Multimedia Search over Integrated Social and Sensor Networks

www.smartfp7.eu

John Soldatos
Athens Information Technology
PO Box 68, Markopoulo Av.
Peania 19002, Athens, Greece
jsol@ait.edu.gr

Moez Draief
Imperial College London
Intelligent Systems and Networks
London SW7 2AZ, UK
m.draief@imperial.ac.uk

Craig Macdonald & Iadh Ounis
University of Glasgow
School of Computing Science
Glasgow G12 8QQ, UK
first.last@glasgow.ac.uk

ABSTRACT

This paper presents work in progress within the FP7 EU-funded project SMART to develop a multimedia search engine over content and information stemming from the physical world, as derived through visual, acoustic and other sensors. Among the unique features of the search engine is its ability to respond to social queries, through integrating social networks with sensor networks. Motivated by this innovation, the paper presents and discusses the state-of-the-art in participatory sensing and other technologies blending social and sensor networks.

Categories and Subject Descriptors: H.3.3 [Information Storage & Retrieval]: Information Search & Retrieval

Keywords: Search Engine, Multimedia, Sensors

1. INTRODUCTION

The Future Internet will include a proliferating number of internet-connected sensors, including cameras and microphone arrays. Based on these sensors, emerging applications will be able to collect, filter, analyze and store large amounts of data captured from the physical world, as well as related metadata captured as part of perceptive multimedia signal processing algorithms. The ability to search this information in a scalable, effective, real-time and intelligent way can empower a wide range of added-value applications in the areas of security/surveillance, smart cities, social networking, e-science and more. The potential is partly manifested in the recent wave of participatory sensing and crowd-sourcing applications [10, 13, 17]. Nevertheless, the vast majority of crowd-sourcing and participatory sensing applications deals with non-AV (Audio/Visual) data and do not provide capabilities for searching and processing multimedia data. Moreover, tools and techniques for searching sensor data [4, 6] are still largely based on the indexing and searching of apriori defined (and usually textual) metadata. Indeed, while they exploit recent advances on sensor ontologies [15] in order to decouple the queries from the low level details of the underlying sensors, they cannot dynamically identify the appropriate sensors for answering queries according to the context of the user and the application domain.

WWW '12, Lyon, France
ACM 978-1-4503-1230-1/12/04.

As a result, there is a pressing need for a next generation multimedia search engine, which will be effective not just for textual search, but also for searching multimedia data (notably audio and video streams) derived from the physical world (i.e. environment generated media and content). This search engine should be able to ask queries to sensors that provide multimedia streams such as cameras and microphone arrays. In order to effectively access and use multimedia data, the search engine should automatically match the application context with sensors and sensor processing algorithms that are appropriate for answering the query at hand. To this end, cutting edge sensor processing algorithms (notably audio and video processing algorithms) can be exploited in order to allow the real-time matching of multimedia data to a given search query or application context.

In addition to integrating multimedia signal processing algorithms, multimedia search engines for environment generated content should be able to support ambient/intelligent synthesis of related content in real-time (i.e. using a low-latency distributed framework). This synthesis could be based on the combination of related environment generated content and metadata, which pertain to the same context [9]. To this end, new indexing and retrieval architectures for multimedia data are required. Another challenge is the need to deal with the dynamism of the sensor web environment, where sensors, multimedia data and application contexts may change dynamically (e.g. in the scope of mobile contexts). Lastly, multimedia search engines for environment generated content must blend with the existing and emerging wave of social networks, in order to enable social queries over sensor networks [2].

The structure of the paper is as follows: Section 2 introduces participatory and environmental sensing that form the SMART (Search engine for MultimediA enviRonment generated contentT) EU project. Section 3 discusses the representation of information from sensor and social networks within SMART. The main technical and technological characteristics of the SMART search engine are presented in Section 4. Furthermore, Section 5 outlines use cases for validating the SMART concept, while Section 6 concludes the paper.

2. PARTICIPATORY & ENVIRONMENTAL SENSING

The SMART search framework will enable the implementation of search services over large scale community environmental and participatory sensing infrastructures, which have recently attracted the interest of cities, communities and individuals. In particular, participatory sensing describes the use of individuals and communities to gather information about their environment. It usually leverages the ubiquity of smart phones as sensing devices, of cloud based services for big data analysis, resource discovery and application delivery, while anticipating the trend towards more powerful sensing and processing capabilities of mobile devices and social networking sites.

Participatory sensing applications tend to be single focus, vertically integrated applications. Users can semantically enrich low level sensor data with high level context capturing qualitative information such as the level of rubbish on a street or their sentiments about a particular topic. For example, Foursquare [10] combines smart phone-based GPS with social networking and mapping services to enable users to track friends, and comment on and describe features of interest. However, the nature of such applications precludes later repurposing and goal adaptation.

