

# Brokers or Bridges?

## Exploring Structural Holes in a Crowdsourcing System

**Qingpeng Zhang**, City University of Hong Kong

**Daniel Dajun Zeng**, University of Arizona and Chinese Academy of Sciences

**Fei-Yue Wang**, Chinese Academy of Sciences

**Ronald Breiger**, University of Arizona

**James A. Hendler**, Rensselaer Polytechnic Institute

**W**ith the advent of Web 2.0 and social computing, social networking has become foundational in shaping communication, collaboration, innovation, and collective behavior,<sup>1,2</sup> and social media platforms have facilitated a number of novel, dynamic social communities and collaboration mechanisms. Among these, *crowdsourcing*—the practice of outsourcing tasks to random large groups<sup>3</sup>—has grown increasingly popular in solving real-world problems, from simple information gathering through Amazon’s Mechanical Turk or Wikipedia to life-saving data that assists in post-disaster rescue missions. The extent of information exchange through crowds of unrelated individuals is understandably taking society in new directions.

Each crowdsourcing system represents a collaboration network, and the nature of that network and contributions associated with particular interactions provide insights into how information and communications technology (ICT) can facilitate collective intelligence. As task-oriented groups form, participants’ positions in the self-organized social group influence task performance, and it is important to understand how that

*A method to measure the contributions of crowdsourcing participants identifies how roles relate to an incident’s investigation and discussion. Using data on the South China tiger incident, the authors evaluate the performance of brokers—those who connect separate groups within a platform and across platforms—and show how results compare with structural hole theory.*

influence occurs. This is only one of many research topics on crowdsourcing. Others include the quality and application of crowdsourced data, user behavior, and system design.

*Structural hole theory* is well known for its ability to relate social network positions to societal influence in many sociological contexts. According to this theory, individuals within a social network collaborate to form clusters. Cluster members have strong ties through internal social connections and thus have access to similar information. However, the clusters are unrelated and noncollaborative, which creates gaps or holes in the

## HOW STRUCTURAL HOLE THEORY RELATES TO WEB USE

Crowdsourcing in any domain, Web or other, begins when individuals form collaborative groups, or clusters. Structural holes are created because unrelated clusters do not connect. Both the established weak-ties hypothesis and structural hole theory address the roles of brokers—people connecting across clusters through weak ties. The assumption is that brokers have an advantage in the breadth, timing, and arbitrage of information and are thus likely to have more innovative ideas and significant contributions than individuals who are strongly interconnected within a group and have access to similar information.<sup>1,2</sup> These theories have been thoroughly validated with case studies in various social contexts.<sup>3</sup>

It is less clear how these theories apply in online communities. Exploring these theories in crowdsourcing online can help define the organizational structure and roles of Web users. Two major questions are central to relating structural hole theory to Web use: How do crowdsourcing participants' positions relate to both the type and extent of their contributions, and do those positions share similar characteristics and principles with offline business mechanisms, such as collaborative work in a large software-development project?

Existing research is addressing similar questions for online communities. In the past five years, researchers have validated structural hole theory in this context through various efforts. One group examined the relation between Slashdot users' rank levels and their social capital measured by structural holes.<sup>4</sup> Others examined strong and weak ties in information propagation on Facebook, showing that weak ties dominate the propagation of novel information.<sup>5</sup>

Because crowdsourcing systems are task-, product-, and innovation-oriented social systems based on massive collaboration among participants, they are more common with traditional business practice than other social media groups. More importantly, the broker role is more significant in crowdsourcing systems than it is in general online communities like Facebook, where information propagation is the broker's main function.

The vital role of structural holes in virtual worlds has also been verified by the founder of structural hole theory, Ronald S. Burt, in his upcoming book tentatively titled *Structural Holes in Virtual Worlds* (<http://faculty.chicagobooth.edu/ronald.burt/research/files/NAVW.pdf>). The study of structural holes raises questions such as, do brokers contribute more than ordinary participants because of their access to information across multiple subgroups? Unfortunately, many efforts to study the role of structural holes in crowdsourcing systems lack sufficiently well-defined data on participants' collaboration. Our measurement method can help solve that problem by providing a way to reliably characterize collaboration data.

