Article

SemConvTree: Semantic Convolutional Quadtrees for Multi-Scale Event Detection in Smart City

Mikhail Andeevich Kovalchuk * , **Anastasiia Filatova** , **Aleksei Korneev** , **Mariia Koreneva, Denis Nasonov *** , **Aleksandr Voskresenskii and Alexander Boukhanovsky**

Faculty of Digital Transformations, Industrial AI Research Lab, ITMO University, Kronverkskii 49,
197101 Saint-Petersburg, Russia; aafilatova@itmo.ru (A.F.); alekseiresearch@gmail.com (A.K.);
mkovalchuk@itmo.ru (M.K.); resurrectiontent@itmo.ru (A.V.); avbukhanovskii@itmo.ru (A.B.)
* Correspondence: mikhail.a.kovalchuk@ya.ru (M.A.K.); denis.nasonov@gmail.com or dnasonov@itmo.ru (D.N.)

Highlights:

What are the main findings?

- Enhanced Event Detection Accuracy: The introduction of the SemConvTree model, which integrates improved versions of BERTopic, TSB-ARTM, and SBert-Zero-Shot, enables a significant enhancement in the detection accuracy of urban events. The model's ability to incorporate semantic analysis along with statistical evaluations allows for discerning and categorizing events from social media data more precisely. This results in approximately a 40% increase in the F1-score for event detection compared to previous methods.

- Semantic Analysis for Event Identification: The SemConvTree model leverages semi-supervised learning techniques to analyze the semantic content of social media posts. This approach helps in understanding the nuanced contexts of urban events, improving the identification process. The model not only recognizes the occurrence of events but also categorizes them into meaningful groups based on their semantic characteristics, which is crucial for effective urban management and planning.

What are the implications of the main findings?

- The increased accuracy in event detection ensures that urban planners and emergency services can respond more effectively to both planned and unplanned urban events. More accurate data leads to better resource allocation, ensuring that services are deployed where they are most needed. This could lead to enhanced safety, improved traffic management, and better crowd control during events, ultimately enhancing urban living conditions.

- By effectively categorizing urban events based on their semantic characteristics, city administrators can gain insights into the types of events that are prevalent in different areas of the city. This can inform more targeted community engagement strategies, help in the planning of public services and facilities, and ensure that urban policies are closely aligned with the actual dynamics of the city. Additionally, this can aid in long-term urban development strategies by identifying evolving trends and shifts in urban activity patterns.

Abstract: The digital world is increasingly permeating our reality, creating a significant reflection of the processes and activities occurring in smart cities. Such activities include well-known urban events, celebrations, and those with a very local character. These widespread events have a significant influence on shaping the spirit and atmosphere of urban environments. This work presents SemConvTree, an enhanced semantic version of the ConvTree algorithm. It incorporates the semantic component of data through semi-supervised learning of a topic modeling ensemble, which consists of improved models: BERTopic, TSB-ARTM, and SBert-Zero-Shot. We also present an improved event search algorithm based on both statistical evaluations and semantic analysis of posts. This algorithm allows for fine-tuning the mechanism of discovering the required entities with the specified particularity (such as a particular topic). Experimental studies were conducted within the area of New York City. They showed an improvement in the detection of posts devoted to events (about 40% higher f1-score) due to the accurate handling of events of different scales. These results suggest

the long-term potential for creating a semantic platform for the analysis and monitoring of urban events in the future.

1. Introduction

Social networks have become vitally important to many people. Over half of the world's population [1] uses social media to express emotions, share their thoughts and support social relationships. Regular users share information about their daily lives, while organizers of mass events broadcast public content through official pages. This trend has made social networks a valuable data source for various tasks related to analyzing urban processes in Smart Cities [2]. Such data allows the creation of recommendation or crime monitoring systems based on detecting multi-scale events. Moreover, previous studies have revealed [3,4] that information about remarkable events, such as hurricanes, earthquakes and floods, appears in social networks faster than in traditional media. Among the diversity of social media, Instagram and Twitter are the most suitable for event detection tasks [5]. Both are highly widespread and continue to increase in popularity [6]. Posts on these platforms may contain text, images, or videos, and some are also geotagged with specific locations and timestamps, facilitating event identification. However, data from social media contains a large amount of noise: posts devoted to food, clothes, spam or advertisements do not reflect information about any event and lead to poor results [7].

In the scope of this work, we rely on one of the most common definitions of an event [8] and adapt it to the urban context:

Definition 1. *An event is a significant thing that happens at some specific time and place.*

Definition 2. *An urban event is an event that happens in an urban environment.*

In order to detect new events, authors of the most advanced solutions usually use historical data and predict the number of new posts for specific locations [9–11]. These algorithms identify event candidates when the actual number of posts exceeds the predicted value. The main disadvantage of this approach is the need to set a high sensitivity threshold and the impossibility of noise processing. This approach yields acceptable results for detecting events represented by many posts, such as stadium soccer games or large concerts. However, these methods often lack the sensitivity needed to detect less popular events that are represented by only a few posts. In order to distinguish miscellaneous types of events, we introduce the two following definitions:

Definition 3. *A high-scale event is an event that is represented in data by more than three posts.*

Definition 4. *A low-scale event is an event that is not a high-scale event.*

We show that semantic analysis of the social media content can help to address the challenging task of multi-scale event detection. Additionally, it can filter out noisy posts that complicate event detection in general.

This work aims to focus on low-scale event detection. Incorporating low-scale events alongside high-scale ones enables precise online detection of urban events. This approach allows for the creation of a real-time, dynamic city map, facilitating the study of various urban processes related to citizens' leisure, work, and recreation, due to the increased number of detected events.

It is important to note that low-scale events may reflect personal aspects of citizens' lives, even though the analyzed sources are publicly available. Therefore, the extracted

knowledge should be handled with care and discretion. Nevertheless, analyzing the distribution of such events within the city can help to provide the appropriate infrastructure for locations where such events are held. For instance, an event detection approach could facilitate the identification of parks frequently used for wedding photoshoots. This information could guide decisions to construct special facilities or cordon off potentially hazardous areas in these parks. Similarly, the proposed algorithm would help handle such low-scale events as car accidents. Extracting this knowledge might lead to traffic stream reorganizing and could have a strong positive social impact.

Our urban event detection approach is based on an anomaly detection mechanism, which identifies the abnormal number of posts relative to historical data. In order to achieve this, we have designed and implemented an early event detection software package comprising the following components:

- Data collectors;
- The semantics extraction and ranking module;
- The adaptive mesh generation module;
- The anomaly detection module;
- The anomaly filtering and event linking module.

Our main contribution is the development of an algorithm capable of detecting urban events at various scales. This opens up opportunities for investigating urban processes dynamically at a granular level. Compared to other event detection approaches, our algorithm significantly increases the number of detectable events, from tens to hundreds of events per day, as demonstrated in the case study of New York City during the experimental evaluation.

2. Related Works

Social networks have become an essential part of our daily lives, documenting people's movements, experiences, and emotions, including information about events in which they are involved. Research has shown that real-life events, including personal ones, leave their mark on social media [12,13]. This occurs either through direct mentions of the event or through changes in users' behavioral patterns on social media platforms. Popular social networks worldwide, such as Twitter and Instagram [5,6,14,15], enable users to share geo-tagged posts with text and pictures. These geo-tagged posts are highly valuable as they can provide significant insights into both high- and low-scale events in users' lives. The online nature of social networks has allowed researchers to identify and track many real-life events, from high-scale events like natural disasters such as hurricanes and floods [16,17] to lower-scale events like traffic accidents and protests [10,18,19].

Event detection through social media data analysis has been the focus of numerous research papers [14,20–24]. One of the popular techniques in this field is anomaly detection, which identifies unusual or unexpected patterns of behavior that deviate from a certain norm. In the context of social networks, anomalies can be caused by various factors, including sudden changes in user behavior and the emergence of new topics or trends. As a result, a wide range of urban events could be captured by the detection of anomalies, i.e., monitoring the social media activity to identify deviation patterns.

