IEREK PRESS Academic Research Community

http://www.ierek.com/press

Print ISSN: 2357-0849

Resourceedings Proceedings of Science and Technology

International Journal on: **Proceedings of Science and Technolgy**

3D Visualizing clout of topics via hashtags on Instagram platform

Guo Xinyi[1], Liu Cuiting[1], Wang Ruisi[1], Raja Majid Mehmood[*,1]

[1] School of Electrical and Computer Engineering, Information & Communication Department, Xiamen University Malaysia

* Corresponding author: rmeex07@gmail.com

Keywords

Information visualization, Instagram,interactive 3D visualization, social media, topic analysis ;

Abstract

This paper intends to develop an interactive, comprehensive information visualization platform of Instagram hashtag analysis. Instagram hashtags has developed themselves into all different kinds of group or communities for users to share hobbies and find similar friends. In order to analyze topic influence and user interest trend from Instagram, which contains billions of end-users and has worldwide influence, hashtag analysis is necessary to gather such information and compare the proportion of people involving in each tags and rank them to visualize.

The visualization is developed in 3D space and consists of time-varying data flow of tags, together with tag comparison analysis, as well as event researches. In the rest of the paper, we mainly discuss the design idea and the development process of the system. An example of the system design work will be shown in the discussion, which involves 4 popular hashtags discussed on Instagram and are shown on the system, displayed as an 3D histogram, together with another comparison histogram to compare different tags, as well as an event view in the back.

1. Introduction

Instagram is one of the most influential social media in the world. Various types of contents such as texts, images and short videos are uploaded on this platform. Users can upload their photos, or videos with description, and give each post corresponding hashtags. It can easily achieve people's reaction from one certain event or topic through posts under hashtag. It's really important to analysis people's reaction of different topics especially in such a huge social media platform, using information visualization method to show the clout of one topic by amount of posts though the timeline.

In Instagram, a hashtag under the topic actually show different perspective of this topic. Therefore, when calculating the clout of one topic, all the posts which are under hashtags related to this subject will be considered in the visualization model. To some extent, these hashtags will show the reason why clout of this topics suddenly increased. For example, when analyzing the clout of football, the hashtag #worldcup will be considered. In world cup period, the clout of football must have a significant increase because of the clout of the world cup. Therefore, these relative hashtags of certain topic of each should be also included in the visualization system. Each topic must have a peak clout, in somehow these peak values reflects the influence of the topics, which should also be considered in the visualization model.

This paper will be based on this data introduces system for visualizing clout of topics via hashtags in Instagram, explain the detailed of system design. Then some example cases applying that visualization system will be presented in the paper, to directly show how the visualization systems visualizing these data.

2. Literature review

BrandMap [1] is a visualization platform which uses a novel approach to visualize complex data. This paper proposes a case study using BrandMap as a visualization tool to measure the distribution of brands in the blogosphere. As many bloggers mentioned their brands, products and services, a huge resource of data requires to be organized. The methodology of visualization is to use objects with different characteristics like colour, size and shapes to represent the key brand dimensions like product attributes, features and themes. The objects are placed in circles with certain angles between them and distances from the centre. The angle between the terms around the centre is computed by hierarchy clustering technique according to their similarities. If two terms are closer to each other, they may be often mentioned and related together in the blog. The distance between the centre and term is calculated depends on the frequency that the term is cited in the blog. The more a term is mentioned in the blog, the closer it is to the centre of circles. This visualization method helps people quickly observe the information about brand dissemination over the Internet.

In Masahiko Itoh [2], social media has been one of the most popular sources for people to acquire information. The goal of this paper is to analyse changes in people's idea, experience and interests through information visualization. A 3D visualization system is introduced in this paper to visualize time-varying topics in multiple media and analyse their future trends. The system design enables people to observe the begin time of the topic, changes in trend of the topic, bursting points, and its lifetime. Different images and events related to the topics are also considered as part of the visualization contents. This visualization system consists of two main part which are Image Flow View and Event View. To visualize the image flow, a three-dimension histogram including stacking images are created. The images are arranged according to their topics and publish time. For the event view, TimeSlice which is a 2D plane is placed in the 3D space to summarize events on the topic keywords. Once a visitor selects a time window, a tree presentation will be displayed on the TimeSlice. In all, this 3D visualization system can be used to explore trends and events in social media.

In Chen et al. [3], as social media becomes more and more popular, a large number of messages are spread over medias every day. This paper aims to explore and analyse social behaviours during the process of message diffusion and propagation. In this case, D-Map which is a comprehensive visualization system is proposed. In D-Map system, social media users are represented by hexagonal nodes with colour and size indicating their behaviours and roles. The users are grouped into different communities according to their behaviours, forming a map metaphor which can visually show the social influence of the centre user. Each community is represented by one colour, and the centre user is highlighted with an outside hexagon. There is also an inside hexagon in each node indicating the number of user's blogs. A centre user with high inner-community influence will be represented with a big size of node. A user with a large number of blogs will be assigned with a large dark hexagon inside the node. This paper collects data from one single user of Sina Weibo with all the reposting blogs, originates all these blogs and then use D-Map to visualize the diffusion process of these blogs. In conclusion, D-Map visualizes users' social behaviours and their influence regarding spreading information on social media during the diffusion process.

