication information of the firm, time, data source, Weibo text (main information), reading number, etc. Note that Weibos that contain only picture(s) or only forward link(s) were discarded as invalid messages. Overall, a total number of 436,310 Weibos were crawled online for the study.

Following the general method in the literature of NLP [61], we truncated the Weibo texts by the punctuations in sentences and then adopted a Chinese NLP tool named Jieba (https://github.com/fxsjy/jieba, accessed on 20 January 2020) to execute the tasks of word segmentation and POS tagging simultaneously.

Some statistics about the dataset are summarized in Table 2.

Table 2. The statistical information of the dataset.

Description	Statistics
Total number of Weibos	436,310
Maximum number of Weibos posted by firms	20,834
Minimum number of Weibos posted by firms	16
Average number of Weibos posted	3605.87
Total words in the dataset	230,771
Total keywords in the dataset	33,043
Maximum number of words in Weibo text	75
Minimum number of words in Weibo text	1
Average number of words in Weibo text	20.83
Maximum number of keywords in Weibo text	52
Minimum number of keywords in Weibo text	1
Average number of keywords in Weibo text	5.14

4.2. Business Events in Firm-Generated Content

Firstly, two experts in the field of business/marketing management were introduced to label the events in the Weibos. The annotation results were mainly based on the experts' understanding of market management, the verbal term(s) of a Weibo, and consistency with previous research results [33,62].

Experts labeled five main business events as *Recruiting, Cooperation, Research, Promotion*, and *Sale* in our dataset. In order to achieve a balanced distribution of data, the experts annotated 600 seed Weibos for each of the five events. Note that we also asked experts to label 600 NULL event data so that the Weibos containing business events can be significantly distinguished from those Weibos without any event information. Some sample triggers extracted for each event are given in Table 3. Apparently, the sample verbs gathered under the same event are semantically similar.

Table 3. The seed triggers extracted from the seed Weibos.

Event Types	Top 5 Event Triggers
Sale	publish/release; buy/purchase; publish and sale; customized; sell
Cooperation	cooperate; unite; join hands; coalesce; build
Promotion	order/subscribe; obtain/gain; share; earn; extract
Research	design; innovate; apply; upgrade; research
Recruiting	participate in; recruit; recruit; join; look for

Next, we carried out trigger enrichment to obtain a better set of event triggers. The calculation results are shown in Table 4 below.

Table 4. Enhancement of the event triggers.

Event	Seed Weibos	Total Terms	Verbs (Initial)	Ratio of Verbs	Verbs (Denoised)	Enriched Triggers
Recruiting	600	4812	969	20.14%	918	388
Cooperation	600	6546	950	14.51%	900	391
Research	600	6628	1171	17.67%	1120	61
Promotion	600	4241	1238	29.19%	1187	77
Sale	600	4729	1060	22.41%	1009	230
NULL	600	2544	1032	40.57%	981	444

4.3. Performance of FAE

This section evaluates the performance of our model in detecting business events from massive firm-generated content in online social media. In the experiments, the FAE method requires setting two parameters. The hyperparameter in Equation (5) is set to $\lambda = 1$; the threshold in Equation (7) is set to $\mu = 0.7$. The other parameters are set to the default values of the Python experimental environment.

The following experiments are based on two basic aspects: firstly, FAE is a feature-extraction-oriented approach, so comparative experiments were carried out between some classical feature representations in the NLP literature and FAE. These methods for feature representation are uni-, bi- and tri-grams [63], TF-IDF [61,64], and the state-of-art method BERT [65]. Secondly, a statistical exploration experiment found that the distribution of different types of business events in our dataset is uneven (see Figure 2). The latest research shows that the random forest (RF) algorithm performs well in the text classification task in unbalanced datasets [66], so we used RF as the classification algorithm in the experiment. In addition, in order to compare the efficiency of different algorithms, we invited three experts to label an additional 3600 tweets as a test set to evaluate the different models, and the differences between experts were resolved by voting.
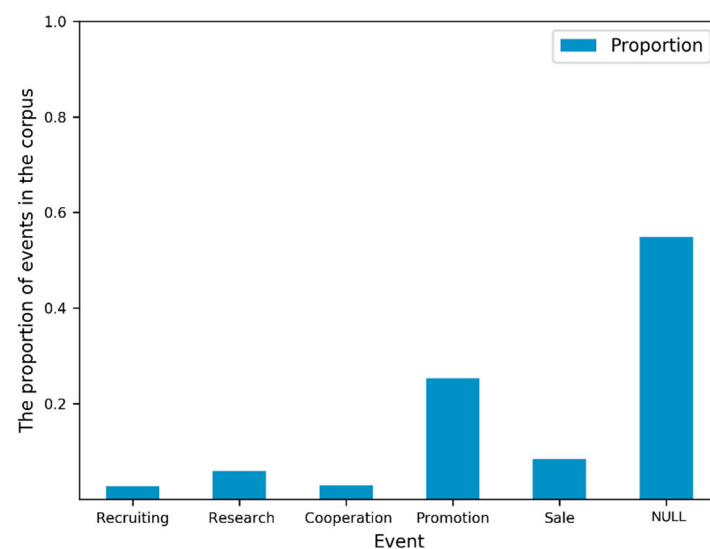
Figure 2. Distribution of business events.

The feature settings regarding the comparison methods are defined as follows: (1) an n-gram is a sequence of n adjacent elements from a string of tokens, which typically are Chinese words or characters; (2) all the methods (except BERT) firstly represent the terms in the corpus into a high-dimensional feature space and then use feature selection [67] to select 4011 features; (3) the focus of the comparison in the experiments is on the representation methods of the text, so BERT does not consider the fine-tuning approach.

In addition, we adopted the three measures of *precision*, *recall* and *F-value*, which are widely adopted in the NLP literature, to evaluate the performance of the different representation methods in detecting business events:

$$Precision = \frac{\text{Number of correctly identified events}}{\text{Total number of identified events}} \tag{12}$$

$$Recall = \frac{\text{Number of correctly identified events}}{\text{Number of true events labeled by experts}} \tag{13}$$