2.1. Frequency-Based Methods

Early research in this field often employed various frequency-based techniques to detect anomalies or novelties in data streams [25–28]. Many authors were interested in using clustering or classification of words to identify new events by the usage of an abnormal lexicon [29]. Various algorithms were used to find anomalous increases in word usage, such as TF-IDF [30–32], N-grams [33,34], and wavelet analysis [35]. In [30], the authors used TF-IDF and hidden Markov mode to find and choose keywords from posts and then graph structures to identify connected components that were considered events. This paper focused solely on keywords, overlooking the spatio-temporal nature of events. The study in [36] proposed searching for clusters of data in both time and space as they expect users

to post more than usual to describe the event. However, they did not consider the content of individual posts. Most authors combined the usage of text features with spatial and temporal features [14,31,37,38]. In [37], researchers used a sliding window and tracked shifts in IDF value against the average value to identify events using temporal information to define a window size. The authors in [31] employed a TF-IDF technique to compare a new post to all previous posts in the collection. Several papers utilized the concept of 'bursts of activity'. Kleinberg [39] proposed an algorithm based on a statistical framework that models the stream using an infinite-state automaton based on hidden Markov models. This article had a significant impact on the field of event detection, inspiring several subsequent research papers [38,40–43]. Zhou et al. [44] utilized spatial features by employing a joint distribution of keywords, named entities, and locations to identify events.

Lee et al. [45] presented a method that utilized both spatial and temporal data features to model the normal state of a target area and its movement patterns. They identified events as deviations from these normal behaviors. However, this method does not filter the noise in social media. Visheratin et al. [11] addressed this problem by utilizing the concept of the city's normal state while proposing a more effective algorithm for area splitting using adaptive geogrids. This approach, combined with consideration of seasonal patterns in user behavior, allowed the authors to achieve high detection accuracy and outperform other methods. Several other articles use a historical grid [46,47] to increase anomaly detection accuracy. By interacting with these systems, users can adjust the map scale, control the timeline, and obtain results with high spatial and temporal granularity. Moreover, this approach allows ranking, which aims to search for burst events given particular time and space limitations.

2.2. Modern Techniques

2.2.1. Modern NLP

Extracting semantic information from social media data is a crucial component of event detection research. By analyzing the meaning and context of words and phrases in social media posts, event detection models can better perceive the content and thus more accurately identify events. Therefore, in this section, we will focus on various NLP methods used for analyzing social media posts and detecting events.

One commonly used approach in NLP is word embeddings [15,15]. Word embeddings are vector representations of words in a high-dimensional space. In this space, words with similar meanings are located close to each other, helping to capture semantic relationships between them. There has been a lot of research on word embeddings in recent years [48–50]. In 2013, Mikolov et al. [48] introduced the Word2Vec algorithm, which uses a neural network to learn word embeddings from large text corpora. Word2Vec has been widely used and has become not only the baseline for many NLP tasks but was also used for event detection [9,51,52]. For instance, in Embed2Detect [52], authors presented an approach based on word embeddings and hierarchical agglomerative clustering. GeoBurst+ [9], an improvement over GeoBurst [7], is one of the most well-known approaches. It performs keyword embedding to capture the subtle semantics of tweet messages. Those systems enable practical and real-time event detection from geo-tagged tweet streams. They use keyword clustering of tweets, followed by the classification of clusters by topics and spatio-temporal coordinates.

Recently, transformer-based models such as BERT, RoBERTa, and SBERT [53–55] have achieved state-of-the-art performance on various NLP tasks, including event detection in social media posts. These models learn to represent words and phrases within the context of the surrounding text, capturing complex semantic relationships. This capability makes them particularly effective for identifying events and their related information. BERT (Bidirectional Encoder Representations from Transformers) is a particularly popular and effective approach among these models. In recent years, there have been several studies demonstrating the effectiveness of BERT for event detection in social media posts. For instance, Wei and Wang [56] combined BERT with Recurrent Neural Networks to detect

events in Chinese text, while Neruda and Winarko [19] integrated BERT with CNN to detect traffic events using Twitter data. Huang et al. [57] proposed combining BERT with an attention-based bidirectional long short-term memory model (BERT-Att-BiLSTM) for detecting emergency-related posts.

Some contemporary methods utilize dependency graphs [58], which represent the grammatical structure of a sentence as a directed graph. In these graphs, words are represented as nodes, and edges denote grammatical relationships. Researchers in [59,60] approached event detection by constructing a dependency graph for each sentence and employing Graph Convolution Networks (GCNs) [61] to extract events. The MOGANED model [62] utilizes Graph Attention Networks [63] to derive multi-order word representations, achieving high precision and F1 scores. Dutta et al. [64] proposed an improvement over previous homogeneous models using Graph Transformer Networks (GTN). This approach produces a heterogeneous graph, whose output is then classified by a multi-layer perceptron with attention, demonstrating improved F1 scores and approaching state-of-the-art performance. However, all those models only use textual features of the data to search for events.

Despite large language models (LLMs) dominating overall NLP benchmarks, their use in the context of this study might be impractical for several reasons [65]. Firstly, LLMs require significant computational resources and storage, which may not be feasible for urban analytics applications, especially in smaller research projects without direct economic impact. Secondly, LLMs often require extensive fine-tuning or preparation of training datasets, particularly when applied to the specific nuances of urban event detection. Additionally, LLMs' black-box nature can produce results that are difficult to interpret or control, posing challenges for accountability in urban data analytics. Moreover, the high costs of deploying and maintaining LLMs can be prohibitive, especially for smaller research projects or municipal budgets.

2.2.2. Multimodal Approaches

Event detection approaches for social networks primarily focus on textual information. However, multimedia items from social media often contain linked metadata and additional features such as images, videos, and audio. Additional features can be derived from texts as well, including emojis, errors, length, and tone. Considering all these features could significantly enhance event extraction from social network data. However, the resulting features are typically heterogeneous, making them challenging to combine for analysis. To cluster such multimedia data, two primary methods exist: early fusion, where features are combined before processing, and late fusion, where each modality is processed separately and then combined [66,67].

In a recent paper, Cui et al. [68] introduced a new technique called Multi-View Graph Attention Network (MVGAN) for event detection in social networks. This approach enhances event understanding by aggregating neighboring nodes and fusing multiple views in a social event graph. The authors first construct a heterogeneous graph by incorporating hashtags, then use graph convolutional networks to learn event-specific representations from different perspectives, such as text semantics and time distribution. Jony et al. [69] employed a late fusion approach to combine visual and scenic image features with textual metadata for identifying flooding in Flickr images. They then utilized the Direct Backpropagation method to train a neural network for classification.

Petkos et al. [70] predicted whether two objects belong to the same event by utilizing multimodal representation vectors derived from picture metadata and computing distances between them. Schinas et al. [71] used this method to construct a multimodal graph-based approach for event detection. The authors constructed a graph by comparing images, where edges indicate that two images belong to the same event, and then applied graph-based clustering to identify events. Tong et al. [72] introduced a new method called Dual Recurrent Multimodal Model (DRMM) for effective feature aggregation of images and

sentences. DRMM employs pre-trained BERT and ResNet models to encode sentences and images, respectively, and uses alternating dual attention for feature selection.

Other researchers have reported using various feature fusions for event detection in social media data, such as time, scene, and objects [73]; geotags, photos, and text [74]; and text and image metadata of photos [75].

Sentiment analysis is another useful feature for event detection. It involves analyzing the emotional tone of social media posts to infer individuals' sentiments toward specific events or topics. The general idea is that during significant events, people may express strong emotions on social media, which can be used to detect and track these events. For instance, Cresci et al. [17] employed sentiment analysis to train Support Vector Machines for identifying and extracting disaster-related posts from Twitter. Other linguistic features used for training include raw and lexical textual data, syntactic features, lexical expansion features, and sentiment analysis. Sufi [16] employed sentiment analysis, among other methods, to identify and locate natural disasters such as floods, bushfires, and earthquakes using social media data. Ali et al. [76] combined sentiment detection techniques with the FastText model and Bi-LSTM for traffic event detection and condition analysis, achieving high-accuracy results.