3. Critical review

Masahiko Itoh [2] has provided a relatively reliable 3D visualization method by combining image stacked histograms from multiple events together with corresponding line charts, as well as an interactive event view displaying aside in 3D space. (Figure 1) Some essential attributes need to be evaluated in its research. The most fundamental one is timeline, which is organized according to the development process of the specific topic stacking with related images and contains proper time interval such as a month, a week, a day. Besides, the stacked images (Figure 2) represent the amount of discussion on social media in terms of this topic, which enables us to find the birth timing, bursting points, changes in popular content, and the lifetime of trends for each topic. Another attribute is topic, displayed by different histograms to classify various topics being discussed, explore differences in bursting timing for every topic, their chronological order, and events with the same timing on different topics. Comparison on reports between mass media and social media are also evaluated as well showing in the difference between histograms and line charts layered together. In addition to the image flow view, event view (Figure 3) is also an indispensable attribute for evaluation, represents by TimeSlices and TimeFluxes, which visualize respectively summarized events on the topic keyword during a selected time window as a tree representation, and changes in the amount of information such as the number of events within a given period of time.

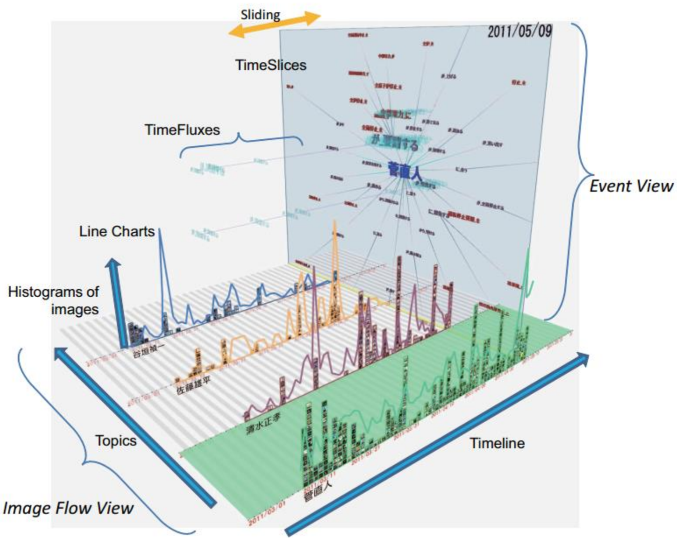

Figure 1 3D visualization system of time-varying information

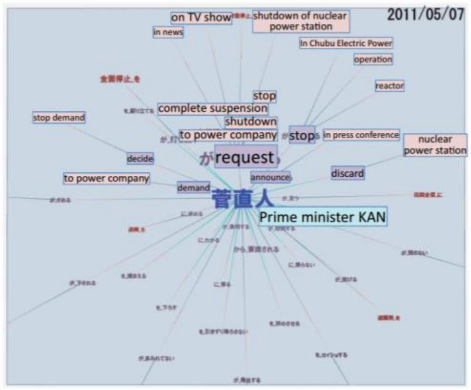

Figure 2 stacked images clustered on timeline

Figure 3 Event view showing via TimeSlice

Similar research in terms of social media content analysis has developed multiple methodologies with respect to provide clear user visual experience. BrandMap method uses a set of layered circles together with variously shaped notations in evaluating angles and distances in between. This method is only helpful in observing topic classification, rather than analyzing trends in information via time-varying data flow, for no time attribute is displayed in BrandMap, and no topic relations is visible in 3D visualization system. D-Map, on the other hand, an easy-visualized form of reposting tree that consist of clustered nodes represented in uniformed colour in the same group, summarizes the diffusion process to illustrate the spreading of messages across different groups of people and reveal the social impact of a central user, while the 3D visualization system only focusing on messages spreading effects--clustered messages forming information trends. In this case, the proposed topic will only become clearer to visualize when applying the 3D model [2].

Limitation of the cited system is quite obvious though. Using stacked image flow is a convenient and plausible method to consist the histogram classify different topics, however, the layered and clustered images will only mess up the visual experience as the number of topics increases. In terms of improving visual effect, the stacked images will be replaced with stacked cubes in a uniformed colour for each single topic in the proposed visualization system. Meanwhile, TimeFluxes that summarizes events on the topic keyword during a selected time window seems to be less meaningful as for TimeSlices will show the appearances every subtopic on selected time. Therefore, the proposed system is going to remove it in order to simplify the visualization.

The proposed 3D visualization system will demonstrate hashtag information only on Instagram. Since no related similar social media is going to be analysed as comparison, line charts (Figure 1) will be removed in order to focusing on the data flow of histograms. As we intend to enhance the analysis of hashtags to illustrate users' interest distribution, a new histogram on y-z plane about comparison between every topic will be summarized by gathering all highest data flow point in each topic and noting the exact time of the occurrence. Specific visualization realization will be mainly discussed in the next phase.