### References

1. M. Granovetter, "The Strength of Weak Ties," *American J. Sociology*, vol. 78, no. 6, 1973, pp. 1360–1380.
2. R.S. Burt, *Structural Holes: The Social Structure of Competition*, Harvard Univ. Press, 1992.
3. R.S. Burt, *Brokerage and Closure: An Introduction to Social Capital*, Oxford Univ. Press, 2007.
4. D. Ganley and C. Lampe, "The Ties That Bind: Social Network Principles in Online Communities," *Decision Support Systems*, vol. 47, no. 3, 2009, pp. 266–274.
5. E. Bakshy et al., "The Role of Social Networks in Information Diffusion," *Proc. 21st ACM Int'l Conf. World Wide Web (WWW 12)*, 2012, pp. 519–528.

social network. Brokers can close these holes by adding novel insights to the crowdsourcing experience or by acting as information bridges in a more passive role. This theory has important implications for crowdsourcing because participants can become brokers with various roles that influence the success of the incident's resolution.

The sidebar "How Structural Hole Theory Relates to Web Use" describes this idea in more detail.

To better understand how structural hole theory relates to crowdsourcing, we developed a method to measure participants' contributions and how they relate to a participant's position. Our focus is on brokers and bridges,

both of which are critical to crowdsourcing but in different ways. To test our method, we collected comprehensive raw data from online forums on the *South China tiger incident*, a crowdsourced investigation into the veracity of a 2007 South China tiger sighting conducted on the Human Flesh Search (HFS) system. HFS is a large-scale

crowdsourcing system popular in China and other East Asian countries, which is used primarily to investigate postings, such as unusual facts or startling pictures. Its curious name arises from an unfortunate translation from its Chinese name, which is closer in meaning to “human-empowered” or “crowd-powered” search. In its singular purpose, it differs somewhat from crowdsourcing in the West, which is not necessarily fixed on ferreting out the truth of a claim but can include awareness raising for a charitable cause or social event.

## INCIDENT AND DATA COLLECTION

The South China tiger incident began as a response to a hunter’s claims of encountering a wild South China tiger, a species that had not been seen in its natural environment since the 1970s. The hunter’s publicized tiger photos prompted online forums and raised questions about the photos’ authenticity. A few Web users initiated discussion threads about the photos and others contributed comments and evidence that could prove or disprove the poster’s claim. Eventually users found proof that the photos and claims were fraudulent, and the incident was resolved. Discussion threads appeared on three platforms: tianya.cn and mop.com, which are nationwide forums on general topics, and xitek.com, a platform for postings by professional and amateur photographers.

We collected six discussion threads: one thread of 1,368 postings from Tianya, three threads with 771 total postings from Mop, and two discussion threads of 3,174 total postings from Xitek.

## PARTICIPANT CONTRIBUTION TYPES

To arrive at contribution types, we created an HFS participant network in which each node represents a unique user ID and each edge represents a citation, or reply-to, relation between two user IDs. This network is consistent with network topology that has been thoroughly documented in research literature.<sup>4,5</sup>

We found that participants could contribute in four ways:

- › *Communication activity*—the number of postings a participant made—is an index of the amount of information that the participant disseminated to others. Previous work showed that most participants contributed one or two postings, leaving less than 5 percent who contributed hundreds of postings.<sup>4,5</sup>
- › *Initiation*—the participant starts a discussion thread.
- › *Investigation and finding*—the participant’s effort to conduct investigations and the key clues that participant found.
- › *Coordination*—the coordinating work performed by a participant.

Among the four contribution types, investigation and finding is the most important, yet the least explored. It is closely related to communication activity because participants rely on postings to distribute information about their investigations and relevant findings. Our method measures the integration of contributions in these two categories, which we refer to collectively as “contribution” in the rest of the article.

## MEASUREMENT METHOD

An HFS episode usually involves thousands of users and postings, making it difficult to manually evaluate an HFS participant’s contribution or performance. On the other hand, HFS episodes have traditionally covered a much broader range of topics than other crowdsourcing systems, so only minor natural language processing (NLP) is needed. To evaluate a participant’s contribution, we combine NLP with our measurement method, which is based on the popular term frequency-inverse document frequency (TF-IDF) weight.

### Assigning an importance weight

TF-IDF is a standard statistical measure used in information retrieval and text mining to evaluate the importance of a word in a document or corpus.<sup>6</sup> TF-IDF is the product of two quantities: TF, which measures the frequency of a word in a document, and IDF, which measures whether the term is common or rare in the whole document:

$$TF(t_i) = \frac{\text{No. of times term } t_i \text{ appears in a posting}}{\text{Total no. of terms in the posting}} \quad (1)$$

$$IDF(t_i) = \ln \left( \frac{\text{Total no. of terms in the posting}}{\text{No. of postings with term } t_i} \right) \quad (2)$$

where  $t_i$  is a word or term.

