2014 IEEE International Conference on Systems, Man, and Cybernetics
October 5-8, 2014, San Diego, CA, USA

A Human-oriented Mutual Assistive Framework Using Collaborative Filtering Towards Disasters

Jing Li and Mingru Zeng
School of Information Engineering
Nanchang University
Nanchang, 330031, China
jing.li.2003@gmail.com, zengmingru@ncu.edu.cn

Abstract—**Originally, collaborative filtering was adopted in purchase recommendation systems (e.g., Amazon.com) based on purchased history. In this paper, we apply collaborative filtering on the basis of accumulated feedbacks of the data extracted from social media from a community of users to build up a knowledge-based framework that can match offers to needs in disaster and emergency situations. This framework is constructed by high-level data fusion, i.e., incorporating text-based natural language processing with image-based processing using long-term relevance feedback, and learns user's preferences and adjusts their needs and offers accordingly. It can be deemed as a fundamental trial for timely mutual assist in disasters.**

Keywords—human-oriented, natural language processing, content-based image retrieval, relevance feedback, collaborative filtering

I. INTRODUCTION

UNEXPECTED natural disasters such as floods and earthquakes always happen without any notice, causing injuries, the loss of life, and property damage. For example, the deadly Sichuan (China) earthquake in 2008 has affected approximately 45.54 million people, which results in 69,195 deaths and damage of around CNY ¥845.2 billion [1]. When confronting a disaster, the governor needs to act immediately, i.e., organizing a group of experts to work out the solutions to keep people safe and avoid further damages. Soon after, local/county governments use their own resources (e.g., police, fire department, emergency medical, public works and other services) to save more lives and protect property from sizeable economic losses. In such disaster situations, timely help is very essential. Besides the above-mentioned resources, whether the stakeholders (e.g., the armed forces, charities, electrical engineers, local authorities, and manufacturers in all industries) could benefit from combining efforts and aligning different sources? During a natural disaster, most of the essential facilities are disrupted, e.g., traffic, gas, water, electricity, etc. Nevertheless, communication infrastructure such as the Internet turns out to be quite robust even right after the disaster happens, which may serve as one of the tools for timely mutual assist among the stakeholders. Via Internet (especially the social media like Facebook and Twitter), information of needs and offers can be posted, which is then immediately exchanged and spread in local communities. In this way, mutual assist that was often trivialized can be brought into effect and a system that matches all kinds of needs and offers from multiple users can contribute to bridge the lack of support organized by government officials. For example, people in earthquake-stricken areas are always short of food, water, tents, sleeping bags, medical facilities, etc. But in real situations, government organizations may not act efficiently with respect to some urgent needs from long distances. Instead, mutual assist around local communities can meet people's expectations for timely help. Moreover, rather than merely responding to the disasters, it can help build long-term resilience through providing further assistance for the guidance regarding post-disaster recovery, e.g., housing, business, agriculture, and education.

In the last few years, data extracted from social media have been widely adopted for applications on a daily basis, e.g., predicting political polling results [2], detecting influenza outbreaks [3, 4], localizing individuals with more precise geographic information [5], and detecting earthquakes' information [6, 7]. Apart from these, in this paper we are going to utilize information transmitted through social media to timely match needs and offers towards the communications and collaborations against disasters. For such a matching task, it faces the following problems: 1) data spasity: the task is mostly conducted between objects, whereas human factors (i.e., the users' information) are ignored; 2) privacy: if a user requests a need or provides an offer, s/he may not want her/his privacy information (such as telephone numbers) to be spread to the public. However, how to fulfill this goal is unclear; and 3) interactivity between users is not taken into account.

To this end, we design a human-oriented mutual assistive matching framework based on collaborative learning to deal with the problems mentioned above, where users are taken as a first-class entity in data modeling and is dubbed 'human-oriented'. Collaborative filtering [8], originally limited to particular uses in E-business, such as product recommendation and webpage search of dating, is adopted to emphasize the collaboration in a group/community during the learning process. Based on constructivism [9], knowledge is constructed, transformed, and shared by the stakeholders so as to meet their individual accountability (i.e., needs and offers), meanwhile group diversity and interaction can contribute positively to the learning process. This matching framework is

