

Hive Open Research Network Platform

Jung Hyun Kim, Xilun Chen,
K. Selçuk Candan
Arizona State University
Tempe, AZ 85287, USA

{jkim294, xilun.chen, candan}@asu.edu

Maria Luisa Sapino
Dipartimento di Informatica
Universita degli Studi di Torino
I-10149 Torino, Italy

marialuisa.sapino@unito.it

ABSTRACT

Did you ever return back from a conference, having met a lot of interesting folks, listened to many inspiring talks, or having your presentation welcomed with a barrage of (of course, constructive!) questions, wishing if only you managed to take record of all these during the event? We are developing the *Hive Open Research Network*¹, a social platform for fostering scientific interactions and reducing friction in scientific exchanges and the underlying integrated services supporting content personalization, preview, and social/scientific recommendations. *Hive* is a conference-centric, but cross-conference platform, where researchers can seed and expand their research networks, keep track of the technical research sessions they are attending, meet new colleagues, share their ideas, ask questions, give and receive comments, or simply keep and/or view records of interactions at a conference they have attended (or wanted to attend, but missed due to other commitments). In its core, *Hive* leverages dynamically evolving knowledge structures, including user connections, concept maps, co-authorship networks, content from papers and presentations, and contextual knowledge to create and to promote networks of peers. These peer networks support each other explicitly through direct communication or indirectly through collaborative filtering. *Hive* provides the following online integrated services: a) understanding the personal activity context through access patterns and analysis of user supplied content, b) context-aware resource discovery, including search, presentation, and exploration support within the scientific knowledge structures, and c) peer discovery, and peer driven resource and knowledge sharing and collaborative recommendations.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information filtering;
H.3.4 [Systems and Software]: User profiles and alert services;
H.3.5 [Online Information Services]: Data sharing, Web-based services; H.3.7 [Digital Libraries]: Collection, Dissemination, Systems issues, User issues; H.3.7 [Group and Organization Interfaces]: Asynchronous interaction, Web-based interaction;

General Terms

Algorithms, Human Factors

Keywords

Social networks, Scientific networks, Peer discovery, Recommendation, Collaborative filtering, Community discovery

¹ This work is supported by NSF Grant 1043583 “MiNC: NSDL Middleware for Network- and Context-aware Recommendations”.

Copyright is held by the authors.
EDBT/ICDT '13, Mar 18-22 2013, Genoa, Italy
ACM 978-1-4503-1597-5/13/03.



Figure 1. A screen shot from the MM'11 edition of the *Hive* open research network

1. INTRODUCTION

We demonstrate the *Hive Open Research Network* (Figure 1), a social platform for fostering scientific interactions and reducing friction in scientific exchanges and the underlying integrated knowledge services for supporting content personalization, preview, and social/scientific recommendations.

Hive is a conference-centric, yet cross-conference platform, where researchers can seed and expand their research networks, keep track of the technical research sessions they are attending, meet new colleagues, share their ideas, ask questions, give and receive comments, or simply keep and/or view records of interactions at a conference they have attended (or wanted to attend, but missed due to other commitments).

Different from social networks, such as Facebook [2], *Hive* focuses on professional research networks. In contrast to professional social networks, such as LinkedIn [1], on the other hand, it is (a) event centric, (b) content (presentations, data, papers, posters) rich, and (c) research oriented. *Hive* also differs from other conference services, such as Pathable [4] and Iamresearcher [3], in many ways. Pathable focuses on a given event and tries to make the physical scheduling at a conference more convenient. Iamresearcher targets individual researchers or conference organizers and helps them create research oriented

websites easily and conveniently. *Hive*, on the other hand, is not event specific and focuses on peer network discovery and management, as well as information sharing and dissemination among researchers within and across conference events.