A term’s weight is represented by

$$TF-IDF(t_i) = TF(t_i) \times IDF(t_i) \quad (3)$$

**TABLE 1.** Top 10 keywords for the South China tiger incident.

| ID | Keyword                    | Type of keyword              | Part of speech | Importance weight |
|----|----------------------------|------------------------------|----------------|-------------------|
| 1  | 年画 (calendar paint)        | Results                      | Noun           | 0.029481          |
| 2  | 攀枝 (Pan-zhi)               | Location and key participant | Noun           | 0.029481          |
| 3  | 鑫 (Xin)                    | Results                      | Noun           | 0.029481          |
| 4  | 平面 (surface)               | Clues                        | Noun           | 0.026533          |
| 5  | 2002                       | Clues                        | Noun           | 0.023585          |
| 6  | 纸老虎 (tiger made of paper)  | Results                      | Noun           | 0.023585          |
| 7  | 纸板 (paper plates)          | Results/clues                | Noun           | 0.023585          |
| 8  | 彩印 (color-print)           | Results                      | Noun           | 0.023585          |
| 9  | 德国 (Germany)               | Results                      | Noun           | 0.023585          |
| 10 | 实地考察 (investigate offline) | Investigation method         | Verb           | 0.023585          |

As Equations 1 through 3 show, the importance of a word increases proportionally to its occurrence in the document. On the other hand, importance is offset by the term's frequency in the corpus, which helps reduce the importance of more common words.

In equation 4,  $K(t_i)$  represents the importance of the word or term  $i$ . If  $i$  is not a keyword, then  $K(i)=0$ .  $C\_Post$ , which represents the contribution of a posting, is the integrated value of TF-IDF and the keywords. A posting  $p$  by a user  $u$  is represented as

$$C\_Post(p) = \sum_i TF(t_i) \times IDF(t_i) \times K(t_i) \quad (4)$$

and the contribution of all postings for that user is represented as

$$C\_User(u) = \sum_p C\_Post(p) . \quad (5)$$

We chose summation instead of the mean for  $C\_Post$  for two reasons. First, empirical studies show that most participants contribute only a small

number of postings.<sup>4,5</sup> Second, the summation of  $C\_Post$  incorporates the frequency of the corresponding user in communicating with other participants, which indicates communication activity.

#### An example

Applying our measurement method to two postings illustrates how it works. Suppose user A posted, "I have started an offline investigation here," and then posted, "Oh! I found clue\_x!" In the first posting, offline investigation refers to the investigation that user A conducted, and in the second posting, clue\_x refers to one of the key clues.

Suppose also that  $TF-IDF(\text{offline\_investigation}) = 0.1$  and  $TF-IDF(\text{clue\_x}) = 0.2$ . The importance weight of the two terms is  $K(\text{offline\_investigation}) = 0.002$  and  $K(\text{clue\_x}) = 0.09$ . Then,  $C\_Post(1) = 0.1 \times 0.002 = 0.0002$ , and  $C\_Post(2) = 0.2 \times 0.09 = 0.018$ . Thus, the contribution of user A is  $C\_User(A) = 0.0002 + 0.018 = 0.0182$ .

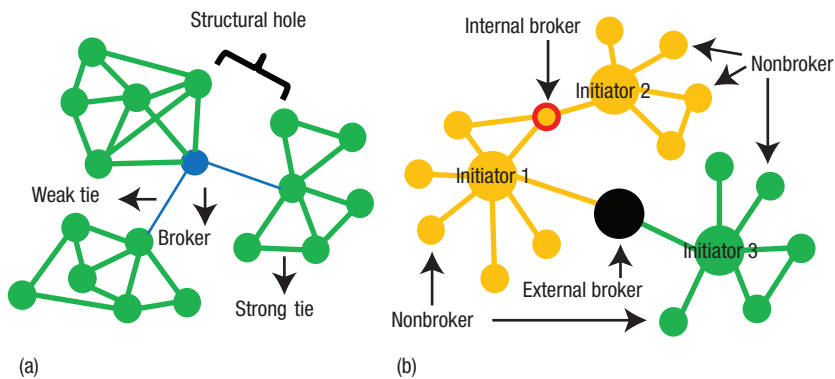
## KEYWORD SELECTION

To select keywords from the HFS postings about the South China tiger incident, we first segmented words from the Chinese characters and then calculated the TF-IDF weight of a word or term across all postings. Then, on the basis of the word's or term's frequency, its TF-IDF rankings, and the judgment of experts, we selected 175 keywords and assigned each one an importance weight. Table 1 shows the 10 most prevalent keywords related to the contribution types of interest (communication activity and investigation and finding).