2.2.3. Filtering Noise

To accurately identify events from social media posts, distinguishing between relevant posts and unrelated noise is crucial. Therefore, in addition to analyzing raw text descriptions using embedding techniques, we applied topic modeling methods to classify the posts. Topic modeling helps identify underlying themes discussed in a set of posts, which can then filter out irrelevant content unrelated to the event of interest. By combining topic modeling with embedding analysis, we can enhance event detection accuracy by reducing data noise.

Latent Dirichlet Allocation (LDA) is a popular topic modeling algorithm for identifying latent topics in a set of posts [77]. LDA assumes each post is a mixture of several topics, with each topic characterized by a probability distribution over the words in the posts. By learning these probability distributions, LDA can identify the most probable topics discussed. Applying LDA to event detection has shown promising results. For instance, Sokolova et al. [78] successfully employed LDA to detect social events in Kenya.

Some researchers employ graph-based approaches for topic detection. These models represent the frequency of word co-occurrence in documents, which can be partitioned into multiple communities. Each community represents a distinct topic, and documents within the corpus are assigned to their most closely related topic. For example, Zhang et al. [79] combined a word co-occurrence graph with a semantic information graph established using LDA. Choi et al. [80] developed a method that creates a keyword graph from social media data and detects local events using a geographic dictionary.

In this work, we apply a more recent topic modeling approach called BigARTM [81], based on Additive Regularization of Topic Models (ARTM) [82]. BigARTM offers a flexible and scalable approach that combines post descriptions with other content types, such as spatial and temporal coordinates. This concept employs principles similar to the multimodal embedding from the TrioVecEvent framework [83]. We also implemented the BERTopic technique [84], which uses the SBERT framework for text embedding, followed by HDBSCAN clustering of embeddings and a class-based TF-IDF to model topic evolution over time.

2.3. Low-Scale Events

Event detection through social media data analysis has been extensively studied, with Afyouni et al. [21] publishing one of the most comprehensive surveys to date, examining hundreds of articles from 2010 to 2021. The survey reveals diverse approaches used by researchers for event detection. However, despite extensive research in this field, there remains a scarcity of articles dedicated to low-scale event detection. Low-scale events

are small, localized occurrences at a specific time and place, such as car accidents, street protests, or localized natural disasters. Detecting and tracking these events is crucial due to their significant impact on individuals and communities.

Krumm and Horvitz [10] identify low-scale events by analyzing tweet time series across different spatial scales and time frames. Their Eyewitness approach achieved high precision using a decision tree classifier but was later outperformed by subsequent research.

GeoBurst+ [9] enabled low-scale event detection, albeit with a modest F1 score. Nevertheless, experiments by the authors demonstrate that GeoBurst+ significantly outperforms GeoBurst and other baseline algorithms.

In 2019, Wei et al. [85] introduced a methodology for automatically detecting recent events from geo-tagged tweets. This prediction method leverages Twitter's general dynamics and location-specific historical data, improving event detection and outperforming GeoBurst and Eyewitness. It employs historical and spatial grids, achieving superior results by using long short-term memory (LSTM) to predict future tweet volumes.

The EvenTweet system [18], developed by Abdelhaq et al. in 2019, detects and tracks low-scale events such as natural disasters, protests, and traffic accidents on Twitter. It identifies bursty words and performs geo-clustering to find anomalous activity bursts. However, EvenTweet's heavy reliance on keyword-based approaches may limit its ability to detect events without clear keywords or hashtags.

In this work, we focus on developing a new approach for handling low-scale events. Our method is based on the historical grid and convolutional quadtrees concept proposed by Visheratin et al. [11], utilizing their key idea of creating a grid with cells that depend on the volume of posts.

3. Semantic Convolutional Quadtree

Anomaly detection approaches vary across research areas, including urban environment analysis and event detection tasks. The ConvTree algorithm addresses anomaly detection, prediction, and clustering problems by processing geospatial and temporal data from social networks. Visheratin et al. [11] demonstrated how to construct a quadtree based on a convolution mechanism, effectively distributing influence between neighboring areas and using frequency characteristics to find anomalies in urban social network data. However, the presented method has limitations in sensitivity and event detection ranges. Analysis of this algorithm revealed high sensitivity to noise, necessitating low sensitivity thresholds to maintain accuracy. Consequently, while effective for high-scale events, it is unsuitable for capturing low-scale ones. In this section, we describe an algorithm that overcomes these limitations and provides semantic interpretation of detected events.

3.1. ConvTree

Traditional frequency-based anomaly detection algorithms struggle to detect events at multiple scales due to predefined splitting criteria unrelated to data. Visheratin et al. [11] proposed an efficient approach using convolutional quadtrees and adaptive geogrids for event detection in geo-data. By employing the Convolutional Quadtree (ConvTree), an advanced quadtree variant, the authors leveraged spatial distribution of data points to subdivide the target region. This innovative use of quadtree enables accurate differentiation between high and low posting frequency regions, enhancing the event detection algorithm's sensitivity.

The authors chose the quadtree data structure for its ability to cover the entire target area while offering scalability for the precise localization of diverse events across various scales. However, standard quadtrees do not consider spatial data distribution during splitting. To address this limitation, the authors enhanced the quadtree by incorporating split point search functionality based on convolutional neural networks (CNNs).

To construct a convolutional quadtree, the authors initially divided the target area into a uniform grid and created a corresponding matrix. Each matrix element is assigned a value equal to the weighted sum of points in the respective grid cell region. They then

performed sequential convolutions on this initial matrix. These convolutions account for the inter-influence of neighboring elements and mitigate adjacency effects resulting from the target area's matrix division.

To select a split point, the authors calculated a descending gradient on the normalized output matrix G, starting from the maximum point. At each step k, they selected an array X_k of values at distance k from the maximum point along any axis and greater than or equal to the gradient value ∇_{k-1} from the previous step. This process continues until X_k becomes empty:

$$X_k = \left\{ g_{i,j} \in G \mid (|i_m - i| = k \vee |j_m - j| = k) \wedge g_{i,j} \geq \nabla_{k-1} \right\}. \tag{1}$$

The gradient is calculated using the equation:

$$\nabla_k = \alpha \times \frac{1}{\|X_k\|} \times \sum_{i=1}^{\|X_k\|} x_i. \tag{2}$$

Here, α is a threshold parameter controlling the algorithm's sensitivity, where $\nabla_0 = \alpha$ as the convolved matrix is normalized. The optimal α value was empirically determined to be 0.8. The authors then selected the coordinates of the square's vertex formed on step $k-1$ closest to the matrix center. This point splits the area into four child elements, and the convolution and split point search process is recursively applied to each child element.

The event detection pipeline consists of four key steps: data collection, historical grid generation, anomaly search, and event detection (Figure 1). In the data collection phase, the authors gathered extensive social network data over at least one year to enable comprehensive statistical analysis. They derived baseline hashtags for each cell, considering every combination of month, day type, and hour, establishing behavioral norms for the city.

