

Hippocampus: Answering Memory Queries using Transactive Search

Michele Catasta[†], Alberto Tonon^{*}, Djellel Eddine Difallah^{*}, Gianluca Demartini^{*},
Karl Aberer[†], and Philippe Cudre-Mauroux^{*}

[†]EPFL—Switzerland
{firstname.lastname}@epfl.ch

^{*}eXascale Infolab, University of Fribourg—Switzerland
{firstname.lastname}@unifr.ch

ABSTRACT

Memory queries denote queries where the user is trying to recall from his/her past personal experiences. Neither Web search nor structured queries can effectively answer this type of queries, even when supported by Human Computation solutions. In this paper, we propose a new approach to answer memory queries that we call *Transactive Search*: The user-requested memory is reconstructed from a group of people by exchanging pieces of personal memories in order to reassemble the overall memory, which is stored in a distributed fashion among members of the group. We experimentally compare our proposed approach against a set of advanced search techniques including the use of Machine Learning methods over the Web of Data, online Social Networks, and Human Computation techniques. Experimental results show that Transactive Search significantly outperforms the effectiveness of existing search approaches for memory queries.

Categories and Subject Descriptors: H.1.2: [User/Machine Systems]: Human information processing

Keywords: Transactive memories; Web search; Web of Data; Crowdsourcing; Social Networks

1. INTRODUCTION

Nowadays, search engines are able to answer a wide variety of queries. Traditionally, Web search queries are classified as navigational, transactional, or informational [3]. In addition, recent studies have shown that many Web search queries are about entities [6]. On the other hand, more complex queries involving joins among different data collections (e.g., birthdate of the mayor of the capital city of Italy) are best answered by databases and Semantic Web systems using declarative query languages (e.g., SQL, SPARQL).

In this paper, we focus on yet another type of queries we call *memory queries*. Memory queries relate to information needs where the user is trying to recall something from the past (e.g., “at what time did I go to sleep yesterday?”). We define *Transactive Search* as the approach where memory

queries are answered by a group of people who *collectively* retrieve knowledge (e.g., “what is the name of the restaurant we all went to last Friday?”).

Traditionally, this type of queries are answered by personal communication and group discussion. In this paper, we focus on *transactive memory* [12] solutions by proposing new methods to answer them and by experimentally comparing existing and novel search techniques to determine which approach is most effective in our context. We propose Web-based methods to answer this type of queries using online Social Networks, the Web of Data, Human Computation, and a novel Transactive Search approach. More specifically, we compare different approaches to answer the following specific query “Who attended the International Semantic Web Conference (ISWC)?”.

Our proposed Transactive Search approach requires the user (i.e., the person who issues the query) to be part of the group of people who is reconstructing the memory (i.e., answering the query) but does not necessarily need to be part of the original event for which he/she is trying to recall information. In our running example, the user did not attend the ISWC conference, but still wants to reconstruct the list of attendees. In this paper, we show how current search techniques are ill-suited for this type of query, whereas our transactive memory search approach performs best but degrades over time. Furthermore, for certain retrieval tasks, we give evidences that tapping into the memories of users’ social ties leads to the best results, notwithstanding the paradigm shift that human memories are incurring in the current “Search Engine era” [8].

The rest of this paper is structured as follows. We briefly review current search approaches in Section 2. In Section 3, we describe a series of search strategies that can be used to answer memory queries on the Web. Section 4 introduces our novel Transactive Search approach. Section 5 reports on the experimental results of our search approaches and, finally, Section 6 concludes the paper.

2. RELATED WORK

Broder [3] provided the first classification of Web search query types. According to his classification, informational queries denote those queries where the user is looking to fill an information gap or to learn something new. Navigational queries are queries made when the user is trying to reach a specific Web page. Finally, transactional queries are those where the user is trying to accomplish a specific task as, for example when buying or downloading something from the Web. Since Broder’s classification, Web Search engines have

become more powerful, and SERPs (Search Engine Result Pages) got richer. Today, many Web Search queries are about entities [6] and SERPs present factual information about them (e.g., in Google’s Knowledge Panel).