978-1-4799-3840-7/14/$31.00 ©2014 IEEE

mainly composed of a content-based image retrieval (CBIR)-based module embedded with human-computer interaction (i.e., how users interact with computers) and collaborative filtering (i.e., how users interact with each other), where the parameters are adapted according to accumulated human feedbacks. In addition, a text-based module using natural language processing is also integrated. The motivation that the framework focuses on the image-based module rather than the text-based one due to the fact that images always provide more information than texts and CBIR describing an image by contents (such as color, texture, shape and so on) can overcome the limitations of the traditional keywords-based image retrieval techniques. Nevertheless, CBIR also entails a problem: there is a semantic gap between the low-level visual features and high-level image concepts [10, 11]. To bridge the gap, relevance feedback – a kind of human-computer interaction technique which takes users' feedback information into account can be adopted.

The contributions of this paper are four-fold:

1) we extend the use of collaborative filtering from product search to timely mutual assist in local communities towards disasters, in which privacy is considered;

2) we embed human-computer interaction and collaborative filtering to consider the interactivity and data sparsity;

3) the matching weight is automatically adjusted according to users' satisfaction; and

4) the framework can be used in a wide range of situations. That is, not only does the framework can be applied in disaster situations (e.g., earthquakes, tornados, landslide, tsunamis, etc), but also it can be used in sudden emergencies, such as car crashes.

In this paper, our scope is about disasters and an exemplar application is set to be under the earthquake disaster situation. This paper is organized as follows. In Section 2, we give the definition and represent needs and offers with a simple structure. Section 3 introduces the newly proposed matching framework. In Section 4, the framework is enhanced by collaborative filtering. The application to disaster management is presented in Section 5. Section 6 concludes.

II. Definition

Before introducing the matching framework, we first present a simple structure of a need and an offer. Each user taking part in the framework holds two types of information: i) a list of needs; and ii) a list of offers. Therefore, a user can be represented by a three-dimensional feature vector, namely, (the user's identity, a list of needs, a list of offers). Fig. 1 describes the relations among users, needs, and offers by a 3D cube, wherein x-axis denotes different users, y-axis represents the needs, and z-axis denotes the offers. Each black dot on the cube is represented by a three-dimensional vector with +1, -1, 0, as its elements, corresponding to whether a user requests a need or provides an offer, or just ignores.

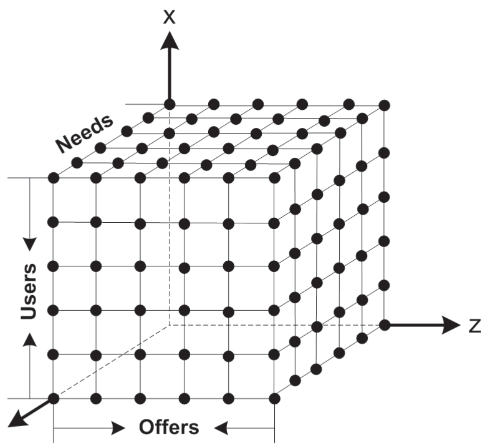

Fig. 1. The 3D cube describes the relations among users, needs, and offers.

However, the simple structure only considers one user's need and another user's offer, which is not always tenable in real disaster situations. Moreover, this may cause that some users only request to address their needs but keep their resources to themselves. What makes a good collaborative activity? Ideally, it should encourage users to listen carefully to the needs from other users in local communities when and if necessary, and one possible way is to respond the users opening their resources to others with useful suggestions that can address his/her needs to an adequate extent. This shows the importance of facilitating discussion and interaction so that users are more willing to provide available offers. Considering this, the matching framework is conducted as follows.

First of all, the framework computes a need of user x and an offer of user y with the maximum matching score, denoted by $(n_{x\_\max}, o_{y\_\max})$. If the number of pairs is more than 1, all of the pairs should be taken into account by the following search steps:

1) find $n_{a\_\max}$ of user a with the highest matching degree with $o_{y\_\max}$;

2) find $o_{b\_\max}$ of user b with the highest matching degree with $n_{x\_\max}$;

3) compare the above-computed degrees to find the higher one, and then follow these steps until a certain minimum matching score is reached.

III. The Matching Framework

More formally, a need or an offer can be represented by its belonging category associated with a list of properties with different paths, as shown in Fig. 2. Usually, a need or an offer belongs to more than one category, and in such a case, the need or offer should be represented with all possible categories.