In the anomaly search phase, the system compares current posting activity in each geogrid cell against the baseline value derived from historical data. The algorithm then proceeds to the event detection phase for cells exhibiting anomalous behavior. Here, hashtag usage analysis is performed to identify events. This involves constructing a post graph within the anomalous cell, with edges representing common hashtags, and then splitting the graph into connected components. A connected component is considered an event if its size exceeds a predetermined threshold value.
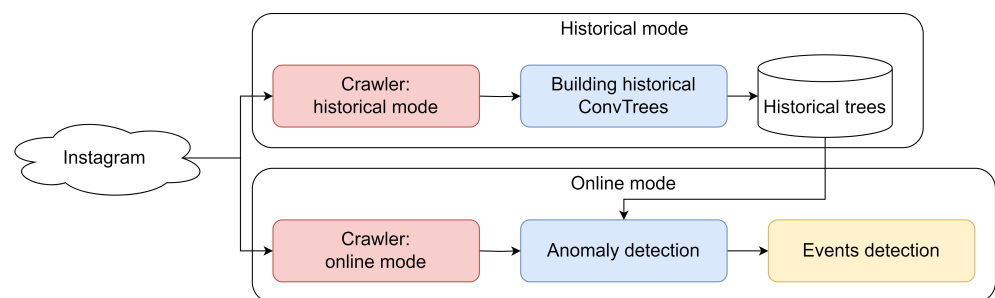

Figure 1. Scheme of events detection pipeline of the ConvTree algorithm.

The authors validated their proposed method through experiments on real-world datasets. The algorithm effectively learns appropriate underlying distributions for diverse datasets and detects anomalous behavior. Experiments demonstrate the method's capability to accurately detect events of various scales, surpassing baseline algorithms in spatiotemporal precision.

However, the method has limitations regarding sensitivity and event detection ranges. An analysis revealed its susceptibility to noise, requiring low sensitivity thresholds to maintain high accuracy. While effective for high-scale events, it may not adequately capture low-scale ones. The following section introduces an algorithm that addresses these limitations and enables semantic interpretation of detected events.

3.2. Semantic-Based Model for Anomaly Detection

The task of event detection most often arises in large cities or districts. Therefore, let us consider the following geogrid G for the investigated urban area:

$$G = \{ <u,v>: u \in \{1 \ldots N_{lat}\}, v \in \{1 \ldots N_{lon}\}\}, \tag{3}$$

where $N_{lat} = \frac{lat_{max} - lat_{min}}{step}$, $N_{lon} = \frac{lon_{max} - lon_{min}}{step}$, and (lat_{min}, lon_{min}) and (lat_{max}, lon_{max}) are latitude's and longitude's min and max values. Geogrid cell $<u,v>$ corresponds to the geographic area, which covered latitude from $lat_{min} + (u-1) \cdot step$ to $lat_{min} + u \cdot step$ and longitude from $lon_{min} + (v-1) \cdot step$ to $lon_{min} + v \cdot step$.

Let us introduce the concept of discrete-time, represented by a set of hourly intervals T. We also introduce the concept of time periods τ as unions of subsets of hourly intervals $t \in T$, which are chosen according to the selected strategy of the algorithm. In our case, we consider time periods, corresponding to each of the 24 h for weekdays and weekends separately for each of the 12 months of the year, constructed during the original tree formation. Thus, the number of time periods τ we consider, following the described logic of aggregation of hourly intervals, is 576.

The set of documents $D^{\tau}_{<u,v>}$ related to the time period τ and the geographic area, represented by grid cell $<u,v>$ is described as follows:

$$D^{\tau}_{<u,v>} = \{d : \lambda(d) = <u,v>, \phi(d) = \tau\}, \tag{4}$$

where $\lambda(d) = <u_d, v_d>$ is a geogrid cell to which the geotag of document d belongs, and $\phi(d) = \tau_d$ is the time zone to which the document's timestamp belongs.

The semantics of each document $d \in D^{\tau}_{<u,v>}$ are described through an extracted and generated vector of L topics: $s^d = \left(s^d_1, s^d_2, \ldots, s^d_L\right)$, where $s^d_l \in [0,1] \forall l \in 1 \ldots L$ and $\sum_{l=1}^{L} s^d_l = 1$.

For each geogrig cell $<u,v>$, we calculate aggregated semantic (topic) vector $\overline{S}^{\tau}_{<u,v>}$ aggregated though time period τ:

$$\overline{S}^{\tau}_{<u,v>} = (\overline{s_1}, \overline{s_2}, \ldots, \overline{s_L}), \tag{5}$$

where

$$\overline{s_l} = \frac{1}{\|D^{\tau}_{<u,v>}\|} \sum_{d=1}^{\|D^{\tau}_{<u,v>}\|} s^d_l, l \in 1 \ldots L. \tag{6}$$

The original ConvTree algorithm was used to detect anomalies by partitioning the geospace into regions g that are subsets of a grid G, each characterized by the less than μ number of posts per time period τ: $g \in P(G) : \|D^{\tau}_g\| \leq \mu$. The value of the hyperparameter μ was chosen empirically by the authors and was equal to 12. To test the effectiveness of the semantic-based model, we should compare the results of identifying anomalies as bursts of post frequency above the threshold value $\mu + 2\sigma$ (according to the original ConvTree algorithm) and the results of anomaly detection using further introduced semantic threshold $\mu^{\tau}_{<u,v>}$.

In SemConvTree, we transformed the frequency attribute μ into the weighted sum of semantic attributes:

$$\mu^{\tau}_{<u,v>} = \sum_{d=1}^{\|D^{\tau}_{<u,v>}\|} \mu^d_{<u,v>}, \tag{7}$$

each of which is calculated using the following formula:

$$\mu^d_{<u,v>} = \frac{1}{L} \sum_{l=1}^{L} e^{\frac{1}{\alpha}\left(\frac{s^d_l}{\overline{s_l}}\right)^{\beta_l}} (s^d_l - \overline{s_l}) \tag{8}$$

Here, α is the adjustment factor, $\alpha \in [0,1]$. It helps to highlight differences in the distribution of topics, and β_l allows to increase or decrease the importance of the topic l in the case of a directional event detection problem. This coefficient works as a regularizer and allows for additional tuning to recognize events on different topics.

In this work, we used the same values of the β_l coefficient for different topics when conducting experimental research to solve the problem of event detection of arbitrary topics and to compare the results with the other methods. Nevertheless, investigating the possibility of adjusting the influence of topics on the event detection task and conducting experiments in this area is on the list of tasks for the near future.

This extension of ConvTree–SemConvTree significantly increases the limits of applicability in the area of social data analysis as we can research different scales of events: high-scale events could be analyzed on the level of $\mu^{\tau}_{<u,v>}$, while low-scale events and events of specific topics could be separately found on the $\mu^{d}_{<u,v>}$ level.

Following this idea, in this paper, we suggest an improvement of the previously developed algorithm by adding a semantic module for two main aims: to decrease data noise (removing ads, reasoning thoughts, etc.), and decrease the overall barrier μ for low-scale event detection.

3.3. Construction Algorithm

This section describes an algorithm for event detection using adaptive geogrids, also known as semantic convolutional quadtrees. These structures partition city space into regions of varying sizes. Regions with more posts have more leaves in the adaptive grid, partitioning the space into cells with approximately equal post counts. Convolutions enhance the quadtree's productivity, while the semantic component ranks posts to reduce noise and advertising influence, increasing the importance of event-related posts. The discovery pipeline comprises two main parts: historical mode and online mode (Figure 2). In historical mode, we collect open posts from the previous year to build historical adaptive grids. These grids, along with real-time post flows, are then fed into the anomaly detection module, which identifies anomalous post bursts in the city space. The detected anomalies are subsequently passed to the event detection module, which performs post-filtering and links posts within anomalies to events.
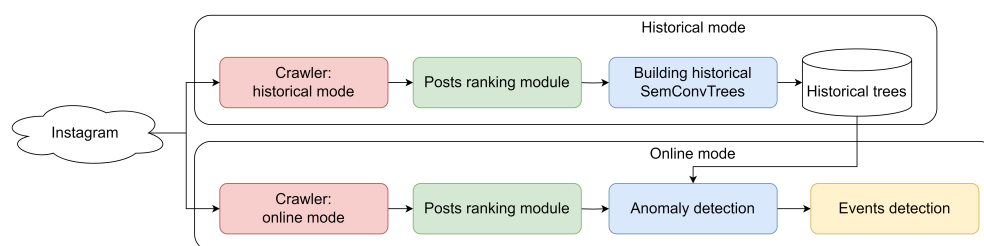

Figure 2. Pipeline of the event detection system SemConvTree.