In Parallel to the evolution of Web Search, the Linked Open Data movement has pushed the availability of large amounts of machine-readable data on the Web. This allows the creation of semantic applications and of semantic search functionalities: Today, it is hence possible to run complex structured queries (e.g., in SPARQL) over multiple datasets describing entities [2], or to leverage the Web of data to answer entity queries [9].

Most recently, new approaches to answer search queries by means of Human Computation have been proposed. Queries that cannot be effectively answered by machines only can now be tackled thanks to crowdsourcing. Research in this domain focused on extending database technologies with Human Computation to improve answers for SQL queries [5]. Also related to the queries we consider in this paper, enumeration queries [10] can be powered by human computations where the crowd is asked to complete a list of elements (e.g., ice cream flavors, US States, etc.). As compared to enumeration queries, this paper focuses on queries that personally concern the user (i.e., memories queries). In this context, a random crowd of workers cannot answer such queries as effectively as the group of people in the user’s social network.

CrowdPlan [7] is a Human Computation algorithm generating simple plans given a high-level mission statement. The authors of CrowdPlan define a series of steps involved in this process, including decomposition, rewriting, assessment, filter, and search. This is a crowd-only approach that can be applied in cases where the user has a complex request for which the answer can be retrieved from the Web. However, such an approach is not always scalable, since in most cases complex requests may require crowd workers to deal with large amount of data. Thus, a hybrid human-machine system would be preferable in many situations in order to enable both scalability and semantic understanding of human intelligence. In this work, we propose and evaluate such a hybrid approach to answer memory queries.

Another related crowdsourcing application is presented in [1], where the creation of sessions for an academic conference (CHI’13) was crowdsourced to the authors and program committee members. In this paper, we contact similar users in order to reconstruct transactive memories.

In summary, different search approaches have been proposed to answer different types of information needs. In this paper, we focus on memory queries only, evaluate existing search approaches for answering such queries, and propose a solution based on Transactive Search to more effectively tackle those queries.

3. SEMANTIC SEARCH APPROACHES FOR MEMORY QUERIES

In the following, we describe different search strategies that may be used to answer memory queries. In particular, we focus on the two following memory queries: “Who attended the ISWC conference in 2012 / 2013?”.

| | |
|--------------------|-----------------------------|
| isFirstAuthor | isConference&WorkshopAuthor |
| isMiddleAuthor | numberOfPapers |
| isLastAuthor | numberOfCoauthors |
| isWorkshopAuthor | hasTweeted |
| isConferenceAuthor | numberOfTweets |

Table 1: Features used in the machine-based search approach. Given a person p and an ISWC edition e , numberOfPapers is the number of papers written by p and appearing in the proceedings of e , while numberOfCoauthors is the number of people who coauthored some paper with p in e .

3.1 Machine-based Search with the Web

The first approach we propose is fully machine-based and uses Web content to answer the query. The goal of this first approach is to estimate the likelihood of a person attending the conference given a list of features. The features we adopt are based on the authorship of papers published at the conference and co-located workshop, their affiliations, and on their presence on online Social Networks relevant to the event. Specifically, the features we used are listed in Table 1. In order to retrieve the data needed to define such features for the ISWC conference, our system runs SPARQL queries over a conference metadata repository storing machine-readable information about ISWC papers and authors¹. Additionally, we use the Twitter API to retrieve tweets relevant to the conference, looking for the hashtags #iswc2012 and #iswc2013 (respectively). If a training set is available (e.g., attendees list from a previous year) we additionally train a regression model (e.g., [11]) to rank candidate attendees by their likelihood of attending the event. Note that this approach is using both the Social Web as well as the Web of Data.