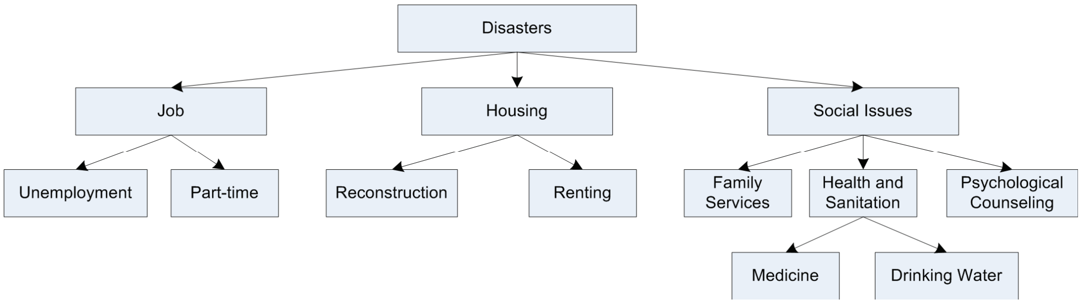

Fig. 2. The hierarchical structure of categories.

A list of property is simply described by (property name, value, weight). Here, the "value" denotes a property's characteristic and the "weight" is denoted by w with $0 \le w \le 1$ of each property quantifying its importance to a user's satisfaction with respect to the other properties. Take a job offer as an example. It can be represented by the following properties: the position name, required qualification, required skills, the income, and the location, which are expressed by {(position name, lecturer of computer science, 1), (qualification, PhD, 1), (skills, academic writing, 1), (income, 35000 pounds, 0.9), (location, London, 1)} with $w = 0.9$ indicating the income is negotiable. The weight adjustment is based on the users' degree of satisfaction with the proposed deal, as described in Section 4.

Therefore, matching between needs and offers consists of two parts: i) category matching C_M ($0 \le C_M \le 1$); and ii) property matching P_M between property values ($0 \le P_M \le 1$), and the total matching score M is given by

$$M\left(n_x, o_y\right) = C_M\left(n_x, o_y\right) * P_M\left(n_x, o_y\right) \tag{1}$$

where $C_M = 1$ if in a category the path of n_x is identical to the path of o_y; $C_M = 0$ if there is no common part of the paths; in other cases, C_M is defined by

$$C_M\left(n_x, o_y\right) = \frac{\text{length}\left(n_x \oplus o_y\right)}{\frac{1}{2}\left(\text{length}\left(n_x\right) + \text{length}\left(o_y\right)\right)} \tag{2}$$

with $\oplus$ denoting the common part of n_x and o_y. For example, given the path $n_x = \left(a, b, c, d, e\right)$ and the path $o_y = \left(a, b, d\right)$, $C_M\left(n_x, o_y\right) = \frac{3}{(1/2)(5+3)} = 0.75$.

$$P_M\left(n_x, o_y\right) = \frac{\sum_{i=1}^{N} w_i * sim\left(P_i\left(n_x\right), P_i\left(o_y\right)\right)}{\sum_{i=1}^{N} w_i} \tag{3}$$

IV. FRAMEWORK ENHANCEMENT BY COLLABORATIVE FILTERING

In this section, the framework is enhanced based on accumulated users' feedback (i.e., ratings from all users), where the property weight is adjusted according to the degree of users' satisfaction of the proposed deal with respect to an increasing number of performed matching.

The process of the weight adjustment using collaborative filtering is conducted as follows. Each user is asked which properties of the item are matched to his/her needs and which items are matched to his/her offers. The weight of the items he/she is satisfied or not satisfied will be increased or decreased, respectively. In the same category, the number of similar ratings N_r is provided for some target items and then the weight between one item m and its neighboring item n is adjusted by Peason correlation coefficient [12], which is

$$w_{m,n} = \frac{\sum_{i=1}^{m}\left(r_{a,i} - \overline{r}_a\right) * \left(r_{u,i} - \overline{r}_u\right)}{\sigma_m * \sigma_n} \tag{4}$$

and then P_M can be revised accordingly.

Since different individuals using social networks have various identity and reputations, we cannot rely on the original social media posts. Therefore, after weight adjustment, we refine the framework based on a reputation management model of identity and reputation so as to avoid biased opinions, wherein credits can be accumulated on the basis of how they have been rated by others in the past. In more details, we analyze the contents of past posts, collecting log information of different users from supporting posts with credits based on relevance feedback. Moreover, credits can be affected by the reputation of a donor, the seriousness of a posted need, and the demand degree of the need by relative positions of community members, extending an individual user to a group of users.

