Analyzing Twitter for Social TV: Sentiment Extraction for Sports

Siqi Zhao and Lin Zhong
Dept. of Electrical & Computer Engineering
Rice University
6100 Main Street
Houston, TX
{siqi.zhao, lzhong}@rice.edu

Jehan Wickramasuriya and Venu Vasudevan
Applied Research
Motorola Mobility
600 N. U.S. Highway 45
Libertyville, IL
{jehan, venu.vasudevan}@motorola.com

Abstract

As TV watchers tweet about how they feel and what they see, they produce valuable information not only about the TV program but also how engaged they are to the program. We have already built a web service, SportSense, which recognizes major events in the US National Football League (NFL) games within 40 seconds after an event takes place by analyzing data retrieved from Twitter in real-time. In this paper, we report our effort to extend SportSense to extract TV watchers' sentimental reaction to major events in live broadcast sports games in *real-time* and present our ongoing work that leverages SportSense for a social TV system that enables TV watchers to better select interesting programs in real-time and to produce personalized program summaries and enables advertisers to customize ads based on recognized events and extracted audience sentiments.

Categories and Subject Descriptors

H5.m [**Information Interfaces and Presentation**]: Miscellaneous

General Terms
Algorithms, Design, Experimentation, Measurement.

Keywords
Twitter, social TV, sentiment extraction.

1. Introduction

Live broadcast TV programs such as sports games, reality shows, and concerts often attract a large number of loyal watchers, creating tremendous commercial opportunities. Key to their attractions is their inherent uncertainty or unpredictability. On the other hand, the unpredictability challenges both TV watchers and advertisers. On one hand, while some TV watchers tune to a live broadcast program for its entire duration, many jump from one channel to another occasionally looking for the most interesting one at a moment because a live broadcast program is not always equally interesting throughout its duration. This is particularly true for sports where a game becomes more interesting when the scores are close. When there are multiple games ongoing at the same time, e.g., up to 10 games take place on Sunday afternoon in the NFL regular season; it is nontrivial for a watcher to determine which game is the most interesting. On the other hand, without knowing what is going on in the program or how engaged the TV audience is to a program at a certain time, an advertiser can only fill advertisement slots based on general demographics of the audience and the overall rating of the program.

We envision a social TV system in which watchers and advertisers can see how other watchers are enjoying a live broadcast program in real-time and what is happening in the program. In this system, a watcher can switch to a program when peers are excited by it or a certain event just takes place. Similarly, an advertiser can pay more to get a time slot when the watchers are highly engaged or even use a commercial that resonates with the event that just happened.

While there are previous efforts based on video analysis toward this goal [1, 2], we are interested in using Twitter to recognize major events and extract audience sentiments in real-time. Compared to video, Twitter has many unique advantages toward the envisioned social TV system. (*i*) Tweets posted by TV watchers directly reflect what they feel and think about the program. While video analysis allows one to learn about the *program*, Twitter analysis directly allows us to learn about the *program audience*. With over 200 million active users, Twitter ensures good coverage of popular TV programs [3]. (*ii*) Second, Live video streaming is available only at a high financial, bandwidth, and computing cost. In contrast, public tweets are free and are easy to retrieve due to their brief, textual nature. (*iii*) Thirdly, the textual nature of tweets makes them amenable to lexicon-based analysis. As a result, end users can easily personalize Twitter-based event recognition and sentiment extraction by using the right keywords. This is significantly easier than personalizing a video analysis tool to recognize a new event. (*iv*) Finally, Twitter allows tweets to be retrieved by tweeters. Therefore, end users of the Twitter-empowered social TV system can personalize their system by tuning to a personalized collection of tweeters, creating a personalized social information channel.

To demonstrate the feasibility of using Twitter for the envisioned social TV system, we have designed and realized a web service, called SportSense that recognizes major events in a sports game in real-time. During the 2010-2011 NFL season, SportSense is able to recognize major events such as touchdowns and interceptions with 90% accuracy within 40 seconds after an event happens. Through a web site, the service visualized the results and provided a popularity "thermometer" for games that were played at the same time.