AnonySense [5] is a framework with privacy protecting measures that enable users to opportunistically contribute data in a reliable and anonymised fashion. The CenceMe project [12] integrates with social networking and only allows sharing of data within defined user groups. PEIR [13] has developed functionality to enable users to introduce noise or realistic synthetic data when they do not wish to share data. Data quality is also an issue due to low quality and miscalibrated sensors and the intrinsic heterogeneity of users.

Fixed-location sensors monitoring environment aspects (e.g. visual, audio, temperature, pollution, windspeed) are being increasingly deployed within communities and cities. For example, the (separate) SmartSantander project [1] envisages a 'smart' city by providing end users with information about environmental aspects such as temperature, or the usage of car parking spaces.

The above demonstrate the strong interest in participatory and environment sensing services, which is empowered by the proliferation of sensor-enabled mobile devices and user generated content. These services underpin the feasibility of the vision of the SMART project, since they provide proof-of-concept implementations for distributed sensor-based search services. Furthermore, they provide ideas for exploiting user generated content (e.g. user annotations) in the SMART platform, which will extend them by (a) exploiting multimedia content derived from the physical world; and (b) providing applications that span several contexts through general technology. SMART will encapsulate both of these features within its planned proof-of-concept applications.

3. MODELLING SENSOR & SOCIAL NETWORKS

SMART aims to combine sensor networks information with social networks information in order to answer sensor based queries in a more social, useful and accurate way. Indeed, information from social networks can be used to enhance the end-users' context and overall understand the context of the query in a much better way. Social networks information can be used to adapt a query for environment generated context to the end-user's daily life. The concept is quite new, yet some motivating use cases have been discussed in [3] and include (a) identifying social acquaintances in localized areas, (b) social sensing based on noise log analysis and (c) improving daily living and health for the elderly. There is a mutual benefit from the convergence of both sensor networks and social networks. Social networks can benefit from the fact that human activity and intent can be directly derived from sensors, which obviates the needs for explicit use input. On the other hand, sensor societies could start their collaboration in a social way (i.e. based on information derived from social networks). However, even though the potential of integrating social networks with sensor networks has been identified, only a few applications exist thus far.

Furthermore, there is a lack of a disciplined framework for modelling and blending sensor information with social information, including a way to deal with privacy and trust issues [8]. Indeed, the specification and modelling of information to be captured, structured and later searched by the SMART search engine is a key concern for the project. The specification depends on the type of applications to be supported, but also on the capabilities of the underlying multimedia processing algorithms, and may include: metadata on the sensor itself (type, model, orientation capabilities); processing output and events (for instance, the identification of a crowd at a given location by a processing algorithm); and social networks information (e.g. current trending topics in a given area). To build a generic search engine for environment generated content, this information should be described in standard and extensible formats, such as SensorML, MPEG-7 and RDF. Deployers – such as smart cities – of SMART sensors and processing components will model their data in such formats at *edge servers*, which are are then indexed by the SMART search engine.

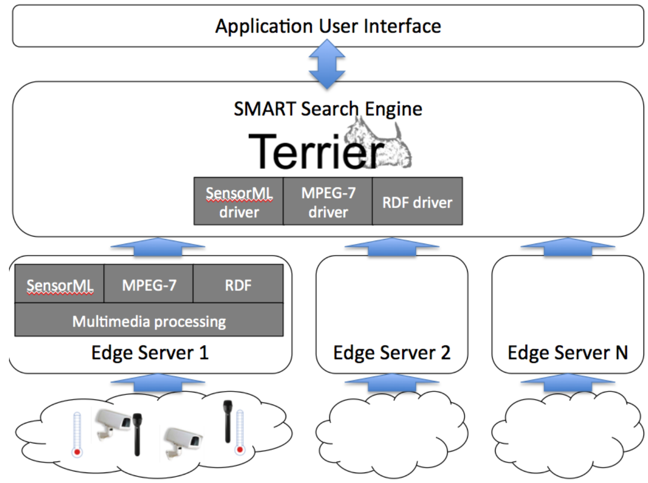

Figure 1: Architecture of the SMART search engine.

4. SEARCH ENGINE CHARACTERISTICS

Figure 1 depicts a high-level overview of the SMART search engine architecture. At the lowest level of the architecture there is a cloud of sensors that provide the physical world data. This data is processed by a number of multimedia processing algorithms, such as video scene analysis, crowd analysis, acoustic event classification and speech processing. Sensor information is collected and processed by edge servers, which constitute "points-of-presence" of the SMART system. Contextual information derived from multimedia processing (at the edge servers) are appropriately represented (SensorML, RDF, etc). The information held in knowledge bases is traversed in order to perform sensor selection, as well as accessing multimedia content pertaining to the end-user's query.