and

$$\textit{F-value} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

Using the aforementioned representations as input in the RF algorithm, the performance of different representations is shown in Figure 3.
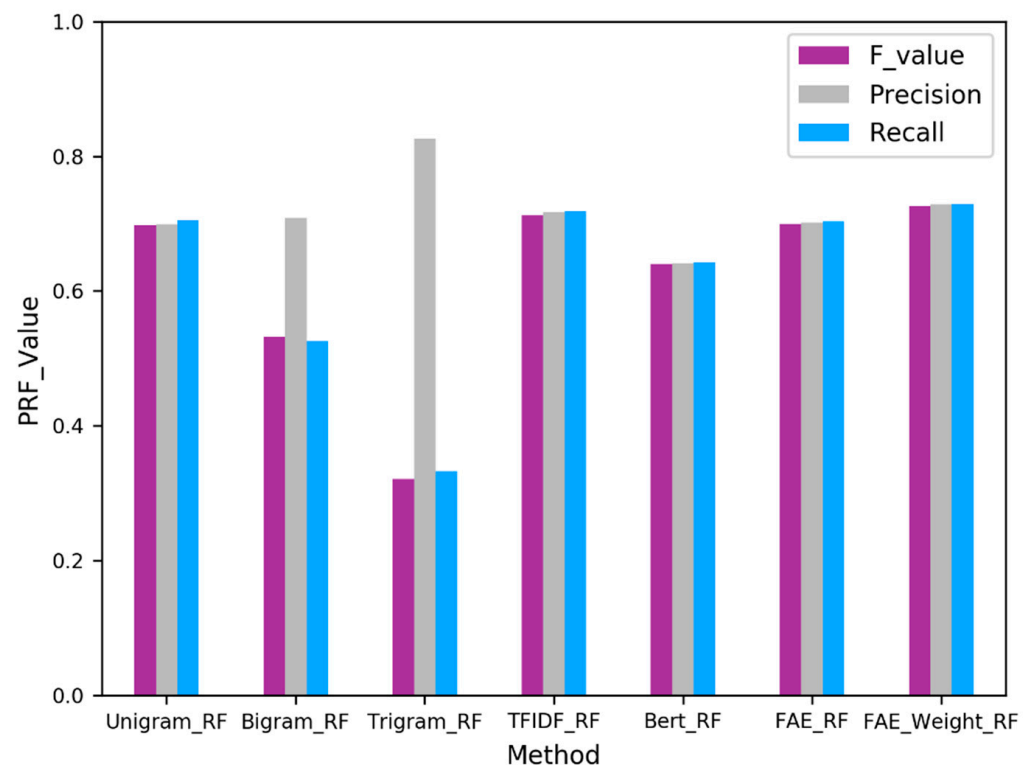
Figure 3. Performance comparison.

Figure 3 illustrates the average performance of all feature representation methods on our text dataset. Some general results can be summarized as follows. First, our semi-supervised method of FAE outperformed the comparison methods when identifying business events from textual social media. Second, the tri-gram representation method had the worst performance in the experiment. What is more surprising is the results of BERT; although its performance is not bad, it is still not as effective as our method.

4.4. Exploring Market Competition

4.4.1. Overall Competitive Landscape in the Market

Because the operating time span of the company's official Weibo account is not the same, we selected firms that had published more than 1000 Weibos from January 2018 to June 2019 (the time period for exploring competition) to carry out the experiment. There were 44 Weibo accounts that met this requirement. We calculated the correlation of these 44 firms as follows. We first constructed a complete match on these 44 firms (each company

must match with the remaining firms); then, we calculated the correlation coefficient of the business events of the two companies in the same pair. The calculation results are presented in Figure 4, where the values in color indicate that the correlation between business events between the two companies is above 0.7. Specifically, the values in the red, yellow and green backgrounds are in the intervals of (0.9, 1], (0.8, 0.9], (0.7, 0.8], respectively.

|	Motor	ShouI	Nokia	Samsu	ZTE	MeiTu	MANN	Shang	MeiZu	Lenov	AiJieK	OPPO	Coolp	Hisen	KaDiI	Hongl	Realm	Gome	Deovo	HuaW	ZUK	FangV	Nubia	XiaoL	HeiSh	8848	MI	HuiW	360	Smart	Doov	KuBi	Konk	Green	Gione	VIVO
Motorola		0.6	0.62	0.51	−0.11	0.63	0.57	0.29	0.55	0.58	0.61	0.58	0.35	0.6	0.24	0.11	0.5	−0.04	−0.01	0.25	0.4	0.37	0.67	0.16	0.61	0.36	0.58	0.37	0.27	0.67	0.48	0.1	0.56	0.01	−0.14	0.62
ShouHuBao			0.73	0.61	−0.18	0.88	0.72	0.39	0.84	0.8	0.81	0.76	0.49	0.85	0.12	0.42	0.61	−0.08	−0.1	0.21	0.65	0.58	0.85	0.23	0.85	0.5	0.79	0.38	0.4	0.74	0.75	0.22	0.74	0	0.22	0.79
Nokia				0.75	−0.29	0.86	0.87	0.55	0.89	0.91	0.86	0.9	0.77	0.89	0.33	0.4	0.72	−0.12	−0.03	0.29	0.55	0.49	0.92	0.25	0.87	0.66	0.87	0.5	0.48	0.79	0.55	0.2	0.76	−0.05	0.12	0.92
Samsung					0.01	0.7	0.75	0.37	0.7	0.71	0.68	0.72	0.58	0.71	0.28	0.38	0.57	−0.05	−0.03	0.41	0.45	0.36	0.7	0.17	0.68	0.46	0.72	0.3	0.34	0.62	0.47	0.19	0.69	0.05	0.15	0.67
ZTE						−0.19	−0.29	−0.06	−0.23	−0.15	−0.25	−0.15	−0.16	−0.17	−0.05	0.07	−0.19	0.26	−0.07	0.65	−0.11	−0.12	−0.23	−0.15	−0.18	−0.21	−0.23	−0.03	−0.04	−0.2	−0.04	0.09	−0.06	−0.2	0.09	−0.27
MeiTu							0.78	0.38	0.93	0.92	0.89	0.88	0.58	0.91	0.26	0.43	0.82	−0.09	−0.11	0.27	0.76	0.52	0.9	0.3	0.95	0.61	0.92	0.39	0.32	0.82	0.72	0.14	0.81	0.04	0.21	0.87
MANN								0.52	0.82	0.81	0.76	0.8	0.69	0.78	0.29	0.38	0.6	−0.15	0.04	0.24	0.48	0.52	0.85	0.14	0.84	0.52	0.78	0.48	0.45	0.67	0.53	0.15	0.72	0.15	0.15	0.78
ShangShan									0.55	0.58	0.33	0.43	0.64	0.55	0.49	0.3	0.13	0.1	−0.06	0.3	−0.17	0.47	0.52	0.12	0.45	0.42	0.47	0.69	0.72	0.27	0.29	0.41	0.47	−0.08	−0.08	0.56
MeiZu										0.92	0.9	0.88	0.73	0.92	0.34	0.45	0.71	−0.13	−0.08	0.26	0.63	0.46	0.91	0.39	0.95	0.64	0.92	0.52	0.48	0.81	0.68	0.21	0.78	0.02	0.19	0.88
Lenovo											0.82	0.89	0.72	0.94	0.38	0.51	0.73	−0.11	−0.08	0.38	0.62	0.57	0.91	0.25	0.93	0.65	0.89	0.5	0.5	0.76	0.69	0.12	0.83	−0.1	0.2	0.87
AiJieKe												0.86	0.86	0.23	0.39	0.71	−0.12	−0.07	0.23	0.68	0.37	0.87	0.39	0.91	0.6	0.9	0.4	0.36	0.83	0.64	0.04	0.74	0.16	0.15	0.88	
OPPO													0.71	0.9	0.33	0.38	0.79	−0.1	−0.04	0.44	0.67	0.51	0.89	0.24	0.92	0.6	0.86	0.44	0.32	0.74	0.67	0.05	0.78	−0.08	0.2	0.85
Coolpad														0.7	0.46	0.25	0.48	−0.13	0.07	0.36	0.31	0.32	0.73	0.07	0.67	0.51	0.71	0.55	0.54	0.32	0.2	0.55	−0.14	0.11	0.66	
Hisense															0.31	0.41	0.72	−0.09	−0.09	0.37	0.64	0.56	0.91	0.3	0.92	0.64	0.91	0.47	0.49	0.78	0.71	0.17	0.8	−0.09	0.18	0.92
KaDiEr																0.06	0.21	−0.04	−0.03	0.14	−0.09	−0.07	0.35	0.31	0.3	0.39	0.3	0.47	0.28	0.29	0.24	−0.04	0.32	−0.04	−0.04	0.28
HongDa																	0.26	−0.02	0.19	0.31	0.24	0.4	0.07	0.47	0.31	0.43	0.2	0.24	0.24	0.46	0.02	0.59	−0.05	0.36	0.38	
Realme																		−0.13	−0.05	0.25	0.79	0.4	0.73	0.16	0.75	0.44	0.78	0.18	−0.15	0.67	0.63	−0.13	0.67	−0.04	0.24	0.71
Gome																			−0.04	0.09	−0.1	−0.07	−0.06	−0.06	−0.11	−0.04	−0.06	−0.09	0	−0.08	−0.08	0.18	−0.07	−0.05	−0.05	−0.04
Deovo																				−0.1	−0.09	0.01	−0.03	−0.06	−0.07	−0.13	−0.08	0.01	−0.05	0	−0.07	−0.04	−0.04	0.26	−0.09	
HuaWei																					0.21	0.24	0.26	−0.09	0.34	0.18	0.27	0.24	0.2	0.15	0.28	0.06	0.47	−0.26	0.06	0.24
ZUK																						0.28	0.62	0.2	0.72	0.35	0.68	−0.05	−0.12	0.68	0.57	−0.1	0.56	−0.11	0.26	0.54
FangWanYuan																							0.55	−0.11	0.52	0.15	0.44	0.27	0.27	0.29	0.49	0.21	0.45	−0.04	−0.08	0.52
Nubia																								0.23	0.91	0.62	0.89	0.51	0.45	0.84	0.61	0.22	0.74	−0.01	0.11	0.89
XiaoLaJiao																									0.3	0.37	0.3	0.11	0.16	0.37	0.35	−0.07	0.2	−0.07	−0.07	0.35
HeiSha																										0.62	0.91	0.46	0.39	0.8	0.75	0.1	0.82	0.02	0.21	0.85
8848																											0.6	0.36	0.39	0.51	0.36	0.12	0.49	0	0.13	0.7
MI																												0.43	0.38	0.79	0.64	0.12	0.79	0.06	0.21	0.86
HuiWei																													0.55	0.4	0.34	0.29	0.49	−0.04	−0.09	0.49
360																														0.32	0.1	0.52	0.32	−0.06	−0.05	0.44
Smartisan																															0.55	0.22	0.65	−0.11	0.06	0.77
Doov																																−0.08	0.77	−0.09	0.28	0.63
KuBi																																	0.01	−0.05	−0.05	0.22
Konka																																		−0.06	0.24	0.75
GreenOrange																																			−0.05	−0.1
Gionee																																				0.15
VIVO																																				

