2015 IEEE International Conference on Multimedia Big Data

On Analyzing the ”Variety” of Big Social Multimedia

Jitao Sang, Changsheng Xu
National Lab of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, Beijing, 100190, China
Email: {jtsang, csxu}@nlpr.ia.ac.cn

***Abstract*—Social media contributes significantly to the arrival of the Big Data era. It is interesting to re-examine “multimedia” in the context of social media. This positioning paper proposes to analyzing into *Variety* of big social multimedia from the perspective of various sources, which is the study of *heterogeneous data generated and consumed in various social media networks* that receives little attention in the literature. Since social multimedia is essentially “user-centric”, which is generated from user and customized for user, a research schema is introduced to exploit the overlapped users as bridges to address challenges and exploit social multimedia data from various sources.**

I. Introduction

A. Social Multimedia and “Variety”

In the past decade, the prominence of social media has revolutionized the way people share and access information, resulting in the social trend in multimedia data generation and consumption. For example, *Facebook* reports 350 million photos uploaded daily as of November 2013; 100 hours of video are uploaded to *YouTube* every minute, resulting in more than 2 billion videos totally by the end of 2013. Moreover, social media gives birth to many new types of multimedia, e.g., image tweet, audio picture, geo-tagged video, etc. This significantly extends the scope and application areas of multimedia.

The multimedia data generated and consumed under social media circumstances is referred to as social multimedia [1]. As illustrated in Figure 1, social multimedia can be simply interpreted as the hybrid of social media and multimedia, with three elements as content, user, and interactions (between both user-user and user-content). As the hybrid of multimedia and social media, social multimedia enjoys advantages of both rich sensory simulation and efficient information access and propagation, thus having great potentials in analysis and utilization. Thanks to the wide prevalence of social multimedia data and the increasing demands for social multimedia services, there has been a growing number of researches on emphasizing the role of social media in multimedia applications, such as image and video understanding [2], [3], [4], multimedia retrieval [5], personalized recommendation [6], travel assistant [7], etc. This trend is evidenced by both the volume of related papers produced each year, and the recent tutorials and special issues in prestigious multimedia conferences and journals [8], [9], [10], [11].

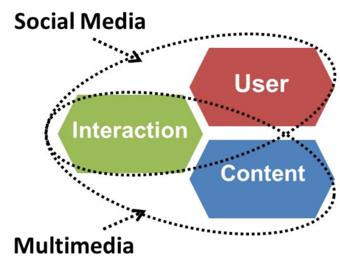

Fig. 1. Social multimedia: hybrid of social media and multimedia.

Social multimedia enjoys significant big data characteristics, with huge volume, exponential growth velocity, various data formats and sources, and low value density. Among the “4V” characteristics, *Volume* concerns the capacity in data storage, *Velocity* concerns the efficiency in data capture and computing, *Variety* concerns the complexity in data analysis, and *Veracity* concerns the data accuracy and quality. Therefore, one fundamental task in social multimedia analysis is to explore the data *variety* for knowledge mining and to facilitate applications, and it is significant for researchers of multimedia analysis to dig into big data problems from the perspective of *variety*.

Variety in social multimedia has two typical interpretations, various modalities and various sources. (1) Various modalities indicate the different data formats, including both the traditional modalities of text, image, audio, video, and the novel combined modalities like image tweet, audio picture, geo-tagged video, etc. In the “small data” era, the topics of exploring multiple traditional and combined modalities, e.g., cross-modality analysis, have already been extensively studied in the multimedia field [12], [13]. Recently, a generalized definition has been introduced and received attentions to address the context as well as the content. For example, the content data of user node and document node, and the context data of user-user link and user-document link, can be explored for collaborative mining and applications. Combining content and context data has long been an important topic in the data mining community [14], [15]. (2) Various sources indicate the different data origins, including both the capturing devices of desktop or mobile, and the issuing channels of official or personal. While not everyone encounters issues of volume and velocity like big companies of Google and Amazon, even the smallest problems involve with variety and

978-1-4799-8688-0/15 $31.00 © 2015 IEEE
DOI 10.1109/BigMM.2015.60

will benefit from exploiting the data from multiple sources. It is noticeable that data from multiple sources may also appear in multiple modalities, which means that exploring various sources actually involves beyond various modalities, and is thus a more challenging problem. In this positioning paper, we emphasize the study into variety of big social multimedia from the perspective of various sources that receives little attention in the literature.