3.2 Hybrid Human-Machine Search

To improve over the Machine Learning (ML) approach described in Section 3.1, we additionally consider Human Computation.

Crowdsourcing for Uncertain Cases.

Specifically, we take the output produced by the machine-based approach described above (i.e., a list of candidate attendees ranked by the likelihood of having attended the conference) and select a set of candidate results to be crowdsourced based on the confidence score of the automatic approach. Note that the size of the candidate set to be crowdsourced depends on the available budget. For each crowdsourced candidate, the system creates a micro-task (also known as a Human Intelligence Task or HIT) on a paid crowdsourcing platform². Each HIT for this query is a binary question about a specific candidate attendee. Such tasks are designed to find evidence that supports the fact that the candidate attended the conference (e.g., on the conference website, Twitter, or using general Web Search).

The advantage of this additional Human Computation step is that the crowd worker can easily read tweets or find relevant blog posts to understand if the candidate is only related to the community (e.g., because he is one of the co-

¹<http://data.semanticweb.org/>

²We currently use Amazon MTurk (<http://www.mturk.com>) and OpenTurk (<http://www.openturk.com>) [4] as crowdsourcing platforms.

authors of one accepted paper), or if he really attended the conference. This step is beneficial to increase the overall effectiveness of the approach as the crowd is able to change the decisions made using our ML techniques.

Note that we do not consider a Human Computation-only approach here because of scalability issues. Previous work [10] has shown that it is possible to compose a list of items from scratch by means of crowdsourcing. However, for our very specific queries it may be too difficult and too expensive. Thus, we use an hybrid human-machine approach to deal with the large set size, by automating the candidate selection step.

Crowdsourcing for Unseen Cases.

In order to improve the ML approach described in Section 3.1, we also use a second type of HIT: We ask crowd workers to find additional names of candidate attendees in addition to the authors and Twitter users we collect programmatically. Specifically, we ask the crowd to visit the event website and to find names of attendees listed there (e.g., keynote speakers, session chairs, organizers, etc.). These names would be difficult to extract automatically due to the specific layout of each event website. In this way, we increase the Recall of the original machine-based approach as new attendees are added to the list.

4. TRANSACTIVE SEARCH

We propose a novel approach, named Transactive Search, to answer the memory queries we consider. In this approach, the user is supported by a system that lets a group of people reconstruct memories in a transactive way, as envisioned in [12]. The user is triggering a Transactive Memory retrieval process by submitting the query to the Transactive Search system.

The first member of the transactive memory group is the user himself. The user has to mention some seed persons who might be related—thus could potentially contribute—to the memory query. In our case, the first user specifies an initial list of persons having attended ISWC. The system then starts the process of contacting the initial seed of attendees to retrieve additional attendees. This process continues iteratively until the system does not obtain any new attendee. Although not a strict requirement, in our experiment the two steps of contributing to the memory contents and nominating new attendees coincide. Nevertheless, it is trivial to generalize the Transactive Search approach in such a way that the two steps are distinct (i.e., first the user contributes to reconstructing the group memory, then nominates new users that have the potential to further improve the group memory).

Memory Cues.

We often store and retrieve memories by associating them with related information (e.g., contextual, spatial and/or temporal information). In order for members of the group to more easily recall elements from the list (i.e., ISWC participants), our system provides them with memory cues. Examples include questions like ‘Who did you meet at the social events?’, ‘Who did you see at the poster/demo sessions?’ and at the workshops? and during the paper presentations?’, etc.

In this way, the system is leading the users to recall less popular elements from the list. Similarly, crowd enumera-

tion approaches [10] ask the crowd about elements starting with a specific letter of the alphabet. This helps the contributor in focusing on potentially less popular elements and hence improves the recall for the query.

Implementation of the Transactive Search system.