The upper part of the engine performs conventional processing of user queries, which includes extracting query terms ("semantics") and using them for the sensor selection and the multimedia content fetching process. The query processing process will be empowered by the SMART multimedia indexing architecture, which will extend the indexing and retrieval architecture used by the Terrier search engine. A detailed description of the full range of technologies comprising the SMART engine is beyond the scope of this paper.

Targeting the innovations listed above, SMART will work towards a multimedia search framework with the following technical characteristics:

Open and Open Source. SMART is designed as an open framework, which is extensible in terms of sensors, ontologies and semantic structures, as well as multimedia processing components (notably video and audio processing algorithms). The main components of the SMART engine will be implemented as open source software building upon the Terrier search engine [16][1]. SMART will form an open source community for sustaining and evolving its technological results.

Multimedia. The SMART search engine will enable query answering based on the real-time processing of multimedia data stemming from the physical environment. Cutting edge multimedia processing components will be researched and adapted, notably in the areas of acoustic event classification and visual scene analysis.

Participatory and Reusable. The very same sensor and multimedia processing algorithms will be able to contribute to multiple concurrent queries of the SMART system. Participatory sensing schemes will be researched along with ways of caching data and queries, while also dealing with mobility and sharing application contexts. Furthermore, a number of Web2.0/Web3.0 [7] mashups will be implemented to allow reuse of sensor queries across multiple applications and searches.

Recent. Input from sensors and multimedia processing algorithms will be collated and indexed in real-time by the SMART search engine. For this purposes, next-generation low-latency indexing architectures will be examined, such as the distributed S4 framework developed by Yahoo! [14].

Relevant. The SMART search engine will deploy state-of-the-art search technologies (e.g. learning to rank [11]), using features identified by multimedia processing algorithms, to identify results from multitudes of sensors, that are relevant to the user at that point in time, at that location.

Smart and anticipatory. Based on machine learning, SMART will be able to anticipate the answers to queries before they are entered. This will empower a level of intelligence, beyond self-learning and ranking algorithms used by existing search engines.

Social. The SMART search engine will seamlessly leverage information and search results from (Web2.0) social networks in order to facilitate the interception of social networks with sensor networks, towards social applications and searches of environment generated content.

Scalable and Dynamic. SMART is designed to be 'massively' scalable. Hence, the project will research a scalable architecture for collecting, searching, caching and combining sensor data in a highly heterogeneous and distributed environment, to dynamically provide the most recent information sensed by the underlying sensor networks.

Context-aware. SMART enables the context-aware orchestration of sensor data and metadata towards accessing data that pertain to a given context. Metadata associated with time, space, location, goals, tasks and more will be used to identify the contribution of a sensor to a particular query. To this end, the project researches sensor selection protocols/algorithms, along with collaborative protocols enabling the orchestration of sensors towards a joint task.

5. USE CASE APPLICATIONS

The SMART project deploys two use cases, namely "live news" and "security and surveillance in urban environments". Each use case will form complete applications comprising multiple configurable and dynamic queries over the SMART engine. They will combine multiple search queries into composite applications.

The live news use case is motivated by the fact that timely sensor-based access to information in the urban environment can be particularly important for news agencies. Hence, a news agency could ask the SMART search engine questions regarding the occurrence and evolution of certain events, i.e. "What is happening now?", "Which places are crowded?", "What are the specific trends in the city?", "Where are riots and fights happening?" and more. The answers to these queries will be provide in the form of multimedia streams mixing multimedia data acquired from the physical world (i.e. sounds/images) along with textual data stemming from sensors and metadata steams (including social networks). Using the presentation layer capabilities of the SMART framework (e.g. reusable Web2.0/Web3.0 mashups) news providers could build, integrate and populate web sites, wikis, blogs or news portals with news/information stemming from the underlying sensing infrastructure. The live news use case will enable end-users to create personalized social news portals containing dynamic life information from sensors deployed within a smart city. End-users will be able to assemble a dynamic news portal, based on a set of queries to the SMART engine and associated mashup components for visualizing them on the portal. Live news applications could typically include correlated queries under a thematic umbrella.

The security and surveillance in urban environments use case is motivated by a number of witnessed tremendous terrorist attacks in urban environments during the last decade (e.g. the collapse of New York's Twin Towers (2001), the bombing of packed commuter trains in Madrid (2004), as well as the London bombings (2005)). These events have led modern cities to deploy numerous sensors (notably cameras) for security purposes. Nevertheless, in such sensor saturated urban environments, it has become difficult to manually observe the sensor streams. SMART can offer a viable and cost effective alternative through enabling the answering of targeted queries, based on sensors and sensor processing algorithms that fulfil certain criteria. The objective is to detect people and/or scenes that could be considered as suspicious across certain times and urban locations. The security and surveillance use cases will leverage sensors' information and AV processing of environment generated content streams with a view to creating wider surveillance applications for the urban environment. The applications will

[1] `http://terrier.org`

be built as compositions of multiple queries to the SMART search engine, which can facilitate the issuance of alarms.