In this workshop paper, we present our ongoing work in extracting audience sentiments from Twitter analysis and realizing Twitter-enabled, social, electronic program guide (EPG) and audience-aware advertisement auction. Although we focus on NFL games in this work, most of the techniques can be readily applied to many other sports games that have a similarly sized fan population and have similar frequencies of major events, e.g., soccer, baseball, and basketball.

The rest of this paper is organized as follows. Section 2 introduces the background of SportSense which is a web service that utilizes Twitter to recognize sports events in real-time. Section 3 provides detailed explanation of our approach on Twitter analysis for sentiment extraction. Then, Section 4 presents the applications that can adopt proposed approach to enhance the TV watching experience. Section 5 demonstrates the effectiveness of our approach for various types of sports, followed by the related work and conclusion in Section 6 and Section 7, respectively.

2. Background of SportSense

We first describe SportSense [4, 5], a web service that recognizes major events from live broadcast NFL games in real-time by analyzing Twitter. SportSense heavily leverages the bounded vocabulary of sports events and employs lexicon-based heuristics to retrieve game-related tweets in real-time and recognize events.

2.1 Lexicon-based Game Tweets Separation

SportSense uses a predetermined lexicon that includes game terminology and team names to retrieve tweets related to NFL games through the Twitter Streaming API. SportSense relies on the team names to collect data when multiple games take place at the same time and attribute these tweets to games based on the mentioned team names. This lexicon-based approach is effective to separate the game related tweets. By manually examining random 5% of the tweets, about 2,000, posted during the 2010 NFL Super Bowl by the NFL followers, we find that extraction by 10 keywords including game terminology and team names achieves a false negative rate below 9% and a false positive rate below 5%. Further, we found that the team names appear in over 60% of the game-related tweets.

2.2 Event Recognition

SportSense recognizes NFL events by analyzing the content of game-related tweets collected using the lexicon-based heuristic and examining the post rate of tweets with keywords related to game events. SportSense employs a simple method to detect an event by calculating the post rate increase as the ratio of the post rate in the second half of a sliding time window to that in the first half. To overcome an undocumented Twitter limit to the Streaming API of 50 tweets per second, SportSense focuses on tweets with a keyword related to the game event only.

The size of the window will have a significant impact on the tradeoff between the delay and accuracy of event detection. A shorter window will lead to a smaller delay but may have a poor performance when the post rate is low and, therefore, there are not many tweets posted in the time window.

To achieve the best tradeoff, SportSense selects the window size adaptively. The sliding window has a variable size of 10, 20, 30, or 60 seconds; and each window is divided into two sub-windows of equal length. At every second, SportSense will start from the shortest window, 10 seconds, to examine the post rate of the event keywords in the window and the post rate ratio of each event type between the two halves. If both the post rate and the ratio exceed their corresponding thresholds, the corresponding event will be recognized; otherwise the window size will increment. The threshold for the post rate of tweets will filter out the random presence of the event keywords when people predict events, or discuss about past events. From our observation, the random presence has a low post rate and low frequency, i.e. at most 2 tweets in a second and usually null. The value is set to half of the window size which is 5, 10, 15, and 30 tweets. The post rate ratio between two sub-windows will examine the increase of the post rate. This ratio is the post rate in the second half window to the post rate of those in the first half. The value is set to 1.3 in this case; that means the post rate in the second half window needs to be at least 1.3 times of the post rate in the first half to proceed.

Half of events can be detected using the window size less than 20 seconds and more than 2/3 of events can be detected using the window size less than 30 seconds. Since we halve the window to detect events, the delay of the system is half of the window size. As a result, this method will introduce less than 10 second delay for half of the events.

2.3 Web Service Realization

We have implemented the solution described above as a real-time web service that visualizes event recogni-