## BROKER ROLES

As Figure 1 illustrates, a broker can be *internal*, connecting clusters within a platform, or *external*, connecting clusters across platforms.

Figure 1a shows connections for an internal broker (blue lines), which are relatively weak compared to the tight interconnection of cluster members



**FIGURE 1.** Internal and external brokers in a social network. (a) An internal broker connects clusters of individuals (nodes) within the same platform (green) who are likely to have similar information. Each cluster represents a discussion thread. The broker's connections (blue lines) are relatively weak because many structural holes remain among clusters. (b) An external broker (large black dot) connects initiators (those who start a discussion thread on a particular platform) and their clusters across multiple platforms—two in this case (yellow and green). Nonbrokers participated in only one discussion on one platform.

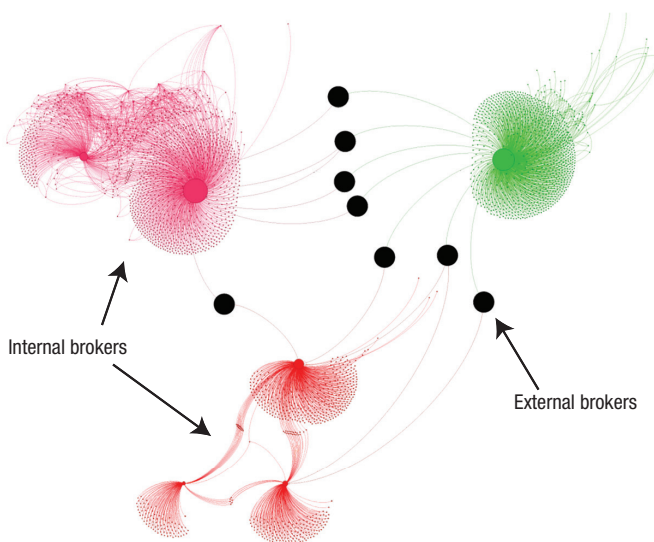
because many other structural holes exist among the three clusters. Because the internal broker governs information flow across the connection, that person is uniquely positioned to use that information in innovative ways.

In terms of discussion threads, internal brokers participated in multiple discussion threads on the same platform, external brokers participated in multiple discussion threads on multiple platforms, and nonbrokers participated in one discussion thread on one platform. The initiator of each discussion thread is the user who posted first. Because all following posts then generate an edge to the initiator, initiators become the hubs of starlike clusters in the network, in which each star-like cluster represents one discussion thread.

Figure 1b shows two platforms (orange and green) and three discussion threads, each of which has an initiator. Initiators can be either nonbrokers or external or internal brokers, depending on whether their edges connect directly to other threads or platforms. An external broker is generally an initiator, although that is not a requirement.

Figure 2 shows the internal and external brokers for the HFS participant network during the South China tiger incident.

The HFS participant network has 8 external brokers, 112 internal brokers in Xitek and 41 internal brokers in Mop, and 2,703 nonbrokers (1147 in Xitek, 925 in Tianya, and 631 in Mop). All external brokers participated in posts on two platforms (3 brokers in Mop and Tianya, 1 in Mop and Xitek, and 4 in Xitek and Tianya). Of the 42 internal brokers, 7 in Mop participated in all three discussion threads, and the other 34 participated in two threads.



**FIGURE 2.** Human Flesh System (HFS) network of participants in the South China tiger incident. The color of a node represents the platform to which the node belongs: green, Tianya; pink, Xitek; and red, Mop. Black denotes multiple platforms. A node's size is proportional to its in-degree, except for external brokers (black nodes).