A crucial component of the algorithm is the post-ranking module. It employs an ensemble of three models to determine a post's potential importance for event detection (Figure 3): BigARTM, BERTopic, and zero-shot classification based on Semantic BERT (SBERT-Zero-Shot). These models are detailed further in the corresponding sections. BigARTM and BERTopic are unsupervised thematic modeling methods, while the zero-shot classification implements a group-supervised approach with predefined text classes. For each post, the label is determined by a majority vote from these three models. The weighted posts are then fed into the remaining algorithm blocks.
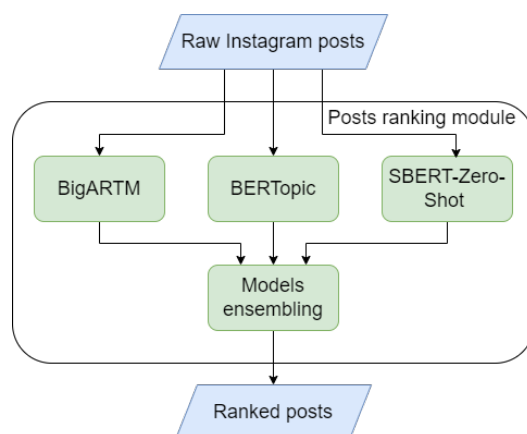
Figure 3. Scheme of the post-ranking module.

4. Semantic Filtering

A crucial hyperparameter of the ConvTree algorithm is the maximum number of posts in a quadtree leaf. Tree construction halts further partitioning into smaller areas when the post count in an area falls below this threshold. A lower threshold increases the algorithm's sensitivity to anomalies. The original article empirically selected a threshold of 12 for noise resistance and high accuracy. Our developed semantic filtering algorithm reduced this value to 3 while maintaining noise resistance and improving the accuracy and completeness of detectable events.

Despite the ability to distinguish common noise and event types, data might be precise for specific areas and times. Classes with small sizes may not appear in training data, even in large labeled datasets. Therefore, unsupervised or semi-supervised approaches offer potentially more flexible solutions for social media post classification. We applied three models: BERTopic, TSB-ARTM, and SBERT-Zero-Shot.

4.1. BERTopic

The filter module, based on the BERTopic approach, assigns tags to posts according to their subject. BERTopic-generated clusters (topics) can be interpreted as groups of event-related posts or noisy posts. We used a marked-up dataset to determine cluster types, assigning labels based on the distribution of labeled posts within. The label with the largest fraction of posts becomes the cluster's overall label. For clusters without labeled posts, the algorithm designates them as event-related. Essentially, this module only filters clusters recognizable as consisting of noisy data. The same logic applies to labeling outlines formed by HDBSCAN. Given the need for a labeled dataset for post classification, this module is considered a semi-supervised filtering approach.

4.2. TSB-ARTM

Additive regularization of topic models (ARTM) [82] evolves from LDA-based models [77]. It combines regularizers to create models with specific properties. This multi-criteria approach optimizes a weighted sum of the primary criterion (log-likelihood) and additional criterion regulators. This method allows simultaneous consideration of multiple optimization criteria and quality metrics for model validation.

ARTM also incorporates additional modalities, such as timestamps or geospatial coordinates. These modalities enable consideration of seasonality and the hypothesis of temporal and spatial connectivity of event posts. Our TSB-ARTM model uses a generalized timestamp (post month) and urban sector information as additional modalities. Like BERTopic, this algorithm performs unsupervised thematic modeling. We employed the same semi-supervised strategy to identify advertising topics for both TSB-ARTM and BERTopic.

4.3. SBert-Zero-Shot

The core concept of applying Sentence-BERT for zero-shot classification involves calculating similarities between embeddings of social media text descriptions and predefined classes.

We utilized two category lists for noise data and events, as defined during the data analysis (detailed in subparagraph 5.1). We then obtained embeddings for these predefined classes and all social media descriptions using the pre-trained sentence-transformers model "paraphrase-multilingual-MiniLM-L12-v2" [86]. This model maps paragraphs to a dense vector space and has proven effective for clustering and semantic search NLP tasks. The final step involved calculating cosine similarities between each social media data embedding and the predefined class embeddings. Consequently, we assigned social media descriptions to their most similar class, converting the label to a binary value indicating the determined text description type: noise or event.

4.4. Models Comparison

We compared the developed models using a manually labeled dataset of New York City, USA posts from 2019. We selected three one-day intervals in February, June, and October to account for seasonal and daily temporal variations. For each day, we examined three 1 h periods: morning (10–11 a.m.), afternoon (4–5 p.m.), and evening (10–11 p.m.).

Table 1 compares the completeness of promotional and non-event themes for each model individually and when used in an ensemble. The results demonstrate that employing an ensemble of the three described models, with majority opinion selection, significantly enhances the completeness of non-event post selection. This improvement subsequently leads to a substantial increase in event selection quality.

Table 1. Filtering step results for posts by time of day: morning (M), afternoon (A), and evening (E), and for posts by month: February (Feb), June (Jun), and October (Oct).

Model	Recall of the Non-Events Posts Detection						
	All	M	A	E	Feb	Jun	Oct
BERTopic	0.42	0.43	0.39	0.41	0.42	0.43	0.39
TSB-ARTM	0.51	0.48	0.5	0.49	0.48	0.52	0.51
SBert-Zero-Shot	0.46	0.44	0.47	0.46	0.45	0.47	0.48
Models ensemble	**0.61**	**0.59**	**0.6**	**0.62**	**0.59**	**0.6**	**0.61**

5. Experimental Evaluation

The key indicators for this problem are the number of different-scale events detected by the algorithm and their precision and recall. We conducted comparisons using data from New York City, which is known for its high Instagram activity levels and is widely used in the context of event detection research [7,9,11,35,76,85].

5.1. DataSet

No publicly available dataset suitable for low-scale event detection exists. Therefore, we collected our own data. We extracted data from Instagram using the Legacy API, credentials, and Golang web-scraping techniques and libraries, without employing search keywords or focusing on specific topics. For crawling, we obtained lists of New York City locations from the Facebook API. We collected data for each location from early 2018 to April 2020. In total, we gathered over 26 million posts from New York City, including text, timestamps, coordinates, and other meta-information used in multimodal models.

To apply semi-supervised filtering modules, obtain candidate embeddings for SBert-Zero-Shot, and evaluate the entire pipeline's results, we defined event categories. We employed crowd-sourced judges to label a portion of the dataset. We manually analyzed the dataset to define categories, identifying the most frequent event types (e.g., festivals, shows, sports competitions) and noise topics (e.g., food, advertisements). To ensure comprehensive

coverage, we added classes for "other private events," "other global events," and "other" for remaining noise data. We extracted posts from three one-hour periods on single days in June, October, and February: morning (10–11 a.m.), afternoon (4–5 p.m.), and evening (10–11 p.m.). This selection aimed to capture seasonality trends in social media user activities. We extracted 5,829 posts, with 646 labeled by judges on the Yandex Toloka platform. Judges were paid $0.01 per answer, with a 250-judgment limit per judge. Given the complex post semantics, we allowed multiple labels per text and added "future events" and "retrospective events" categories. Table 2 presents the defined categories and the number of labeled posts for each category.

Table 2. Categories and posts number.

