

Received April 2, 2019, accepted April 9, 2019, date of publication April 23, 2019, date of current version May 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2912009

# 'A Wisdom of Crowds': Social Media Mining for Soccer Match Analysis

YOUNG-SEOK KIM<sup>1</sup>, (Member, IEEE), AND MIJUNG KIM<sup>2</sup>, (Member, IEEE)

<sup>1</sup>Department of Creative IT Engineering, Pohang University of Science and Technology (POSTECH), Pohang, South Korea

<sup>2</sup>Korea Sport Industry Development Institute (KSIDI), Pohang University of Science and Technology (POSTECH), Pohang, South Korea

Corresponding author: Mijung Kim (mijungkim@postech.ac.kr)

This work was supported in part by the Sports Promotion Fund of the Seoul Olympic Sports Promotion Foundation from the Ministry of Culture, Sports and Tourism, and in part by the Ministry of Science and ICT (MSIT), South Korea, under the ICT Consilience Creative Program supervised by the Institute for Information and Communications Technology Planning and Evaluation (IITP) under Grant IITP-2019-2011-1-00783.

**ABSTRACT** The advent of social media has allowed channeling of the voice of sports fans that have essentially lead to gathering and storing fan-generated, large-scale opinions about sports match and team performance. Although research utilizing social media data for the purposes of supporting consumer market research has been increasing throughout the recent decade, there is a lack of studies using social media mining approach to improve team performance. In this paper, an opportunity mining approach is proposed to identify opportunities to improve team performance based on text mining and cluster analysis. A case study of the 2018 Fédération Internationale de Football Association (FIFA) World Cup final qualification of Korea, Korea versus Uzbekistan, was conducted to explain how the proposed method works. Fan comment data collected in the study revealed 16 different opportunities that would satisfy fans with regard to the team performance, and of those, two main extreme opportunities were identified.

**INDEX TERMS** Text-mining, clustering methods, opportunity algorithm, social network services.

## I. INTRODUCTION

Aside from great talent, exceptional teamwork, and dedicated training, numerous sports athletes, coaches, managers, teams, and leagues take advantage of Big Data as it has potential to provide insights regarding the critical factors associated with winning or losing.

One of the most popular and state-of-art methods for team performance analysis is to use video data that are captured from on-field cameras and crunched into thousands of data points per second by providing each players' performance metrics, such as player speed, position and possession time. However, this method of using video data is not cost efficient as data processing and analysis are complicated, computationally burdensome, and slow. More recently, quantitative approaches are applied such as using wearable device including GPS to measuring and calculating team performance based on running speed, distance, time, etc. However, this study proposes new approach to analyze team performance, especially focusing on the fans' perspective. It focuses on the value of the fans as a strategic partner in addressing some of

challenges that a team has. In the fields of business and management, the expertise, ingenuity, and creativity of individual members of the public are harnessed as an innovative problem solving approach. In this regards, this study suggests finding the factors associated with winning or losing from outside of the field.

Before, during and after a sporting event, fans share their opinions and unique analysis of the match. Moreover, evaluation and analysis by the experts and sports analysts provide not only useful information and knowledge of team performances but also wisdom to win. Thus, the comments, Op-Eds, and even tweets written by the public can be considered as a unique source of data to assess team performance which is essentially equivalent to "Wisdom of the Crowd". In this respect, this study analyzed online public opinions of the Korean National football team as a case study. In particular, this study uses Outcome-Driven Innovation (ODI) methodology, a strategy and innovation process of making product and marketing decision, to determine potential 'opportunities' with regards to team performance to help teams reach their goals (i.e. providing satisfaction to fans through winning) based on the "Wisdom of the Crowd."

The associate editor coordinating the review of this manuscript and approving it for publication was Chunsheng Zhu.

