International Journal of Multimedia and Ubiquitous Engineering
Vol.8, No.5 (2013), pp.361-374
http://dx.doi.org/10.14257/ijmue.2013.8.5.36

How Did Reviewers Give Their Reviews? Characterizing Review Activity in Academic Conferences

Zhiwen Yu[1], Xinxin Zhang[1], Yunji Liang[1] and Daqing Zhang[2]

[1]School of Computer Science, Northwestern Polytechnical University
Xi'an, Shaanxi, P. R. China
[2]Telecommunication Network and Services Department
Institut TELECOM SudParis
Evry Cedex, France

[1]zhiwenyu@nwpu.edu.cn, [2]daqing.zhang@it-sudparis.eu

Abstract

Human activities have been investigated and applied in various fields, such as context-aware computing, search engine, social network services, location-based services, automated visual surveillance, and multimodal human–computer interaction. However, the human activities in the review process of academic conferences have seldom been explored. There is no doubt that review process plays an important role in deciding whether a paper can be accepted or not. In this paper, we present our work to understand the review activities by analyzing the anonymized review data of two conferences (ACM SIGCOMM and UIC). The descriptive statistics and the data mining technology are adopted in the analysis. We got some interesting knowledge, which is significant for interpreting how the reviewers give their reviews in academic conferences, such as the relationships between the score, confidence and review length, and reviewer activity patterns.

Keywords: *review activity, academic conference, statistical method, data mining*

1. Introduction

Human activity is one of the most important characteristics for explaining individual or group variability. Recognizing human activity benefits many applications, such as context-aware computing [13], search engine [4], social network services [8, 3], location-based services [12, 11], automated visual surveillance [9], and multimodal human–computer interaction [1]. For example, through analyzing and predicting human browsing behavior, the searching quality of search engine can be improved [4].

Current activity analysis mainly focuses on human physical activity, such as primitive actions (*e.g.*, walking, running, and jumping) and activities of daily living (*e.g.*, making coffee, brushing teeth, and eating meal) [5], whereas little research has been conducted on their intellective activity. Analyzing human intellective activity is useful to understand how they conduct intellective works. Reviewing journal or conference papers is a typical intellective activity.

During the review process of an academic conference, reviewers are asked to assess the quality of the assigned papers, and the results will be used to determine whether a paper can be accepted or not. However, the objectivity of the review results can not be always guaranteed. One reason is that it is difficult to set an absolutely objective standard to evaluate a paper. Furthermore, reviewers may have different opinions on the evaluating metrics.

Therefore the evaluation metric setting and review activity are important in academic conferences. Through analyzing paper reviews, we can find underlying relationship between different metrics and typical review activity patterns. Such knowledge can be used as suggestions for conference organizers in setting review parameters in order to make the review process more efficient and fairer. Moreover, the discovered review patterns can be utilized to compare two academic conferences.

In this study, we intend to understand the review activities by analyzing the anonymized review data of two conferences. Through the descriptive statistical method, we obtain an intuitive picture of the data illustrated with charts, graphics, etc. With the data mining technology, specifically clustering, we discover some implicit knowledge, such as the review activity patterns and reviewer patterns.

The rest of this paper is organized as follows. In Section 2, we discuss previous related studies. Section 3 describes the data used in our study. The methods for analyzing are presented in Section 4. Section 5 elaborates the results. Finally, we conclude the paper in Section 6.

2. Related Work

Activity recognition has received much attention in the past decades. Human activities can broadly be categorized into individual activity and group activity. Individual activity refers to the actions or behavior of a single person. Rashidi *et al.*, [20] mined and tracked human activities by using the sensing data collected in physical smart environments, and then detected changes in an individual's patterns and lifestyle. Kim *et al.*, [10] classified different individual activities, such as running, walking, crawling, and sitting, based on micro-Doppler signatures. In [16], head tracking is used to recognize individual actions in a meeting room like entering, leaving, walking, getting up, sitting down, and being at whiteboard.

Group activities are performed by more than one person. The changing number of people and interaction between them make the recognition of group activity much more complicated. Ni *et al.*, [17] recognized human group activities with localized causalities (self-causality, pair-causality, and group-causality) through analyzing surveillance video database. Morita *et al.*, [15] proposed a mining method for multimodal interactions to extract important patterns of group activities. McCowan *et al.*, [14] recognized group actions in meetings by modeling the joint behavior of participants based on a two-layer HMM framework.

Different from the above studies, which focus on analyzing human physical activity, this paper aims to understand human intellective activity during the paper reviewing process for academic conferences.