4. Proposed method

The method we want to use is 3D visualization method. In the visualization model, we want to have a timeline to record the clout (amount of posts) of each topic. Therefore, there will be billions of timelines in such a widely-used social media. If we just use different colored timeline to represent different topics, it will be in a mess. In this situation, using 3D visualization method is a good choice to design the visualization model. We can construct a 3D coordinate, using topics and timeline as x-axis and y-axis, the clouts of each topics following time as z-axis. What's more, to differentiate the topics in visual effect, we will use different colors for different topics. The projection of clout for each topic on x-z plane will be shown as histogram. And corresponding related hashtags will be shown when clicking certain topic at certain time in y-z plane as tree structure. In a world, we will sufficiently use the 3D system to visualize the clout of topics in Instagram in different perspectives.

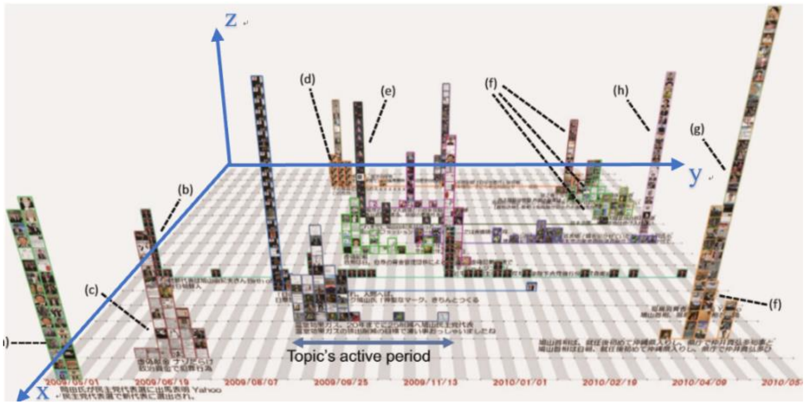

Figure 4 sample model to explain

The proposed visualization system contains 3 main parts, hashtag time-varying view, event view, and hashtag comparison view. (Figure 5). Users can observe the trend of the hashtag in time sequence, in the meantime to explore the related interesting hashtags, as well as to compare the popularity between each irrelevant hashtag in order to explore majority users' interests. In the hashtag time-varying view, we utilized 3D space with 3D histogram and stacked them on a timeline, which makes visualizing the heat of the hashtags easier and simultaneously see the 3 views feasible. Users can zoom, rotate, and pan the 3D space to interactively change the region being focused on and to avoid problems with occlusion.

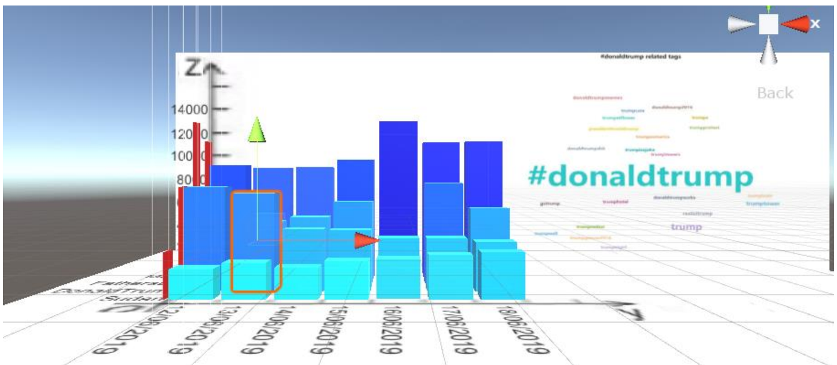

Figure 5 3D visualization system model of hashtag analysis platform

Visualizing hashtag time-varying view

Visualizing hashtag time-varying flow is adapted to multiple 3D histograms divided according to the hashtags, resembling to the representation of bitmap in 3D form. (Figure 6) In the y-axis, timeline is presented in a selected time interval, which can be a year, a month, a week, or even a day. Users can determine how long they would like to visualize for the heat of the discussion. We then arrange multiple histogram flows in the 3D space to compare multiple hashtags. In the x-axis, hashtags represented by the 3D histograms are arranged, labelled by the name of hashtags. Users can manually or automatically define the order of hashtags on the x-axis by using their rankings if they have them, whether it is on alphabetic or categorizing order. This allows us to explore differences in bursting timing for every hashtag, their chronological order, and hashtags with the same timing on different topics. In the z-axis, the amount of discussion in the selected hashtag of corresponding time is represented. It is clearer to visualize with gradient colour so that the levels of heat (discussion amount) in the selected hashtag are presented by the shade of the colour, which provides easy comparison among the heat between different hashtags in different time.

The system supports the interaction to explore the detail information and exact content of the selected hashtags. Users can access the original hashtag page by clicking the tag name, and the original webpage from the web browser will pop out automatically and show all the information included in it and related side topics.