B. *Re-examination of "Multimedia"*

Within the multimedia research community, the definition of "multimedia" has been around *heterogeneous data*, e.g., "the field of multimedia refer to all problems that require the use of heterogeneous data types for a satisfactory solution". In early years, *heterogeneous data* shows narrow focus on different modalities, which is consistent to the first interpretation of variety as mentioned above. Recently, the community is active to extend this understanding and has recognized the potentials of leveraging data from various sources. For example, human event analysis is introduced as a significant multimedia problem by recording and analyzing the personal data from different sources (e.g., physiological sensors, social media activities, logical data etc.) [16], [17], [18].

Research in experiential media systems has claimed that, "Unlike computer vision or audition whose core ideas are tied to how human perceived light or sound, multimedia appears to lack a connection to how we experienced the world" [19]. We have been very interested in re-examining "multimedia" by returning to its *experiencing* interpretation and emphasizing *heterogeneous* from the perspective of various sources. In the context of social media, we are taking an exploratory look at the definition of "multi-media":

- **"Media"** is related to how people *experience* the world. Under social media circumstances, we experience the world by contributing to different Online Social Network (OSN) services. According to the UGC nature of social media, people record and broadcast their experiences through various social media activities. The new interpretation of "media" is consistent to its original meaning of "medium", i.e., the user-contributed OSNs are obviously the most comprehensive mediums in social media.
- **"Multi"** is related to *heterogeneous*. OSNs are equipped with different functions or have different focuses. Following the interpretation to "media", new interpretation of "Multi" indicates the heterogeneous data people interact with on these OSNs, ranging from multimedia content in media sharing networks, information stream in microblog networks, to location and POI in check-in networks, commodity in social transactional networks, and social communications in private/professional social networks.

Therefore, it is interesting to interpret "multimedia" in social media as **the heterogeneous data created and consumed in various OSNs** (illustrated in Fig. 2). These heterogeneous data record people's online activities from versatile angles, and reflect the physical world at the same time. Exploring the new "multimedia" is critical to connect between the

Fig. 2. "Multimedia" in the context of social media: the heterogeneous data created and consumed in various OSNs.

separated data island and facilitate value mining from big social multimedia. From the perspective of content analysis, understanding the association among heterogeneous data is fundamental to advanced social media analysis and applications. For example, collective understanding of social streams, online comments, and user-generated images and videos will enable comprehensive tracking to social event progress and evolution. From the perspective of user modeling, exploiting the available user data on different OSNs contributes to an integrated online user profile and thus improved customized social media services. For example, connecting between users' video watching pattern and consuming history will lead to advanced user modeling and facilitate cross-OSN targeted advertising.

II. Challenges

The challenges in exploring the various sources of big social multimedia lie in two-fold.

(1) **Organization**. At very high speed, social multimedia generates a torrent of unorganized and unstructured data. The heterogeneous data are unevenly distributed among various OSNs, and presented as a state in chaotic conglomeration. Simple organization on a daily basis or depending on the networks where the data originate is not efficient enough, in data capture, data analysis, and turning data into actionable information. *Hashtag* is a featured social media organization tool to help track conversations. We have conducted a preliminary data analysis in a self-collected cross-OSN dataset, and examined hashtags in OSNs of Instagram, Google+, Tumblr and Twitter, respectively. We observed that there exists vocabulary gap between different OSNs: only 5.4%, 7.7%, 11.2% of the hashtags on Instagram, Google+, and Tumblr also exist in Twitter. The low overlapping ratio makes hashtag non-efficient in data organization across different OSNs.