V. THE APPLICATION OF DISASTER MANAGEMENT

This section applies the proposed mutual assistive matching framework in an earthquake-stricken situation. The 2008 Sichuan earthquake disasters showed that government officials did not always implement to carry out help in time and were not able to organize mutual assist in local areas. In that scenario, the stakeholders include: i) government officials; ii) people in disastrous areas, such as farmers and students; iii) companies that provide help items or services, such as manufacturers and merchants; and iv) academic experts in earthquakes and medical treatments. Fig.3 illustrates the main issues of various stakeholders.

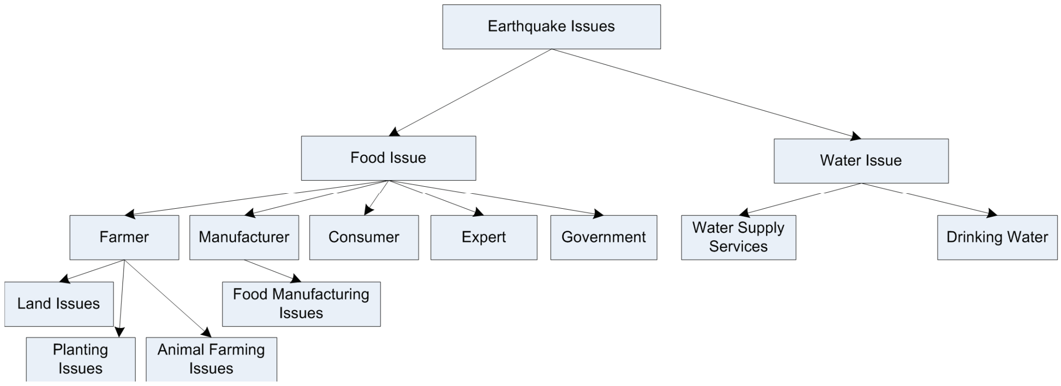

Fig. 3. Categorization of food and water issues.

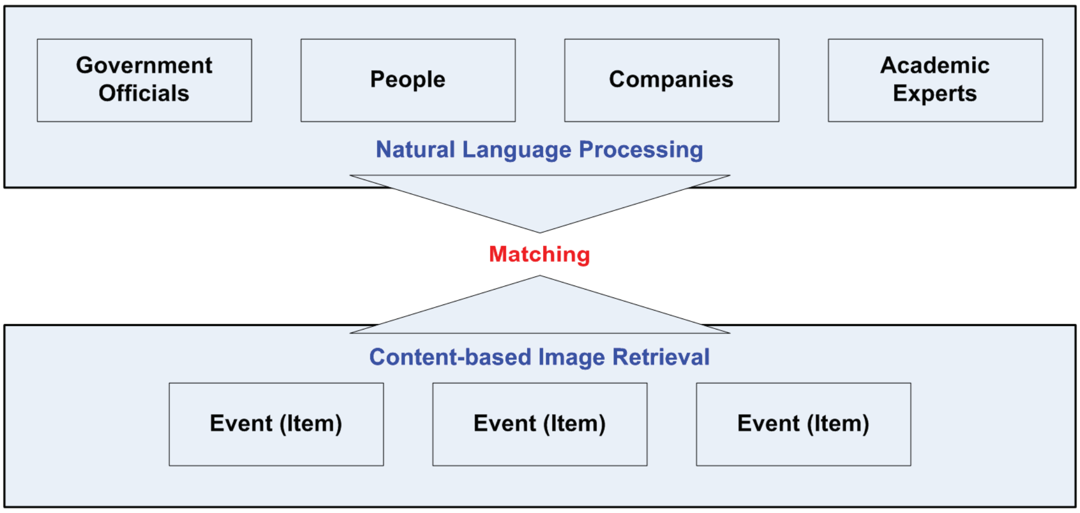

Fig. 4. The matching framework consists of two modules.

For example, an earthquake victim may have food and water issues, i.e., limited food and water supply. The government officials take a wide range of responsibilities, such as initiating a panel of topical experts with all necessary expertise to figure out the way to supply food and water to disastrous area, where each expert could define his expertise as an offer and the government could define their competence needs. However, several questions remain unanswered. Which category does the food issue belong to? How long does it take to form the expert panel? Nevertheless, the above hierarchical structure (as shown in Fig. 3) is still not the hierarchy of needs and offers since it contains stakeholders who are not a part of the issue that they care about. To develop the hierarchy, we categorize needs and offers in a more general way by merging all the sub-trees with a stakeholder in the root to one common tree of all issues. During the process, an internet environment should help and to provide a comprehensive description of matching problem.