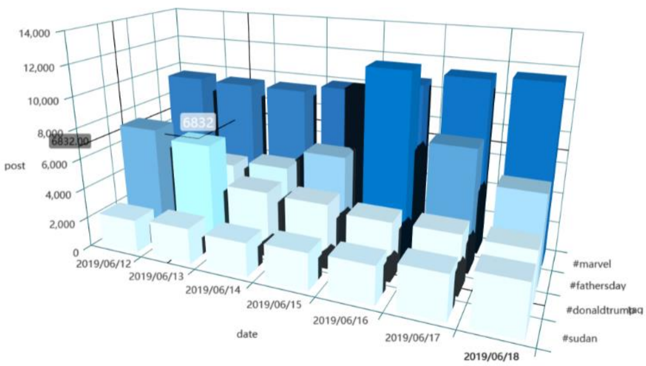

Figure 6 visualization of the hashtag time-varying view in 3D space

Visualizing hashtag comparison view

In this section, each hashtag will be gathered in the same dimension to be analysed for comparison on summarized all highest data flow point in each hashtag and represented as a 2D histogram displayed on the x-z plane. (Figure 7) We display each summarized hashtag in the position aligning to the x- arrangement of hashtags on x-axis in hashtag time-varying view. As each highest data will be presented in this view, it is easier to generate the most representative data in each hashtags and deal with their comparison using these high data, and generalized people's most interested topics and their tendency of focusing area. The colour of this histogram will be distinguished from the one on hashtag time-varying view, in gradient representation also in order to differentiate the level of the information amount. Labels on each attribute will not be provided since the presentation has already been displayed in 3D histogram, and they have matched with each other distinctly.

This view provides a further conclusion of the discussion heat comparison among every available hashtag. It not relatively easy to generalize the highest amount of the discussion in each tag though the gradient colour in 3D histogram directly, therefore conclude the heat level between tags. However, with this viewport by putting every heat level conclusion into the same dimension and analyse them together, users will have better and easier user experience during the research.

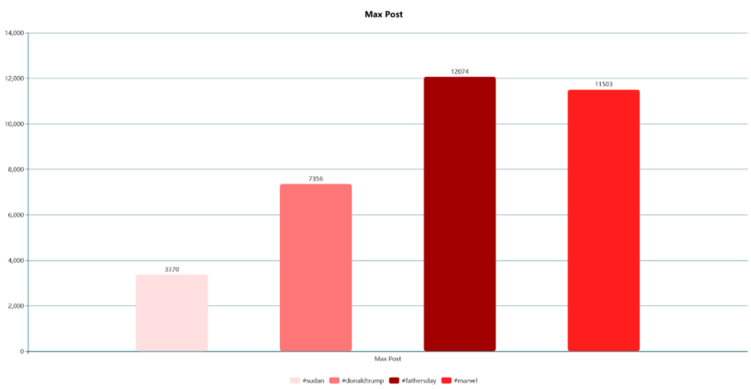

Figure 7 visualization of the hashtag comparison view on x-z plane

Visualizing event view

The event view is adapted from Masahiko Itoh [2], which could be a good visualization of exploring related interests and is defined as a set of dependency relations on the name of hashtags, to explore detailed information about a selected hashtag and timing. User will need to choose a hashtag in the hashtag time-varying view, and select a specific time point, the Event View retrieves tags related to the main hashtag and automatically moves to the point of timing on the timeline to display events belonging to the time window. Users can hide it if it is not necessary. The visualization that is going to be implemented is word cloud chart. it is a very simple, clear visualization for the displacement of all the tags together comparing to the tree chart or any other method. It provides relations in between each tag as well as their discussion heat comparison, in which the most frequently researched hashtags will be displayed in the largest size of the word, and the size will reduce accordingly due to the rank of the heat of the tags. The chosen hashtag is placed in the middle with the most eye-catching size and colour, and the title of the chart (in figure 8, showing as “donaldtrump related tags”) is placed on the top of the view as well as the exact amount of its posts.

This event view allows users to explore the real-time events happened to the corresponding tags. For example, in the hashtag “donaldtrump”, we can tell all the related tags from the event view is about Donald Trump, and the recent behavior he did, like the one “trumpamerica” that shows him participating 2020 American president campaign; as well as people’s attitude towards him, such as the one “gotrump”, which shows people’s resistance to Donald Trump and wish him goes off stage.

Figure 8 visualization of the event view on y-z plane

5. Results & Discussion

The final result of our proposed visualization method is shown at figure 13. Compared with the visualization prototype at figure 14, we have already improved all the three main sections of the model.