## II. RELATED WORKS

### A. TEXT MINING AND SPORTS MATCH ANALYSIS

There has been extensive work done on text mining related to sports using web news articles, webcast, or social media such as Twitter. One part of the research focuses on event detection or game analysis [1]–[4]. In particular, Xu *et al.* [1] used text information provided by soccer webcasts to detect soccer events and searched for corresponding soccer event images with the time information of the detected events. Jung *et al.* [2] proposed a method of association rule mining, event growth index, and pathfinder network analysis to extract team performance-related knowledge by using text information provided by webcasts.

The other part of the research using text mining is sentiment extraction from fans' responses and expressions to sports games or teams and athletes on social media [5]–[9]. For example, Zhao *et al.* [5], [6] and Zhao *et al.* [7] extracted real-time fan sentiments toward sports games by analyzing Twitter and found that the limited vocabulary of sports games made lexicon-based analysis methods highly effective. Park *et al.* [8] analyzed fan sentiments toward sports players as well as games by developing a sports sentiment dictionary.

Despite the growth of research fields applying text mining tools to analyze social media data and fan comments, currently there are no research on team performance analysis that utilizes fan requirements of sports team performance. This study focuses on fans' perspectives on team-performance as a 'Wisdom of the Crowd,' and examines the opportunity to improve team performance by satisfying fans' requirements through outcome-driven innovation (ODI) approach.

### B. OUTCOME-DRIVEN INNOVATION

The purpose of the opportunity algorithm is to identify the extent to which each team performance-related topic is a potential opportunity for improvement from a fan-centered view. Previously, the opportunity algorithm was used to prioritize the unmet needs of customers. Originally, the concept, developed by Ulwick [10], was proposed along with ODI for assisting businesses in the creation of breakthrough products and services. The underlying philosophy of ODI assumes that an opportunity for innovation exists when there is a highly important need yet the need is not well satisfied; the opportunity for creating value increases as the (customer's) need increases in its importance and as customer satisfaction decreases. Based on this philosophy, the opportunity algorithm computes opportunity (on a scale from 0 to 10 from a customer perspective) with importance and satisfaction as the key factors. The most important needs, but least satisfied receive the highest priority. The functional form of opportunity is defined as such:

$$O = I + \text{Max}(I - S, 0) \quad (1)$$

where  $O$  stands for opportunity;  $I$  for importance;  $S$  for satisfaction. Using such algorithm allows firms to find business opportunities for growth by ranking these opportunities in a priority sequence [11], [12]. Previously, the opportunity

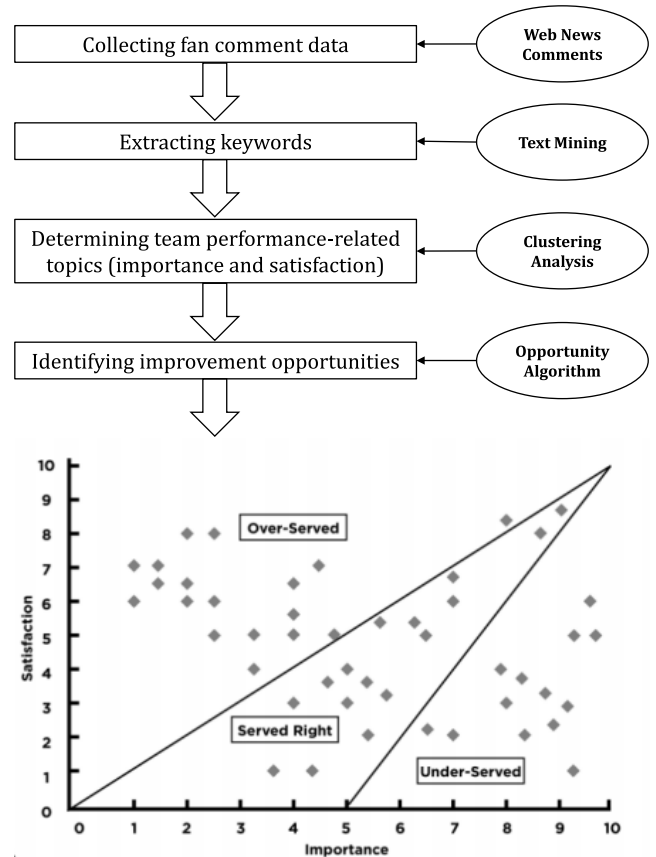


FIGURE 1. Overview of the proposed approach.

algorithm has been used to identify potential opportunities for product topics [11]–[13].