Figure 4. Overall competitive landscape.

We can see from Figure 4 that, in the targeted time interval, the competitive relationship between enterprises is still very fierce. After pairing the 44 firms, it can be seen that more than 25.7% of the paired relationships have a high correlation value (greater than 0.7) of their business events. Among them, the competitive behavior of the firm "IVVI" is quite incredible, and its business events have a very high correlation (greater than 0.7) with 48% of other firms' events. This is followed by KuNuo and AiJieMO; their business events are also highly related to the business events of more than 40% of other firms. On the contrary, social media information released by XiaoLaJiao suggests it is the least willing to compete in the market. The business events it expresses in the market are extremely low (less than 0.3) in comparison to 50% of other businesses' events. The company that followed XiaoLaJiao was ZUK, whose business events were unrelated to 45% of the other firms' events.

4.4.2. Exploring Behavior-Based Competition

In order to better observe the micro-competitive relationship between firms, we selected 9 companies with the top market share among China's mobile handset manufacturers (i.e., Huawei, Xiaomi, Meizu, ZTE, Lenovo, Coolpad, OPPO, Vivo, and HTC) to show their business behaviors as well as their competition relationships on the market. The results are shown in Figure 5 below.

The results in Figure 5 illustrate the proportion of the same action taken in different companies in the market. In each sub-figure, the value of the y-axis indicates the proportion of a particular event that happened in a company. Therefore, the larger the value on the y-axis, the greater the event of the corresponding company.

(**a**) Promotion

(**b**) Sale

(**c**) Research

(**d**) Cooperation

(**e**) Recruiting
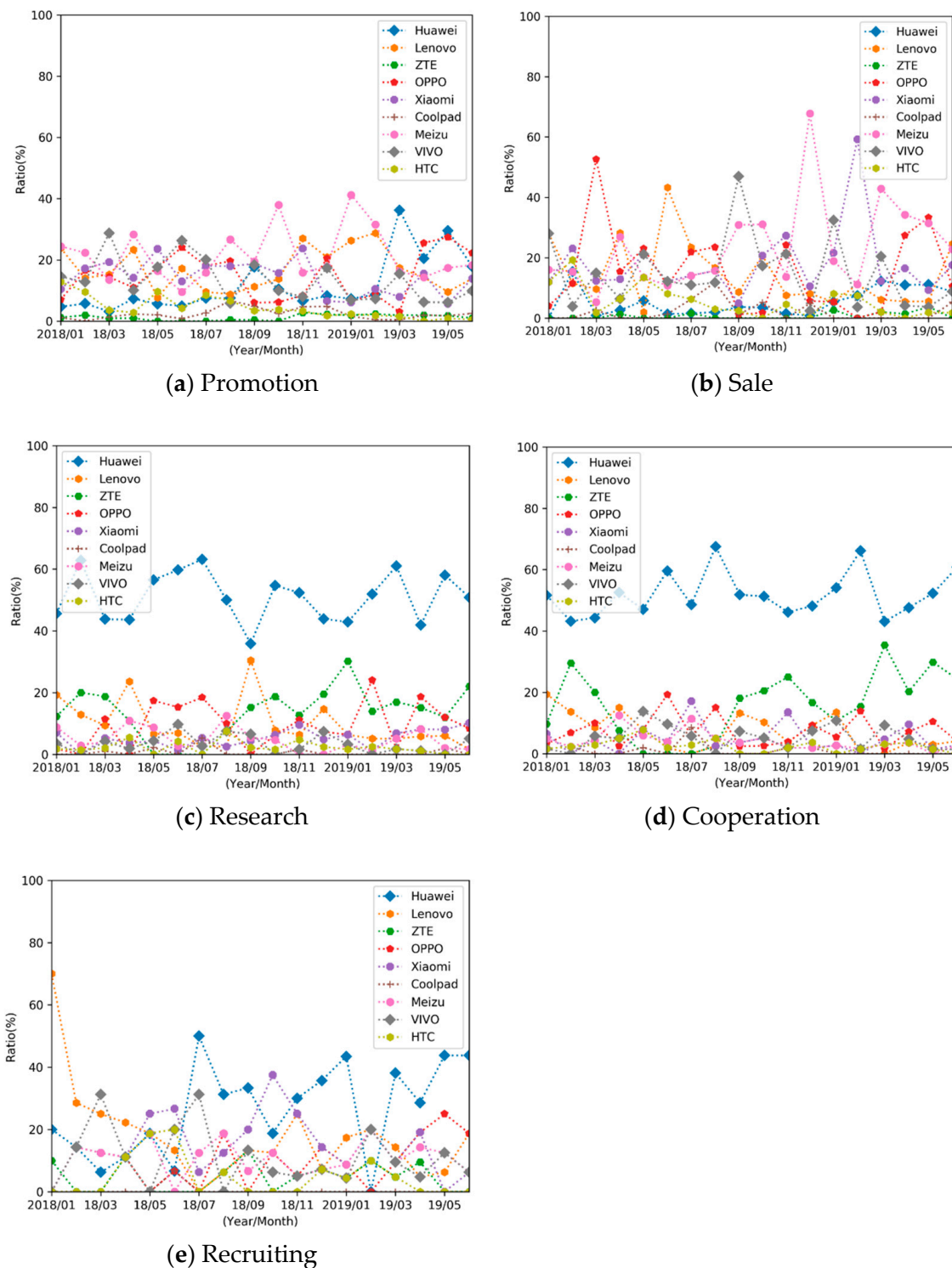
Figure 5. Event competition of firms.