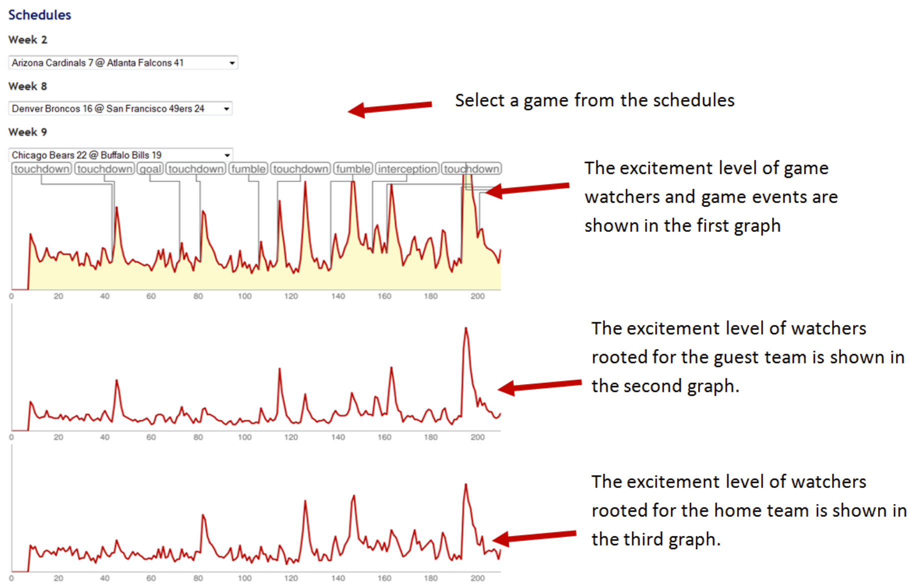

Figure 1: Visualized results of real-time event recognition and sentiment extraction at SportSense.com

tion results through a website throughout the 2010-2011 NFL season. The implementation is coded in PHP and consists of the backend for data collection and analysis and the frontend for web visualization.

The backend consists of two parallel modules and a MySQL database. The data collection module collects game-related tweets through Twitter Streaming API. Collected tweets are saved in the MySQL database. The event recognition module will retrieve tweets from database, separate tweets to games, recognize events, and generate the results in Google Chart format.

As analyzed above, the backend can introduce many seconds of delay to event recognition. To minimize this delay, we created multiple threads to maximize the parallelism of data analysis. The data collection module employs one thread to retrieve tweets from Twitter, decode and save the tweets into the MySQL database. The event recognition runs as another thread that retrieves data from the database and analyze them for event recognition.

The frontend visualizes the analysis results using Google Charts API through a website. For ongoing games, the website shows a color-coded bar chart for the "hotness" of all concurrent games according to the post rate of tweets related to each game. For each game, the website provides three line charts that draws post rate of tweets related to the game as the *excitement level* and denotes recognized events, as shown in Figure 1. The first chart shows the excitement level of all game watchers; the second and third charts show that of game watchers rooted for each side of the game, respectively. The line charts and recognized events for past games can be retrieved from the website by either team name or game schedule.

2.4 Recognition Performance

The web service has been active since Week 8 of the 2010-2011 NFL season. Using data from the NFL website as the ground truth we examined the events recognition performance of SportSense for games in the last two weeks (27 games in Week 16-17), playoffs (5 games), and the Super Bowl. We considered not only how many events have been correctly recognized but also how many are missed. Note that four events are targeted: touchdown, interception, field goal, and fumble. SportSense recognized all the events in the Super Bowl game without false positives and perform decently for other games. In particular, SportSense recognized 92% (151 out of 163) touchdowns, 75% (51 out of 68) interceptions, 74% (84 out of 113) fumbles, and 67% (61 out of 91) field goals in these 33 games. The average delay in event recognition is about 40 seconds which is acceptable compared to a delay of about 90 seconds for the ESPN web page in updating the score.

3. Sentiment Extraction

Our ongoing work seeks to extend SportSense for sentiment extraction. That is, we are interested in knowing how excited game watchers are and how positive their feelings are toward the game and each team of the game. We leverage techniques created by sentiment analysis research. Our current approach encompasses four steps: pre-process game-related tweets, detect sentimental tweets, and classify sentimental polarity. They are summarized in Figure 2 and described in details below.

```
    Sentiment Extraction
1.  Data Collection
    Collect tweets in real-time using Twitter
    Streaming API for 2010-2011 season NFL
    games.
2.  Pre-Processing
    Remove URLs, @username, and stop words.
    Identify all capital words, emoticons,
    question marks, and exclamation marks.
3.  Extracting Sentimental Tweets
    Produce a list of 20 frequent sentimental
    words based on 50 games data.
    Identify the sentiment orientation of these
    20 words and take them as seed words.
    Grow the positive and negative lexicon by
    mining synonyms and antonyms of the 20 seed
    words in the WordNet.
4.  Classifying Sentiment Polarity
    If (the tweet contains one of the senti-
    mental words)
      Tweet orientation = word orientation
    If (but clause or negation words appear)
      Flip the tweet orientation.
```

Figure 2: Lexicon based Sentiment Extraction algorithm.