Recently, natural language text analysis approach has been used to determine the precise customers' needs in ODI research. For instance, Joung [14] analyzed customer complaint data from a firm and discovered high-priority customer requirements through the opportunity algorithm. Jeong [11] used the opportunity algorithm to identify product opportunities based on topic modeling and sentiment analysis of social media data.

This study suggests the use of textual data mining analysis and ODI method using sports fan comments to identify under-served fan requirements for team performance improvement. The proposed approach can evaluate the potential opportunity of sports team performance topics for improvement using online fan comments. This feature could help prioritize team performance directions for fan-centered game strategy development.

## III. PROPOSED METHODOLOGY

Our approach for identifying opportunity of team performance using online text mining (i.e., online fan comments) is built on three theoretical concepts: text mining, clustering, and opportunity algorithm (Fig. 1). First, team performance-related keywords discussed by fans are identified by applying

term frequency – inverse document frequency (TF-IDF) measure of information retrieval. TF-IDF is a numerical statistic which reflects how important a word is to a document in a corpus. It is the most common weighting method used to describe documents in the Vector Space Model [15]. Second, the team performance-related topics as well as the importance value of the topics are identified by clustering algorithm. Lastly, the opportunity score of each team performance-related topic is evaluated by the opportunity algorithm composed of the team performance-related topics' importance and satisfaction values. As a result, the directions for team performance improvement are deduced from the team performance-related topics that were highly unsatisfactory with a high opportunity score.

### A. STEP 1: GATHERING AND PREPROCESSING DATA

Web news comments related to a target sports match should be collected as the primary step of the proposed approach. The collected material should consist of online comment data generated by sports fans on web news; our approach identifies major team performance-related topics that are currently being discussed by fans and the satisfaction analysis of the team performance-related topics is performed from a fan-centered perspective therefore comments from fans should be the primary source of data. Collection of online comments may be conducted with techniques used to collect large-scale social media data such as web crawling via interfacing with application programming interfaces (APIs). An example API one can utilize to collect comment from is Reddit API (<https://www.reddit.com/dev/api>).

After the extraction/collection of a set of fan comments of the target sports match, the next step of our approach is to extract keywords (or key phrases) from the fan comments to structure each of the comments. Although a list of keywords can be obtained by calculating the TF-IDF, some of the keywords may be irrelevant, grammatically incorrect, or too generic for textual analysis. Identification of keywords can be facilitated by applying TF-IDF while excluding onomatopoeic words (e.g., 'haha', 'OMG'), emoticons (e.g., '^\_^', ':)') and irrelevant words (e.g., other matches rather than the target one) from the final keyword list. The final step in our proposed approach is to structure the fan comments as an array of keywords and their frequency that appear in their corresponding comments.

### B. STEP 2: IDENTIFICATION OF TEAM PERFORMANCE-RELATED TOPICS AND ASSESSMENT OF IMPORTANCE AND SATISFACTION

After preprocessing the data, the next step in the approach is to define team performance-related topics and compute the importance value of each topic using clustering algorithm. Clustering algorithms such as hierarchical clustering or k-means [16] are applied to a similarity matrix. The results of clustering analysis help ODI experts identify analogous fan requirements, and in turn, fan requirements can be identified by using the identified topics with the assigned importance

value. When analyzing clustering results, fan requirements must be inferred from the unambiguous evidence. Good clustering assigns one or two customer requirements per cluster.