## RESULTS AND IMPLICATIONS

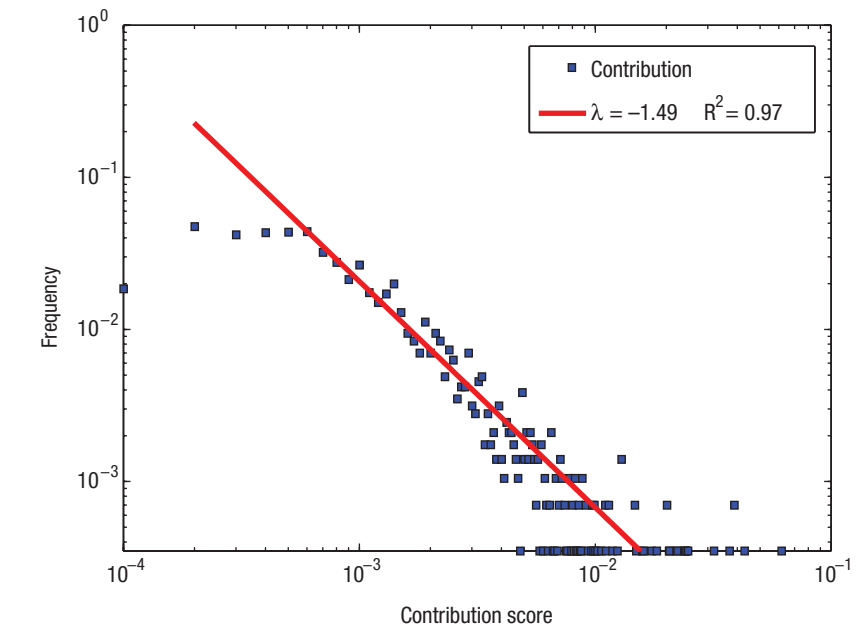
According to structural hole theory, an external broker is likely to have a bigger contribution than an internal broker or a nonbroker, because the external broker has access to the information from different discussion threads involving multiple platforms. The theory also supports the idea that an internal broker is likely to contribute more than a nonbroker, because of access to information from different discussion threads.

To test these theoretical assumptions, we calculated the average contribution of HFS participants in each discussion thread on each platform, compared differences among the three platforms, and then compared the contribution of external and internal brokers and nonbrokers against the roles associated with structural hole theory.

### Average contribution of participants

In the distribution of contribution values for all discussion threads on all three platforms, shown in Figure 3, the value for most HFS participants is very small. Values larger than 0.001 ( $10^{-3}$ ) follow a power-law distribution. Thus, most participants did not contribute much to the South China tiger incident, while a few made significant contributions. This finding matches the findings in previous studies on the distribution of humans' activities and innovation, and in turn partially validates the proposed measure's effectiveness.

Figure 4 shows the average contribution score for HFS participants according to the discussion thread on a particular platform. The average weight for Xitek participants is significantly higher than for Tianya and



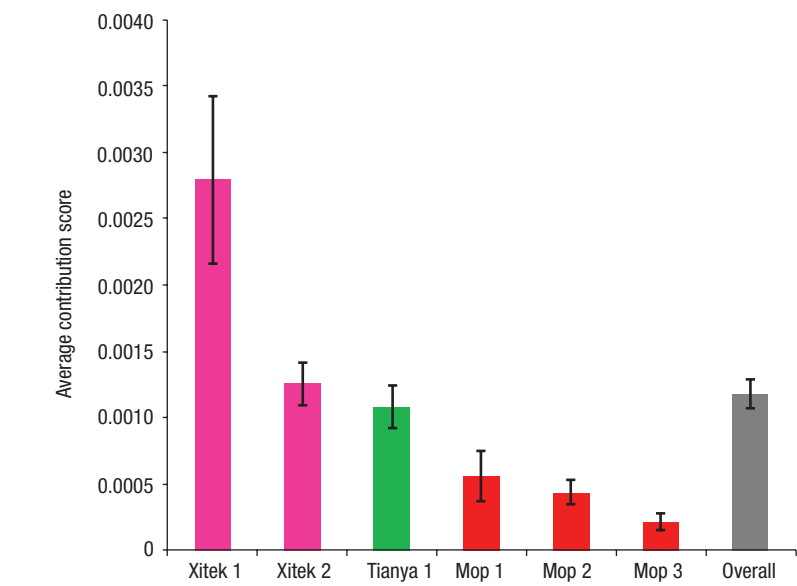
**FIGURE 3.** Distribution of participants' contributions during the South China tiger incident on HFS. Most participants did not contribute or contributed very little (left of the red line), but a few made significant contributions (had a high score, as represented by the outlying blue squares to the right of the red line).

Mop participants. The large contribution from Xitek users is understandable given that it is a platform specifically for photographers, while Tianya and Mop are forums for general topics.