Concept map and personalization services	<ul style="list-style-type: none"> Learn key concepts to bootstrap concept map from a given set of contextually-relevant documents
Peer network services	<ul style="list-style-type: none"> Select peer network Locate similar peers (subject to peer's preferences) Send request/reply to peers
Discovery, context- and collaborative e-recommendation and preview services	<ul style="list-style-type: none"> Search peers and resources based on context; <ul style="list-style-type: none"> alternative context representations that the user can specify include concept maps, access history, and knowledge from peers (collaborative filtering) Rank peers and resources based on context Request resource recommendations based on context Relationship discovery and explanation among peers and other resources Community discovery and tracking Generate summary previews and highlights for updates and resources based on based on context
Personal activity history services	<ul style="list-style-type: none"> Search and visualize personal, group, or community activity history based on current context

Table 1. List of sample *Hive* services

1.1 Use Scenario

A second year PhD student, Zach, is attending the EDBT 2013 conference to present a paper. He had also attended the last year's edition of the same conference, where his advisor and groupmates had a paper, as well as the ACM Sigmod 2012 conference, where he has published his first paper on social media. In both conferences, he had met and connected with researchers who were working in topics that were of interest to him.

- Before leaving for EDBT'13, Zach logs in to *Hive* and uploads his presentation slides; he also takes a quick look at the already uploaded presentation slides of a few relevant papers he had noticed in the published program. Zach also checks which of the researchers who he had connected in the past are coming to EDBT'13. In addition, *Hive* proposes five other researchers that Zach may want to connect during the event and for each provides a list of sessions that the researcher may most likely attend. Zach highlights the set of researchers whose (session check-in, question, comment, answer) activities he would like to follow and instructs *Hive* to provide real-time updates regarding these during the conference.
- While Zach is checking the list of presentations in his session, *Hive* reminds Zach that the chair of his session is one of the authors whose paper he had cited in his Sigmod12 paper and that one of the other authors in the same session was a co-author with his advisor a few years back. Zach decides to follow these as well. In addition, he places the avatars of these two researchers into his "session" workpad for quick access. This workpad serves both as a bookmark

and, when selected to be the active workpad, also gives context to other recommendations *Hive* provides to Zach.

- In the first day's keynote presentation, Zach uses *Hive* to follow the comment, question, and answer traffic, including the related twitter activity. He notices that one of the questions posted on *Hive* about the presentation is in fact in his own research area and shares his opinion on the topic.
- The next sessions in the conference are not in areas of direct interest to Zach. But, Zach notices that a few of the researchers he is following are checking-in into a session on large scale graph processing. He also notices that a paper in the session has cited a few of the papers he had cited in his own work and that the author of the paper is a frequent co-author of one of the researchers in his workpad. So Zach decides to attend this session.
- The presentation indeed raises his curiosity and he finds himself posting a few questions about the details not clarified in the presentation. A few minutes later, one of the co-authors of the presenter, who happens to follow the discussion remotely, confirms his opinion. While this exchange occurs in *Hive*, the exchange is also broadcasted in twitter with the session's hashtag. Zach adds this presentation, the presenter's avatar, and the link to the paper into his "to investigate later" workpad. Zach also happily notices that a few of the attendees of the session decide to "follow" his questions.
- In the break, Zach receives an update from *Hive* reporting that there is already a question posted regarding the presentation he had uploaded a few days back. He notices that one of the researchers who were following him, named Aaron, questions an equation he has included in one of his slides. Indeed, when he checks the slides he notices that there was a typo and he corrects the slide. Zach also sends a thank you note to the researcher who highlighted the error in the slide. After a brief exchange in which he learns that Aaron is also working on a related problem, Zach sends a connection request to Aaron and receives an acknowledgement a few minutes later.
- The next day, Zach presents his paper in an afternoon session. After his presentation, the session chair takes a few questions from the audience and then reads a few questions posted on *Hive*, some of which through Twitter. Zach answers to the questions as best as he could within the given time limit. He then makes a note to answer these questions on *Hive* in more detail after his presentation.
- After his presentation, he notices that one of the audience members, named Ann, who has asked questions to him had a related paper in EDBT 2010, which he had cited and another paper accepted to SIGMOD'13, which has not been published yet. He also notes that she has check-in in a session on community detection. He adds Ann's avatar to his workpad and then goes to the session on community detection.
- Once he is back to the University, his advisor (who had missed the conference due to another commitment) and Zach spend some time in their first meeting, discussing his activities and the connections he has made in EDBT'13.