Our Transactive Search system was developed as a Web application, aimed at reconstructing the attendees’ lists of two conferences, while recording at the same time all users interactions. We put a special focus on the user experience to make sure that the members of the group memory would spend as little time as possible to input the data. We deem this aspect of key importance because (in its current implementation) our Transactive Search system lacks any form of incentive for the contributors (contrary to most of the crowdsourcing platforms in use nowadays, which often rely on monetary incentives). On the other hand, we made sure that the user knew in real-time if his/her input was novel, or had already been contributed by a previous user (hence introducing a gamification element). Manual input is prone to errors and typos, so we also employed fuzzy matching throughout the whole stack, to keep duplicate entries at a minimum. Finally, we kept track of all duplicate mentions, to identify the list elements that were easily remembered by the group, while at the same time analyzing the decay of novel contributions over time.

New potential contributors in the group were contacted automatically via email, unless the mentionee did not specify any email address (which was left as an optional field, to speed up the input phase). The missing email addressees were retrieved by a domain expert, but the same task could have been accomplished by the crowd. For the given task, we would have had to crowdsourcing about 300 addresses, totaling an expense of around \$45.

5. EXPERIMENTAL EVALUATION

5.1 Experimental Setting

To highlight the difference between the three types of search approaches discussed above (i.e., Web-based, Human Computation, and Transactive Search), we picked a scenario where all systems could be used without being biased towards a clear winner. In our experiment, the target query is to retrieve the participants list of an event. The requested information has the property of being partly documented on the event website and on the social Web. In this case, automatic and crowdsourcing approaches can be used, hence we can perform a sensible comparison with our Transactive Search approach.

In the following experimental evaluation, we compare three fully automatic approaches, three hybrid approaches, and our Transactive Search approach. The simplest method we test is *Authors and Tweets*, that merely returns the list of all authors of papers accepted at the conference and related workshops as well as people who tweeted during the edition of the ISWC conferences taken into consideration. This method is expected to perform well in terms of Recall but poorly in terms of Precision. *SVM* and *MP5 Regression* [11] are two ML approaches that use the set of features shown in Table 1 to compute a list of participants. The former uses Support Vector Machines to classify people as attendees or non-attendees, while the latter uses a decision tree model to give a ranked list of people based on their likelihood of

| Approach | Precision | Recall | F-measure |
|-------------------------|---------------|---------------|---------------|
| Authors and Tweets | 0.3293 | 0.7277 | 0.4534 |
| SVM | 0.6947 | 0.3784 | 0.4899 |
| M5P Regression | 0.8151 | 0.2474 | 0.3796 |
| Hybrid_uncertain | 0.6743 | 0.4262 | 0.5223 |
| Hybrid_unseen | 0.6979 | 0.4179 | 0.5228 |
| Hybrid_uncertain_unseen | 0.6362 | 0.5780 | 0.6057 |
| Transactive Search | 0.8677 | 0.3410 | 0.4896 |

Table 2: Effectiveness of machine-based, hybrid, and Transactive Search approaches using Crowdsourcing for ISWC 2012.

attending the conference.³ In both cases, in order to separate training data and test data, we use cross validation between the two editions (2012 and 2013) of the conference taken into consideration.

The first hybrid approach we take into account, namely *Hybrid_uncertain*, is an enhancement of MP5 Regression that uses crowdsourcing to confirm uncertain cases. Specifically, we ask the crowd to find evidence of attendance for all uncertain attendees, i.e., for people with a score between 0.2 and 0.5, selected by manual inspection and cross-validation over the two datasets.

The second hybrid approach (*Hybrid_unseen*) extends the list of people who authored a contribution and those who tweeted by asking the crowd to search the Web for attendees (e.g., event organizers on the event website) and provide a list of names. Finally, with *Hybrid_uncertain_unseen* we enhance MP5 Regression even further by modifying *Hybrid_uncertain* to use the crowd to increase the list of attendees, as described in Section 3. The first crowdsourcing task (check the Web to find whether a person has attended the event) was rewarded \$0.05 and assigned to 3 workers while the second crowdsourcing task (find names of event attendees on the Web) was rewarded \$0.5 and assigned to 5 workers. Final decisions based on the crowd input was based on majority votes.