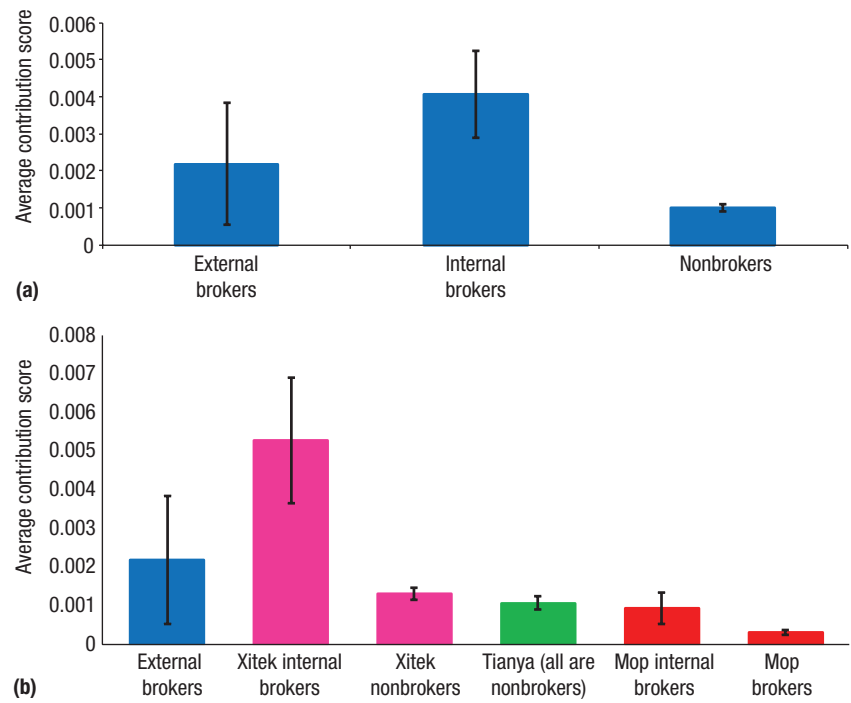
### Contributions of brokers and nonbrokers

Figure 5 shows the contributions among internal and external brokers and nonbrokers. Average contribution scores for both internal and external brokers, in Figure 5a, are significantly higher than those for nonbrokers—a finding that partially validates structural hole theory. However, the average contribution score for internal brokers is almost twice that for external

brokers—a surprising finding that contradicts structural hole theory, which maintains that external brokers should contribute more because of their unique access to information across platforms. The results in Figure 5b show that internal brokers in Xitek made by far the largest contributions, with an average contribution score twice that of the average weight for all external brokers. Indeed, nonbrokers on Xitek contributed more than the internal brokers and nonbrokers on the other platforms combined. This finding is additional evidence that HFS participants with backgrounds related to the nature of the crowdsourcing experience (in this case photography) contribute more



**FIGURE 4.** Average contribution score for posts about the South China tiger incident on the Xitek, Tianya, and Mop platforms. Numbers designate discussion threads. Because the incident centered on determining the authenticity of the South China tiger photo, contributions are higher for Xitek, a forum dedicated to photographers. This finding proves that participants with a background related to the incident contribute more than those without that background. (Averages are with a 95 percent confidence interval.)



**FIGURE 5.** Average contribution score of brokers and nonbrokers in the HFS network for the South China tiger incident. (a) Average weight across three platforms and (b) average contribution score by platform. Internal broker weights are generally higher than external broker weights, which contradicts structural hole theory, but the weights are skewed by the Xitek platform, where posters had a background strongly related to analyzing photos—a skill that was highly relevant to the crowdsourcing investigation. (Averages are with a 95 percent confidence interval.)

than those without such backgrounds.

We also found some interesting relationships among internal brokers. One of the most important clues was found by an internal broker who posted only a few times without receiving much attention from others, but this broker sent findings to another internal broker who posted many times and was regarded as a leader in the investigation. This result is an example of how the weight for internal brokers can increase in unexpected ways, and it supports the supposition that internal brokers are important to the success of the crowdsourcing experience.

These results led us to explore why internal brokers performed better than external brokers. After reading all the brokers' posts, we found that the internal brokers (those only on Xitek) were discussing various issues related to the South China tiger incident and sharing their expertise and information on their latest discovery. In contrast, external brokers did not involve themselves in the discussions and investigations, but rather acted as bridges. In other words, they did not take advantage of their unique cross-platform position to be more innovative and productive, but rather restricted their activities to information dissemination.