The analysis of reviewer activity has seldom been reported in literatures. Numerous studies have been done to evaluate the paper review process. Douceur *et al.*, [7] proposed an approach of ranking papers rather than the traditional way of rating, when evaluating paper individually or aggregately. The goal is to ensure that the quality of every accepted paper is higher than every rejected paper. Crowcroft *et al.*, [6] identified several problems with the current review process. They proposed a grand unified mechanism that gives incentives to authors, reviewers, and the community to do "right things".

Several works used the review data to investigate the review process. Anderson [2] conducted a study showing significant differences among reviewers. Some factors leading to a less efficient and less effective conference review process were investigated. Papagiannaki [18] proposed to use the review data and author feedback to evaluate the reviews. Through data analysis, some questions about reviewers and the interpretations were identified. Papagiannaki and Rizzo [19] described the whole review process of the conference of SIGCOMM2009 and presented some discussion about the review form.

Our work is partially related to [19], both use the review data of SIGCOMM2009. However we aim to discover the review and reviewer activity patterns based on the review data other than discussing the review process.

3. Data

3.1. Data Overview

To discover the unobserved behavior patterns of reviewers with academic papers, we got the review dataset of the premium conference, ACM SIGCOMM 2009. For comparison, we also collected the reviews of UIC 2009 and 2010 (International Conference on Ubiquitous Intelligence and Computing). The dataset of UIC includes the reviewer information which is used to analyze the variability between reviewers. Both datasets were anonymized in order to protect the reviewer privacy. We use SIGCOMM and UIC to refer to the two datasets in the following description.

SIGCOMM dataset contains 820 records while UIC dataset has 536 records. The properties of the dataset of SIGCOMM and UIC are not totally consistent, but they share three available properties score, confidence, and character count. Score refers to the quality of the paper in the reviewer's opinion. Confidence is used to evaluate how well the reviewer believes the score he or she gives to the paper. It also reflects the reviewer's expertise on the paper. Character count is the length of a particular review.

In the SIGCOMM dataset, the score of submissions ranges from 1 to 5, where 1 means the poorest and 5 means the best, while in UIC, the score ranges from -3 to 3. The confidence interval scale is from 1 to 4 in the SIGCOMM and from 0 to 4 in the UIC conference. Character count does not have any specific interval. Within the two datasets, the shortest review has no characters while the longest one has nearly 20000 characters.

3.2. Data Preprocessing

As the data format in the two datasets is not consistent, we need to prepare the data with universal format before it can be analyzed. Three preprocessing methods are involved: discretization, normalization, and representation.

3.2.1. Discretization:

Data like character count with variable length should be discretized at first. We sort the data in ascending order, and then assign three points to split the data into 4 intervals which are labeled as Short, Medium, Long and Very long, as shown in Table 1. For instance, 8.9% in the line of SIGCOMM means the number of reviews that fall into *Short* occupies 8.9 percentages in the SIGCOMM conference.

Table 1. The discretized character count of two datasets

Conference	Short	Medium	Long	Very long
SIGCOMM	8.9%	51.1%	27.56%	12.44%
UIC	2.8%	58.02%	29.10%	10.08%

3.2.2. Normalization

When analyzing review pattern with clustering method, we need to calculate the similarity between different reviews. The first step is to calculate the standard dissimilarity between the three properties (score, confidence, and character count) of different reviews. Standard dissimilarity means the dissimilarities we get are all in an interval from 0 to 1. To ensure that, we need to normalize the raw data to make sure all of them are in a range of [0, 1]. The normalization equation is shown as below.

$$x' = \frac{x - \min}{\max - \min} \tag{1}$$

where x is the raw value of a property, max and min refer to the maximum value and minimum value of raw data, and x' means the corresponding value of x in the new standard scale from 0 to 1. For example, we can normalize the score set {1, 2, 3, 4, 5} in SIGCOMM to a new set of {0, 0.25, 0.5, 0.75, 1}.

3.2.3. Data representation

To analyze the patterns of different reviewers, the dataset of UIC needs to be organized in a new format. Take the score as an example (see Table 2), the table is two-dimensional where the horizontal axis shows how many papers carry a specific score and the vertical axis represents different reviewers. The total number of the reviewers is 145. For instance, Reviewer 2 gave the score of '-2' for one paper, '-1' for one paper, and '2' for two papers. Similar reorganization is performed on the property of confidence and the discretized character count.

Table 2. The reorganized form of score

Reviewer id	Score						
	-3	*-2*	*-1*	*0*	*1*	*2*	*3*
1	0	1	0	1	1	1	0
2	0	1	1	0	0	2	0
3	0	1	1	1	0	0	0
…	…	…	…	…	…	…	…
145	0	0	2	0	2	0	0

For easy processing in the data analysis, we represent the transformed data with vectors, such as $reviewerScore(i) = (x(1), x(2), x(3), x(4) \cdots)$, where i is the reviewer id, $x(k)$ represents the frequency of a specific item of Score. For example, $reviewerScore(1) = (0,1,0,1,1,1,0)$ is the score vector of Reviewer 1, which indicates that he or she reviewed four papers with the scores of -2, 0, 1 and 2 (refer to TABLE II), respectively. $reviewerConfidence(1) = (0,0,2,2,0)$ means that Reviewer 1 gave the confidence of 2 to two papers and 3 to another two papers (confidence in UIC ranges from 0 to 4).