Importance is measured by the frequency of fan comments in each cluster. If many fans mention about a factor, then developing a solution to the factor is an important fan requirement [17]. In this respect, the number of comments can be used as the surrogate for importance.

For computing the level of satisfaction regarding the topic of interest (i.e. team performance-related topic), the level of satisfaction may be categorized/labeled as satisfied, dissatisfied and neutral. The method for categorizing each comment may be based reaching a consensus (i.e. Delphi method) on the satisfaction level through content analysis or through an algorithmic procedure such as sentiment analysis (as done in Jeong *et al.* [11]). Due to the lack of high performing and accurate methods for categorizing/labeling the level of satisfaction regarding each topic, the Delphi method is recommended as of now. Specifically, given a set of topics (i.e. clusters) denoted as  $T_i$ , for  $i = 1, \dots, N$ , where  $N$  denotes the total number of topics, the average satisfaction of each topics, denoted  $TS_i$  as indicated in (2) is calculated as:

$$TS_i = \frac{\sum_{j=1}^{J_i} CS_{i,j}}{J_i} \quad (2)$$

where  $CS_{i,j}$  denotes the satisfaction level of the  $j$ -th comment in the  $i$ -th topic,  $TS_i$  denotes the average satisfaction for the  $i$ -th topic, and  $J_i$  denotes the total number of comments in the  $i$ -th topic. For the values imposed on the satisfaction labels,  $CS_{i,j} = 0$  for dissatisfied,  $CS_{i,j} = 5$  for neutral, and  $CS_{i,j} = 10$  for satisfied.

With the proposed methodology of clustering and opportunity score calculation in the framework of the ODI method, the opportunity algorithm may be used to identify team performance-related topics and their importance and satisfaction on a specific sports match in order to identify potential technical solutions for improving team performance.

### C. STEP 3: UTILIZING THE OPPORTUNITY ALGORITHM FOR IDENTIFICATION OF PERFORMANCE OPPORTUNITIES

As the final step, the opportunity algorithm is applied to the topics' importance and satisfaction scores, thereby allowing for the evaluation of the opportunity potential of each team performance-related topic and generation of an opportunity landscape map. Based on (1), the most important team performance-related topics that are least satisfied have the highest opportunity; it is these game-related topics that could be explored and researched for further performance improvement. Additionally, visualization through the opportunity landscape map allows for ease in understanding and evaluating the opportunity potential of each team performance-related topic. Specifically, the opportunity landscape map is represented by three partitions: served-right, over-served, and underserved (Fig. 1). The needs in the partition represented as "served-right" are considered appropriately satisfied, the needs in the partition represented as "over-served" are

considered satisfied in excess, and the needs in the partition represented as “underserved” are under-satisfied, with respect to the importance underlying the needs. Therefore, innovation opportunities may be considered through the topics that have been categorized as underserved needs based on the opportunity algorithm.

#### IV. CASE STUDY

A case study of 2018 FIFA World Cup final qualification of the Korean national football team, namely Korea versus Uzbekistan, was conducted to explain how the proposed method works. The authors chose this game because it includes the largest number of fan comments on Korea's top news portal website among 10 matches for 2018 FIFA World Cup qualification of the Korean national football team.

A total of 111,767 comments were gathered from Naver news, Korea's top news portal website (<http://sports.news.naver.com>) and only fan comments posted on the game day were collected and then translated from Korean to English. In this process, fan comments without team performance-related words such as names of the players and coaches were eliminated to extract more relevant keywords with the match. Moreover, considering characteristics of Korean language, specifically comments online, all grammatically incorrect words were excluded. To translate large-scale text, Google translation API was used. After translating, the authors divided each sentence into words to create a TF-IDF vector. The whole process was done using Python.

After preprocessing of the data, k-means clustering was used to identify team performance-related topics. The optimal number of clusters was found by using the elbow method.