Generally, if different firms adopt similar strategies, it can be expected that their behavior in the market should also be similar. The evolution lines in Figure 5a,b are relatively dense and show little variation, which indicates that the firms competed fiercely in the two behaviors of *Promotion* and *Sale*. This finding is interesting for the results of *Research* and *Cooperation* events: since mobile phone manufacturing is considered a high-tech industry in China, for example, the Huawei Company is likely to show the public that it is very active at *Research* and *Cooperation* events.

Figure 5d also highlights the fact that Huawei and ZTE had some confrontations when implementing their *Cooperation* events. As seen in the figure, the trend and direction of their *Cooperation* events moved in almost the opposite direction during the same time period (see Figure 5d; the blue line indicates Huawei events, and the green line indicates ZTE events).

In Figure 5b, the competition related to the variable *Sale* is more complicated. Overall, OPPO and Meizu have a strong preference for using the competitive behavior of *Sale*. Interestingly, Vivo, Lenovo, and Xiaomi also preferred the *Sale* strategy in the second half of 2018. HTC and Lenovo gradually reduced the proportion of Sale behavior after November 2018. In contrast, Meizu continuously strengthened its *Sale* behavior after January 2019.

4.4.3. Exploring Time-Based Competition

Figure 6 illustrates the evolving chain of the selected 9 companies on the 5 business behaviors (see Table 4). The observed time interval is from January 2018 to May 2019, with a total of 17 (continuous) time slices. In each sub-figure, the value on the y-axis indicates the strength of the corresponding event at a time slice, and the value on the x-axis indicates the persistence of the evolution of an event.
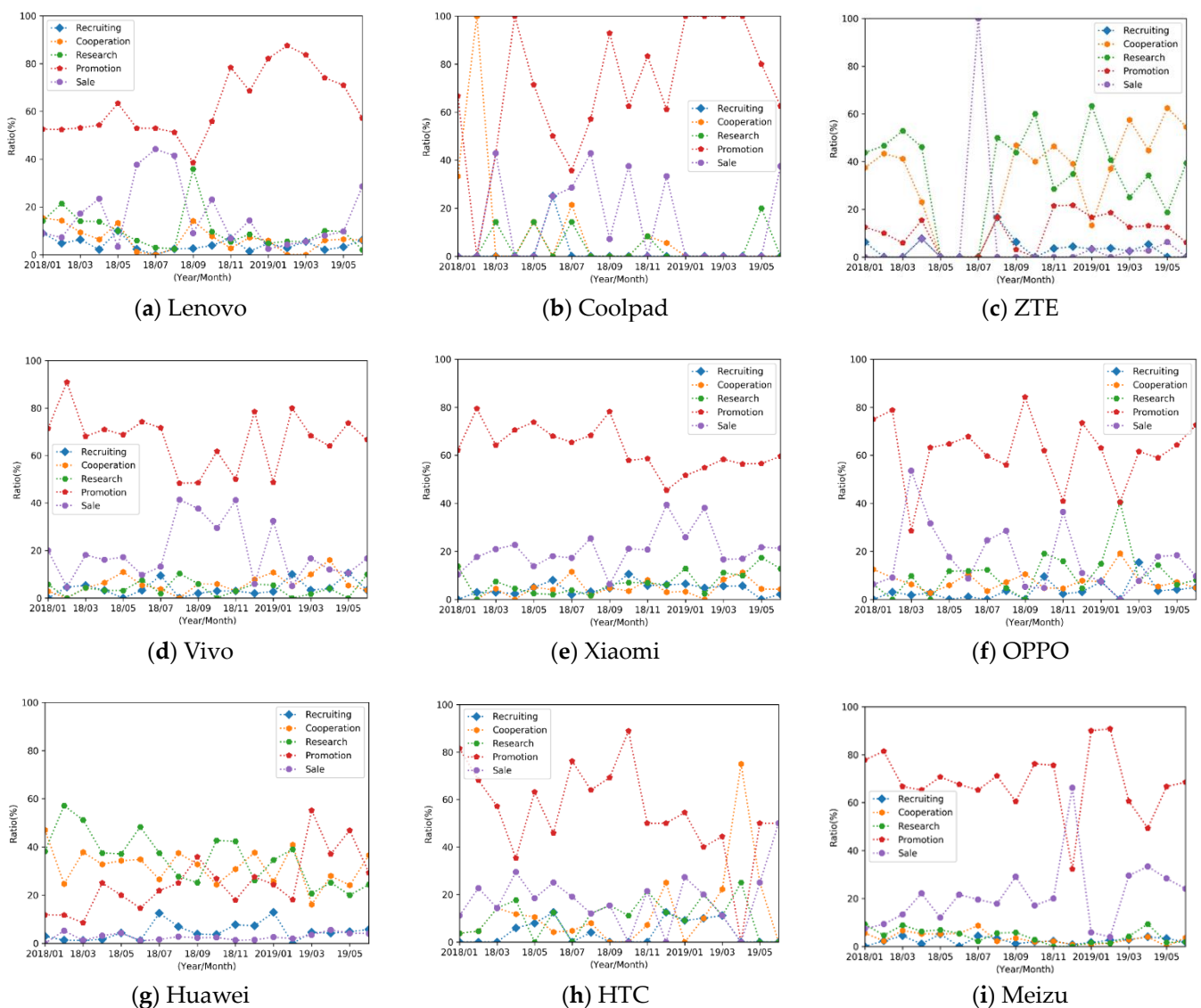
Figure 6. Event evolution of firms.

On the basis of the analysis of the events of these companies and their evolution chains, we identify some "strategic groups" [68]. First of all, with the exception of Huawei and ZET, the most positive business event for companies is *Promotion*. This shows that firms have attached great importance to the role of Weibo, a social media platform, in promotion. Second, the event lines of ZTE, Huawei, and Xiaomi were more concentrated, indicating that they were relatively balanced when implementing these five business events. Third, Huawei and ZTE were more willing to show strong and continuous *Cooperation* and *Research* behaviors in the market. Interestingly, Coolpad was quiet in the market for the time slices of January 2019 to March 2019. HTC suddenly became quite active in the market after September 2018, which was a remarkable phenomenon for a marketer.

From the perspective of market competition, in addition to ZET and Huawei, the remaining firms used *Promotion* as the main event published on social media. In particular, Lenovo, CoolPad, Vivo, Xiaomi, and Meizu have been in strong competition in terms of market promotion.

