Mobilizing Search of the Here and Now

Jonas Michel

The University of Texas at Austin, Department of Electrical and Computer Engineering
The Center for Advanced Research in Software Engineering
Email: jonasrmichel@mail.utexas.edu

TR-ARiSE-2011-013

Mobilizing Search of the Here and Now

Jonas Michel
Department of Electrical and Computer Engineering
University of Texas at Austin
Email: jonasrmichel@mail.utexas.edu

***Abstract*—The escalating number of digitally-accessible devices pervading our everyday environments gives rise to the availability of tremendous amounts of human- and device-generated data. This data possesses strong spatial and temporal semantics, it captures phenomena and states of the environment, and is extremely volatile, being created, moved, stored, and deleted on-demand at rapid rates. The requisite support for general-purpose expressive search of the "here" and "now" has eluded realization due to the complexities of indexing, storing, and retrieving relevant information within a vast collection of highly ephemeral data. We aim to address this gap in the research (i) through search mechanisms that are sensitive to the bearings of space and time on the relevance of search results and (ii) more fundamentally, through the design of a general-purpose data model that facilitates information availability in pervasive computing environments and exposes spatial and temporal relationships between digital data and physical phenomena. Finally, this paper outlines simulation and deployment evaluations of the proposed research.**

I. Introduction

The density of digital devices woven into the fabric of our daily lives is rapidly increasing. An escalating presence of technologies like mobile devices, sensor networks, smart objects [1], and RFID tags in our everyday spaces gives rise to the availability of unprecedented amounts of digitally-accessible sensory, contextual, and user data. This data is *ephemeral*, representing changing real-world conditions of the environment, social interactions among humans, etc. Envisaged pervasive computing environments of the near future will enable opportunistic wireless access to nearby resources and services embedded in the environment. As in the Internet, the high volumes of data in such *Personalized Networked Spaces* (PNetS) motivate the need for mechanisms that help a human user *search* for the information she needs *here* and *now* as she moves through such densely populated and rapidly changing information spaces.

Human users in PNetS need information that is immediate and localized. This tight integration of the user and her immediate surroundings introduces novel search requirements. Using an Internet search engine, one may find timetables of trains from Zurich; however, when hurrying to board a crowded train, one may need to search for the closest available second class seat. When planning a trip to an amusement park, one may use the Internet to search for directions, hours of operation, reviews, etc. While at the park, a visitor may wish to know which rides her friends are on, where the closest bench in the shade is, the closest ride with a short wait time, etc. These situations demand expressive capabilities for searching the user's *here* and *now*. We argue that such information is most effectively collected on-demand, directly from the environment. In other words, search *of* the here and now must be performed *in* the here and now.

Developing the requisite support for general-purpose and expressive search over highly dynamic data in a distributed fashion with limited resources unearths novel challenges. Large volumes of extremely volatile data coupled with the dynamic and heterogeneous nature of PNetS precludes traditional information retrieval techniques. Moreover, a user's highly spatially- and temporally-dependent information needs demand more expressive compositions of relevance. In this paper, we describe research to address these imminent challenges, first, through the development of search mechanisms for PNetS and second, with a novel general-purpose data model for pervasive computing applications. Finally, we outline an evaluation plan to quantitatively measure the performance of our mechanisms and qualitatively evaluate their utility.

II. Search of the Here and Now

We believe developing the requisite support for performing search within pervasive computing environments calls for a new paradigm of information retrieval. This work aims to provide expressive mechanisms in this new paradigm. Furthermore, we argue that spatiotemporal information needs are present in *every* pervasive computing application. Therefore, our work also aims to design a general-purpose formal data model for pervasive computing applications that explicitly tags information with spatial and temporal semantics.

A. Search Mechanisms

Developing mechanisms for performing search directly within unpredictable and highly dynamic data-rich environments calls for a new paradigm of search that explicitly separates search from advanced indexing of data. This radical shift is motivated by three factors: (i) Searchable data in PNetS is often supplied by human users, who are likely unwilling to publicly share information (i.e., on the Internet), but willing to share with other nearby, potentially unknown, users [2]; (ii) Large volumes of transient data cannot easily be centrally

This material is based upon work supported by the National Science Foundation under Grant No. 844850. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

indexed and associated with a relative spatiotemporal context, as is traditionally done in Internet [3] and Geographic Information Retrieval [7], [9]; (iii) Devices that comprise PNetS may be far from reliable connectivity, and connectivity to the Internet may be sparse or unattainable; therefore, resources cannot be polled and statistically ranked [8]. These factors necessitate a decentralized and scalable approach to localized search.

To this end, we have developed *Gander*, a search engine for PNetS [4], and *myGander*, a prototype user interface and mobile application for Gander [5]. Gander is founded on a novel conceptual model that does not index data relative to its context, but instead performs a query *in* the context. Gander's query processing uses mobile ad hoc network query protocols as spatial sampling strategies to incrementally construct and adjust a result set. Acquired results are dynamically ranked using multidimensional combinations of relevance metrics.