(2) **Association**. Analogous to cross-media analysis, a fundamental problem in exploring the social multimedia data from various sources is to establish their association. Traditional cross-media analysis assumes that different modalities are associated in semantic-level, e.g., a common semantic subspace exists where image, audio and text describing the same

concept distribute closely. However, the association between cross-OSN data is more complex as the OSNs have quite different focuses. One single metric cannot fit to all cases. Other hubs to associate the heterogeneous data are explicit links, including *associated link* within the same multimedia object (e.g., tweet and photo in image tweet, video and tag in annotated video), *internal link* in the same OSN (e.g., reshare in Twitter, reblog in Tumblr), and *external link* between OSNs (e.g., cross-posting). Within the collected dataset, 32.8% of the examined activity data has internal link and 18.7% is externally connected. Although the linked ratio seems high, association by the explicit links can only yield low-level knowledge. Moreover, there is no ready unified solution to integrate the three types of explicit links for efficient association.

III. Research Schema

A. User-centric Characteristic and Solution

Social multimedia is quite different from the *content-centric* traditional and web multimedia. The analysis focus of traditional multimedia is the multimedia content, and the goal is content understanding and application, e.g., media content analysis, semantic classification and annotation, structured media authoring. Web multimedia is heavily related to the WEB1.0 environment, which is dominated by broadcast media developed by professional designers and for passive users. The transition from WEB1.0 to WEB2.0 is largely marked by the stronger user involvement and participation. It is widely recognized that the center of WEB2.0 is the user, who has become a creator, writer, and producer [20]. Rising under the participatory WEB2.0, social multimedia has an obvious *user-centric* characteristic:

- **From User: user is the basic data collection unit.** UGC is dominating the Internet these days among various fields. Viewing each user as a data sensor, social multimedia is constituted by what users see, listen, think, feel and speak. By aggregating the individual user data, we can obtain very useful crowded knowledge.
- **For User: user is the ultimate information service target.** Social multimedia services are typically user-oriented and have a customized trend. The importance of end users are emphasized in an unprecedented level. Traditional "one-to-all" strategy is no more adequate towards users' customized demands. The critical component for the personalized information services is user modeling.

One solution to address the above challenges in cross-OSN social multimedia data organization and association is from a user-centric perspective. Specifically, the solution is inspired by the fact that people usually engage in many different OSNs simultaneously for different purposes. For example, the same individual may communicate with his/her friends on Facebook, follow real-time hot events on Twitter, subscribe and watch videos on YouTube, share and discuss favorite restaurants on Yelp, etc. Report shows that different OSNs share remarkable percentage of overlapped users [21]. The overlapped users serve as both index for organization, and bridge for association between cross-OSN data. On one hand, the cross-OSN data actually derives from the same group of users. The overlapped user is a natural and efficient index for organization, which will then facilitate user modeling and personalized information services. On the other hand, the overlapped users' interactions with the social multimedia data provide important clues for association mining. The interactions can be viewed as high-level supervision, where the wisdom of crowds are exploited.

B. Cross-OSN Data Organization

As discussed above, user data distribute among various OSNs, which need to be jointly analyzed towards comprehensive user modeling and personalized services. This calls for the necessity and demonstrates the reasonability to organize cross-OSN social multimedia data along the overlapped users.

One organization scheme is *complementary organization*, which directly aggregates the disparate data around the same overlapped user. User data on different OSNs reflect user status, interest and preference from versatile perspectives. Take health model for example, the cross-OSN data may include the number of steps you walk tracked by $Fitbit$, how often you check in to local gym using $Foursquare$, and what you eat based on the pictures of your meals that you post on $Instagram$. Each piece of information, by itself, may be inconsequential. However, organized and aggregated by the overlapped users, the heterogeneous social health data will make up for the shortage of physical health records, and significantly facilitate health insurance and smart healthcare. Following this scheme, we have aggregated user profiles from Google+ and YouTube, and user social networks from Flickr and Twitter to enable complementary cross-OSN user modeling and facilitate applications in friend and video recommendation [22], [23].