Category	Posts Number	Category	Posts Number
Festival	64	Concert	115
Sport event	317	National holiday	214
Show/ Flashmob/ Pride	55	Exhibition	46
Stroll/ Camping	120	Accident	2
Lectures/Conferences	3	Other	2289
Other private event	135	Private celebration	157
Food	594	Other public event	164
Event advertisement	80	Other advertisement	205
Future event	17	Retrospective event	36
Unsure	2031		

5.2. Experimental Studies

For event detection experiments, we divided the dataset into two parts. The first part, comprising Instagram posts from 2018, was used to build historical data and adaptive grids. The second part, covering early 2019 to April 2020, was used to search for events in New York City. Through ranking, we significantly increased the quadtree's sensitivity to anomalous bursts. This enhancement allowed us to detect not only more high-scale events (Table 3) but also a substantial number of low-scale events. It is important to note that due to different data sources (Instagram and Twitter) and varying time intervals for event searches, the comparison between algorithms is not entirely fair. The lack of large, generally accepted datasets with proper markup for algorithm alignment remains a challenge in event detection tasks, as highlighted by Korneev et al. [87].

Our primary objective was to enhance the algorithm's ability to detect low-scale events. However, we also succeeded in improving accuracy and completeness for high-scale events. Notably, we identified significantly more low-scale events, increasing the total number of detected events from 10,000 to 177,000, a 16-fold increase (Table 4).

Table 5 clearly illustrates the quality differences in event extraction between the ConvTree model [11] and our developed multimodal algorithm with additional filtering. The substantial improvement in precision and recall metrics results from comparing the algorithms on a dataset rich in low-scale event posts. Our filtering-based model effectively identifies such events, while the original algorithm [11] primarily focuses on global events.

Table 3. Precision and recall comparison with other approaches.

Method	Precision	Recall	Avg. Events per Day
Eyewitness [10]	70%	-	-
GeoBurst+ [9]	35%	48%	-
TrioVecEvent [83]	78%	60%	-
ConvTree [11]	77%	18%	22.2
SemConvTree	**86%**	**64%**	365.6

Table 4. Comparison of the original algorithm with low sensitivity, with high sensitivity (with a lot of advertising and noise), and the finalized algorithm with filtered advertising and noise.

Method	Count of Events	Count of Event Posts
ConvTree	10,757	151,084
ConvTree with high sensitive and noise events	263,533	803,454
SemConvTree	177,315	538,628

Table 5. Comparison of multi-scale event detection results across all periods and separately for February (Feb), June (Jun), and October (Oct). P—Precision, R—Recall.

Model	All		Feb		Jun		Oct	
	P	R	P	R	P	R	P	R
ConvTree [11]	0.77	0.18	0.78	0.21	0.77	0.17	0.77	0.17
SemConvTree	0.86	0.64	0.87	0.58	0.85	0.63	0.86	0.64

6. Conclusions and Future Works

In this paper, we developed the semantic convolution tree algorithm for low-scale event detection, extending existing solutions by incorporating topic modeling and selecting the most expressive features for event detection. We demonstrated the construction of a semantic module based on an unsupervised learning mechanism using an ensemble of topic models. We also refined the BERTopic, TSB-ARTM, and SBert-Zero-Shot models to consider temporal and spatial modalities. Our event detection algorithm significantly increased the number of detected events (more than 16-fold) and improved detection accuracy.

However, we identify several avenues for future research. We plan to process multidimensional event probability weights (textual and spatial semantics) and integrate data from other social networks and regions for cross-cultural analysis. Moreover, we aim to extract anomalies from both the overall message distribution and topic-specific message subsets. Obtaining separate metrics for specific categories would allow for interesting comparisons from a detectability perspective. Another idea is to include additional modalities, such as retrospective and future event labels. Finally, there is a considerable research gap in comparing different event detection methods. We plan to create and publish a universal dataset allowing researchers to conveniently compare their approaches.

7. Compliance with Ethical Standards

Research involving social media data analysis and processing carries the risk of revealing personal and sensitive information. We wish to clarify our perspective on event detection systems in this context. Our research extends beyond event detection to the broader study of urban processes. While personal events should not be shared, analyzing their distribution within a city can inform infrastructure development for such events. For instance, our framework can identify parks frequently used for wedding photoshoots, providing valuable information for infrastructure planning and enhancing public interest. Additionally, our algorithm can handle other small-scale events, such as car accidents, potentially highlighting dangerous road sections. This knowledge could lead to traffic flow reorganization, resulting in a significant positive social impact.

8. Research Data Policy and Data Availability Statement

For ethical reasons, we cannot publish the dataset used in this study as we have not fully anonymized the data for publication. However, we are currently developing a comprehensive anonymized multimodal dataset for event detection approaches, which will be publicly available upon completion.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Dixon, S.J. Number of Social Media Users Worldwide from 2017 to 2028 (in Billions), May 2024. Available online: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ (accessed on 5 June 2024).
2. Wolniak, R.; Stecuła, K. Artificial Intelligence in Smart Cities—Applications, Barriers, and Future Directions: A Review. *Smart Cities* **2024**, *7*, 1346–1389. [CrossRef]
3. Earle, P.S.; Bowden, D.C.; Guy, M. Twitter earthquake detection: Earthquake monitoring in a social world. *Ann. Geophys.* **2012**, *54*. [CrossRef]
4. Osborne, M.; Moran, S.; McCreadie, R.; Lunen, A.V.; Sykora, M.; Cano, E.; Ireson, N.; Macdonald, C.; Ounis, I.; He, Y.; et al. Real-Time Detection, Tracking, and Monitoring of Automatically Discovered Events in Social Media. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Association for Computational Linguistics, Baltimore, MD, USA, 22–27 June 2014. [CrossRef]
5. Lim, B.H.; Lu, D.; Chen, T.; Kan, M.Y. #mytweet via Instagram. In Proceedings of the Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, Paris, France, 25–28 August 2015; ACM: New York, NY, USA, 2015. [CrossRef]
6. Giridhar, P.; Wang, S.; Abdelzaher, T.; Amin, T.A.; Kaplan, L. Social Fusion: Integrating Twitter and Instagram for Event Monitoring. In Proceedings of the 2017 IEEE International Conference on Autonomic Computing (ICAC), Columbus, OH, USA, 17–21 July 2017; IEEE: New York, NY, USA, 2017. [CrossRef]
7. Zhang, C.; Zhou, G.; Yuan, Q.; Zhuang, H.; Zheng, Y.; Kaplan, L.; Wang, S.; Han, J. GeoBurst: Real-Time Local Event Detection in Geo-Tagged Tweet Streams. In Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '16, New York, NY, USA, 17–21 July 2016; pp. 513–522. [CrossRef]
8. McMinn, A.; Moshfeghi, Y.; Jose, J. Building a large-scale corpus for evaluating event detection on twitter. In Proceedings of the International Conference on Information and Knowledge Management, Proceedings, San Francisco, CA, USA, 27 October–1 November 2013; pp. 409–418. [CrossRef]
9. Zhang, C.; Lei, D.; Yuan, Q.; Zhuang, H.; Kaplan, L.; Wang, S.; Han, J. GeoBurst+: Effective and Real-Time Local Event Detection in Geo-Tagged Tweet Streams. *ACM Trans. Intell. Syst. Technol.* **2018**, *9*, 34. [CrossRef]
10. Krumm, J.; Horvitz, E. Eyewitness. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, Bellevue, WA, USA, 3–6 November 2015; ACM: New York, NY, USA, 2015. [CrossRef]
11. Visheratin, A.A.; Mukhina, K.D.; Visheratina, A.K.; Nasonov, D.; Boukhanovsky, A.V. Multiscale event detection using convolutional quadtrees and adaptive geogrids. In Proceedings of the 2nd ACM SIGSPATIAL Workshop on Analytics for Local Events and News, Seattle, WA, USA, 6 November 2018; ACM: New York, NY, USA, 2015. [CrossRef]
12. Saha, K.; Seybolt, J.; Mattingly, S.M.; Aledavood, T.; Konjeti, C.; Martinez, G.J.; Grover, T.; Mark, G.; De Choudhury, M. What Life Events Are Disclosed on Social Media, How, When, and By Whom? In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21, New York, NY, USA, 8–13 May 2021. [CrossRef]
13. DiCarlo, M.F.; Berglund, E.Z. Use of social media to seek and provide help in Hurricanes Florence and Michael. *Smart Cities* **2020**, *3*, 1187–1218. [CrossRef]
14. Becker, H.; Naaman, M.; Gravano, L. Beyond Trending Topics: Real-World Event Identification on Twitter. *Proc. Int. AAAI Conf. Web Soc. Media* **2021**, *5*, 438–441. [CrossRef]
15. Khodabakhsh, M.; Kahani, M.; Bagheri, E.; Noorian, Z. Detecting life events from Twitter based on temporal semantic features. *Knowl.-Based Syst.* **2018**, *148*, 1–16. [CrossRef]
16. Sufi, F.K. AI-SocialDisaster: An AI-based software for identifying and analyzing natural disasters from social media. *Softw. Impacts* **2022**, *13*, 100319. [CrossRef]