One explanation for this limited role is that HFS is an open source crowdsourcing system. With all information publicly available, participants can easily access other platforms without building an explicit social relationship with a participant in another platform. In addition, the South China tiger incident evoked many traditional media reports, which helped to disseminate information, albeit later than

crowdsourcing. In the Web 2.0 era of open information, the external broker apparently has a lesser part in shaping the crowdsourcing experience than the major contributor role described in structural hole theory.

### Contribution score and network topology

To characterize the importance of HFS participants relative to other criteria that might define the participant’s importance, such as perspectives, we examined how popular topological properties of a social network and by extension the HFS network<sup>2,7</sup> relate to a participant’s contribution score. Specifically, we calculated the Pearson correlation coefficient between these properties, which include brokerage measures such as efficiency, constraint, and hierarchy,<sup>8</sup> and contribution score.

As Figure 6 shows, the brokerage measures correlated more negatively with contribution score than did most other properties (row 10, columns 6 through 8), and out-degree had a significantly stronger correlation for contribution score and brokerage measures—a strong positive correlation for contribution score (row 10, column 5) and negative correlations for brokerage measures (rows 6 through 8, column 5).

To explain this interesting finding, we plan to explore causality through a more in-depth study of temporality—how the passage of time affects correlation patterns.

Our research provided considerable evidence that positions in social networks influence the contribution of crowdsourcing participants. The results of

| Topological property      | 1     | 2     | 3     | 4     | 5     | 6     | 7     | 8     | 9    | 10   |
|---------------------------|-------|-------|-------|-------|-------|-------|-------|-------|------|------|
| 1. Betweenness centrality | 1.00  |       |       |       |       |       |       |       |      |      |
| 2. Closeness centrality   | −0.06 | 1.00  |       |       |       |       |       |       |      |      |
| 3. Clustering coefficient | 0.02  | −0.30 | 1.00  |       |       |       |       |       |      |      |
| 4. In-degree              | 0.48  | −0.06 | −0.01 | 1.00  |       |       |       |       |      |      |
| 5. Out-degree             | 0.35  | −0.23 | 0.25  | 0.02  | 1.00  |       |       |       |      |      |
| 6. Efficiency             | −0.10 | 0.32  | −0.73 | −0.04 | −0.44 | 1.00  |       |       |      |      |
| 7. Constraint             | −0.21 | 0.16  | 0.04  | −0.04 | −0.67 | 0.10  | 1.00  |       |      |      |
| 8. Hierarchy              | −0.11 | 0.32  | −0.63 | −0.03 | −0.53 | 0.86  | 0.46  | 1.00  |      |      |
| 9. Eigenvector            | 0.53  | −0.06 | 0.00  | 1.00  | 0.05  | −0.05 | −0.06 | −0.04 | 1.00 |      |
| 10. Contribution score    | 0.26  | −0.14 | 0.13  | 0.05  | 0.64  | −0.21 | −0.41 | −0.26 | 0.07 | 1.00 |

**FIGURE 6.** Pearson correlations between well-known topological properties and participants’ contribution score. Each cell is colored according to correlation coefficient scale, starting with solid (−1) and lightening to white (0), and continuing with positive values from light to solid blue (+1). The numbers in the columns correspond to the property number.

applying our method to data from the South China tiger incident challenge accepted notions in structural hole theory. We found that participants with a shared background contribute more than those without one and that the average contribution of internal brokers within a platform is significantly larger than the average contribution of external brokers across platforms, who tend to be mostly information disseminators. Our analysis of data from the Hangzhou drag racing incident—13,490 postings on two platforms—support these findings and conclusions (see supplemental material: <http://personal.cityu.edu.hk/~qingzhang4/hfs-computer2016/Supplement.pdf>).

This inconsistency, which stems from differences in real- and virtual-world collaboration, suggests the need for new theories to explain and model features unique to the collaborative mechanisms in crowdsourcing

systems. To that end we plan to develop advanced NLP algorithms to automatically discover participants’ activities (investigating offline, disseminating information, organizing working groups, and so on) and to measure multiple contribution types. We also plan to create a conceptual model of participants’ contributions and statistical models of the relationship between contributions and a social network’s topological properties. Future efforts might also build on work to connect social science theories other than structural hole theory to big data on the Web.<sup>8</sup> Developing new theories will be essential to understand Web 2.0 more deeply and to fully realize its potential. ■

### ACKNOWLEDGMENTS

Zhang’s work was supported by National Natural Science Foundation of China (NSFC) grant 71402157 and City University of Hong Kong grant 7200399. Zeng’s work was supported in part by Important National



## ABOUT THE AUTHORS

**QINGPENG ZHANG** is an assistant professor in the Systems Engineering and Engineering Management Department at City University of Hong Kong. His research interests include social networks, complex systems, and data analytics for healthcare. Zhang received a PhD in systems and industrial engineering from the University of Arizona. He is a member of IEEE, ACM, and the Institute for Operations Research and the Management Sciences (INFORMS). Contact him at qingpeng.zhang@cityu.edu.hk.