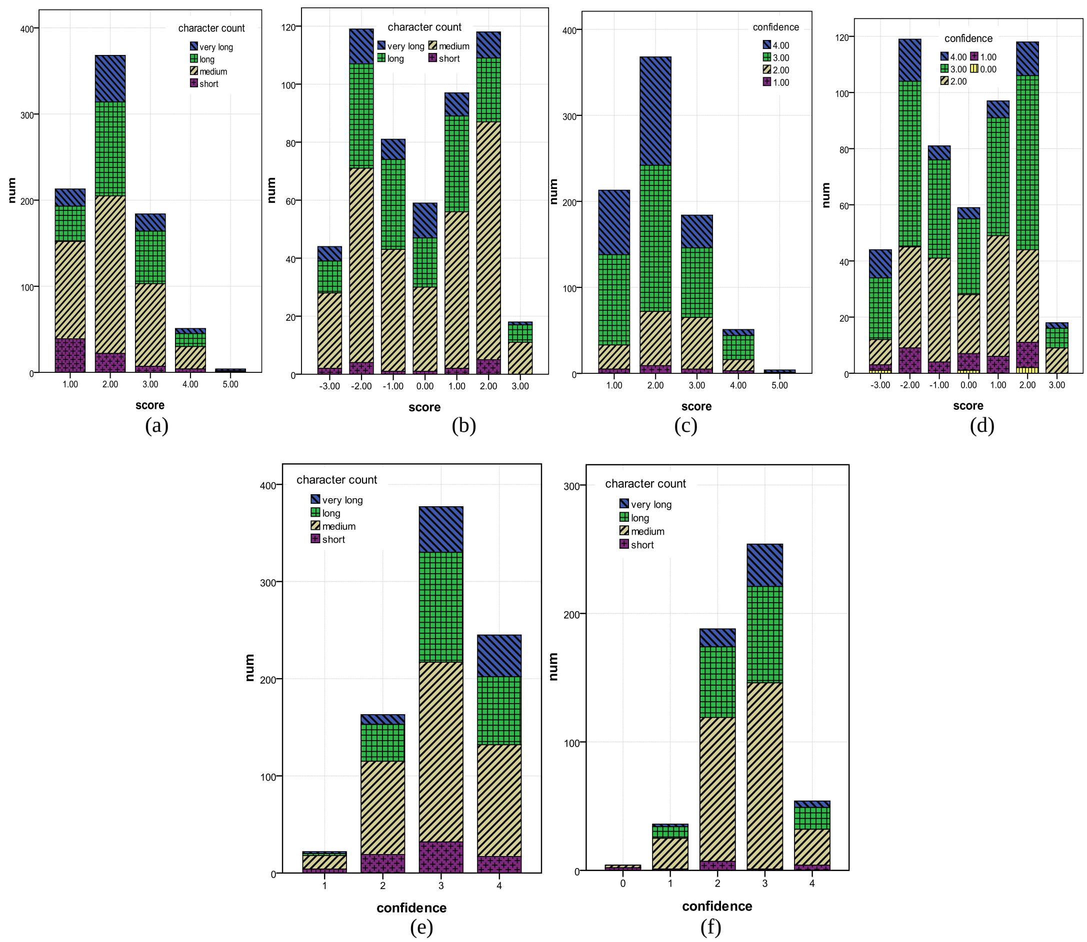


Figure 1. Correlation: (a) score and character count in SIGCOMM, (b) score and character count in UIC, (c) score and confidence in SIGCOMM, (d) score and confidence in UIC, (e) confidence and character count in SIGCOMM, (f) confidence and character count in UIC

3.3. Correlation between Score, Confidence, and Character Count

We illustrate the correlations among score, confidence, and the length of review characters in Figure 1.

3.3.1. Score and character count

SIGCOMM: As shown in Figure 1(a), reviewers have a tendency to score papers in the low and medium area (from 1 to 3), and only 55 papers carry scores of 4 or 5. Furthermore, most of reviewers' feedbacks are in a range of medium to long. We can also observe that the lowest score tends to be correlated with a short review (above 50% of short reviews) and the score 2 is likely with 'very long' review.