When using a clustering-based approach to identifying importance, a frequency based method to compute the importance score for each topic may be employed. Specifically, given a set of topics (i.e. clusters) denoted as  $T_i$ , for  $i = 1, \dots, N$ , where  $N$  denotes the total number of topics, the importance of each topics, denoted  $I_i$  as indicated in (4) is calculated as:

$$TP_i = \frac{C_i}{\sum_{i=1}^N C_i} \quad (3)$$

$$I_i = \left( 9 \times \frac{TP_i - TP_{min}}{TP_{max} - TP_{min}} \right) + 1 \quad (4)$$

where  $C_i$  denotes the total number of comments in the  $i$ -th topic and  $TP_i$  in (3) denotes the proportion of topics mentioned for the  $i$ -th topic. The importance score is normalized on the scale of 1-10 based on the ratio of the difference in the maximum topic proportion (i.e.  $TP_{max}$ ) and the minimum topic proportion (i.e.  $TP_{min}$ ) also known as Min-Max normalization. Based on this approach, importance score was normalized between 1 and 10.

The satisfaction score was obtained by content analysis. Specifically, content analysis for fan satisfaction were divided into three categories (i.e., satisfied, not satisfied, and neutral). Fan comments including words to express their satisfaction with each team performance-related requirement (i.e. topic)

TABLE 1. A list of what fans requested.

No.	Underserved fan requirements	Opportunity level
1	Reforming Korea Football Association	Extreme opportunity
2	Replacing coach Shin Taeyong	Extreme opportunity
3	Offering the coach position to Hiddink Including Kim	Solid opportunity
4	Minjae to starting lineups	Solid opportunity
5	Improving players' performance for 2018 Russia Worldcup	Appropriately served
...		
16	Picking up new faces such as Lee Seungwoo and Baek Seunggho	Overserved

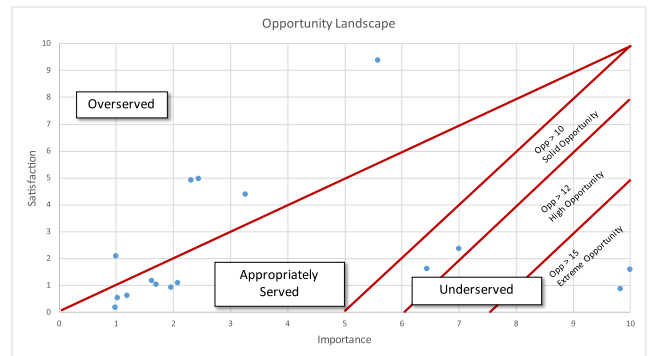


FIGURE 2. The opportunity landscape of Korea National football team performance.

were given 10 points and those including words to express their dissatisfaction were given 0 point. Comments with neutral expression were given 5 points. Three research assistants conducted independent content analysis on all comments and agreement was reached for comments with more than 1 differing labels. There were no comments that had greater than two different labels. The overall satisfaction for each team-related requirement were calculated in accordance with (2).

Lastly, the opportunity score was calculated based on the abovementioned importance and satisfaction scores for each fan requirement (i.e. topics).

#### V. RESULTS

The initial preprocessing of the fan comment data resulted in a total of 48,786 comments, and 18,398 features. The optimal number of clusters, based on k-means clustering and the elbow method, was 16 (Table 1). The highest satisfaction score in team performance-related fan requirements was 9.4 and the lowest was 0.2.

The analysis identified two extreme opportunities and two solid opportunities (Table 1, Figure 2). The two extreme opportunities were related to the governance of the Korea Football Association (KFA) and the role of the head coach,