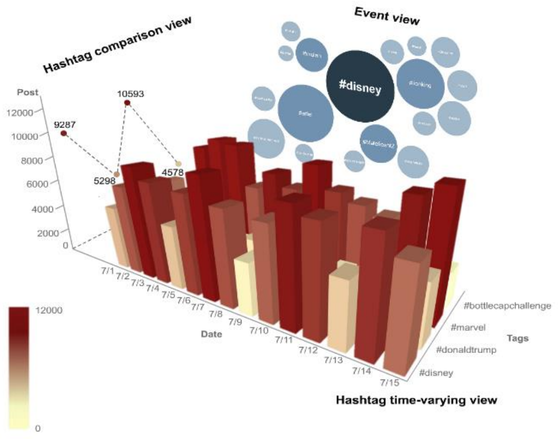

Figure 5 Final visualization model

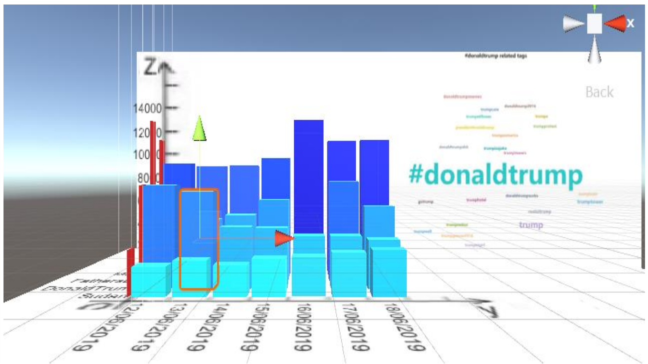

Figure 6 Visualization Prototype

Hashtag time-varying view improvement

Firstly, the gradient colour of columns were changed. For the prototype, we only used different levels of blue colour to represent different amounts of discussion which was confusing to distinguish. Thus, we decided to apply two different colours which were light yellow and dark red to the gradient colour range so that the levels of heat (discussion amount) could be identified quickly. The top view of the 3D histogram at figure 15 was also more convenient for users to compare the levels of heat among topics.

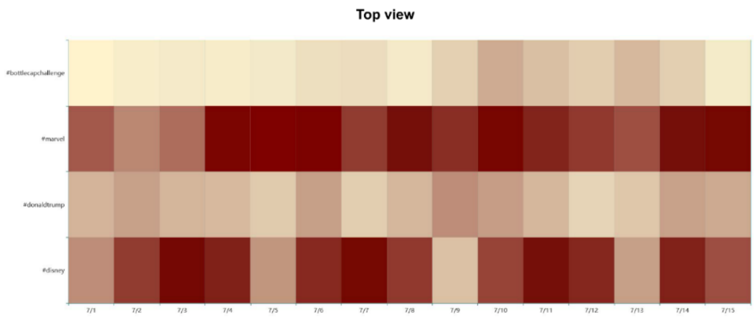

Figure 7 bit-map from top view

Then, we enlarged the dataset. At first, we only used a week to display the discussion amount of each hashtag. However, we found that seven days were not enough to show obvious changes in trend of topic discussion. Therefore, we increased the evaluate time from 7 days to 15 days. The changes of discussion amounts could be clearer compared with before. The dataset collected on Instagram platform is shown at figure 16.

Tag	7/1	7/2	7/3	7/4	7/5	7/6	7/7	7/8	7/9	7/10	7/11	7/12	7/13	7/14	7/15	Total
#disney	5290	7290	9020	8079	5087	7808	9287	7360	4047	7070	8567	7887	4870	8050	6798	316000
#donaldtrump	4356	4832	4342	4243	3836	4865	3746	4290	5298	4909	4305	3580	3908	4802	4590	5500000
#marvel	6520	5420	6049	9860	10593	10320	7308	8594	7684	9574	7965	7355	6746	8574	9355	23700000
#bottlecapchallenge	1208	2380	2805	2439	3020	3320	3390	2809	3689	4578	4086	3780	4259	3706	2887	30500000

Figure 8 dataset for 4 topics during 15 days

Hashtag comparison view improvement

At first, we used a 2D bar chart placed at the left side of the model for comparison of the highest data flow points of each hashtag. Anyway, the columns of the bar chart were similar to those of the 3D histogram. In order to distinguish these two sections, the bar chart was replaced with a line chart. What's more, the colors of points referred to the same gradient color criterion of 3D histogram. The number of discussion amount at each point were marked.

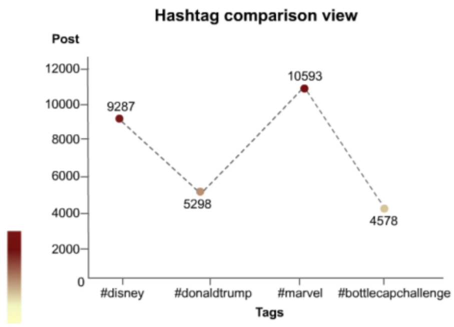

Figure 9 line chart of peak value

Event view improvement

The event view we used at first showing the relationships of the hashtag and its related tags was shown at figure 18. The visualization randomly displayed all the related words of tags, and used the size of words to represent for the discussion amounts. However, this visualization method looked so messy that users were difficult to understand the randomly-placed words soon. Different sizes of words also made the view not so nice, and some tags were too small to recognize. What's more, this view was too colourful to be placed in the visualization model. The colours which were redundant may catch users' attentions.