UIC: According to Figure 1(b), nearly 80% of papers score in a range from -2 to 2, especially many papers are with the score of -2 or 2. In addition, papers with 'very long' review usually do not appear with a score of 5. Meanwhile, when the review is too short, the paper will not get the highest score.

Generally speaking, the reviewers are reluctant to rank papers with the extremes of the score scale. What's more, long reviews rarely appear in papers with the highest scores and more papers are with the character count of 'very long' when their scores are in the low or medium area.

3.3.2. Score and confidence

SIGCOMM: As shown in Figure 1(c), only 3% papers get the lowest confidence and less than 15% of those get the highest score. Over 50% of papers that carry a medium score (*e.g.*, 2 and 3) get the highest confidence.

UIC: From Figure 1(d) we can see that no papers with a low confidence like 0 or 1 get the highest score.

In a word, the most diffident reviews seldom appear with the highest score. In addition, most confident reviews appear in the papers with a medium score.

3.3.3. Confidence and character count

SIGCOMM: Figure 1(e) shows that in SIGCOMM, more papers (nearly 80%) are having the top two confidence values. It means the reviewers of SIGCOMM were very confident in their review. In addition, papers with the longest feedbacks usually appear with the confidence of 3 or 4.

UIC: According to Figure 1(f) we can observe that more than 80% papers get a confidence of 2 or 3, which means the UIC reviewers were not as confident as those of SIGCOMM. Furthermore, around half papers are with Very long or Long reviews. However, the papers with the lowest confidence did not get longer reviews.

Generally we get the following observations: (1) fewer papers are provided with a low confidence, (2) higher confidence always appears with longer reviews, and (3) the lowest confidence seldom appears with a long review.

4. Data Analysis

To understand the review activities, we analyze the anonymized review data of the two conferences by using two methods: descriptive statistics and data mining. Through descriptive statistical method, we can get an intuitive understanding of the data. Through data mining technology, some implicit analysis results like the review and reviewer patterns can be obtained.

4.1. Descriptive Statistical Method

Descriptive statistical method [21] is useful to reveal the general characteristics of a dataset through classification, generalization and calculation. It can be used to obtain potential features in review activities. Important information behind the data can be intuitively illustrated with charts, graphics, etc. Two specific descriptive statistical methods are adopted in this paper as follows:

- Frequency can be used for score and confidence statistics. For example, how many papers get a score of 1, 2, or others, and can tell us which score the reviewers prefer to mark. Histogram and scatter plot are used to show the frequency. The divided bar graph can not only show the frequency of one attribute but also show the relationship between two continuous variables.

- A scatter plot can also be used to visually illustrate a relationship between two continuous variables. The subtle connections between score and character count, or confidence and character count can be easily obtained from this kind of plot through certain processing like jittering or fitting.

4.2. Data Mining Method

Descriptive statistics is a simple and convenient method to obtain intuitive characteristics from data; however, it can not offer us more comprehensive analysis results. Data mining is a powerful method of discovering new knowledge. We first use a graph to represent the reviews and their relationship and then present a mining method for analyzing the graph.

Definition 1 (Review Graph): A graph is used to represent the reviews and their relationship, denoted as $G=(V,E)$, where $V=\{0,1,...,n\}$ is the set of nodes representing individual reviews, and $E=\{(v_i,v_j)\,|\,v_i,v_j\in V, v_i\neq v_j\}$ is the set of edges connecting the reviews. Each node has three properties: score, confidence, and character count, denoted as $V_i=\{score_i, confidence_i, character\ count_i\}$. Whether two nodes should be connected relies on their similarity.

We designed a graph-based review clustering method to analyze different review activity patterns. The method is depicted as Algorithm 1. The algorithm first calculates the similarity between two reviews (Step 1). If the similarity is larger than a threshold r, it connects the two nodes in the graph (Step 2). The algorithm then scans the graph with Breadth First Search (BFS) based on a threshold Lambda to find different clusters (Step 3). It finally outputs the clusters of reviews (Step 4). The similarity calculation and clustering are described in detail in the following subsections.

Algorithm 1 Graph-based review clustering
Input: a review dataset, threshold *r* and *Lambda*
Output: clusters of reviews
Procedure:
(1) calculate the similarity between every two reviews
(2) if the similarity is larger than *r*, link the two reviews in the review graph
(3) scan the graph with Breadth First Search (BFS) based on *Lambda* to find different clusters
(4) output the clusters of reviews

4.2.1. Review similarity calculation

To get the similarity between two reviews, we first calculate the dissimilarity of each property $diss_k(x,y)$, and then use equation (2) to get the total dissimilarity between two reviews. And finally, the similarity can be obtained according to (3).

$$dissimilarity(x,y)=\frac{\sum_{k=1}^{n} diss_k(x,y)}{n} \tag{2}$$

$$similarity(x,y)=e^{-dissimilarity(x,y)} \tag{3}$$

In this paper, we use (4) to get the dissimilarity of score, confidence, and character count between every two reviews.

$$diss_k(x, y) = x' - y' \tag{4}$$

Here x and y are score, confidence or character count of two objects, x' and y' are the normalization of x and y.