As shown in Fig. 4, the matching interface contains two modules: i) natural language processing (NLP) to deal with text data; and 2) relevance feedback-based content-based image retrieval (CBIR) to deal with image data. It can be accessed by all users. On one hand, text information is semantically analyzed and translated by NLP, including dependency parsing, semantic enrichment, tense/modality/sentiment extraction, and so on. On the other hand, the contents of images uploaded on the social networking platform are analyzed by content-based image retrieval technique based on accumulated relevance feedback, which consists of feature extraction, feature description, and similarity matching.

VI. Conclusions and Future Works

This paper introduces a human-oriented mutual assistive matching framework based on collaborative learning to deal with the matching problems in disaster situations, which are data spasity, privacy, and interactivity between users. This matching framework is composed of a text-based module using natural language processing and a content-based image retrieval (CBIR)-based module embedded with human-computer interaction (i.e., how users interact with computers) and collaborative filtering (i.e., how users interact with each other). Its advantages are: 1) the matching weight is automatically adjusted according to users' satisfaction; and 2) it can be used in a wide range of situations. Since it is not easy to test the framework in real disaster situations, the proposed framework can be deemed as a fundamental trial for timely mutual assist in disasters.

Although many issues remain to be explored, our ultimate goal is to collect all computational powers from resources spread over networks both in time and space to accomplish a large-scale image retrieval task with respect to the mutual assistive matching problem. In future, we intend to extend the proposed framework by embedding sources of human-centric contextual information: human poses, community activity, preferences, behavior modeling, and location.

ACKNOWLEDGEMENTS

This work is supported by the grant 20142BAB217030 approved by Jiangxi Provincial Department of Science and Technology, China.

REFERENCES

[1] [Online]. Available: http://en.wikipedia.org/wiki/2008_Sichuan_earthquake

[2] L. Ampofo, N. Anstead, and B. O'Loughlin, "TRUST, CONFIDENCE, AND CREDIBILITY Citizen responses on Twitter to opinion polls during the 2010 UK General Election," *Information Communication & Society,* vol. 14, pp. 850-871, 2011 2011.

[3] E.-K. Kim, J. H. Seok, J. S. Oh, H. W. Lee, and K. H. Kim, "Use of Hangeul Twitter to Track and Predict Human Influenza Infection," *Plos One,* vol. 8, Jul 24 2013.

[4] A. Signorini, A. M. Segre, and P. M. Polgreen, "The Use of Twitter to Track Levels of Disease Activity and Public Concern in the US during the Influenza A H1N1 Pandemic," *Plos One,* vol. 6, May 4 2011.

[5] L. Backstrom, E. Sun, and C. Marlow, "Find me if you can: improving geographical prediction with social and spatial proximity," presented at the Proceedings of the 19th international conference on World wide web, Raleigh, North Carolina, USA, 2010.

[6] P. Earle, "Earthquake Twitter," *Nature Geoscience,* vol. 3, pp. 221-222, Apr 2010.

[7] P. S. Earle, D. C. Bowden, and M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world," *Annals of Geophysics,* vol. 54, pp. 708-715, 2011 2011.

[8] X. Su and T. M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques," *Advances in Artificial Intelligence,* vol. 2009, 2009.

[9] M. Dooly, "Constructing Knowledge Together," in *Telecollaborative Language Learning: A guidebook to Moderating Intercultural Collaboration*, M. Dooly, Ed., ed: Verlag Peter Lang, 2008, pp. 21-45.

[10] J. Li and N. M. Allinson, "Long-term learning in content-based image retrieval," *International Journal of Imaging Systems and Technology,* vol. 18, pp. 160-169, 2008 2008.

[11] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," *Ieee Transactions on Circuits and Systems for Video Technology,* vol. 8, pp. 644-655, Sep 1998.

[12] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," in *Sigir'99: Proceedings of 22nd International Conference on Research and Development in Information Retrieval*, M. Hearst, F. Gey, and R. Tong, Eds., ed, 1999, pp. 230-237.