Shin Taeyong, respectively: fans wanted the KFA (1) to be reformed (Importance score 10.0, satisfaction score 1.6) and (2) to change the head coach (Importance score 9.8, satisfaction score 0.9). The solid opportunities were to offer Hiddink, the former head coach of the Korean National football team, the head coach position again (Importance score 7.0, satisfaction score 2.4) and to include Kim Minjae in the starting lineup (Importance score 6.4, satisfaction score 1.7). On the other hand, there were five overserved items: (1) Picking up new faces such as Lee Seungwoo and Baek Seungho (Importance score 1.0, satisfaction score 2.1), (2) Changing strikers from Hwang Heechan and Sohn Heungmin to Lee Dongguk (Importance score 2.3, satisfaction score 4.96), (3) The best player was Yeom and the worst player was Kwon (Importance score 2.5, satisfaction score 5.0), (4) Lee Dongguk missed some chances but would shoot better in the next game (Importance score 3.3, satisfaction score 4.4), and (5) The best defense player was Kim Minjae (Importance score 5.7, satisfaction score 9.4).

As a result, by identifying fan requirements and prioritizing them, the Korean National football team can apply a selection and concentration strategy in developing an innovative strategy for improved team performance. As demonstrated through the case study, sports teams can reduce energy wasted on performance related to overserved fan requirements while paying close attention to performance and strategy related to underserved fan requirements. The underserved fan requirement can provide 'wisdom' in developing strategies that meet fans' needs, and may potentially lead to winning games based on the importance and dissatisfaction expressed by fans.

## VI. CONCLUSION

This study proposed the utilization of online fan comments for the identification of sports team performance improvement opportunities through combining text mining and opportunity algorithm as the core methodology. Regarding the specific steps of the approach, each team performance-related topics from fans' perspective was defined by clustering of the online fan comment data. Following the clustering, each topic's measure of importance and satisfaction were computed. Specifically, the concept underlying the computation of importance (of the topic) is contribution (i.e., number of times mentioned) and the concept underling the computation of satisfaction (of the topic) is sentiment (i.e., positive vs. negative). As the final step, the opportunity algorithm is utilized in order to assess the opportunity value and improvement direction of team performance-related topics from the perspective of sports fans. The approach was demonstrated using the 2018 FIFA World Cup final qualification of the Korean National football team matching against Uzbekistan in September 5<sup>th</sup>, 2017. The performance improvement directions based on the top 16 topics for the Korean National football team were found through this case study. The topic with the highest opportunity was about the role of the sports governance body; specifically, the management practice and policies of the Korean National football team yielded the

largest opportunity value. The proposed approach contributes to the identification and assessment of new strategy opportunities across various sports domains, using fan comments and social media data. It thereby assists team coaches, athletes, and sports governance bodies to identify areas for improving performance by capturing potential opportunities in the perspective of sports fans. Furthermore, given the underlying method provides "opportunities" to be mined as long as there are the "Wisdom of the Crowd", the proposed method is generalizable to other domains. For instance, utilizing google and yelp reviews as the "Wisdom of the Crowd" may allow restaurant owners to improve aspects of their hospitality management without having to spend an arm and a leg on management consultants. As such, if there is an opportunity for improvement in a certain area in addition to the "Wisdom of the Crowd", the proposed method should be able to reveal opportunities based on the concept of importance and satisfaction.

## APPENDIX

TABLE 2.

Cluster No.	Keywords
1	Lee Seungwoo, Baek Seungho, young players, pick, use, player, line, generation, need, old
2	Shin Taeyong, director, resign, time, finals, football, interview, soccer, Korean national game, learn
3	Come, Hiddink, coach, football association, money, Korea director, good, want
4	Football, association, soccer, Korean, money, real, change, team, problem, new, director
5	Yeom Kihoon, Kwon Changhoon, today, pass, game, Lee Dongguk, cross, time, national, class, best, real
6	Hiddink, bring, comes, coming, petitions, Korean football, like, answer, people, soccer, Korea, director
7	Coach, Shin Taeyong, new, learn, team, national football, world, senior director, sorry
8	Ki Sungyong, know, really, play, team, coach, make, director, change, member, fans, right, lose, sorry
9	Worldcup, finals, football, time, team, Russia, national Korean, play, able, Iran, experience, advance
10	Heungmin Sohn, Tottenham, national team, really, today, real play, shot, goal, did, ace, player, army
11	Hwang Heechan, Sohn Heungmin, change, striker, good, better, like, young, shot, game, pass, goal, Lee Dongguk
12	Kim Minjae, good, like, real, football, defense, people, better, Korean, want, know, national association, best, let
13	Korea football association, let, people, make, coach, bring, money, change, director, resign Hiddink, stop, leave
14	Park Jisung, Lee Youngpyo, hiddink coach, like, miss, national player, team, players, come, want, really, think, 2002, good, retirement, great, legend
15	Lee Dongguk, good, shot, did, minutes, change, time, player, goal, real, world cup, shooting, best, striker, effective
16	Chang Hyunsoo, Kwon Changhoon, injured, game, better, good, today, replacement, injuries, look clam, replaced