Figure 10 event view of visualization prototype

Consequently, we chose to use a bubble chart to present the event view. The tags were divided into three different layers according to their degrees of relevance indicated by bubble colors. Take topic "Disney" as an example, the event view is shown as figure 19. The root hashtag "#disney" was placed as the first layer in the middle in darkest color with largest text. For the second layer, "#mulan", "#lionking", "#ariel" and "#maleficent2" were put in lighter color with smaller text. These four tags respectively represented for four films produced by Disney company. Thus, they belong to direct relative tags of "#disney" and should be put in the second layer. The remaining tags put in lightest color with smallest text were the third layer. These tags were related topic with those four films in the second layer instead of directly deriving from "#disney" hashtag. What's more, tags in the third layer were put near to their parent tag in the second layer. For example, "#mushu" was the main character of film mulan and "#liuyifei" was the one who acted mushu. Both of these two tags were directly related to tag "#mulan", so these two bubbles were placed near bubble "#mulan".

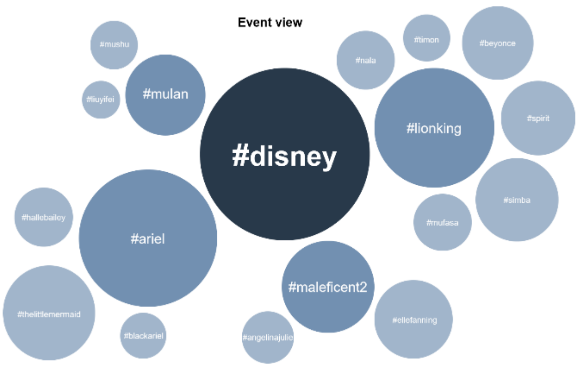

Figure 11 Event view of final visualization model

Compared with the first visualization, the outlook of bubble chart was neater and clearer to understand. Tags were represented with bubbles in gradient color placed in order, and size of texts were fixed for each layer. What's more, users can easily realize the relationship like parent or grandparent between tags. The dataset collected on Instagram platform for event view of "#disney" harshtag is shown at figure 20.

#disney	23700000
#ariel	3800000
#hallebailey	17200
#thelittlemermaid	1600000
#blackariel	2000
#lionking	2400000
#beyonce	15100000
#nala	1300000
#spirit	8200000
#simba	2300000
#mufasa	340000
#timon	264000
#mulan	731000
#mushu	153000
#liuyifel	42300
#maleficent2	1000000
#ellefanning	148000
#angelinajulie	5000

Figure 12 Dataset of event view

6. Evaluation

In this section, we intend to use the user evaluation method, questionnaire, to conduct the research on how the 3D visualization model performs. We have found 10 people who are not related to the development of the visualization system at all to the finish 2 questionnaires displayed down below, and the results are concluded as well. In the development of the first questionnaire, 24 questions are composed and are considered the most comprehensive, representative ones that can fully reflect the problem of the amended system. The second questionnaire is designed for analysing people's view about the comparison between the model created initially and the one amended, and is consist of 7 questions in order to give an objective analysis about whether the changes is worth it or not.

The completed result as the degree of satisfaction for the questionnaire experiment is displayed below. The calculation is based on the average of the value that the number of people in each degree multiply the degree value, which is represented by 0, 25%, 50%, 75% and 100% corresponding to strongly disagree, disagree, neutral, agree, strongly agree respectively. Here are some aspects that we have concluded from the result.

The model performs generally well in organization, complexity and accusation. The first 2 questions reflect the superiority by giving the result of exceeding 80% from the feedback. For the hashtag time-varying view, the general performance is 85%, which is considered fairly well in visualization in general. However, there is still a relatively low result in telling the trend of each tags, which will be better in visualizing it by replacing the 3D histogram into line chart, which is not a better solution in this case. In hashtag comparison view, the result is still a pleasant one in general performance especially in organization and differentiating data trends. In the meantime, the line chart is considered a good representation method for this view. Nevertheless, the difficulty of telling people's interested topics is still exist, which may be resolved by providing more hints through the diagram. In event view, the general performance is fair enough to be considered as an indispensable part of the model, and bubble chart can make the visualization much easier. The colour and size selection are helpful for visualizing heat differences among each tag.

question	Strongly disagree	disagree	Neutral	agree	Strongly agree	result
1			1	5	4	82.50%
2				4	6	90.00%
3				6	4	85.00%
4				4	6	90.00%
5			2	2	6	85.00%
6				6	4	85.00%
7			2	4	4	80.00%
8			3	4	3	75.00%
9			2	5	3	77.50%
10			3	4	3	75.00%
11			3	4	3	75.00%
12			2	4	4	80.00%
13			2	5	3	77.50%
14		1	2	5	2	70.00%
15			3	5	2	72.50%
16		2	1	4	3	70.00%
17			4	3	3	72.50%
18				4	6	90.00%
19				6	4	85.00%
20				5	5	87.50%
21				6	4	85.00%
22			3	4	3	75.00%
23				6	4	85.00%