3.1 Data collection

For the 2010-2011 season NFL games, we collected tweets during game time using the Streaming API and game keywords identified from the 2010 Super Bowl. We collected the tweets and their metadata such as tweet source, created time, location, and device. These tweets were analyzed for event recognition in real-time through a web service described previously. For the regular season games and playoffs, we collected more than 19 million game-related tweets over a period of 9 weeks including 100 games, from 3.5 million users. We collected about 1 million game-related tweets from over half a million users for 2011 Super Bowl. The evaluation of our solutions was performed in real-time when a game was ongoing and was repeated with trace-based emulation off-line if necessary.

3.2 Pre-Processing

Tweets are very noisy to extract sentiments because the tweets are short, informal, and ungrammatical by its nature. We leveraged pre-processing procedures in [6, 7] to extract the relevant signals from the tweets. First of all, URLs and replies to users (@username) often appear in tweets. We remove them as being irrelevant to the sentiment of a tweet. Moreover, we identify the all capital words, emoticons, question marks, and exclamation marks. Capital words are commonly used in tweets to represent emotions; emoticons and exclamation marks are important indicators of the presence of sentiments. We include the emoticons in the sentiments lexicon to determine the sentiment polarities. We label the all capital words and exclamation marks as emphasis of the sentiments in the tweets. We label question marks as uncertainty about events or sentiments. Finally, we remove other punctuations and stop words to avoid noise. Such removal brings more benefit than harm to our analysis because tweets are brief and usually semantically simple.

Table 1: Occurrences of words indicating positive and negative sentiments in 2 million game tweets

Word	Mentioned	Word	Mentioned
Go	**186,451**	Lose	**62,710**
Win	**164,654**	Beat	**45,406**
Like	**81,762**	F***	**42,046**
Lol	**80,703**	Love	**35,823**
Good	**64,733**	Great	**27,184**

Table 2: Occurrences of POMS words

Word	Mentioned	Word	Mentioned
Good	**64,733**	Sad	**5,215**
Ready	**18,230**	Fight	**5,016**
Sorry	**11,096**	Active	**4,884**
Blue	**8,005**	Angry	**1,363**

3.3 Extracting Sentimental Tweets

To decide whether a tweet is sentimental, we study the lexicon-based approach reported in prior work [8]. The approach forms a lexicon with sentimental words and detect if any of the sentimental words appears in the text.

First we apply information retrieval techniques to generate a list of 20 most frequent words using over two million game-related tweets from 50 games played in 4 weeks of the 2010-2011 NFL season. This list contains the most frequent appeared words according to the term frequency.

We then identify the part of speech of each word, i.e. noun, verb, adjective, etc. Previous work only considered adjectives, e.g., [8], and sometimes sentimental verbs, e.g., [9]. We expand the sentiment verbs in sports domain because action words are an essential part of sports. The most frequently mentioned action verbs are *"go"*, *"win"* and "*beat"*, etc. For example, *go team A* and *beat team A* emphasize sentiments toward team A through the verbs.

Finally, we produce a list of sentimental words using adjectives and verbs from the list of most frequent words as seed. Similar to sentiment detection in product review [8], we assume synonyms share the same sentiment orientation while antonyms share the opposite sentiment orientation. That means each sentimental word has either positive or negative sentiment orientation. We manually decide the sentiment orientation of the most frequent appeared words and utilize a large lexical database, WordNet [10], to search the synonyms and antonyms of the seed words to grow the lexicon. In addition, we also include the most frequent appeared slang and abbreviations in the sentiment lexicon such as *lol* and *wtf*, etc, which are heavily used and have clear senti-