## ACKNOWLEDGMENT

The authors would like to thank Dokyun Kim, Jaewon Shin, and Junsoo Jo for satisfaction analysis.

## REFERENCES

- [1] C. Xu, J. Wang, K. Wan, Y. Li, and L. Duan, "Live sports event detection based on broadcast video and web-casting text," in *Proc. 14th ACM Int. Conf. Multimedia*, 2006, pp. 221–230.
- [2] H.-S. Jung, J.-U. Lee, J.-H. Yu, H.-S. Lee, and D.-H. Park, "In-depth analysis of soccer game via webcast and text mining," *J. Korea Contents Assoc.*, vol. 11, no. 10, pp. 59–68, Nov. 2011. [Online]. Available: <https://www.dbpia.co.kr/Journal/ArticleDetail/NODE01713822>
- [3] J. Nichols, J. Mahmud, and C. Drews, "Summarizing sporting events using twitter," in *Proc. ACM Int. Conf. Intell. User Interfaces*, 2012, pp. 189–198.
- [4] C. Xu, Y.-F. Zhang, G. Zhu, Y. Rui, H. Lu, and Q. Huang, "Using webcast text for semantic event detection in broadcast sports video," *IEEE Trans. Multimedia*, vol. 10, no. 7, pp. 1342–1355, Nov. 2008. [Online]. Available: <http://ieeexplore.ieee.org/document/4668533/>
- [5] 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*. Houston, TX, USA: Rice Univ., 2011. [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1106/1106.4300.pdf>
- [6] S. Zhao, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "Analyzing twitter for social TV: Sentiment extraction for sports," in *Proc. 2nd Int. Workshop Future Telev.*, 2011, pp. 11–18.
- [7] S. Zhao, L. Zhong, J. Wickramasuriya, V. Vasudevan, R. LiKamWa, and A. Rahmati, *SportSense: Real-Time Detection of NFL Game Events from Twitter*. Houston, TX, USA: Rice Univ., 2012. [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1205/1205.3212.pdf>
- [8] S. G. Park, G. S. Won, and S. W. Lee, "Web news comment-based sentiment analysis of the South Korean national team members in the 2014 Brazil World Cup," *Korean J. Sport Manage.*, vol. 20, no. 2, pp. 13–28, Apr. 2015. [Online]. Available: <http://kiss.kstudy.com/thesis/thesis-view.asp?key=3317731>
- [9] C.-M. Chen and L.-H. Chen, "A novel approach for semantic event extraction from sports webcast text," *Multimedia Tools Appl.*, vol. 71, no. 3, pp. 1937–1952, Aug. 2014. doi: [10.1007/s11042-012-1323-6](https://doi.org/10.1007/s11042-012-1323-6).
- [10] A. W. Ulwick, *What Customers Want: Using Outcome-Driven Innovation to Create Breakthrough Products and Services*. New York, NY, USA: McGraw-Hill, 2005.
- [11] B. Jeong, J. Yoon, and J.-M. Lee, "Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis," *Int. J. Inf. Manage.*, to be published. doi: [10.1016/j.ijinfomgt.2017.09.009](https://doi.org/10.1016/j.ijinfomgt.2017.09.009).
- [12] T.-M. Yen, Y.-C. Chung, and C.-H. Tsai, "Business opportunity algorithm for ISO 9001: 2000 customer satisfaction management structure," *Res. J. Bus. Manage.*, vol. 1, no. 1, pp. 1–10, Jan. 2007. [Online]. Available: <https://scialert.net/abstract/?doi=rjbm.2007.1.10>
- [13] C. P. Killen, M. Walker, and R. A. Hunt, "Strategic planning using QFD," *Int. J. Qual. Rel. Manage.*, vol. 22, no. 1, pp. 17–29, Jan. 2005. doi: [10.1108/02656710510572968](https://doi.org/10.1108/02656710510572968).
- [14] J. Joung, "Customer complaints analysis for strategic product development using textual data mining and the outcome-driven innovation method," *Int. J. Intell. Enterprise*, Dec. 2017.
- [15] M. S. T. Deokar, "Text documents clustering using K means algorithm," *Int. J. Technol. Eng. Sci.*, vol. 1, no. 4, pp. 282–286, Jan. 2013. [Online]. Available: <http://ijtes.com/upload/10.text%20clustering.pdf>
- [16] M. S. G. Karypis, V. Kumar, and M. Steinbach, "A comparison of document clustering techniques," in *Proc. KDD Workshop Text Mining*, 2000, pp. 1–20.
- [17] J.-C. Chebat, M. Davidow, and I. Codjovi, "Silent voices: Why some dissatisfied consumers fail to complain," *J. Service Res.*, vol. 7, no. 4, pp. 328–342, May 2005. [Online]. Available: [https://www.researchgate.net/profile/Jean-Charles\\_Chebat/publication/258158576\\_Silent\\_Voices\\_Why\\_Some\\_Dissatisfied\\_Consumers\\_Fail\\_to\\_Complain/links/0c96052c6ef5864455000000/Silent-Voices-Why-Some-Dissatisfied-Consumers-Fail-to-Complain.pdf](https://www.researchgate.net/profile/Jean-Charles_Chebat/publication/258158576_Silent_Voices_Why_Some_Dissatisfied_Consumers_Fail_to_Complain/links/0c96052c6ef5864455000000/Silent-Voices-Why-Some-Dissatisfied-Consumers-Fail-to-Complain.pdf)



