Synergistic User $\leftrightarrow$ Context Analytics

Andreea Hossmann-Picu
Zan Li
Torsten Braun
Univ. of Bern, Switzerland

C. Marios Angelopoulos
Orestis Evangelatos
José Rolim
Univ. of Geneva, Switzerland

Michela Papandrea
Kamini Garg
Silvia Giordano
SUPSI, Switzerland

Aristide C. Y. Tossou
Christos Dimitrakakis
Aikaterini Mitrokotsa
Chalmers Univ., Sweden

ABSTRACT

Various flavours of a new research field *(socio-)physical* or *personal analytics* have emerged in the past couple of years, with the goal of deriving semantically-rich insights from people's low-level physical sensing combined with their (online) social interactions. In this paper, we argue for more comprehensive data sources, including environmental (e.g. weather, infrastructure) and application-specific data, to better capture the interactions between users and their context, in addition to those among users. To illustrate our newly proposed concept of synergistic user $\leftrightarrow$ context analytics, we first provide some example use cases. Then we present our ongoing work towards a synergistic analytics platform: a testbed, based on mobile crowdsensing and the Internet of Things (IoT), a data model for representing the different sources of data and their connections, and a prediction engine, for analyzing the data and producing the insights.

Categories and Subject Descriptors

H.3.4 [**Systems and Software**]: Information Networks

Keywords

crowd-sensing; information fusion; crowd-sensing analytics

1. INTRODUCTION

Various flavours of a new research field *(socio-)physical* or *personal analytics* have emerged in the past couple of years and are getting more and more attention, both from researchers [**?**, **?**, **?**, **?**] and from practitioners (mostly startups). As the name suggests, the goal is to derive semantically-rich insights (high-level activity, preferences, intentions) from people's low-level physical sensing (location, type of activity etc), from their (online) social interactions, or more compellingly from a combination of these. The results of such analytics could be valuable in a variety of contexts, from better and more accurate customer engagement for businesses to space and event planning that accounts for the self-organising phenomena in pedestrian crowds, and including higher value location-based services for users.

We argue here that people's mobility and behavior are substantially influenced by their broader environment as well. Conditions such as weather, infrastructure state, air quality, food quality, radiation level etc do determine to a great extent the way a person moves and acts. For example: on a *rainy day*, someone may decide to take the bus to work, rather than cycle; or during *pollen peak times*, an allergic person may skip their regular jogging sessions. Moreover, depending on the application scenario, specialized information (e.g., in retail, items currently on sale or number of shop assistants currently available) may also be very meaningful. For example, when analyzing shopper behavior, if a *big sale* is announced, someone may reschedule their regular shopping to be able to attend the sale. The ever wider availability and usage of mobile devices, sensors, as well as their supporting infrastructures means that such data could be just as readily available as the more personal, user-centric type, for example via crowdsourcing and/or crowdsensing.

In light of this, we argue here for building a much more comprehensive user context. We propose the concept of *synergistic user $\leftrightarrow$ context analytics*, illustrated in Fig. **??**, as a way to promote the generalizability of an analytic initiative. Synergistic analytics is a modular construction, consisting of the above-cited personal analytics core (based on smartphone and online media data), enriched with extra layers of additional information, such as environmental, infrastructure-related or specialized data (e.g. retail). It is a shift from individual analytic disciplines (e.g. prediction of next place or activity, of the next device interaction, mining (crowd-)sensed context data etc) towards a more holistic, yet still user-centric perspective. True to its name, the results of synergistic analytics will be much more than the sum of its parts: instead of isolated predictions of limited scope, deeper, semantically richer inferences are possible.

In addition, collecting and processing user data poses significant privacy challenges. While in some of the related research, privacy is more of an afterthought, we strongly believe that privacy-protection must be developed alongside and in full synergy with the other components of a system.

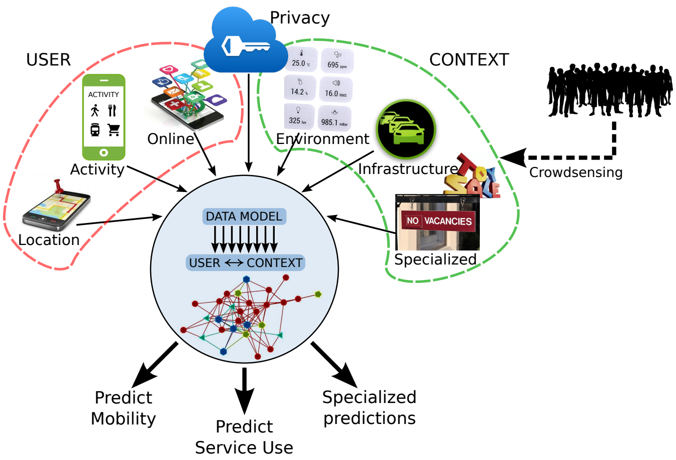

Figure 1: Synergistic User ↔ Context Analytics

To illustrate our proposal of synergistic analytics, we present here our ideas and ongoing work on a platform for providing privacy-preserving, location- and context-based services to users. The platform should support a variety of applications, such as highly personalized navigation, user-optimized coupon dispensing, dissemination of localized, user-centric recommendations, etc., discussed more in detail in Sec. **??**.