17. Cresci, S.; Tesconi, M.; Cimino, A.; Dell'Orletta, F. A Linguistically-Driven Approach to Cross-Event Damage Assessment of Natural Disasters from Social Media Messages. In Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion, New York, NY, USA, 18–22 May 2015; pp. 1195–1200. [CrossRef]

18. Abdelhaq, H.; Sengstock, C.; Gertz, M. EvenTweet: Online localized event detection from twitter. *Proc. VLDB Endow.* **2013**, *6*, 1326–1329. [CrossRef]

19. Neruda, G.A.; Winarko, E. Traffic Event Detection from Twitter Using a Combination of CNN and BERT. In Proceedings of the 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Virtual, 23–26 October 2021; pp. 1–7. [CrossRef]

20. Timokhin, S.; Sadrani, M.; Antoniou, C. Predicting venue popularity using crowd-sourced and passive sensor data. *Smart Cities* **2020**, *3*, 42. [CrossRef]

21. Afyouni, I.; Aghbari, Z.A.; Razack, R.A. Multi-feature, multi-modal, and multi-source social event detection: A comprehensive survey. *Inf. Fusion* **2022**, *79*, 279–308. [CrossRef]

22. Said, N.; Ahmad, K.; Regular, M.; Pogorelov, K.; Hassan, L.; Ahmad, N.; Conci, N. Natural Disasters Detection in Social Media and Satellite imagery: A survey. *arXiv* **2019**, arXiv:1901.04277. [CrossRef]

23. Atefeh, F.; Khreich, W. A Survey of Techniques for Event Detection in Twitter. *Comput. Intell.* **2015**, *31*, 132–164. [CrossRef]

24. Saeed, Z.; Abbasi, R.; Maqbool, O.; Sadaf, A.; Razzak, I.; Daud, A.; Aljohani, N.; Xu, G. Twitter: A Survey and Framework on Event Detection Techniques. *J. Grid Comput.* **2019**. [CrossRef]

25. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv.* **2009**, *41*, 15. [CrossRef]

26. Markou, M.; Singh, S. Novelty detection: A review—part 1: Statistical approaches. *Signal Process.* **2003**, *83*, 2481–2497. [CrossRef]

27. Ada, I.; Berthold, M.R. Unifying Change—Towards a Framework for Detecting the Unexpected. In Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops, Vancouver, BC, Canada, 11 December 2011; pp. 555–559. [CrossRef]

28. Dries, A.; Rückert, U. Adaptive concept drift detection. *Stat. Anal. Data Min. ASA Data Sci. J.* **2009**, *2*, 311–327. [CrossRef]

29. Liu, N. Topic Detection and Tracking. In *Encyclopedia of Database Systems*; Liu, L., Özsu, M.T., Eds.; Springer US: Boston, MA, USA, 2009; pp. 3121–3124. [CrossRef]

30. Zhang, X.; Chen, X.; Chen, Y.; Wang, S.; Li, Z.; Xia, J. Event detection and popularity prediction in microblogging. *Neurocomputing* **2015**, *149*, 1469–1480. [CrossRef]

31. Brants, T.; Chen, F.; Farahat, A. A System for New Event Detection. In Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, SIGIR '03, New York, NY, USA, 28 July–1 August 2003; pp. 330–337. [CrossRef]

32. Kaleel, S.B.; Abhari, A. Cluster-discovery of Twitter messages for event detection and trending. *J. Comput. Sci.* **2015**, *6*, 47–57. [CrossRef]

33. Aiello, L.M.; Petkos, G.; Martin, C.; Corney, D.; Papadopoulos, S.; Skraba, R.; Göker, A.; Kompatsiaris, I.; Jaimes, A. Sensing Trending Topics in Twitter. *IEEE Trans. Multimed.* **2013**, *15*, 1268–1282. [CrossRef]

34. Lampos, V.; Cristianini, N. Nowcasting Events from the Social Web with Statistical Learning. *ACM Trans. Intell. Syst. Technol.* **2012**, *3*, 72. [CrossRef]

35. Weng, J.; Yao, Y.; Leonardi, E.; Lee, B.S. Event Detection in Twitter. *Proc. Int. AAAI Conf. Web Soc. Media* **2011**, *5*, 401–408. [CrossRef]

36. Cheng, T.; Wicks, T. Event Detection using Twitter: A Spatio-Temporal Approach. *PLoS ONE* **2014**, *9*, e97807. [CrossRef] [PubMed]

37. Weiler, A.; Grossniklaus, M.; Scholl, M. Event Identification and Tracking in Social Media Streaming Data. In Proceedings of the CEUR Workshop Proceedings, Athens, Greece, 28 March 2014; Volume 1133.

38. He, Q.; Chang, K.; Lim, E.P. Analyzing Feature Trajectories for Event Detection. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '07, New York, NY, USA, 23–27 July 2007; pp. 207–214. [CrossRef]

39. Kleinberg, J. Bursty and Hierarchical Structure in Streams. *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* **2002**, *7*, 91–101. [CrossRef]

40. Fung, G.P.C.; Yu, J.X.; Yu, P.S.; Lu, H. Parameter Free Bursty Events Detection in Text Streams. In Proceedings of the 31st International Conference on Very Large Data Bases, VLDB Endowment, VLDB '05, Trondheim, Norway, 30 August–2 September 2005; pp. 181–192.

41. He, Q.; Chang, K.; Lim, E.P.; Zhang, J. Bursty Feature Representation for Clustering Text Streams. In Proceedings of the SDM, Minneapolis, MN, USA, 26–28 April 2007.

42. Kumar, R.; Novak, J.; Raghavan, P.; Tomkins, A. On the Bursty Evolution of Blogspace. In Proceedings of the 12th International Conference on World Wide Web, WWW '03, New York, NY, USA, 20–24 May 2003; pp. 568–576. [CrossRef]

43. Mei, Q.; Zhai, C. Discovering Evolutionary Theme Patterns from Text: An Exploration of Temporal Text Mining. In Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, KDD '05, New York, NY, USA, 21–24 August 2005; pp. 198–207. [CrossRef]

44. Zhou, D.; Chen, L.; He, Y. An Unsupervised Framework of Exploring Events on Twitter: Filtering, Extraction and Categorization. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI '15, Austin, TX, USA, 25–30 January 2015; AAAI Press: Washington, DC, USA, 2015; pp. 2468–2474.

45. Lee, R.; Wakamiya, S.; Sumiya, K. Discovery of unusual regional social activities using geo-tagged microblogs. *World Wide Web* **2011**, *14*, 321–349. [CrossRef]

46. Feng, W.; Zhang, C.; Zhang, W.; Han, J.; Wang, J.; Aggarwal, C.; Huang, J. STREAMCUBE: Hierarchical spatio-temporal hashtag clustering for event exploration over the Twitter stream. In Proceedings of the 2015 IEEE 31st International Conference on Data Engineering, Seoul, Republic of Korea, 13–17 April 2015; IEEE: New York, NY, USA, 2015. [CrossRef]

47. Rehman, F.U.; Afyouni, I.; Lbath, A.; Basalamah, S. Understanding the Spatio-Temporal Scope of Multi-scale Social Events. In Proceedings of the 1st ACM SIGSPATIAL Workshop on Analytics for Local Events and News, Redondo Beach, CA, USA, 7–10 November 2017; ACM: New York, NY, USA, 2017. [CrossRef]

48. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781.