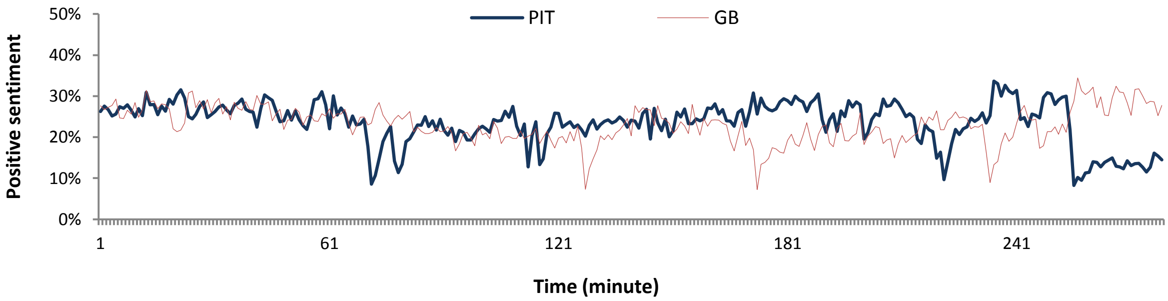

Figure 3: The trend of positive sentiments for both teams in the Super Bowl between Pittsburgh Steelers (PIT) and Green Bay Packers (GB).

ments. We create our lexicon based on millions of game related tweets. Therefore we believe it is accurate and comprehensive to determine the sentiment orientation in sports domain.

In all, we extract and collect 104 sentimental words (61 positive, 43 negative) and 16 emoticons. In our experiment including 50 games in 4 weeks, we find more than one third of the tweets contain an entry from the list of sentimental words and, therefore, are considered sentimental. Among the sentimental tweets, 87% are positive-only, 11% negative-only and only 2% contain both. The top 10 frequently mentioned sentimental words are either positive or negative as listed in Table 1.

3.4 Classifying Sentiment Polarity

We are only interested in if a tweet reflects positive or negative sentiments. Positive and negative sentiments include positive and negative emotions, evaluations, and stances, respectively. The sentiments are obvious in most tweets. For example: *"Good defence Titans!!!!!"* and *"TOUCHDOWN COLTS!!"* convey the positive sentiments. In contrary, *"The Titans are undisciplined and bad on Defence"* expresses the negative sentiments directly.

The primary reason that we only consider positive and negative sentiments is that the most frequent appeared sentimental words are either positive or negative. An apparent alternative to our positive-negative categorization is POMS, a well-established psychometric instrument. POMS (profile of mood states) assessment is a factor-analytically derived inventory that measures six identifiable mood or affective states. For example, Bollen *et al* [18] discovered the correlation between public mood and events in the social political, cultural and economic sphere by applying POMS to Twitter analysis. We examine the occurrences of POMS terms in the same 50 games that we extract positive and negative lexicons. Although POMS contains 6 factors that may provide more information about sentiments, we found that the POMS terms are rarely mentioned comparing to other frequent sentimental words except "Good" (see Table 2).

When we determine the sentiment orientation of a tweet, we consider special conditions including *but* clause, negation, and multiple sentimental words. When the tweet contains the *but* clause which starts with "but" or "however", the sentiment orientation in the clause is regarded as the tweet sentiment. When the negation word, i.e. "no" or "not", appears within a threshold distance, e.g. 3 words, around the sentimental words, the sentiment orientation is the opposite of its original [8]. When multiple sentimental words appear in the tweet, the number of positive/negative words determines the sentiment orientation. If positive or negative words dominate the tweet, the orientation of the tweet is regarded as positive or negative. If the tweet contains equal number of sentimental words, we consider the tweet as mixed of positive and negative sentiments. But we show that very few, only 2%, sentimental tweets have mixed sentiments.

3.5 Results

Next, we demonstrate our sentiments extraction results. Since positive sentiment is the dominant polarity which weighs almost 90% of sentimental tweets, we present the results using the percentage of positive sentiments in each minute. However, we still count the negative sentiment. When we calculate the percentage of positive sentiment, we first subtract the number of negative tweets from the number of positive tweets. Since the number of the negative sentiment is stable and infrequent, the negative sentiment does not impact the general trend of the sentiment.