5. Discussion

5.1. Generalizability of the Study

We were also concerned with the generalizability of this study. We addressed this issue from the following two aspects. First, this paper presented a complete solution to the problem of organizational event detection from massive amounts of short texts (see Figure 1). The solution consisted of several steps, and in each step, we presented or adopted some ad hoc or heuristic techniques. Technically, these methods can be commonly used in NLP tasks. Second, we expanded our approach to another dataset, in which 48,595 abstracts of business articles were crawled from a website.

In particular, to impose our method on the new dataset, at first, experts labeled 3800 abstracts as the seed texts, and we annotated four types of business events, *Marketing, Financing, Mergers and Acquisitions (MA)*, and *Stock Market Performance (SMP)*. Then, we used the generated events to monitor the evolution of business events between January 2010 and February 2018 (in the Chinese market) for four Internet companies, namely Alibaba, JD, Amazon (China), and eBay (China). The results are shown in Figure 7a–d. It is particularly evident that the main marketing actions of the four companies listed, as major e-commerce platforms, are focused on marketing events.

Furthermore, the competition events of *Marketing*, *Financing*, *MA,* and *SMP* are shown in Figure 8a–c, respectively. As can be seen in Figure 8a, *Marketing* events in the market were adopted primarily by Amazon, followed by Alibaba and JD. The *Marketing* events of Alibaba, JD, and Amazon had different trends during the same time period, which indicated that these three companies were trying to avoid positive confrontation. In addition, it is worth mentioning that eBay is far less active in the market for all events than the three aforementioned companies. This is also in line with the fact that eBay has faded out of the Chinese market due to its poor performance during the time period presented in the data.

In conclusion, these experimental results indicated that the proposed method, i.e., FAE, in this paper works well on the new dataset.

5.2. Research Implications

The value of the proposed method in this work lies in the use of efficient text mining methods to obtain the appropriate representation of a firm's events from massive data and in bridging the gap between massive social media and latent market information for decision-making, in particular in the marketing area. For example, managers can develop optimal marketing strategies based on the extracted events implemented by competitors.

Thus far, studies for exploring the business value of social media have been focused mainly on using intelligence methods for brand/product marketing [21,27], customer analysis [9], and their mixture [10]. Researchers, however, have also realized that MI based on social media analysis provides "competitive intelligence" by helping businesses understand their environments, competitors, and overall business trends [69]. Unfortunately,
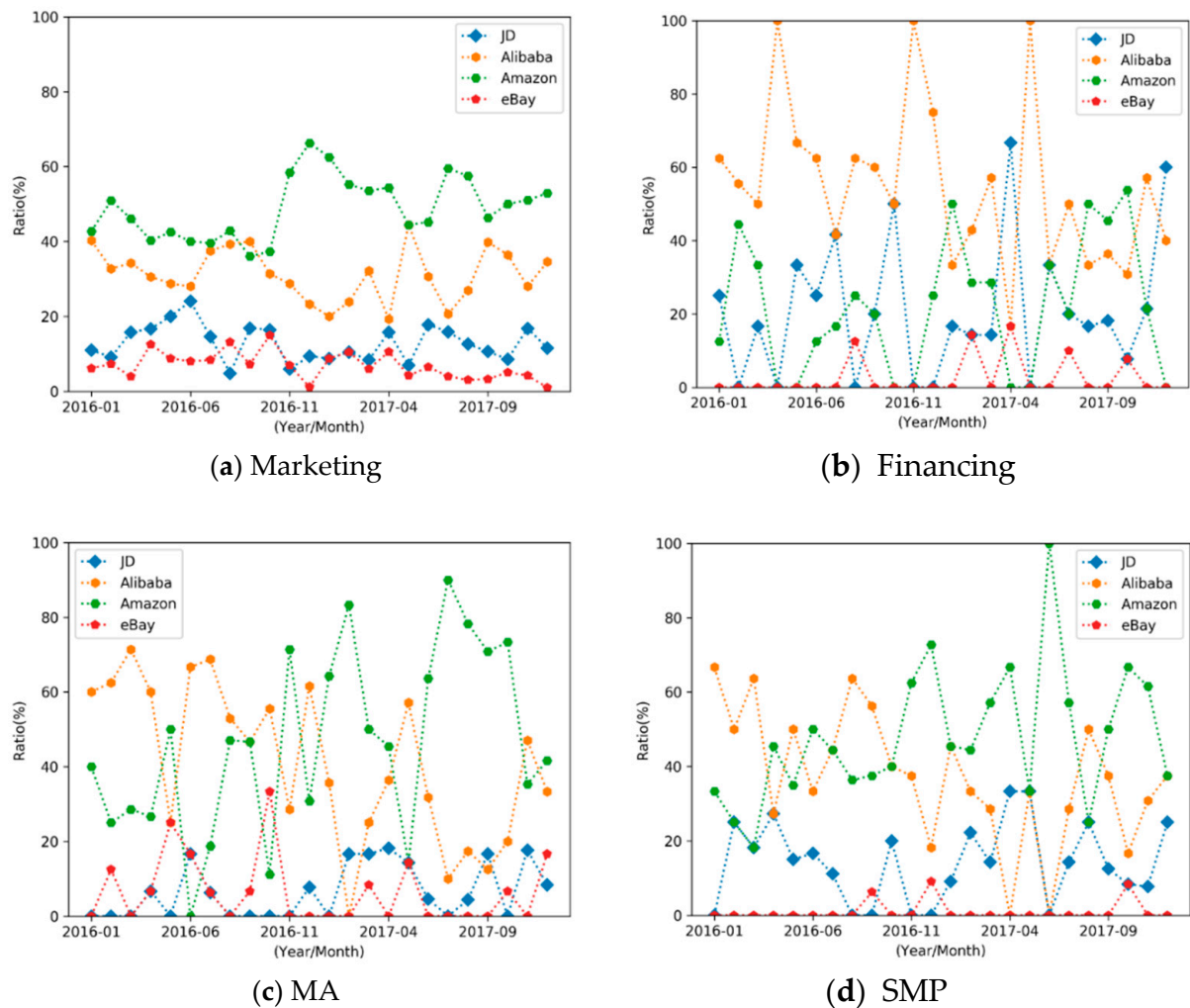
no definitive research results tell us how these wonderful things can happen. This work proposes a method to fill the application gap by revealing competitive information from massive firm-generated content in online social media.

Thus, the contribution of this paper is to provide intelligent methodological ideas on the application of large amounts of social data to management practices. In fact, for not only the competition problem but any behavioral correlation between firms that can be found in the data, the approach proposed in this paper can provide ideas for its analysis and discussion.

The prior literature is rich in studies of firm competition theory, and these studies have improved our understanding of how firms compete with each other in the marketplace. However, competition among firms in the market is very complex and usually changes and evolves over time, regions, user groups, etc. This leads to a tendency to bias (misidentification) and omission (ignoring competitors) in identifying firms' competitive relationships from traditional managerial dimensions, and knowledge about inter-firm competitive relationships is limited or even ambiguous if not analyzed from the perspective of firms' market actions (events). The research in this paper fills this gap by refining firms' market actions and constructing associations between these actions and competition.