Another scheme is *collaborative organization*, where collaborative characteristics between the heterogeneous cross-OSN data of the same user are explored to motivate cross-OSN personalized services. Along this scheme, we have examined the temporal characteristics of overlapped users' social activities between Twitter and YouTube. With observation that user response on Twitter is faster than that on YouTube, we develop a real-time personalized YouTube video recommender by employing the observed Twitter activity as auxiliary knowledge to predict his/her YouTube activities and preferences [24].

C. Cross-OSN Data Association

Cross-OSN data association aims to discover correlations between heterogeneous social knowledge [1], which can then be applied to unseen users. As discussed above, the cross-OSN data association is more complex and not necessarily semantic-based. Instead of designing a bottom-up association rule, one possible way is to exploit the overlapped users' collaborative activities, and resort to a crowdsourcing-like solution.

[1] Social knowledge indicates a typical pattern in users' social relation or social activity data, e.g., the SNS patterns in Facebook, the video watching patterns in YouTube, the consuming patterns in Amazon, etc.

In our previous work along this line, we assume that, *if a set of users collaboratively involve in social knowledge A_i on OSN A, and social knowledge B_j on OSN B, it is very likely that A_i and B_j are closely correlated* [25]. To distinguish it from the traditional semantic-based correlation, we call this associated social knowledge as "crowd-perceptive correlated". In this work, we associated between the friend-following patterns on Twitter and the video interaction activities on YouTube, and discovered some interesting results impossible to stand out in semantic-based solutions. For example, it is observed that the users who like watching *cute animal* YouTube videos largely follow *Australian official account* on Twitter, and the users who subscribed frequently to *war & political* YouTube videos are likely to follow *U.S. famous actors* on Twitter. While it is difficult to interpret, the observation is exactly the motivation for a data-driven solution and much consistent to the recent trending big data motto: "causality is not necessary, but correlation is important" [26].

The suggested procedures for cross-OSN data association is summarized as: (1) Determine the heterogeneous social knowledge involved in different OSNs. (2) Extract heterogeneous topics on each network and conduct cross-OSN topic association based on the observed overlapping users. (3) Design collaborative applications based on the derived heterogeneous knowledge association.

IV. Conclusion

Exploring data variety, especially the heterogeneous data from various sources, is critical to effectively utilization of the big social multimedia data. In this paper, we have argued why it is important and necessary to research into cross-OSN social multimedia data organization and association. We hope that this positioning paper could serve as a good chance to emphasize the collective utilization of various social multimedia sources and further the agenda of cross-OSN analysis and application in the multimedia community.

Acknowledgment

The authors would like to thank Prof. Ramesh Jain of UCI for his insightful discussion contributed to this paper. This paper is supported in part by the National Basic Research Program of China (No. 2012CB316304), National Natural Science Foundation of China (No. 61225009, 61332016, 61303176), and Beijing Natural Science Foundation (No. 4131004). This work is also supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

References

[1] M. Naaman, "Social multimedia: highlighting opportunities for search and mining of multimedia data in social media applications," *Multimedia Tools and Applications*, vol. 56, no. 1, pp. 9–34, 2012.

[2] L. Cao, J. Yu, J. Luo, and T. S. Huang, "Enhancing semantic and geographic annotation of web images via logistic canonical correlation regression," in *Proceedings of the 17th ACM international conference on Multimedia*, 2009, pp. 125–134.

[3] J. Sang and C. Xu, "Browse by chunks: Topic mining and organizing on web-scale social media," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 7, no. 1, p. 30, 2011.

[4] J. Sang, C. Xu, and J. Liu, "User-aware image tag refinement via ternary semantic analysis," *Multimedia, IEEE Transactions on*, vol. 14, no. 3, pp. 883–895, 2012.

[5] J. Sang and C. Xu, "Right buddy makes the difference: an early exploration of social relation analysis in multimedia applications," in *ACM Multimedia*, 2012, pp. 19–28.

[6] Y. Dong, J. Tang, S. Wu, J. Tian, N. V. Chawla, J. Rao, and H. Cao, "Link prediction and recommendation across heterogeneous social networks," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, 2012, pp. 181–190.