4.2.2. Clustering

We adopt Breadth First Search (BFS) method to traverse all the nodes and divide them into several clusters. Parameter *Lambda* is defined to determine whether a non-classified node belongs to the current cluster. The specific algorithm is shown in Algorithm 2.

To determine whether the current node V_i belongs to cluster A_k, we first calculate the summation of the connectivity between V_i and each existing node in A_k (Steps 4-5). Then calculate the average connectivity, and if the average connectivity is larger than the threshold *Lambda*, we put node V_i into A_k, otherwise, return to Step 3 to determine whether V_i belongs to A_{k+1} (Steps 6-9). Traverse all the nodes in the graph with BFS and finally each node will fall into one of the clusters.

```
Algorithm 2  Clustering algorithm
(1)  for  each node Vi  in the graph
(2)  if  Vi is non-classified
(3)      for  each cluster Ak
(4)          for  every node Ak [i] in cluster Ak
(5)              sum += adjacent(Vi, Ak [i])
(6)          if (sum/Ak.length)>Lambda
(7)              Vi belongs to Ak
(8)          else
(9)              k=k+1, return to step(3)
(10)     i=i+1, return to step(1)
(11) End
```

We use graph-based clustering method to cluster the reviews of SIGCOMM and UIC. A simple K-means clustering method is utilized to discover the reviewer patterns of UIC, and compared with the results of statistical classifying method.

5. Results

With the methods described above, we analyze the two datasets and present the results in this section. Data analysis tools, such as SPSS (Statistical Product and Service Solutions) and MATLAB are utilized here. Raw data is preprocessed with both the two tools. We use SPSS to implement the K-means clustering. Through the graph-based clustering algorithm implemented with MATLAB, we get the review clusters.

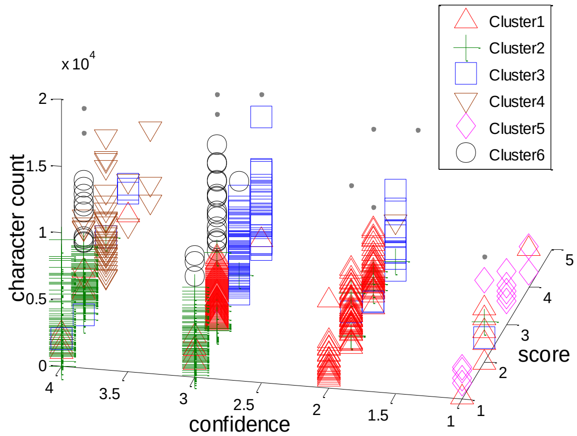


Figure 2. Review clusters in SIGCOMM

Table 3. Cluster characteristics in Sigcomm

Cluster	Percentage	Characteristics
1	35.24%	Low to medium score medium confidence short to medium character count
2	36.22%	Low score High confidence Short to medium character count
3	15.37%	Medium to high score Medium confidence
4	5.85%	Medium to high score The highest confidence
5	1.59%	The lowest confidence Short character count
6	4.51%	Low to medium score Medium to high confidence Medium to long character count

5.1. Review Patterns

The graph-based mining method is utilized to cluster reviews in SIGCOMM and UIC. The results are described in the following subsections.

5.1.1. Review pattern in SIGCOMM

When the parameter r is set as 0.75 and Lambda as 0.95, the reviews of SIGCOMM are clustered into 6 classes, as shown in Figure 2. Table 3 shows the percentage and characteristics of the reviews in each cluster. For instance, the reviews in Cluster 1 are likely with low to medium score, medium confidence, and short to medium character count. Cluster 1 and Cluster 2 occupy more than 70% of all the reviews. Very few reviews are put into Clusters 4, 5, and 6. For example Cluster 5 includes only 1.59% reviews, which are featured with the lowest confidence and short character count.

Ten reviews (represented as black node in Figure 3) are not included in any clusters. We can observe that their character count is relatively longer than others.

5.1.2. Review pattern in UIC

When setting the parameter r as 0.72 and Lambda as 0.96, the graph-based mining algorithm clusters the reviews of UIC into 9 classes. We find that the last four clusters share the same characteristics of long character count, so we merge them into one and finally get 6 clusters as shown in Figure 3. The percentage and characteristics of the reviews in each cluster are shown in Table 4. We observe that Cluster 2 has the largest number of reviews with medium confidence and short to medium character count. Very few reviews are included in Cluster 4 (merely 1.50%) with a both low score and confidence.