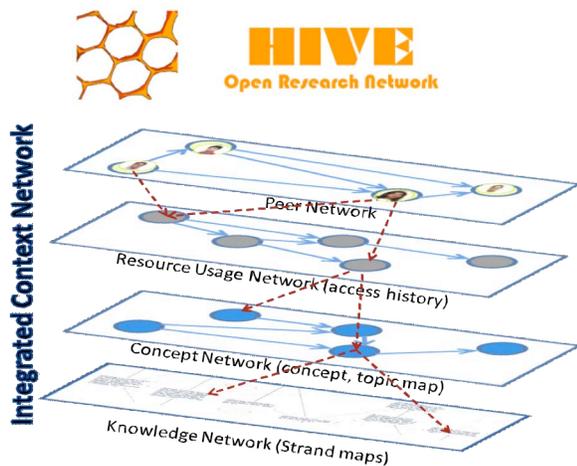


Figure 3. Multiple integrated layers of the dynamic *Hive* knowledge network

A preliminary version of *Hive* has been made available to the attendees of the ACM Multimedia 2011 conference in November 2011 (this version of *Hive* is available to the public at <http://hive.asu.edu>), and of ACM Sigmod 2012 conference in May 2012 (<http://hive.asu.edu/sigmod12>). *Hive* is also being used as a learning support platform for the CS515 (graduate level) and CS408 (senior level) courses at ASU, to encourage students' collaboration and studying material exchange.

2. OVERVIEW of HIVE

Basic functionalities of *Hive* are built using JomSocial [5], a Joomla-based platform for building social networks. These basic functionalities include many of the Facebook like social features.

Unlike other JomSocial-based social networks, however, *Hive* also provides various knowledge rich services and functionalities. In particular, we note that accessing scientific content effectively requires a proper understanding of the personal activity context, context-aware resource discovery, and peer-network driven resource and knowledge sharing and collaborative recommendations. Therefore, in its core, *Hive* leverages dynamically evolving knowledge structures, including concept maps, co-authorship networks, content from papers and presentation, and contextual knowledge to create and to promote networks of peers (**Figure 3**).

These peer networks support each other explicitly through direct communication or indirectly through collaborative filtering. *Hive* provides the following online integrated services:

- understanding the personal activity context through access patterns and analysis of user supplied content,
- context-aware resource discovery, including search, presentation, and exploration support within the scientific knowledge structures, and
- peer discovery, and peer driven resource and knowledge sharing and collaborative recommendations.

In particular, *Hive* uses the following evidences for discovering and explaining relationships between individuals (peers) and for recommending new peers or resources [6][7] (**Figure 2**):

- profile and declared interest,
- current and past affiliation, group membership,

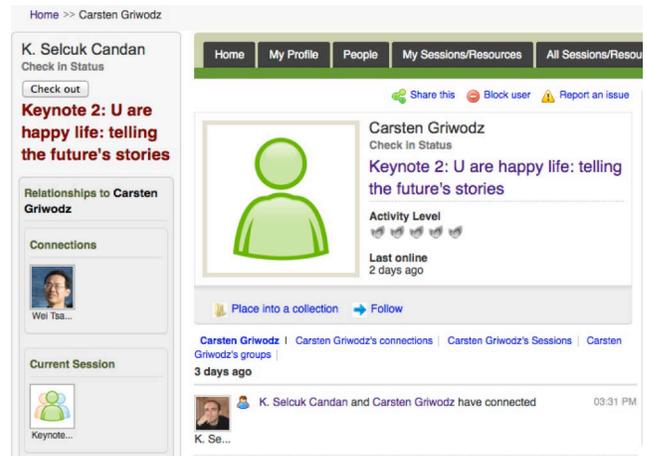


Figure 2. Relationships between the users “K. Selcuk Candan” and “Carsten Griwodz” are shown on the right column