We take the results of 2011 Super Bowl, which is shown in Figure 3, as an example to examine the effectiveness of our approach. In the beginning of the game, the percentages of positive sentiment of both teams are about the same. Along with the game progress, the trends of positive sentiment start to fluctuate. Note the trends of two teams tend to fluctuate apart when the events take place because the events impact oppositely on fans' sentiments of two teams. More interestingly, the percentage of positive sentiment towards the scoring or the winning team will not increase a lot but percentage towards the opponent will drop considerably. It is because the number of tweets related to the scoring or winning team also increases significantly that stabilizes the percentage. When the game is over, it is not surprise that the positive sentiment of the winning team dominates.

We obtained the similar results in another 30 games in the NFL regular season and playoffs. Thus, we conclude that our approach is adequate to extract the sentiments, especially when the events happen.

4. Twitter-Enabled Applications

Program event recognized and sentiments of TV watchers extracted by SportSense can be used in various ways to enhance the TV watching experience. We next discuss a few from our ongoing work.

4.1 Personalized EPG Overlays (e*PG)

Users consume large amounts of video content, ranging from broadcast TV to Internet video sources. Along with the increasing selection of content, discovery of that content is becoming a pain point for users. The general problem is given a dilemma of selecting a program (channel) among a large number of live programs, how does the user decide what is worth watching? This problem is particularly prevalent in the sports domain where it is not uncommon for fans to track multiple games that are occurring simultaneously. For example, up to 10 games are played on every Sunday afternoon at the same time during the NFL regular season. Furthermore, generically switching between games may cause users to miss particular moments of interest. Even though static EPG information can provide a certain level of information in regard to programming, this doesn't apply to live programming where events of interest are dynamic and vary among users.

Given the tools developed for SportSense, we utilize the real-time excitement level of each program to enable users to switch to the most exciting ongoing program (game). Existing work has applied visual, audio or webcast analysis to learn about the program content and its excitement level. Our work opens a novel, orthogonal direction by utilizing social data to infer about the audience reaction in real-time and in turn augment that information in the form of overlays for traditional EPG as shown in Figure 4. These overlays are generated using a combination of sentiment extraction and event recognition from the collected Twitter data (which is performed in real-time) as well as external information about the specific program/user as it is available. For example, if it's known who my favourite teams/players are, the overlays can indicate when any of them are participating in active games being displayed on the EPG. Furthermore, we can utilize general excitement levels on Twitter to indicate if games are 'interesting' – i.e. in overtime, or the scores are close. The general notion of mining Twitter for audience sentiment around TV and then translating that to specific information that can be associated with programming is a useful one due to the fact that there is a lot of information that is not captured by traditional metadata services (EPG being a concrete example). Tweets become the direct information reflecting the game events and audience sentiments. In addition, Twitter users spread out all over the broadcasting area such that the result is more representative.

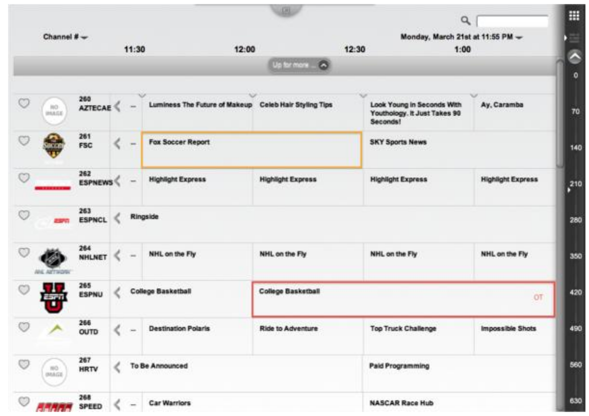

Figure 4: Screen snapshot of e*PG

4.2 Dual-Screen Advertising

Any additional knowledge regarding TV watchers (a.k.a. audience measurement) is important to advertisers, networks and stations. Advertisers want to know the size and characteristics of the audience they are reaching when they purchase ads on a particular program. The network or station cares about audience size and composition because they determine the amount it can charge for commercial time [11]. Methods such as programming rating and commercial rating are applied to measure TV and commercial audience. They also require deploying people meter to collect measuring samples [11]. Not only are such methods expensive, but also they rate a large chunk of a program or usually an entire program because commercials are delivered in the pre-planned break time of a program. As more and more commercials are embedded during the program without a break, there is a need to determine how engaged the audience is at a much finer grain. Because many TV watchers tweet while watching TV, we can estimate the audience excitement level by analyzing tweets from TV watchers in real-time. The analysis not only tells us how engaged the audience is overall but also tell which sector of the audience is more engaged. For example, in a sports game, we are able to tell what events are taking place and which side of fans is exciting. Advertisers can bid on the time slots based on the sentiments of their target customers.