[7] Q. Fang, J. Sang, C. Xu, and K. Lu, "Paint the city colorfully: Location visualization from multiple themes," in *Advances in Multimedia Modeling*. Springer, 2013, pp. 92–105.

[8] S. D. Roy, G. Lotan, and W. Zeng, "Social multimedia signals: Sense, process, and put them to work," *MultiMedia, IEEE*, vol. 20, no. 1, pp. 7–13, 2013.

[9] P. Cui, L. Xie, J. Sang, and C. Xu, "Social multimedia computing," in *Proceedings of the ACM International Conference on Multimedia, MM*, 2014, pp. 1237–1238.

[10] S. Boll, R. Jain, J. Luo, and D. Xu, "Introduction to special issue on social media," *TOMCCAP*, vol. 7, no. Supplement, p. 25, 2011.

[11] Y. Tian, S. Yu, C.-Y. Lin, W. Gao, and W. Zhou, "Special issue on social multimedia computing: Challenges, techniques, and applications," *Journal of Multimedia*, vol. 9, no. 1, pp. 1–3, 2014.

[12] N. Rasiwasia, J. Costa Pereira, E. Coviello, G. Doyle, G. R. Lanckriet, R. Levy, and N. Vasconcelos, "A new approach to cross-modal multimedia retrieval," in *Proceedings of the international conference on Multimedia*. ACM, 2010, pp. 251–260.

[13] F. Wu, X. Lu, Z. Zhang, S. Yan, Y. Rui, and Y. Zhuang, "Cross-media semantic representation via bi-directional learning to rank," in *ACM Multimedia*, 2013, pp. 877–886.

[14] G.-J. Qi, C. C. Aggarwal, Q. Tian, H. Ji, and T. S. Huang, "Exploring context and content links in social media: A latent space method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 850–862, 2012.

[15] P. Goyal, L. Behera, and T. M. McGinnity, "A context-based word indexing model for document summarization," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 25, no. 8, pp. 1693–1705, 2013.

[16] V. K. Singh, M. Gao, and R. Jain, "Situation recognition: an evolving problem for heterogeneous dynamic big multimedia data," in *Proceedings of the 20th ACM Multimedia Conference, MM '12, Nara, Japan, October 29 - November 02, 2012*, 2012, pp. 1209–1218.

[17] R. Jain and L. Jalali, "Objective self," *IEEE MultiMedia*, vol. 21, no. 4, pp. 100–110, 2014.

[18] R. Jain, L. Jalali, S. Pongpaichet, and A. Gupta, "Building social life networks," *IEEE Data Eng. Bull.*, vol. 36, no. 3, pp. 91–98, 2013.

[19] H. Sundaram, "Experiential media systems," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 9, no. 1s, p. 49, 2013.

[20] S. Boll, "Multitube - where web 2.0 and multimedia could meet." *IEEE MultiMedia*, vol. 14, no. 1, pp. 9–13, 2007.

[21] M. Duggan and A. Smith, "Social media update 2013," *Pew Internet and American Life Project*, 2013.

[22] Z. Deng, J. Sang, and C. Xu, "Personalized video recommendation based on cross-platform user modeling," in *Multimedia and Expo (ICME), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–6.

[23] M. Yan, J. Sang, T. Mei, and C. Xu, "Friend transfer: Cold-start friend recommendation with cross-platform transfer learning of social knowledge," in *Proceedings of the 2013 IEEE International Conference on Multimedia and Expo, ICME 2013, San Jose, CA, USA, July 15-19, 2013*, 2013, pp. 1–6.

[24] Z. Deng, M. Yan, J. Sang, and C. Xu, "Twitter is faster: Personalized time-aware video recommendation from twitter to youtube," *TOMCCAP*, vol. 11, no. 2, p. 31, 2014.

[25] M. Yan, J. Sang, and C. Xu, "Mining cross-network association for youtube video promotion," in *Proceedings of the ACM International Conference on Multimedia*, 2014, pp. 557–566.

[26] V. Mayer-Schönberger and K. Cukier, *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013.