**DANIEL DAJUN ZENG** is a professor in the Management Information Systems Department at the University of Arizona, and a professor at the State Key Laboratory of Management and Control for Complex Systems of the Chinese Academy of Sciences. His research interests include intelligence and security informatics, spatial-temporal data analysis, infectious-disease informatics, social computing, recommender systems, software agents, and applied operations research and game theory in e-commerce and online advertising systems. Zeng received a PhD in industrial administration from Carnegie Mellon University. He is a Fellow of IEEE. Contact him at zeng@email.arizona.edu.

**FEI-YUE WANG** is the director of and a professor at the State Key Laboratory of Management and Control for Complex Systems of the Chinese Academy of Sciences. His research interests include social computing, economic computing, parallel management, Web and services science, and agent-based intelligent systems. Wang received a PhD in computer and systems engineering from Rensselaer Polytechnic Institute. He is a Fellow of IEEE, the American Association for the Advancement of Science (AAAS), the International Federation of Automatic Control (IFAC), the International Council on Systems Engineering (INCOSE), and the American Society of Mechanical Engineers (ASME). Contact him at feiyue.wang@ia.ac.cn.

**RONALD BREIGER** is a professor in the School of Sociology at the University of Arizona. His research interests include social and adversarial networks, network stratification, mathematical models and theory, and measurement issues in cultural and institutional analysis. Breiger received a PhD in sociology from Harvard University. Contact him at breiger@email.arizona.edu.

**JAMES A. HENDLER** is a professor in the Computer Science Department at Rensselaer Polytechnic Institute and director of the Rensselaer Institute for Data Exploration and Applications. His research interests include Web science, big data, and artificial intelligence. Hendler received a PhD in artificial intelligence from Brown University. He is a Fellow of IEEE, the Association for the Advancement of Artificial Intelligence (AAAI), AAAS, and the British Computer Society (BCS). Contact him at hendler@cs.rpi.edu.

*Social Network Analysis and Mining*, R. Alhajj and J. Rokne, eds., Springer, 2014, pp. 1879–1892.

3. A. Doan, R. Ramakrishnan, and A.Y. Halevy, "Crowdsourcing Systems on the World Wide Web," *Comm. ACM*, vol. 54, no. 4, 2011, pp. 86–96.
4. F.-Y. Wang et al., "A Study of the Human Flesh Search Engine: Crowd-Powered Expansion of Online Knowledge," *Computer*, vol. 43, no. 8, 2010, pp. 45–53.
5. Q. Zhang et al., "Understanding Crowd-Powered Search Groups: A Social Network Perspective," *PLoS ONE*, vol. 7, no. 6, 2012; <http://dx.doi.org/10.1371/journal.pone.0039749>.
6. H.C. Wu et al., "Interpreting TF-IDF Term Weights as Making Relevance Decisions," *ACM Trans. Information Systems*, vol. 26, no. 3, 2008, pp. 13–1–13–37.
7. R.S. Burt, *Brokerage and Closure: An Introduction to Social Capital*, Oxford Univ. Press, 2007.
8. J.W. Mohr et al., "Graphing the Grammar of Motives in National Security Strategies: Cultural Interpretation, Automated Text Analysis and the Drama of Global Politics," *Poetics*, vol. 41, no. 6, 2013, pp. 670–700.

Science and Technology Specific Projects grant 2013ZX10004218 and NSFC grants 91024030, 91224008, and 71025001. Wang's work was supported in part by NSFC grants 61533019, 71232006, and 61233001. Breiger's work was supported in part by US National Science Foundation (NSF) grant 1314631.

## REFERENCES

1. A.C. Weaver and B.B. Morrison, "Social Networking," *Computer*, vol. 41, no. 2, 2008, pp. 97–100.
2. Q. Zhang, D. Difranzo, and J.A. Hendler, "Social Networking on the World Wide Web," *Encyclopedia of*



Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.