Given the fact that we can utilize Twitter data to extract sentiment/events about programming that is being consumed by users, we applied this information as part of a framework for coordinating concurrent dual-screen

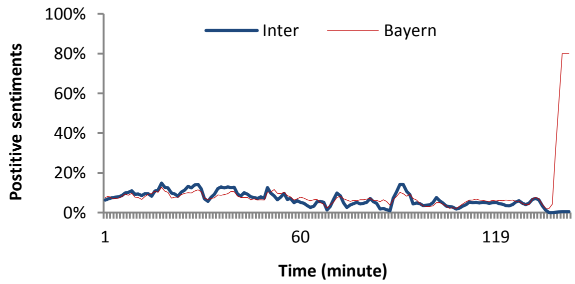

Figure 5: The trend of positive sentiments for both teams in a UEFA Champions League game between Inter Milan and Bayern Munich.

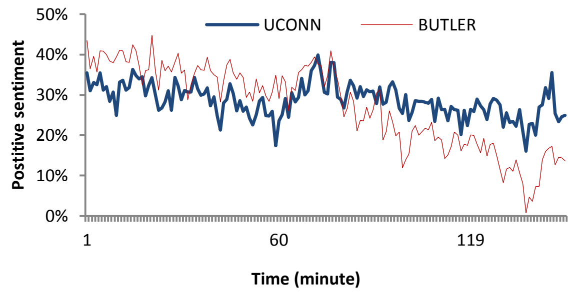

Figure 6: The trends of positive sentiment for both teams in the NCAA men's basketball tournament final between Univ. of Connecticut and Butler Univ.

campaigns for advertising. In a world where proliferation of second screens is increasing substantially and where a brand advertiser still has a sizable investment in the primary screen, e.g., TV, and uncontrolled second screen can be a liability or distraction. Our system for synchronized, dual-screen advertising offers TV advertisers brand and message protection through cross-screen context sharing & content coordination. Combining the data from aggregated social networks like Twitter, analyzing and extracting things like sentiment and events allows us to monetize any TV-related context as a biddable keyword for advertisers. For example Goldman & Sachs can synchronize the second screen experience, e.g., tablets and smartphones, with an event corresponding to the primary screen, e.g., touchdown = success+fame, thus playing an ad on the companion device when such an event has occurred in the game being watched on the primary screen. The system also enables advertisers to prevent things like brand dissonance between screens and provide brand/message protection. The system architecture provides enhancements to a traditional campaign manager to support the extra intra-program information gleaned from Twitter, as well as a messaging and notification system to execute second-screen campaigns on desired devices such as tablets or smartphones.

5. Other Sports Games

As we have already shown the effectiveness of the sentiments extraction on the NFL games, we next briefly demonstrate the generality of our approach on various types of sports. We apply our approach on two of the most popular sports, soccer and basketball as well. We choose one game from UEFA (the Union of European Football Associations) Champions League between Inter Milan and Bayern Munich, and one game from NCAA (National Collegiate Athletic Association) men's basketball tournament between University of Connecticut and Butler University. Both games are broadcasted on TV and have a large number of watchers.

Figure 5 shows the sentiment trends extracted our real-time Twitter analysis for both teams in the Champions League game. The only goal from Bayern Munich happened at the last minute of the game. Therefore, the sentiments regarding to Bayern Munich became extremely positive in the end. Although there was no goal in other part of the game, events such as yellow card or goal opportunity could impact the sentiments. Figure 6 illustrates the sentiments of both teams in the NCAA men's basketball tournament final. Since scoring events are common in the basketball game, the sentiments are changing smoothly during this game. When the University of Connecticut kept leading, its sentiments became more positive until the end of the game.

Although these two games are insufficient to prove the effectiveness of our approach on soccer and basketball, the results illustrate the feasibility of applying our approach on various sports games.