Figure 2113 Results of Questionnaire 1

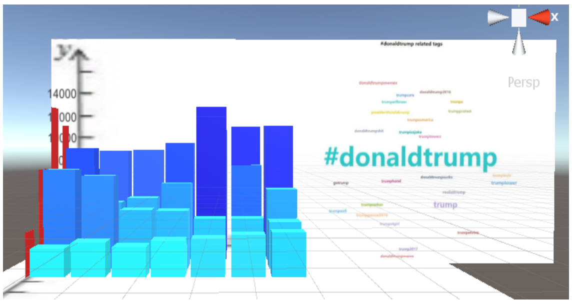

Figure 22 Model 1

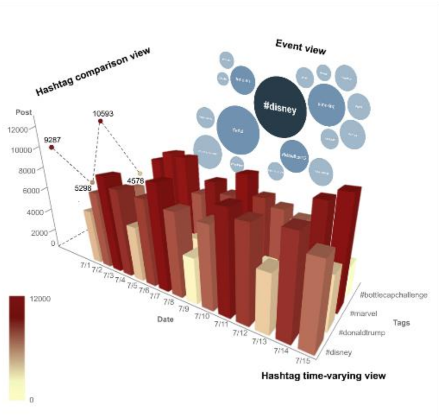

Figure 143 Model 2

In questionnaire 2 result that are displayed below using the same calculation method as the previous one, the overall comparison between the initial model and the amended one is quite obvious: the latter one is better in general visualization absolutely, especially in organization aspect, which has clearly been a huge development in revising the model. Besides, the hashtag time-varying view and the event view both display significant improvement. However, most people that have conducted the research considered that the finalized model does not convey more information than the initial one, which is reasonable since the amendment did not focus much on the information supplement. Moreover, many people think using line chart rather than histogram on hashtag comparison view does not seem to be a significant improvement because both can convey the trend among tags fairly well.

question	Strongly disagree	disagree	Neutral	agree	Strongly agree	result
1					10	100.00%
2		2	4	3	1	57.50%
3				4	6	90.00%
4				3	7	92.50%
5		3	2	4	1	57.50%
6				2	8	95.00%

Figure 154 Result of Questionnaire 2

7. Comparison with existing work

The visualization method we used is inspired from Masahiko Itoh [2]. But we choose different topics, we focus on the Instagram - a huge social media platform, trying to represent the clout of topics in same time-vary to analyse the trend of different topics, and represent the reason why these topics was mentioned in user's Instagram. Masahiko Itoh [2] focused on multiple media, such as TV, Blog and so on, trying to represent the clout of topics in same time varying in different platforms, and represent the topics keyword in different time. Comparing visualization model, our visualization model consists of 3 part: hashtag time varying view, hashtag comparison view and event view, their visualization model contains 2 part: Image Flow View and Event View. The hashtag comparison view that we have directly represents the peak value of different topics using line chart. The image flow view and our hashtag time varying view are quite similar, and the both event view actually has similar functionalities.

Hashtag time varying view & Image flow view

Masahiko Itoh [2] use histogram diagram to represent the trend of different topics in blog, but they use the image square sequence instead of the square in traditional histogram, which not only shows how many blogs related to this topic, and also the images inside. But they ignore the number of blogs, if the highest number of blogs related to one topic is more than 1 hundred, the size of image square unit will be extremely small. Therefore, these images won't be received by users. Otherwise, there are million users in a huge social media, there is no possibilities that hottest topic has less than 1 hundred blogs. The image histogram is not really useful in this scenario. Their image flow view also contains the line chart to represent the trend of different topics in other media.

Our hashtag time varying view also uses the histogram diagram. We use the traditional one-color cube instead of the image design. Furthermore, we differentiate the number of posts by color. It will be more clearly to represent the difference of amount. The original histogram just differentiate amount by the height of square, we can also represent it by color. We also change the 2D square in traditional histogram into 3D cube, considering the characteristics of 3D visualization model to avoid no information presented from other angles, especially top view. The different color cube design will make a bit-map from the top view, which also time- varying shows the clout of different topics.

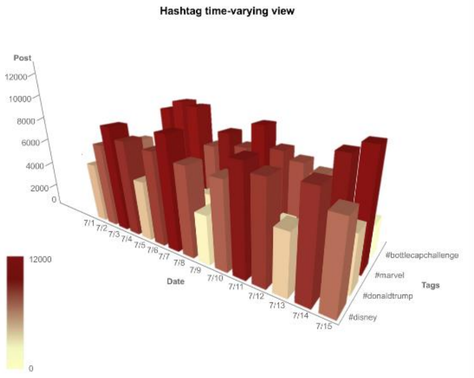

Figure 165 Hashtag time varying view

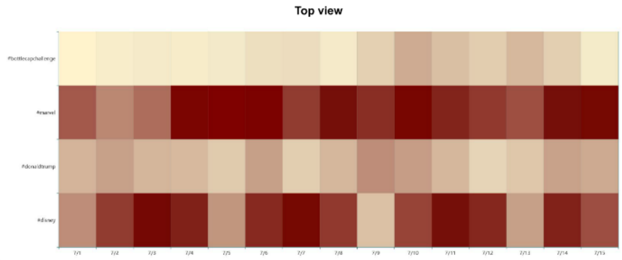

Figure 176 bit-map from top view

Comparison between 2 event view

The event view of Masahiko Itoh [2] uses the tree diagram, which only shows the related words. But we also want to show the amount of each subtopic of one topic to analyzing the main reason that this topic was suddenly hot in these days in our scenario. We use bubble chart instead of the tree diagram.