In the scope of the above use cases, the project will pursue the integration with real-life social networks, to demonstrate the merits of the SMART approach along with the described advancements over the state of the art. SMART will take advantage of the ESKUP social network[2], which is managed by the prominent Europen media group and project partner PRISA Digital. The ESKUP platform includes information and users' profiles that can be directly correlated to information stemming from sensors deployed in Spanish smart cities.

6. CONCLUSION

This paper has presented work in progress, which will be carried out in the scope of the FP7 EU co-funded project SMART. Given that the work is in its infancy, the paper has provided the main concepts and ideas underpinning SMART, along with a brief description of some of the background work and technologies, as well as how the project will be demonstrated to interested attendees. The work is focused on the development of a multimedia search framework facilitating the development of search applications that access, process and visualize information stemming from the physical world. Among the main characteristics of the framework is its ability to integrate information stemming from social networks, to endow query results with a social dimension. This feature is in line with emerging applications that blend social networks with sensor networks, as well as with the wave of participatory sensing applications. SMART puts emphasis on processing and indexing multimedia information data, which is a key distinguishing characteristic comparing to conventional participatory sensing applications. To this end, SMART is working on a novel multimedia indexing architecture, along with leading edge components for audio and visual processing of physical world information.

Acknowledgements

Part of this work has been carried out in the scope of the EC co-funded project SMART (FP7-287583). The authors acknowledge valuable help and contributions from all partners of the project.

7. ADDITIONAL AUTHORS

Additional authors: Dimitris Drakoulis (Telesto), Tomas Garcia Fresno (City of Santander), Marta Gonzalez (City of Santander), Zvi Kons (IBM), Javier Lasa (Prisa Digital), Walter Matta (S3LOG), Jose Miguel Garrido (ATOS), Paul Moore (ATOS), Karim Moumene (Prisa Digital), Aristodemos Pnevmatikakis (AIT).

8. REFERENCES

[1] J. Bernat. SmartSantander : the path towards the smart city vision. In *ETSI M2M Workshop*, 2010.

[2] J. G. Breslin and S. Decker. The future of social networks on the internet: The need for semantic. *IEEE Internet Computing*, 11(6):86–90, 2007.

[3] J. G. Breslin, S. Decker, M. Hauswirth, and et. al. Integrating social networks and sensor networks. In *W3C Workshop on the Future of Social Networking*, 2009.

[4] J. Camp, J. Robinson, C. Steger, and E. Knightly. Measurement driven deployment of a two-tier urban mesh access network. In *MobiSys '06*, 2006.

[5] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. Anonysense: privacy-aware people-centric sensing. In *MobiSys '08*, 2008.

[6] D. Guinard and V. Trifa. Towards the web of things: Web mashups for embedded devices. In *Proceedings of the 18th International World Wide Web Conference (WWW '09)*, 2009.

[7] J. Hendler. Web 3.0 emerging. *IEEE Computer*, 42(1):111–113, 2009.

[8] A. Hogan, A. Harth, and A. Polleres. Saor: Authoritative reasoning for the web. In *Proceedings of the 3rd Asian Semantic Web Conference (ASWC '08)*, 2008.

[9] M. Ingebretsen. Making sense of sensor data. *IEEE Intelligent Systems*, 25(3):3–5, 2010.

[10] M. Kirkpatrick. The era of location-as-platform has arrived. Technical report, ReadWriteWeb, 2010.

[11] T.-Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3(3):225–331, 2009.

[12] E. Miluzzo, N. D. Lane, S. B. Eisenman, and A. T. Campbell. Cenceme: injecting sensing presence into social networking applications. In *Proceedings of the 2nd European conference on Smart sensing and context (EuroSSC'07)*, 2007.

[13] M. Mun, S. Reddy, K. Shilton, and et. al. Peir, the personal environmental impact report, as a platform for participatory sensing systems research. In *MobiSys '09*, 2009.

[14] L. Neumeyer, B. Robbins, A. Nair, and A. Kesari. S4: Distributed stream computing platform. In *Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW '10)*, 2010.

[15] D. O'Byrne, R. Brennan, and D. O'Sullivan. Implementing the draft W3C semantic sensor network ontology. In *8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2010.

[16] I. Ounis, G. Amati, V. Plachouras, B. He, C. Macdonald, and C. Lioma. Terrier: A high performance and scalable information retrieval platform. In *Proceedings of the 2nd workshop on Open Source Information Retrieval at SIGIR 2006*, 2006.

[17] S. Reddy, D. Estrin, and M. Srivastava. Recruitment framework for participatory sensing data collections. In *Proceedings of the 8th International Conference on Pervasive Computing*, 2010.

[2] `http://eskup.elpais.com`