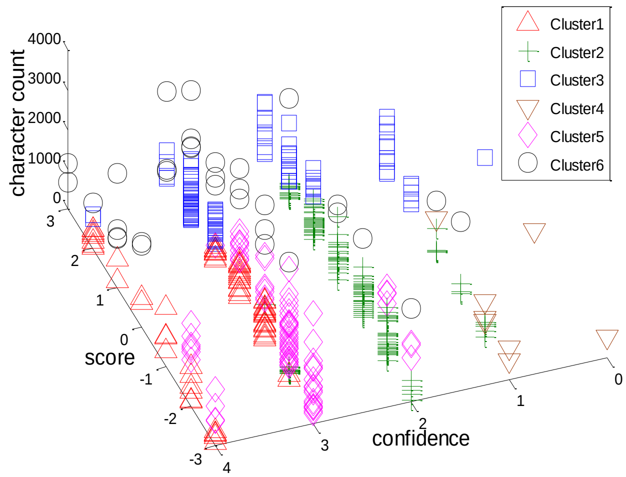


Figure 3. Review clusters in UIC

Table 4. Cluster characteristics in UIC

Cluster	Percentage	Characteristics
1	15.14%	High confidence Short to medium character count
2	31.96%	Medium confidence Short to medium character count
3	26.73%	High score
4	1.50%	Low score Low confidence
5	18.50%	Low to medium score High confidence
6	6.17%	Long character count

5.2. Reviewer Activity Patterns

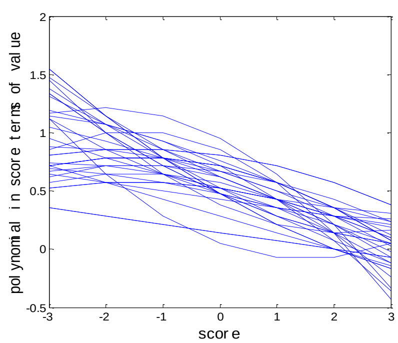
In order to study reviewer activity patterns, we must know the reviewer information, e.g., the exact papers reviewed by a reviewer. As the reviewer information is only available in the UIC dataset, this section's results are therefore restricted to UIC.

5.2.1. Descriptive statistics

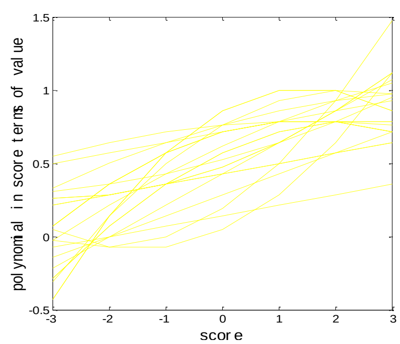
The reviewer data is represented with vectors as mentioned in Section 3.2.3. A scatter plot and fitting curves are generated. Based on the fitting curves, the reviewers are categorized into four classes according to score, confidence, and character count, respectively. Class 1 refers to the group of reviewers who give low score, low confidence, or short review. Class 2 is opposite to Class 1. Class 3 includes the reviewers who prefer to give the extreme score, confidence, or character count. Class 4 is opposite to Class 3. TABLE V shows the percentage of each class.

Figure 4 shows the four different types of polynomial fitting curves according to score, which represent four different groups of reviewers in ranking papers. The percentage of each class is 22.76%, 15.17%, 14.48%, and 47.59% (see TABLE V). Figure 4(a) shows the polynomial fitting curves of Class 1 who rank papers with lower scores, and Class 2 is the opposite, as shown in Figure 4(b). Figure 4(c) illustrates the reviewers (Class 3) who prefer to give the extreme scores, *i.e.*, very high and very low scores. Many reviewers (Class 4, nearly 48%) tend to give the papers medium scores rather than the extreme scores as shown in Figure 4(d).

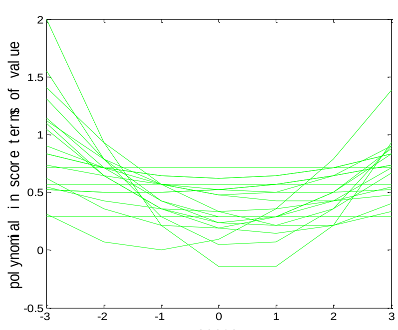
For the classes grouped in terms of confidence, a majority of reviewers, about 81%, give more medium or second top confidence, and this is probably because of their nature of modesty. The classes based on character count are similar with those of confidence. Most reviewers tend to input a medium length feedback, rather than the shortest or longest reviews. From another point of view, we can find that for nearly 40% reviewers, the length of their reviews is in the same level. For example, Reviewer 2 reviewed four papers and the review lengths are all in the medium level, while Reviewer 16 merely reviewed two papers, but all are in the long level.