5.2 Experimental Results

Fully Automatic, Hybrid Human-Machine, and Transactive Search approaches.

Table 2 and 3 show the effectiveness results for our baseline approaches and for the Transactive Search approach, respectively for the ISWC 2012 and 2013 attendees lists⁴.

When comparing automated techniques with our Transactive Search approach, we note that in the 2013 experiment (where we asked people to remember an event that happened about 1.5 months earlier) Transactive Search outperforms the best automatic approach by 46% in terms of F-measure (see Table 3).

Transactive Search Statistics.

The starting point of our Transactive Search approach is the set of event attendee names the user could remember in the first place. For the 2012 experiment we had 33 seed elements, while for the 2013 experiments we had 31 initial

³During the experiments, we considered as selected attendees all people with a score greater than 0.5.

⁴The ground truth data consists of the official attendees list gently provided by the conference organizers.

| Approach | Precision | Recall | F-measure |
|-------------------------|---------------|---------------|---------------|
| Authors and Tweets | 0.3048 | 0.6906 | 0.4229 |
| SVM | 0.6632 | 0.4532 | 0.5385 |
| M5P Regression | 0.6599 | 0.4652 | 0.5457 |
| Hybrid_uncertain | 0.5864 | 0.4964 | 0.5377 |
| Hybrid_unseen | 0.4884 | 0.6043 | 0.5402 |
| Hybrid_uncertain_unseen | 0.4592 | 0.6211 | 0.5280 |
| Transactive Search | 0.9006 | 0.7136 | 0.7963 |

Table 3: Effectiveness of machine-based, hybrid, and Transactive Search approaches using Crowdsourcing for ISWC 2013.

names. Hence, the bias incurred by the initial conditions across the two experiments is minimal.

In total, our system sent 399 e-mails to potential attendees, among which 266 have been read. Out of all contacted persons, 149 participated to the experiment for at least one of the two events. On average, in 20% of the cases people provided e-mail addresses in addition to the names themselves. In the other cases, we asked a domain expert to complete the e-mail addresses.

The Effect of Time on Human Memories.

By comparing Table 2 and 3, we note that in the 2012 experiment (where we asked people to remember something that happened about 13 months earlier) Transactive Search loses heavily in terms of Recall as compared to 2013, while still keeping the same high level of Precision. This suggests that contributors correctly entered names of attendees but could remember less as more time had passed. Manually checking the errors, we notice that some are due to incorrect spelling of names and others to “false positives” like Michael Stonebraker, who participated to a panel session at the event (thus, contributors listed him) but was not in the official list of event attendees. For 2012, our Transactive Search could in the end come up with 164 out of the actual 481 attendees, while for 2013 a total of 299 attendees were determined out of 419.

Comparison among Automatic and Hybrid Approaches.

By comparing the other approaches among each other, we notice that for 2012 the best performing approach in terms of F-measure is the hybrid human-machine approach using two steps of crowdsourcing (to correct automatically-assigned labels and to enhance the original list of authors and people who tweeted). On the other hand, for 2013 the best performing approach does not take advantage of crowdsourcing but simply uses regression models [11] over the proposed features to estimate people attendance. We also note that taking the entire list of authors and people who tweeted obtains very high Recall but an unacceptable level of Precision: not all the authors are actually going to attend the conference.