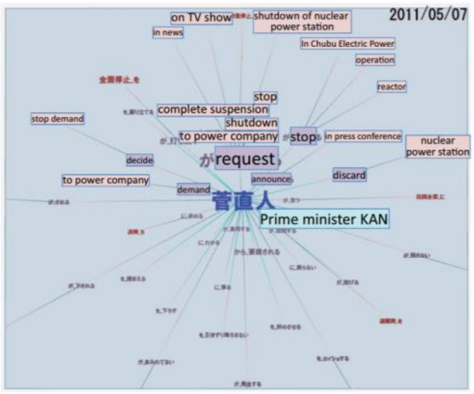

Figure 187 Event view [2]

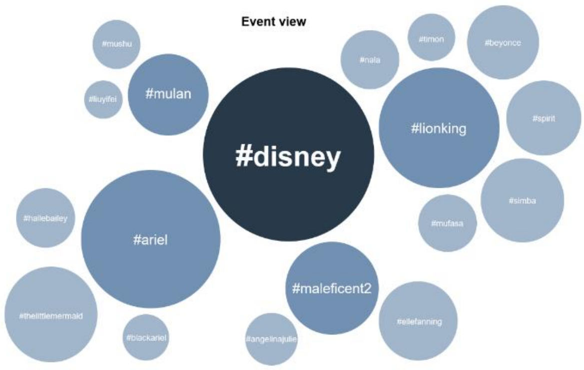

Figure 2819 Event view of our model

8. Conclusion

Above all, the visualization method proposed in this study is mainly used for evaluating the levels of heat (discussion amount) of topics and related information. There may be still some weaknesses that the visualization has, anyway, these approaches will be improved in the future work.

Appendix A. Questionaire 1

No.	Question	Strongly disagree	disagree	Neutral	agree	Strongly agree
1	I am able to tell what kind of topic that this model is analysing					
2	I am able to tell that there are 3 significantly different views					
3	I think the hashtag time-varying view is well organized in visualization and 3D histogram model is a good representation					
4	I am able to tell that there are 4 hashtags showing in the visualization model					
5	I am able to tell that the time in analysis is 15 days in the visualization model, and the time interval is 1 day					
6	I am able to tell which tag is more popular-discussed and which is not					

7	The organization of the hashtag time-varying view is tidy and can easily get me involved					
8	I can easily tell the trend of each tag in time sequence in hashtag time-varying view					
9	I can tell the colour is appropriately arranged and significantly differentiable in hashtag time-varying view					
10	I can easily tell that the data in hashtag comparison view is closely related to the one in hashtag time-varying view					
11	I am able to tell that data in hashtag comparison view represents the highest data point in each tag in hashtag time-varying view					
12	I think the hashtag comparison view is well organized in visualization and line chart is a good representation					
13	I am able to tell that the data comparison between each tag showing in hashtag comparison view is significant					
14	I can tell people's interests towards real-time topics by visualizing hashtag comparison view					
15	I can tell that the event view shows the detailed information of each hashtag					
16	I can tell what each bubble represents					
17	I am able to tell what is the main purpose of adding the event view					
18	I think the event view is well organized in visualization and bubble chart is a good representation					
19	I think the event view is helpful in visualizing and analysing hashtag related events					
20	I can tell which bubble represents the main hashtags and which bubbles are related tags					
21	I can tell that the bubble size and colour represent the heat of discussion on that topic					
22	I can tell that there are significant differences in sizes among each bubble					
23	I can tell that there are significant differences in colours among each bubble					
Do you think there is any other suggestions for the improvement of this model?						

Appendix B. Questionaire 2

No.	Question	Strongly disagree	disagree	Neutral	agree	Strongly agree
1	I think model 2 is better in visualization in general than model 1					

2	I think model 2 conveys more information than model does					
3	I think model 2 is better in visual organization than model 1					
4	I think the hashtag time-varying view in model 2 is better in organization and easier in visualization than the one in model 1					
5	I think the hashtag comparison view in model 2 is better in organization and easier in visualization than the one in model 1					
6	I think the event view in model 2 is better in organization and easier in visualization than the one in model 1					
7	Do you have any other suggestions on improving the visualization after comparing the 2 models above?					

Acknowledgements

This research was supported by Xiamen University Malaysia. We thank our supervisor Prof. Raja Majid Mehmood from Xiamen University Malaysia who provided insight and expertise that greatly assisted the research.

References

[1] Amadeu Sa de Campos Filho, F. F. (2012). Brandmap: an Information Visualization Platform for Brand Association in Blogosphere. *16th International Conference on Information Visualisation*, (pp. 316-320).

[2] Masahiko Itoh, M. T. (2013). VisualizingTime-VaryingTopicsviaImagesandTextsforInter-MediaAnalysis. *17th International Conference on Information Visualisation*, (pp. 568-576).

[3] Siming Chen, S. C. (2016). D-Map: Visual Analysis of Ego-centric Information Diffusion Patterns in Social Media. *IEEE Conference on Visual Analytics Science and Technology (VAST)* , (pp. 41-50). Baltimore, Maryland, USA .