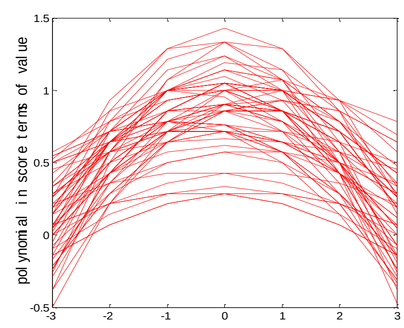
(a) Polynomial fitting curve of Class 1



(b) Polynomial fitting curve of Class 2



(c) Polynomial fitting curve of Class 3



(d) Polynomial fitting curve of Class 4

Figure 4. The classes of reviewers categorized according to score

Table 5. The reviewer percentage of each class according to score, confidence, and character count (reviewer total number:145)

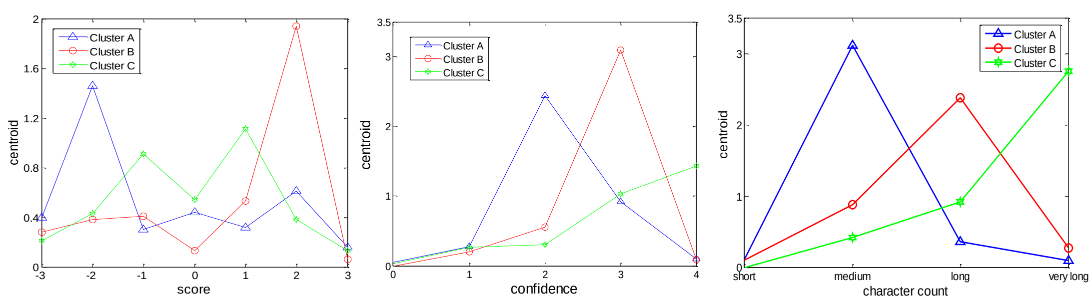
Score		Confidence		Character count	
Class	*Percentage*	*Class*	*Percentage*	*Class*	*Percentage*
1	22.76%	1	1.38%	1	2.76%
2	15.17%	2	13.79%	2	4.83%
3	14.48%	3	3.45%	3	3.45%
4	47.59%	4	81.38%	4	88.96%

5.2.2. K-means clustering

Descriptive statistical method is an intuitive way to classify the reviewer patterns by using chart, histogram, and figure, but it lacks quantification details in categorization. To overcome the weakness of the descriptive statistical method, we adopt the K-means clustering to extract the patterns of different reviewers.

Comparing the clustering results, we decide the reasonable cluster number as 3. All reviewers' activities of evaluating papers with score, confidence and review length are clustered into three groups marked as A, B and C, as shown in Figure 5. Curves below are produced by the centroids of each group, which represent the characteristics of each group.

Figure 5(a) shows the score behavior clusters. Reviewers in Cluster A, nearly 40% of reviewers, prefer to give papers score under 0, near -2. By contrary, Cluster B has a tendency of high score and Cluster C is in the middle level. From Figure 5(b), the medium and the second top confidence are likely to be given separately by two groups A and B with larger number of reviewers, and the reviewers of each group are all around 40%. Cluster C presents an incremental curve and indicates that reviewers in this cluster are more likely to give papers with the highest confidence. The cluster centroid trends of character count depicted in Figure 5(c) are similar to those of confidence. The cluster distribution is presented in Table 6. We can see that more than half (58%) of the reviewers (Cluster A) give most of the papers they review medium length feedbacks, and the reviews of Cluster B are likely to be long, while only 8% of reviewers (Cluster C) give more papers the longest reviews.



(a) (b) (c)

Figure 5. Clusters based on (a) score, (b) confidence, and (c) character count

Table 6. Cluster distribution according to score, confidence and character count (reviewer total number:145)

Score		Confidence		Character count	
Cluster	*Percentage*	*Cluster*	*Percentage*	*Cluster*	*Percentage*
A	39.31%	A	42.07%	A	58.62%
B	22.07%	B	37.24%	B	33.1%
C	38.62%	C	20.69%	C	8.28%

6. Conclusion

In this paper, we conduct a study of investigating review activity in academic conferences by using descriptive statistics and data mining method. Three properties of reviews are used such as score, confidence, and review length. We got some interesting results. For example, more than half of reviews in SIGCOMM are featured with low to medium score, medium or high confidence and short to medium character count. However, in UIC many reviews (above 30%) are corresponding to high confidence and short to medium character count. With the method of descriptive statistics, the reviewers are clustered into four classes according to the three properties separately, and the reviewers were likely to give reviews with medium score, confidence or character count. Three clusters of reviewer were obtained when using K-means clustering method. Take the cluster distribution according to score for example, nearly 40% of reviewers prefer to give paper a higher score. For character count, only 8% of reviewers evaluate papers with 'very long' comments. All these results are useful for interpreting how the reviewers give their reviews in academic conferences. In the future, we plan to enrich the datasets to achieve more statistically significant results.