(**a**) Alibaba   (**b**) JD

(**c**) Amazon (China)   (**d**) eBay (China)
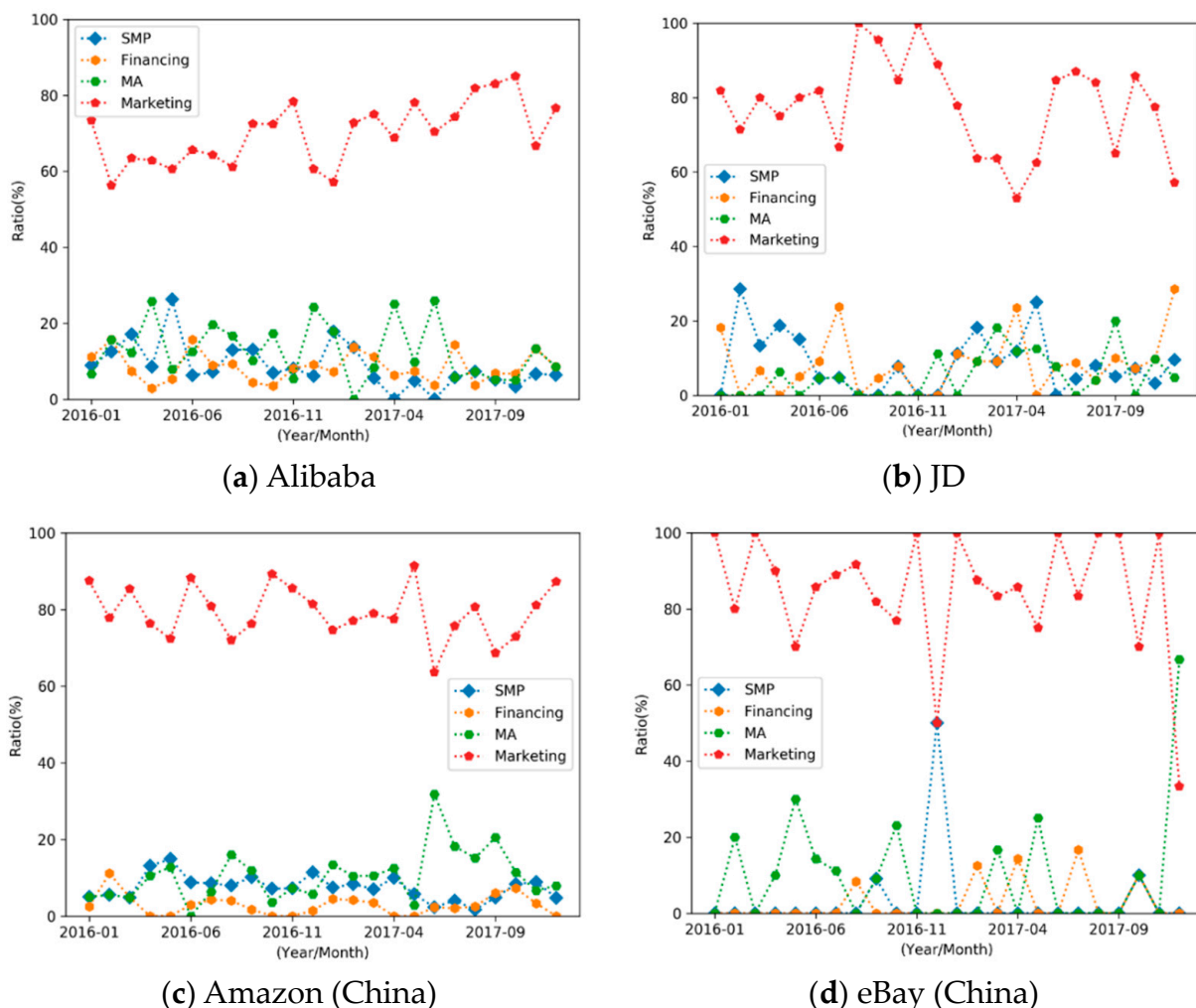
Figure 7. Event evolution of the Internet companies.

J. Theor. Appl. Electron. Commer. Res. **2023**, *18*

923

Figure 8. Event competition of the Internet companies.

5.3. Managerial Implications

On the basis of our proposed FAE method, we effectively extracted the *business events* from massive firm-generated contents, which is useful information for marketers. The foremost application of our work is the exploration of competitive behavior through open social media, which enables marketers to summarize their competitors' online microblogs with salient competitive events. In particular, the method proposed in this work can be used to discover and visualize the temporal evolution of business behaviors as they overlap or can be discriminated from those of competitors. As such, marketers can further analyze the motivations for companies to adopt the same (or different) strategies in the same period with the help of organizational behavior theories. In this way, this framework can enable marketers to effectively represent the trend of corporate–industry behavior to avoid potential decision-making risks.

In addition to identifying firms' events, our approach also harvests rich event triggers. This information can be used by investors to subdivide the (semantic) differences at the event level when different companies follow the same behavior. In particular, by mining the evolution information of firms' events over a long period of time, investors can better identify major changes in the management of the business as well as false behaviors in the market. For example, some traditional enterprises may have exhibited unreasonable investment behaviors in the Internet industry during the Internet boom.

5.4. Limitations

We are aware of several limitations of this study. First, the selection bias between experts may have resulted in different annotations on the same seed microblogs. To reduce the impact of this difference on the final result, we should be more precise in the steps of labeling, such as improving the capabilities of experts, and we should design strategies to cross-validate the results of these different experts. Second, it would be interesting to extend the event detection method to characterize firm competition. We believe, however, that future research of theoretical analysis will be key to understanding why firms undertake the actions that we observed and how those actions may have affected their performance.

In addition, if the corpus is obtained from a limited number of websites, the method proposed in this paper may be faced with the problem of information completeness for identifying a firm's events. Future work should benefit from an enhanced corpus to address this limitation.

Finally, there is no theoretical depth in the discussion of corporate market competition. The initial idea of this paper is to use intelligent methods (big data processing) and new data sources (social media) to explore the competitive behavioral relationships between firms. However, while the "event" information from the large amount of data analyzed provides a clue for analysis, it requires further theoretical discussion as to whether it can truly account for the competitive relationships that exist between firms. This is the part of our future work that needs to be strengthened.

6. Conclusions

Social media data contain valuable information that can be used in decision-making. The challenges of a lack of definition, mixed topics, and various types of documents, however, have prevented the efficient extraction of business events from massive amounts of social media data.

To address these challenges, theoretically, in contrast to the traditional view of analyzing firms' competitive relationships based on market segmentation and strategy, we propose that firms' competitive market strategies may be continuously planned, but at the implementation level, they can only be reflected by their targeted "events" in the marketplace. Moreover, from a practical point of view, we demonstrate that the online social media content generated by firms is timely data reflecting their "behavior" in the marketplace. The proposed method has two main parts. The first involves crawling textual data online, labeling event types, and enriching event triggers. It results in a set of extracted events as well as their triggers, which can then be used as a classifier to identify event information from texts. The second mainly includes identifying business event information from massive firm-generated content and exploring the potential competitive relationship between any two enterprises in the market by organizing their extracted business events into a firm–event matrix. The experimental results with real data extracted from Weibo.com show that the method proposed in this work can efficiently extract *business event* information from social media, and such information may enable marketers to summarize their competitors' online microblogs with salient competitive events.