**YOUNG-SEOK KIM** (M'16) received the Ph.D. degree in sports science from Seoul National University, Seoul, South Korea, in 2014. He is currently the Director of the Korea Sport Industry Development Institute, POSTECH. For ten years of his professional career, he not only developed broad understanding of global sport industry but also collaborated with various industry stakeholders and demonstrated his leadership by working with the International Olympic Committee (IOC)

and the Fédération Internationale de Football Association (FIFA). Corresponding to the innovative government policy of sport industry, he discovered a valuable vision and future of sports IT convergence and has joined one of the leading engineering schools, POSTECH, to pursue interdisciplinary research. His research interests include sports engineering, human–computer interface/interaction (HCI), sports, and sports science and technology. He is currently a Member of the Editorial Board of the *Journal of Global Sports Management* and the Chair of the Sports IT Convergence Division in the Korean Society for Sports Management. He was a recipient of the Achievement Award for Dedication to the Development for Sport IT Industry, Minister of Culture, Sports and Tourism, South Korea, in 2017.



**MIJUNG KIM** (M'16) received the Ph.D. degree in telecommunication, information studies and media from Michigan State University, East Lansing, USA, in 2014. She is currently a Senior Research Associate with the Korea Sport Industry Development Institute, POSTECH. She has been collaborating with great engineering scholars to address interactive technology in sports at POSTECH, since 2015. Specifically, her team's research aims to explore the technical potential of sensors (e.g., wearable computing) and the use of ICT to motivate people to move or exercise by providing social support or through gamification, and create innovative interaction techniques. Her research interests include social media analysis, media habit formation, and new media effects based on her academic background of social sciences.

...