Acknowledgements

This work was partially supported by the National Basic Research Program of China (No. 2012CB316400), the National Natural Science Foundation of China (61222209, 61103063), and the Specialized Research Fund for the Doctoral Program of Higher Education (20126102110043).

References

[1] G. Abowd, E. Mynatt and T. Rodden, "The human experience", IEEE Pervasive Computing, **(2002)**, January-March, pp. 48–57.
[2] T. Anderson, "Conference reviewing considered harmful", ACM SIGOPS Operating Systems Review, vol. 43, no. 2, **(2009)** April.
[3] A. Ankolekar, Y. Luon, G. Szabo and B. Huberman, "Friendlee: A Mobile Application for Your Social Life", MobileHCI, **(2009)**, Bonn, Germany.
[4] J. Attenberg, S. Pandey and T. Suel, "Modeling and predicting user behavior in sponsored search", KDD **(2009)**, pp. 1067-1076.
[5] L. Chen, J. Hoey, C. Nugent, D. Cook and Z. Yu, "Sensor-Based Activity Recognition", IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 46, no. 6, **(2012)** November, pp. 790-808.
[6] J. Crowcroft, S. Keshav and N. McKeown, "Viewpoint Scaling the academic publication process to internet scale", Communications of the ACM, vol. 52 no. 1, **(2009)** January.
[7] J. R. Douceur, "Paper rating vs. paper ranking", ACM SIGOPS Operating Systems Review, vol. 43 no. 2, **(2009)** April.
[8] N. Eagle, A. Pentland and D. Lazer, "Inferring friendship network structure by using mobile phone data", Proceedings of the National Academy of Sciences, vol. 106, no. 36, **(2009)**, pp. 15274–15278.
[9] M. Ekinci and E. Gedikli, "Silhouette Based Human Motion Detection and Analysis for Real-Time Automated Video Surveillance", Turkish Journal of Electrical Engineering and Computer Sciences, vol. 13, no. 2, **(2005)**, pp. 199-230.
[10] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine", IEEE Trans. Geosci. Remote Sens., vol. 47, **(2009)**, pp. 1328.
[11] M. -P. Kwan, "GIS methods in time-geographic research: Geocomputation and geovisualization of human activity patterns", Geografiska Annaler B, vol. 86, no. 4, **(2004)**, pp. 205–18.
[12] M. -P. Kwan and J. Lee, "Geovisualization of human activity patterns using 3-D GIS: a time-geographic approach", In GOODCHILD **(2004)**, M.F. and JANELLE, D.G. (eds): Spatially Integrated Social Science.
[13] L. Liao, D. Fox and H. Kautz, "Location-based activity recognition using relational Markov networks", Proc. of the International Joint Conference on Artificial Intelligence, IJCAI, **(2005)**.

[14] I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard and D. Zhang, "Automatic Analysis of Multimodal Group Actions in Meetings", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, **(2005)**, pp. 305-317.
[15] T. Morita *et al.*, "A Pattern Mining Method for Interpretation of Interaction", Proc. 7th Int"l Conf. Multimodal Interfaces, ACM Press, **(2005)**, pp. 267–273.
[16] H. Nait-Charif and S. J. McKenna, "Head tracking and action recognition in a smart meeting room", Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Los Alamitos, CA **(2003)**, pp. 24-31.
[17] B. Ni, S. Yan and A. Kassim, "Recognizing human group activities with localized causalities", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), **(2009)**, pp. 1470-1477.
[18] K. Papagiannaki, "Author feedback experiment at PAM 2007", ACM SIGCOMM Computer Communication Review, vol. 37, no. 3, **(2007)** July.
[19] K. Papagiannaki and L. Rizzo, "The ACM SIGCOMM 2009 technical program committee process", SIGCOMM Comput. Commun. Rev., vol. 39, no. 3, **(2009)** July, pp. 43-48.
[20] P. Rashidi, D. Cook, L. Holder and M. Schmitter-Edgecombe, "Discovering activities to recognize and track in a smart environment", IEEE Transaction on Knowledge and Data Engineering, **(2010)**.
[21] A. Spriestersbach, B. Röhrig, J. -P. du Prel, A. Gerhold-Ay and M. Blettner, "Descriptive statistics: the specification of statistical measures and their presentation in tables and graphs", part 7 of a series on evaluation of scientific publications. Dtsch Arztebl Int., vol. 106, no. 36, **(2009)**, pp. 578–583.