6. Related Work

Several concurrent projects also study tweets about sports games, however they do not provide real-time event detection. Hannon *et al* [12] used post rate of tweets to produce video highlights of the World Cup offline. They did not recognize game events nor did they produce highlights in real-time. Chakrabarti and Punera [13] assumed that a game event is already recognized and focused on describing the event using Hidden Markov Models trained with tweets collected from events happened in the past. Therefore, our focus on real-time event recognition is complementary, and addresses a more difficult and fundamental problem.

Existing works on sentiments measurement and opinion detection focus on product review and tweets moods modelling. Hu and Liu [8] mined and summarize the customer reviews of a product. Pang et al [14] and Zhuang [15] focus on sentiments classification in movie reviews. Jansen et al [16] investigated Twitter as a form of electronic word-of-mouth for sharing consumer opinions concerning brands. Extracting sentiments from sport games-related tweets are significantly different from product or movie reviews, because reviews have formal format and rich context information but tweets are colloquial without context. Furthermore, as people are emotional during sports games, their sentimental expressions are diverse and unexpected.

Bollen *et al.* [17] modelled public mood and emotion according to people's Twitter posts. Pandey and Iyer, Barbosa and Feng [18] proposed machine learning approaches to classify sentiments on tweets. They focus on tweets with certain expressions over a long time, e.g. one year in [17]. For sentiment extraction from game-related tweets, we must extract sentiments in real-time without the expression or structures leveraged by [17].

7. Conclusion

In this work, we described our ongoing effort in extracting real-time audience sentiments by analyzing Twitter. We showed that the limited vocabulary of sports games makes lexicon-based analysis methods highly effective. For several major sports games, including the US NFL, UEFA Champions League, and NCAA basketball tournament, SportSense is able to not only recognize major game events in real-time but also capture the sentiment toward each side of a game during the game. We described our ongoing work in leveraging the real-time event recognition and sentiment extraction of SportSense for a social TV system with socially informed electronic program guide (e*PG) and dual-screen advertising. Our work demonstrates the potential of Twitter as a key information source toward social TV systems.

8. References

[1] L. He, E. Sanocki, A. Gupta, and J. Grudin, "Auto-summarization of audio-video presentations," in *Proc. ACM Multimedia*, 1999.

[2] S. Nepal, U. Srinivasan, and G. Reynolds, "Automatic detection of 'Goal' segments in basketball videos," in *Proc. ACM Multimedia*, 2001.

[3] M. Shiels, "Twitter co-founder Jack Dorsey rejoin company, http://www.bbc.co.uk/news/business-12889048," 2011.

[4] S. Zhao, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "Human as Real-Time Sensors of Social and Physical Events: A Case Study of Twitter and Sports Games," Technical Report, Rice University, 2011.

[5] SportSense, "http://sportsense.us."

[6] V. Pandey and C. Iyer, "Sentiment Analysis of Microblogs."

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

[8] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. ACM SIGKDD*, 2004.

[9] T. Nasukawa and J. Yi, "Sentiment analysis: capturing favorability using natural language processing," in *Proc. ACM K-CAP*, 2003.

[10] G. A. Miller, "WordNet: a lexical database for English," in *Commun. ACM*. vol. 38, 1995, pp. 39-41.

[11] M. B. George Belch, *Advertising and Promotion: An Integrated Marketing Communications Perspective*: McGraw-Hill Irwin, 2009.

[12] J. Hannon, K. McCarthy, J. Lynch, and B. Smyth, "Personalized and automatic social summarization of events in video," in *Proc. ACM IUI*, 2011.

[13] D. Chakrabarti and K. Punera, "Event Summarization using Tweets," in *Proc. AAAI ICWSM*, 2011.

[14] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *Proc. ACM EMNLP ACL-02 Volume 10*, 2002.

[15] L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in *Proc. ACM CIKM*, 2006.

[16] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *Journal of the American Society for Information Science and Technology,* vol. 60, pp. 2169-2188, 2009.

[17] J. Bollen, A. Pepe, and H. Mao, "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena," in *Proc. AAAI ICWSM*, 2011.

[18] L. Barbosa and J. Feng, "Robust Sentiment Detection on Twitter from Biased and Noisy Data," in *Proc. ACM COLING*, 2010.