Looking at the crowdsourcing contribution we note the following. For both 2012 and 2013, asking the crowd to improve the results yields higher Recall and lower Precision. In the first crowdsourcing task, where the crowd worker had to confirm the attendance of a target person, a gain in Recall means that workers correctly modified from 0 to 1 the label of many candidates. This is the case when they found positive evidence of attendance on the Web while the ML approach could not find it. On the other hand, the loss in

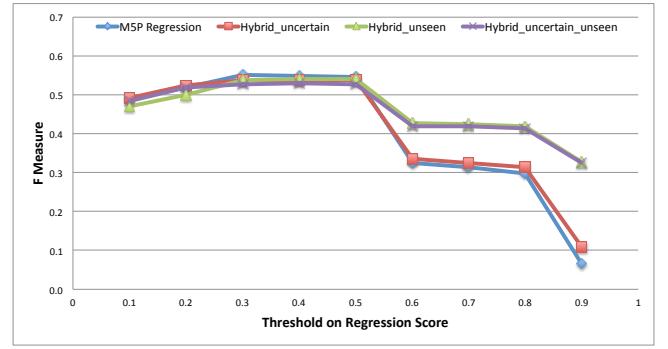
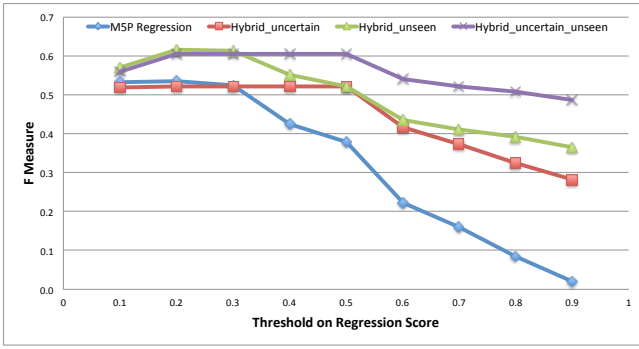


Figure 1: Variation of F-measure for ISWC 2012 (left) and 2013 (right) by changing the binarization threshold on the classification score for both the regression based approach and the crowdsourcing enhancements.

Precision is explained by the fact that crowd workers incorrectly modified some labels from 1 to 0. This is probably the case because they could not find attendance evidence even when the person actually attended the event. In that case, workers should rather have selected the option ‘Don’t Know’.

Figure 1 shows how effectiveness varies by changing the classification threshold. We observe that the Human Computation approaches are more robust as compared to the ML regression baseline when varying the threshold. Specifically, for threshold values in the crowdsourced interval (i.e., 0.2 – 0.5) the effectiveness reaches its maximum. For the other experimental results, we decided to use 0.5 as the threshold.

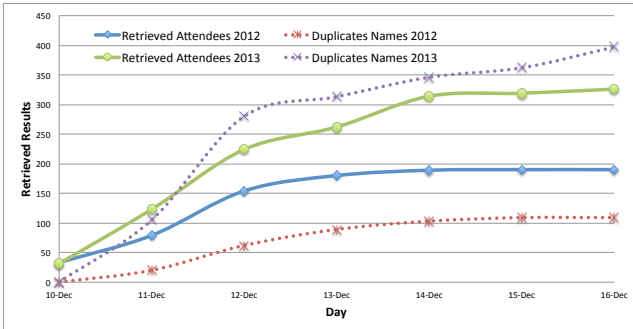


Figure 2: Number of attendees found by the Transactive Search approach over time.

Attendee Set Size over Time with Transactive Search.

Figure 2 shows how many attendees have been collected at each step of the Transactive Search approach that we run for seven iterations in total. As we can see from both experiments, in the first 2-3 days most of the attendees are collected. Then, a plateau is reached. We hypothesize that the number of recalled attendees is higher in 2013 because it is closer in time to the experiment date (i.e., December 2013). Similarly, the lower number of duplicate names reported in 2012 hints to the effects of human memory decay.

The Effect of Hubs and Authorities in the Transactive Memory Graph.

Using the data collected from the Transactive Search experiment, we can construct a graph which we call the *Trans-*

active Memory Graph. Specifically, each node in this graph is a mentioned attendee while there is an edge from node i to node j whenever a person i reported a person j as an attendee during the experiment. A visualization of the 2013 graph is shown in Figure 5.