- co-authorship, direct citation, or indirect citation (e.g., citing the same paper or transitive citation),
- online “follow”ing,
- conference participation (related conferences, same conference different years, same conference same year),
- session participation/check-in (related sessions or same session/same time),
- reciprocal question, comment, and answer activities,
- user-provided content (publication, presentation, other supporting material) similarity, and
- activity (e.g. browsing, commenting) similarity.

These relationships are contextualized based on user's own activity history, including the current session s/he is checked in as well as the content of his or her active workpad (**Figure 4**). The workpad interface is a tool to help the user keep record of the things that attract his or her interest in the conference. When the user sees a presentation that he/she likes, a question that tickles his/her mind, or a colleague that she wants to connect later, she can simply drag-and-drop the corresponding avatar into the current workpad. The user can also name and save workpads and can choose from different saved workpads, each corresponding to a different context or state of mind. The content of the currently active workpad defines the user's activity context and all the searches and recommendations are contextualized according to this active workpad. The user can export workpads as collections accessible to others or import a collection as active workpad.

For these knowledge management, content annotation, and peer recommendation services, *Hive* relies on our Middleware for Network- and Context-aware Recommendations (MiNC) engine, which provides services that help minimize the extraneous load on users, while they search, share, and access digital resources and peer networks [8].

2.1 Understanding and Representing the Personal Activity Context

The domain knowledge captured by the usage context includes concepts, their significance in the learning domain, as well as the strength of the inter-relationships between concepts and resources [9]. To capture the structure of a knowledge domain as well as the relationships between domain concepts, we leverage general purpose ontologies as well as domain specific concept maps.

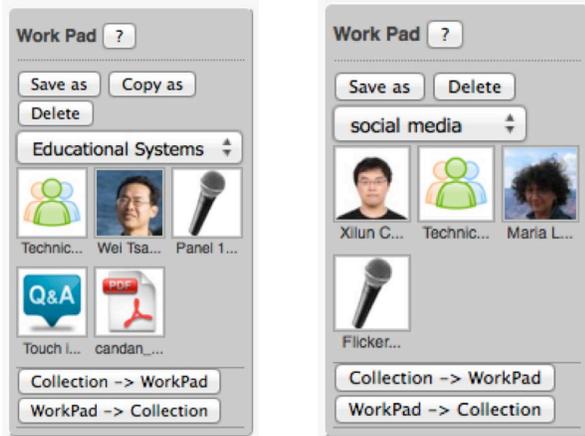


Figure 4. Two different work pads of the same user: the work pads can contain many different types of resources, including avatars of the users, presentations, sessions, questions and answers, documents, and collections. The work pads serve both as a book mark as well as context for search and recommendations

To support services where the activity context is determined by external materials, we apply novel concept map bootstrapping algorithms that rely on user highlights, bookmarks, notes, or documents. These algorithms, including [10], extract, in a semi-automated manner, dominant concepts and their relationships specific to a given material.

2.2 Network Layer Alignment and Integration

Hive uses the multiple context layers of the “context network”, shown in **Figure 3**, in an integrated manner search and recommendation. Integration of layers starts with an alignment phase, which requires identification of mappings between concepts and relationships among different layers. In *Hive*, since the original layers are likely to match partially and since layers can conflict or reinforce each other, the result of the alignment process is imprecise. For example, structures of the underlying concept-maps and concept significances may differ from one context layer to the other. The *Hive* services that rely on these layers function in the presence of such imperfect alignments. For weighted graph data management, *Hive* relies on our R2DB, weighted RDF data management system [11][12].