Author Contributions: Conceptualization, H.Y. and B.M.; Data curation, W.D.; Formal analysis, W.D.; Funding acquisition, H.Y. and B.M.; Investigation, Y.Q.; Methodology, H.Y. and Y.Q.; Project administration, B.M.; Resources, H.Y. and B.M.; Software, W.D.; Supervision, B.M.; Validation, W.D. and H.Y.; Visualization, W.D.; Writing—original draft, H.Y.; Writing—review & editing, B.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

References

1. Hitt, M.A.; Ireland, R.D.; Hoskisson, R.E. *Strategic Management: Competitiveness and Globalization-Concepts and Cases*, 10th ed.; Cengage Learning: Boston, MA, USA, 2012.
2. Clark, B.H.; Montgomery, D.B. Managerial Identification of Competitors. *J. Mark.* **1999**, *63*, 67–83. [CrossRef]
3. Wei, C.P.; Chen, L.C.; Chen, H.Y.; Yang, C.S. Mining suppliers from online news documents. In Proceedings of the Pacific Asia Conference on Information Systems, PACIS, Jeju Island, Republic of Korea, 18–22 June 2013; p. 261.
4. Coyne, K.P.; Horn, J.L. Predicting Your Competitor's Reaction. *Harv. Bus. Rev.* **2009**, *87*, 90–97.
5. Hsieh, K.-Y.; Vermeulen, F.J.O.S. The Structure of Competition: How Competition Between One's Rivals Influences Imitative Market Entry. *Organ. Sci.* **2014**, *25*, 299–319. [CrossRef]
6. Ramaswamy, V.; Gatignon, H.; Reibstein, D.J. Competitive Marketing Behavior in Industrial Markets. *J. Mark.* **1994**, *58*, 45–55. [CrossRef]
7. Porter, M.E. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*; Free Press: New York, NY, USA, 1980.
8. Pant, G.; Sheng, O.R.L. Web Footprints of Firms: Using Online Isomorphism for Competitor Identification. *Inf. Syst. Res.* **2015**, *26*, 188–209. [CrossRef]
9. Holsapple, C.W.; Hsiao, S.-H.; Pakath, R. Business social media analytics: Characterization and conceptual framework. *Decis. Support Syst.* **2018**, *110*, 32–45. [CrossRef]
10. Lee, D.; Hosanagar, K.; Nair, H.S. Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook. *Manag. Sci.* **2018**, *64*, 5105–5131. [CrossRef]
11. Rishika, R.; Kumar, A.; Janakiraman, R.; Bezawada, R. The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation. *Inf. Syst. Res.* **2012**, *24*, 108–127. [CrossRef]
12. Opesade, A.O. Twitter-Mediated Enterprise–Customer Communication: Case of Electricity Distribution Services in a Developing Country. *Soc. Sci. Comput. Rev.* **2021**, *40*, 1578–1594. [CrossRef]
13. Martín-Rojas, R.; Garrido-Moreno, A.; García-Morales, V.J. Fostering Corporate Entrepreneurship with the use of social media tools. *J. Bus. Res.* **2020**, *112*, 396–412. [CrossRef]
14. Kumar, A.; Bezawada, R.; Rishika, R.; Janakiraman, R.; Kannan, P.K. From Social to Sale: The Effects of Firm-Generated Content in Social Media on Customer Behavior. *J. Mark.* **2016**, *80*, 7–25. [CrossRef]
15. Hogenboom, F.; Frasincar, F.; Kaymak, U.; de Jong, F.; Caron, E. A Survey of event extraction methods from text for decision support systems. *Decis. Support Syst.* **2016**, *85*, 12–22. [CrossRef]
16. Lefever, E.; Hoste, V. A Classification-based Approach to Economic Event Detection in Dutch News Text. In Proceedings of the International Conference on Language Resources and Evaluation; European Language Resources Association (ELRA): Paris, France, 2016; pp. 330–335.
17. Sheng, J.; Lan, H. Business failure and mass media: An analysis of media exposure in the context of delisting event. *J. Bus. Res.* **2019**, *97*, 316–323. [CrossRef]
18. Hartmann, J.; Huppertz, J.; Schamp, C.; Heitmann, M. Comparing automated text classification methods. *Int. J. Res. Mark.* **2019**, *36*, 20–38. [CrossRef]
19. Sprugnoli, R.; Tonelli, S. One, no one and one hundred thousand events: Defining and processing events in an inter-disciplinary perspective. *Nat. Lang. Eng.* **2017**, *23*, 485–506. [CrossRef]
20. Kunneman, F.; Van Den Bosch, A. Open-domain extraction of future events from Twitter. *Nat. Lang. Eng.* **2016**, *22*, 655–686. [CrossRef]
21. Salminen, J.; Yoganathan, V.; Corporan, J.; Jansen, B.J.; Jung, S.-G. Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type. *J. Bus. Res.* **2019**, *101*, 203–217. [CrossRef]
22. Glance, N.; Hurst, M.; Nigam, K.; Siegler, M.; Stockton, R.; Tomokiyo, T. Deriving marketing intelligence from online discussion. In Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, Chicago, IL, USA, 21–24 August 2005; pp. 419–428.
23. Netzer, O.; Feldman, R.; Goldenberg, J.; Fresko, M. Mine Your Own Business: Market-Structure Surveillance through Text Mining. *Mark. Sci.* **2012**, *31*, 521–543. [CrossRef]
24. Xu, K.; Liao, S.S.; Li, J.; Song, Y. Mining comparative opinions from customer reviews for Competitive Intelligence. *Decis. Support Syst.* **2011**, *50*, 743–754. [CrossRef]
25. Wu, S.; Rand, W.; Raschid, L. Recommendations in social media for brand monitoring. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 345–348.