A first analysis of the graph shows a decreasing distribution of in-degree for both experiments: many people are mentioned just once while few popular attendees are mentioned many times (see Figure 3). The top in-degree nodes

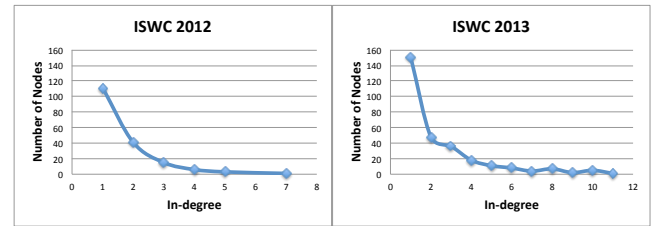


Figure 3: In-degree of nodes in the Transactive Memory Graph.

(i.e., the attendees who were mentioned most often) are ‘Oscar Corcho’ for 2013 and ‘Natalya F. Noy’ for 2012, who are both prominent members of the Semantic Web community.

Figure 4 shows the cumulative distribution function for the nodes out-degree in the Transactive Memory Graph. We observe that the value grows rapidly in both cases, having 80% of the nodes within degree 20 in 2012 and 30 in 2013.

In terms of efficiency, the compared approaches clearly have different execution timescales: The automatic ML approaches are in the below-one-second execution time range, the hybrid approaches with crowdsourcing take 1-2 days to complete, while the Transactive Search approach we run was

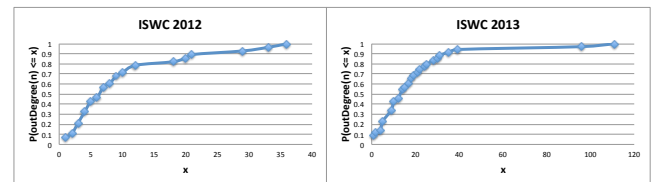


Figure 4: Cumulative distribution function for the nodes out-degree in the Transactive Memory Graph.