2.3 Context-Aware Resource Discovery, Search, and Exploration

Hive relies on the underlying integrated context network to filter, summarize, and rank alternatives and adapt according to their relevance. Context-aware ranking and preview services include (a) relevant snippet extraction from documents, (b) key concept extraction for automated annotations, and (c) content summarization documents and update reports. Once all the concepts are extracted and ranked (based on the context), *Hive* propagates the concepts within the relevant neighborhoods of the knowledge network using adaptation strategies, based on the current active context (defined by the workpad). These annotations are then used for ranking, recommendations, and summarization tasks. Summarization of the scheduled update reports are performed relying on hierarchical table summarization

techniques, which preserve maximal information while minimizing the footprint of the reported information [13], and context-aware snippet extraction algorithms [14].

2.4 Dynamic Peer-Networking and Collaborative Recommendation Support

Hive provides peer-network services in two ways: (a) peer recommendation, where the system locates other peers with similar interests or activity contexts and (b) peer-network based resource recommendation. We have developed SCENT, an innovative, scalable spectral analysis framework for internet scale monitoring of multi-relational social media data, encoded in the form of tensor streams. SCENT focuses on the computational cost of structural change detection in tensor streams and extends compressed sensing (CS) to tensor data. Through the use of randomized tensor ensembles, SCENT is able to encode the observed tensor streams in the form of compact descriptors and detect significant changes in the underlying structure faster and more accurately than the other methods 0.

3. CONCLUSIONS

In this paper, we demonstrate *Hive*, a social platform for fostering scientific interactions and improving scientific exchanges. *Hive* is a conference-centric, but cross-conference platform, which leverages dynamically evolving knowledge structures, including user connections, concept maps, co-authorship networks, content from papers and presentation, and contextual knowledge to create and to promote networks of peers.

4. ACKNOWLEDGMENTS

We thank Mijung Kim and Profs. Hari Sundaram and Hasan Davulcu for their inputs on *Hive* and the underlying *Middleware for Network- and Context-aware Recommendations (MiNC)*.

5. REFERENCES

- [1] <http://www.linkedin.com>
- [2] <http://www.facebook.com/>
- [3] <http://www.iamresearcher.com>
- [4] <http://www.pathable.com/>
- [5] <http://www.jomsocial.com/>
- [6] J.H. Kim, K.S. Candan, and M. L. Sapino. Impact Neighborhood Indexing (INI) in Diffusion Graphs. Proc. CIKM'12.
- [7] K.S. Candan, W.-S. Li: Reasoning for Web Document Associations and its Applications in Site Map Construction. DKE 43(2), 2002.
- [8] <http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=1043583>
- [9] J. W.Kim, K. S. Candan, J. Tatemura. Efficient Overlap and Content Reuse Detection in Blogs and Online News Articles. WWW'09
- [10] L. Di Caro, K. S. Candan, and M. L. Sapino. Navigating within News Collections using Tag-Flakes. J. Vis. Lang. and Comp., 2011.
- [11] J. P. Cedeño and K. S. Candan. R2DF framework for ranked path queries over weighted RDF graphs. Proc.WIMS, 2011.
- [12] S. Huang, X. Li, J. P. Cedeño, K. Se. Candan, M. L. Sapino, Representing and Querying Weighted Semantic Web with R2DB. Under review for demonstration in EDBT'13
- [13] K. S. Candan, H. Cao, Yan Qi, M. L. Sapino: AlphaSum: size-constrained table summarization using value lattices. EDBT'09.
- [14] Q. Li, K. S. Candan, Y. Qi. Extracting Relevant Snippets for Web Navigation. AAAI 2008.
- [15] Yu-Ru Lin, K. S. Candan, H. Sundaram, L. Xie. SCENT: Scalable Compressed Monitoring of Evolving Multi-Relational Social Networks. ACM TOMCCAP, 7:29, 2011