26. Klostermann, J.; Plumeyer, A.; Böger, D.; Decker, R. Extracting brand information from social networks: Integrating image, text, and social tagging data. *Int. J. Res. Mark.* **2018**, *35*, 538–556. [CrossRef]
27. Hays, S.; Page, S.J.; Buhalis, D. Social media as a destination marketing tool: Its use by national tourism organisations. *Curr. Issues Tour.* **2013**, *16*, 211–239. [CrossRef]
28. Schivinski, B. Eliciting brand-related social media engagement: A conditional inference tree framework. *J. Bus. Res.* **2021**, *130*, 594–602. [CrossRef]
29. Onishi, H.; Manchanda, P. Marketing activity, blogging and sales. *Int. J. Res. Mark.* **2012**, *29*, 221–234. [CrossRef]
30. Colicev, A.; Kumar, A.; O'Connor, P. Modeling the relationship between firm and user generated content and the stages of the marketing funnel. *Int. J. Res. Mark.* **2019**, *36*, 100–116. [CrossRef]
31. Kaiser, C.; Ahuvia, A.; Rauschnabel, P.A.; Wimble, M. Social media monitoring: What can marketers learn from Facebook brand photos? *J. Bus. Res.* **2020**, *117*, 707–717. [CrossRef]
32. Shone, A.; Parry, B. *Successful Event Management: A Practical Handbook*, 4th ed.; Cengage Learning: Boston, MA, USA, 2013.
33. Morgeson, F.P.; Mitchell, T.R.; Liu, D. Event system theory: An event-oriented approach to the organizational sciences. *Acad. Manag. Rev.* **2015**, *40*, 515–537. [CrossRef]
34. McAfee, R.P.; McMillan, J. Competition and Game Theory. *J. Mark. Res.* **1996**, *33*, 263–267. [CrossRef]
35. Ailawadi, K.L.; Kopalle, P.K.; Neslin, S.A. Predicting Competitive Response to a Major Policy Change: Combining Game-Theoretic and Empirical Analyses. *Mark. Sci.* **2005**, *24*, 12–24. [CrossRef]
36. Chevalier-Roignant, B.; Trigeorgis, L. *Competitive Strategy Options and Games*; The MIT Press: Cambridge, MA, USA, 2011.
37. Han, S.; Ye, Y.; Fu, X.; Chen, Z. Category role aided market segmentation approach to convenience store chain category management. *Decis. Support Syst.* **2014**, *57*, 296–308. [CrossRef]
38. Smith, W.R. Product Differentiation and Market Segmentation as Alternative Marketing Strategies. *J. Mark.* **1956**, *21*, 3–8. [CrossRef]
39. Montgomery, D.B.; Moore, M.C.; Urbany, J.E. Reasoning About Competitive Reactions: Evidence from Executives. *Mark. Sci.* **2005**, *24*, 138–149. [CrossRef]
40. Bergen, M.; Peteraf, M. Competitor identification and competitor analysis: A broad-based managerial approach. *Manag. Decis. Econ.* **2002**, *23*, 157–169. [CrossRef]
41. Bloodgood, J.M.; Bauerschmidt, A. Competitive Analysis: Do Managers Accurately Compare Their Firms to Competitors? *J. Manag. Issues* **2002**, *14*, 418–434.
42. Clark, B.H. Managerial identification of competitors: Accuracy and performance consequences. *J. Strateg. Mark.* **2011**, *19*, 209–227. [CrossRef]
43. Gilbert, R.A. Bank Market Structure and Competition: A Survey. *J. Money Credit Bank.* **1984**, *16*, 617–645. [CrossRef]
44. Reger, R.K.; Palmer, T.B. Managerial Categorization of Competitors: Using Old Maps to Navigate New Environments. *Organ. Sci.* **1996**, *7*, 22–39. [CrossRef]
45. Shubik, M.; Levitan, R. *Market Structure and Behavior*; Harvard University Press: Cambridge, MA, USA, 2013.
46. Ciliberto, F.; Murry, C.; Tamer, E. Market Structure and Competition in Airline Markets. *J. Political Econ.* **2021**, *129*, 2995–3038. [CrossRef]
47. Hooley, G.; Greenley, G. The resource underpinnings of competitive positions. *J. Strateg. Mark.* **2005**, *13*, 93–116. [CrossRef]
48. Urban, G.L.; Johnson, P.L.; Hauser, J.R. Testing Competitive Market Structures. *Mark. Sci.* **1984**, *3*, 83–112. [CrossRef]
49. Fabrizio, C.M.; Kaczam, F.; de Moura, G.L.; da Silva, L.S.C.V.; da Silva, W.V.; da Veiga, C.P. Competitive advantage and dynamic capability in small and medium-sized enterprises: A systematic literature review and future research directions. *Rev. Manag. Sci.* **2022**, *16*, 617–648. [CrossRef]
50. Peteraf, M.A.; Bergen, M.E. Scanning dynamic competitive landscapes: A market-based and resource-based framework. *Strateg. Manag. J.* **2003**, *24*, 1027–1041. [CrossRef]
51. Roberts, P.W.; Amit, R. The Dynamics of Innovative Activity and Competitive Advantage: The Case of Australian Retail Banking, 1981 to 1995. *Organ. Sci.* **2003**, *14*, 107–122. [CrossRef]
52. Liu, Y.; Jiang, C.; Zhao, H. Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media. *Decis. Support Syst.* **2019**, *123*, 113079. [CrossRef]
53. Liu, Y.; Qian, Y.; Jiang, Y.; Shang, J. Using favorite data to analyze asymmetric competition: Machine learning models. *Eur. J. Oper. Res.* **2020**, *287*, 600–615. [CrossRef]
54. Zhao, K.; Cong, G.; Chin, J.-Y.; Wen, R. Exploring market competition over topics in spatio-temporal document collections. *VLDB J.* **2019**, *28*, 123–145. [CrossRef]
55. Jayachandran, S.; Gimeno, J.; Varadarajan, P.R. The Theory of Multimarket Competition: A Synthesis and Implications for Marketing Strategy. *J. Mark.* **1999**, *63*, 49–66. [CrossRef]
56. Fülöp, M.; Orosz, G. State of the Art in Competition Research. In *Emerging Trends in the Social and Behavioral Sciences*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2015; pp. 1–16.
57. Bai, X.; Marsden, J.R.; Ross, W.T.; Wang, G. How e-WOM and local competition drive local retailers' decisions about daily deal offerings. *Decis. Support Syst.* **2017**, *101*, 82–94. [CrossRef]
58. Mikolov, T.; Chen, K.; Corrado, G.S.; Dean, J. Efficient Estimation of Word Representations in Vector Space. In Proceedings of the International Conference on Learning Representations Workshop Track, Scottsdale, AZ, USA, 2–4 May 2013.

59. Ji, H.; Grishman, R. Refining event extraction through cross-document inference. In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-08: HLT), Columbus, OH, USA, 15–20 June 2008; pp. 254–262.
60. Li, X.; Nguyen, T.H.; Cao, K.; Grishman, R. Improving Event Detection with Abstract Meaning Representation. In Proceedings of the First Workshop on Computing News Storylines, Beijing, China, 26–31 July 2015; pp. 11–15.
61. Wan, Y.; Peng, Z.; Wang, Y.; Zhang, Y.; Gao, J.; Ma, B. Influencing factors and mechanism of doctor consultation volume on online medical consultation platforms based on physician review analysis. *Internet Res.* **2021**, *31*, 2055–2075. [CrossRef]
62. Jacobs, G.; Lefever, E.; Hoste, V. Economic Event Detection in Company-Specific News Text. In Proceedings of the First Workshop on Economics and Natural Language Processing, Melbourne, Australia, 20 July 2018; pp. 1–10.
63. Wang, X.; Song, Y. Viral misinformation and echo chambers: The diffusion of rumors about genetically modified organisms on social media. *Internet Res.* **2020**, *30*, 1547–1564. [CrossRef]
64. Lo, W.H.; Lam, B.S.Y.; Cheung, M.M.F. The Dynamics of Political Elections: A Big Data Analysis of Intermedia Framing between Social Media and News Media. *Soc. Sci. Comput. Rev.* **2021**, *39*, 627–647. [CrossRef]
65. Sun, C.; Qiu, X.; Xu, Y.; Huang, X. How to Fine-Tune BERT for Text Classification? In Proceedings of the China National Conference on Chinese Computational Linguistics; Springer: Cham, Switzerland, 2019; pp. 194–206.
66. O'Brien, R.; Ishwaran, H. A random forests quantile classifier for class imbalanced data. *Pattern Recogn.* **2019**, *90*, 232–249. [CrossRef]
67. Yuan, H.; Deng, W. Doctor recommendation on healthcare consultation platforms: An integrated framework of knowledge graph and deep learning. *Internet Res.* **2022**, *32*, 454–476. [CrossRef]
68. Czepiel, J.A.; Kerin, R.A. Competitor analysis. In *Handbook of Marketing Strategy*; Shankar, V., Carpenter, G.S., Eds.; Edward Elgar Publishing: Northampton, MA, USA, 2011; pp. 41–57.
69. Fan, W.; Gordon, M.D. The power of social media analytics. *Commun. ACM* **2014**, *57*, 74–81. [CrossRef]