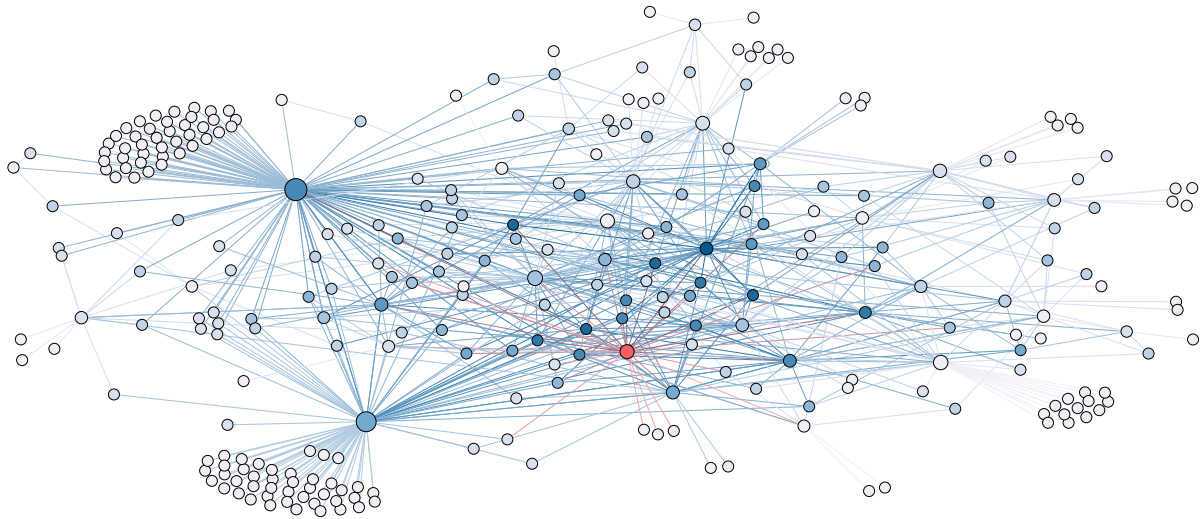


Figure 5: Transactive Memory Graph emerging from the Transactive Search process run for ISWC2013. Nodes represent attendees, and edges represent nominations (i.e., who mentioned whom in our system). The node size is proportional to the out-degree (i.e., how many mentions were made), while the color saturation is proportional to the in-degree (i.e., how many mentions were received). The source node is colored in red, and it is interesting to notice how central and well-connected it is in the graph. Conversely, some of the top contributors (on the left) are well-connected with the central component, but introduced also new, isolated clusters (pointing to how effective the Transactive Search approach is, since it could introduce new items disregarded by the majority of the community). Lastly, the diameter of the graph is 9, meaning that different nodes mentioned by the source node introduced new, isolated nodes until the end of the experiment.

stopped after 7 days in total (during which we decided to iterate every 24 hours). However, we did not focus on efficiency for this paper, as the Transactive Search iterations could have been executed more promptly, e.g., new attendees could have been contacted as soon as they were discovered, rather than batching the emails in daily iterations.

6. CONCLUSIONS

Not all queries can be answered using Search Engines. In this paper, we addressed a specific type of memory queries, which can be answered using a Transactive Search approach: A group of people collectively remembers the list of people attending an event that happened in the past. We experimentally compared traditional automatic Web-based approaches based on Machine Learning and modern Human Computation approaches combining regression models with selective crowdsourcing, and a novel approach based on Transactive Search. Experimental results show that the Transactive Search system remains the best way to answer queries for which the required information is not fully documented on the Web, yielding an improvement up to 46% over the best alternative approach.

Acknowledgements We are grateful to all the participants of our experiment, who contributed fundamentally to the performance of our Transactive Search system. Furthermore, we thank Armin Haller for providing the ground truth of ISWC2013. This work was supported by the Swiss National Science Foundation under grant number PP00P2 128459, and by the Haslerstiftung in the context of the Smart World 11005 (MEMORIES) project.

7. REFERENCES

- [1] P. André, H. Zhang, J. Kim, L. Chilton, S. P. Dow, and R. C. Miller. Community clustering: Leveraging an academic crowd to form coherent conference sessions. In *First AAAI Conference on Human Computation and Crowdsourcing*, 2013.
- [2] C. Bizer, T. Heath, and T. Berners-Lee. Linked data—the story so far. *IJSWIS*, pages 1–22, 2009.
- [3] A. Broder. A taxonomy of web search. *SIGIR Forum*, pages 3–10, 2002.
- [4] D. E. Difallah, G. Demartini, and P. Cudré-Mauroux. Pick-a-crowd: Tell Me What You Like, and I’ll Tell You What to Do. *WWW ’13*, pages 367–374, 2013.
- [5] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin. CrowdDB: Answering queries with crowdsourcing. *SIGMOD ’11*, pages 61–72, 2011.
- [6] R. Kumar and A. Tomkins. A characterization of online browsing behavior. *WWW ’10*, pages 561–570, 2010.
- [7] E. Law and H. Zhang. Towards large-scale collaborative planning: Answering high-level search queries using human computation. In *AAAI*, 2011.
- [8] B. Sparrow, J. Liu, and D. M. Wegner. Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, pages 776–778, 2011.
- [9] A. Tonon, G. Demartini, and P. Cudré-Mauroux. Combining inverted indices and structured search for ad-hoc object retrieval. In *SIGIR*, pages 125–134, 2012.
- [10] B. Trushkowsky, T. Kraska, M. J. Franklin, and P. Sarkar. Crowdsourced enumeration queries. *ICDE ’13*, pages 673–684, 2013.
- [11] Y. Wang and I. H. Witten. Induction of model trees for predicting continuous classes. In *Poster papers of the 9th European Conference on Machine Learning*, 1997.
- [12] D. M. Wegner. Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior*, pages 185–208. 1987.