49. Peters, M.E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. Deep contextualized word representations. *arXiv* **2018**, arXiv:1802.05365.

50. Pennington, J.; Socher, R.; Manning, C. Glove: Global Vectors for Word Representation. *EMNLP* **2014**, *14*, 1532–1543. [CrossRef]

51. Zhang, Y.; Shirakawa, M.; Hara, T. A General Method for Event Detection on Social Media. In Proceedings of the Symposium on Advances in Databases and Information Systems, Tartu, Estonia, 24–26 August 2021.

52. Hettiarachchi, H.; Adedoyin-Olowe, M.; Bhogal, J.; Gaber, M.M. Embed2Detect: Temporally clustered embedded words for event detection in social media. *Mach. Learn.* **2021**, *111*, 49–87. [CrossRef]

53. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805. [CrossRef]

54. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv* **2019**, arXiv:1907.11692. [CrossRef]

55. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Hong Kong, 3–7 November 2019.

56. Wei, Z.; Yongli, W. Chinese Event Detection Combining BERT Model with Recurrent Neural Networks. In Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 25–27 December 2020; pp. 1625–1629. [CrossRef]

57. Huang, L.; Shi, P.; Zhu, H.; Chen, T. Early detection of emergency events from social media: A new text clustering approach. *Nat. Hazards* **2022**, *111*, 851–875. [CrossRef] [PubMed]

58. McDonald, R.; Nivre, J. Analyzing and Integrating Dependency Parsers. *Comput. Linguist.* **2011**, *37*, 197–230. [CrossRef]

59. Nguyen, T.; Cho, K.; Grishman, R. Joint Event Extraction via Recurrent Neural Networks. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 5 January 2016; pp. 300–309. [CrossRef]

60. Liu, X.; Luo, Z.; Huang, H. Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Brussels, Belgium, 31 October–4 November 2018. [CrossRef]

61. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv* **2017**, arXiv:1609.02907. http://arxiv.org/abs/1609.02907.

62. Yan, H.; Jin, X.; Meng, X.; Guo, J.; Cheng, X. Event Detection with Multi-Order Graph Convolution and Aggregated Attention. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 5766–5770. [CrossRef]

63. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph Attention Networks. *arXiv* **2018**, arXiv:1710.10903. [CrossRef]

64. Dutta, S.; Ma, L.; Saha, T.K.; Lu, D.; Tetreault, J.; Jaimes, A. GTN-ED: Event Detection Using Graph Transformer Networks. *arXiv* **2021**, arXiv:2104.15104.

65. Raiaan, M.A.K.; Mukta, M.S.H.; Fatema, K.; Fahad, N.M.; Sakib, S.; Mim, M.M.J.; Ahmad, J.; Ali, M.E.; Azam, S. A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access* **2024**, *12*, 26839–26874. [CrossRef]

66. Snoek, C.G.M.; Worring, M.; Smeulders, A.W.M. Early versus late fusion in semantic video analysis. In Proceedings of the MULTIMEDIA '05, Besancon, France, 6–9 February 2005.

67. Sukel, M.; Rudinac, S.; Worring, M. Multimodal Classification of Urban Micro-Events. *arXiv* **2019**, arXiv:1904.13349. [CrossRef]

68. Cui, W.; Du, J.; Wang, D.; Kou, F.; Xue, Z. MVGAN: Multi-View Graph Attention Network for Social Event Detection. *ACM Trans. Intell. Syst. Technol.* **2021**, *12*, 27. [CrossRef]

69. Jony, R.I.; Woodley, A.; Perrin, D. Fusing Visual Features and Metadata to Detect Flooding in Flickr Images. In Proceedings of the 2020 Digital Image Computing: Techniques and Applications (DICTA), Melbourne, Australia, 29 November–2 December 2020; pp. 1–8. [CrossRef]

70. Petkos, G.; Papadopoulos, S.; Kompatsiaris, I. Social event detection using multimodal clustering and integrating supervisory signals. In Proceedings of the 2nd ACM International Conference on Multimedia Retrieval, ICMR 2012, Hong Kong, 5–8 June 2012. [CrossRef]

71. Schinas, M.; Papadopoulos, S.; Petkos, G.; Kompatsiaris, I.; Mitkas, P. Multimodal Graph-based Event Detection and Summarization in Social Media Streams. *Int. J. Multimed. Inf. Retr.* **2015**, 189–192. [CrossRef]

72. Tong, M.; Wang, S.; Cao, Y.; Xu, B.; Li, J.; Hou, L.; Chua, T.S. Image Enhanced Event Detection in News Articles. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 9040–9047. [CrossRef]

73. Guo, C.; Tian, X. Event recognition in personal photo collections using hierarchical model and multiple features. In Proceedings of the 2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP), Xiamen, China, 19–21 October 2015; pp. 1–6.

74. Kaneko, T.; Yanai, K. Event photo mining from Twitter using keyword bursts and image clustering. *Neurocomputing* **2016**, *172*, 143–158. [CrossRef]

75. Zaharieva, M.; Zeppelzauer, M.; Breiteneder, C. Automated Social Event Detection in Large Photo Collections. In Proceedings of the 3rd ACM Conference on International Conference on Multimedia Retrieval, ICMR '13, New York, NY, USA, 16–29 April 2013; pp. 167–174. [CrossRef]

76. Ali, F.; Ali, A.; Imran, M.; Naqvi, R.A.; Siddiqi, M.H.; Kwak, K.S. Traffic accident detection and condition analysis based on social networking data. *Accid. Anal. Prev.* **2021**, *151*, 105973. [CrossRef]

77. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.

78. Sokolova, M.; Huang, K.; Matwin, S.; Ramisch, J.J.; Sazonova, V.; Black, R.; Orwa, C.; Ochieng, S.; Sambuli, N. Topic Modelling and Event Identification from Twitter Textual Data. *arXiv* **2016**, arXiv:1608.02519.

79. Zhang, C.; Wang, H.; Cao, L.; Wang, W.; Xu, F. A Hybrid Term-Term Relations Analysis Approach for Topic Detection. *Knowl.-Based Syst.* **2016**, *93*, 109–120. [CrossRef]

80. Choi, D.; Park, S.; Ham, D.; Lim, H.; Bok, K.; Yoo, J. Local Event Detection Scheme by Analyzing Relevant Documents in Social Networks. *Appl. Sci.* **2021**, *11*, 577. [CrossRef]

81. Vorontsov, K.; Frei, O.; Apishev, M.; Romov, P.; Dudarenko, M. BigARTM: Open Source Library for Regularized Multimodal Topic Modeling of Large Collections. In *Communications in Computer and Information Science*; Springer International Publishing: Berlin, Germany, 2015; pp. 370–381. [CrossRef]

82. Vorontsov, K.V. Additive regularization for topic models of text collections. *Doklady Math.* **2014**, *89*, 301–304. [CrossRef]

83. Zhang, C.; Liu, L.; Lei, D.; Yuan, Q.; Zhuang, H.; Hanratty, T.; Han, J. TrioVecEvent. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; ACM: New York, NY, USA, 2017. [CrossRef]

84. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv **2022**, arXiv:2203.05794.

85. Wei, H.; Zhou, H.; Sankaranarayanan, J.; Sengupta, S.; Samet, H. DeLLe. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Analytics for Local Events and News, Chicago, IL, USA, 5 November 2019; ACM: New York, NY, USA, 2019. [CrossRef]

86. Chaffey, D. *Global Social Media Statistics Research Summary 2022*; Smart Insights: Leeds, UK, 2022.

87. Korneev, A.; Kovalchuk, M.; Filatova, A.; Tereshkin, S. Towards comparable event detection approaches development in social media. *Procedia Comput. Sci.* **2022**, *212*, 312–321. [CrossRef]