B. A Pervasive Computing Data Model

A key challenge in pervasive computing is capturing and managing information subject to the high levels of dynamics in reality. In this vein, we believe that *every* application targeting pervasive computing environments possesses strong spatiotemporal underpinnings, not strictly search. The abundance of existing work in pervasive computing focuses almost exclusively on how to *query* data (e.g., through query abstractions and publish/subscribe mechanisms), neglecting questions related to how data is created, moved, stored, and destroyed. We believe these shared concerns must be extrapolated into a cohesive data model to facilitate composition, reduce developer responsibilities, and enable data to exploit contextual dependencies.

In this light, we have proposed a straw-man data model for pervasive computing applications [6]. The crux of our data model is the notion of *spatiotemporal trajectories*, which enable applications to reason about relationships between spatiotemporal phenomena and the digital data that represents those phenomena as they both evolve over time. For example, a digital *datum*, representing air quality (AQI) at a particular place and time as sensed by a user's smartphone, moves in physical space as it is carried by the user and may also move by being shared with other devices. The datum's spatiotemporal trajectory records these dynamics of the observation (the sensor reading) and the physical phenomenon (the AQI), allowing applications to create rules governing how datums capturing the phenomenon should move, change, live, and die. A rule may, for example, degrade datums that are further in space and time from their points and times of origin.

III. Evaluation

To validate our approaches and demonstrate their potential impact, we will conduct two tracks of evaluation in parallel: (i) performance evaluations in simulated PNetS and (ii) real-world deployments of our mechanisms and data model in mobile applications.

We will perform quantifiable measurements within simulated large-scale PNetS. Using a network simulation framework like OMNeT++ [10] we may create real-world pervasive computing scenarios and evaluate our mechanisms' performance (e.g., in terms of query overhead, latency, memory and battery requirements, etc.) against Internet (i.e., centralized) search. Furthermore, simulation-based evaluations enable a global view of PNetS to be maintained. This global view provides a baseline against which to quantifiably measure the *quality* of results returned by our search mechanisms. Insights from our simulations will drive implementation of these mechanisms as general-purpose middleware.

To demonstrate the real-world impact of our approaches, we will implement our mechanisms and data model as middleware components to be realized in mobile applications. Mobile devices provide a powerful sensing, storage, and communication medium though which human users may be inserted to close the feedback loop. As an example, we will develop a search application for use on a university campus during crowded events. Students and visitors may use the application to find nearby friends, read live reviews about demonstrations, and obtain information about nearby events.

IV. Conclusions

The rapidly increasing density of wireless and smart technologies pervading our daily environments yields a rich and highly dynamic information space extending our physical surroundings. We argue that novel requirements in this search space necessitate a new paradigm of information retrieval for pervasive computing. This paper identifies the requirements for search mechanisms in this new paradigm and proposes work to investigate the beginnings of such mechanisms. Complementary to these goals, we further propose research to design a general-purpose data model for pervasive computing to facilitate data visibility, an appropriate separation of concerns, and expressive spatiotemporal reasoning. Finally, we outline a methodology for evaluating our approaches through simulated scenarios and real-world deployments.

V. Acknowledgement

I would like to thank my advisor, Christine Julien, for her invaluable guidance and support and Jamie Payton and Gruia-Catalin Roman for their continued collaborative contributions.

References

[1] L. Atzori, A. Iera, and G. Morabito. The internet of things: A survey. *Computer Net.*, 54(15):2787–2805, 2010.

[2] Q. Jones, S. Grandhi, S. Karam, S. Whittaker, C. Zhou, , and L. Terveen. Geographic place and community information preferences. *CSCW*, 17(2–3):137–167, 2008.

[3] C. D. Manning, P. Raghavan, and H. Schutze. *Intro. to Information Retrieval*. 2009.

[4] J. Michel, C. Julien, J. Payton, and G.-C. Roman. Gander: Personalizing search of the here and now. In *MobiQuitous*, 2011.

[5] J. Michel, C. Julien, J. Payton, and G.-C. Roman. The gander search engine for personalized networked spaces. Technical Report TR-ARiSE-2011-008, The University of Texas at Austin, 2011.

[6] J. Michel, C. Julien, J. Payton, and G.-C. Roman. A spatiotemporal model for ephemeral data in pervasive computing environments. Technical Report TR-ARiSE-2011-012, The University of Texas at Austin, 2011.

[7] D. Mountain and A. MacFarlane. Geographic information retrieval in a mobile environment: evaluating the needs of mobile individuals. *J. of Info. Science*, 33:515–530, 2007.
[8] B. Ostermaier, K. Römer, F. Mattern, M. Fahrmair, and W. Kellerer. A real-time search engine for the web of things. In *IOT*, 2010.
[9] T. Reichenbacher. Geographic relevance in mobile services. In *LocWeb*, 2009.
[10] A. Vargas. OMNeT++ Web Page. http://www.omnetpp.org, 2009.