The illustrating scenarios for synergistic analytics underscore several scientific challenges, to be addressed by relying on the following four intertwined research pillars, also shown in Fig. **??**: (i) location and activity prediction; (ii) context (environment, infrastructure etc) awareness via crowdsensing analytics; (iii) social profile and behavioral analytics; and (iv) privacy-preservation methods for each of the above. The first and third pillar correspond to the aforementioned personal analytics core, while the second pillar represents a first layer of information on the broader context.

The interactions among the four components are the key to illustrating our concept, we hence focus on them in this paper. In particular, after a brief review of related efforts (Sec. **??**) and a more detailed description of our motivating use cases (Sec. **??**), we present our ongoing work on three main aspects of the proposed privacy-preserving location- and context-based platform: a testbed with two units that we aim at integrating (a crowdsensing unit with smartphones and an Internet of Things (IoT) unit with sensors/actuators) in Sec. **??**; a data model and storage solution, for efficiently representing and processing the highly heterogeneous information collected from the smartphones and from the sensors/actuators in Sec. **??**; and a predictive analytics engine **??**.

2. RELATED WORK

In this section, we briefly summarize the recent efforts in the area of (socio-)physical analytics, as well as research on individual components of physical analytics systems (e.g., crowdsensing, privacy etc).

(Socio-)physical analytics. In [**?**], the authors present their ongoing efforts in creating a system that can integrate mobile sensing data with data from online social networks, to provide insights into user mobility, and more importantly their interactions (both online and physical). A similar system, *SocialFusion*, is presented in [**?**]. However, SocialFusion focuses on the immediate context of individuals, rather than on their interactions. A third similar work [**?**] proposed the PA (personal analytics) engine to generate high-level user states (e.g., emotions, preferences, engagements), that can be used to helpfully intervene in the user's actions.

Finally, [**?**] goes more in the direction of our current work, in the sense that the authors recognize the importance of a more comprehensive user context (including, weather, light or sound level, scenario-specific data etc). However, the paper mainly offers a nice collection of highly specialized use cases, along with preliminary ideas on how to technically realize each of them separately. In contrast, we aim at describing a more unifying and complete system.

Mobile crowdsensing. Besides location, a user's context consists of a variety of other factors, both immediate, personal (activity, heart rate etc) and of a broader nature (weather, pollution etc). Traditionally, these factors are measured via standalone, specialized sensors, which may be costly and hard to install. However, recently, the concept of *mobile crowdsensing (MCS)* has emerged [**?**], whereby the current wide availability of smart mobile devices (smartphone, smart watches, gaming systems etc) is exploited for measuring the above factors, via the numerous sensing capabilities of the devices (inertial, GPS, light, camera etc). While this solves the issue of acquiring and installing new sensors, it does come with a variety of other challenges: from resource limitations (energy, bandwidth and computational power) to privacy issues and including the lack of a unifying architecture (to optimize the cross-application usage of sensors on the same or even across multiple devices). In our current work on MCS [**?**], we have already defined the fundamental components of a mobile crowdsensing system – *crowd*, *server*, *task*) – and their interactions, and we have identified incentives for engaging the *crowd*. Our proposed synergistic analytics platform will tackle the additional issues raised by integrating and jointly analyzing data from different MCS sensors to extract very comprehensive patterns and predictions about user behavior and/or their context.

Privacy and security. Mobile crowdsensing (including location and activity sensing) raises many privacy and security concerns. First, as explained above, the *crowd* provides sensed data to a *server*, which may or may not be trusted. If the server is not trusted, computation must be performed on encrypted data, which can be achieved via homomorphic encryption [**?**] or, more generally, through secure multi-party computation [**?**]. Even if the server is trusted, private information may still leak, for example, when a third party constructs clever queries that, if answered truthfully, cause the server to divulge private information. A characterisation of resistance to this is given by the concept of differential privacy [**?**]. To our knowledge, these issues have not yet been addressed in the context of mobile crowdsensing, and it is our goal to design efficient algorithms, fitted for these cases.

3. SYNERGISTIC ANALYTICS USE CASES

As argued in the introduction, our privacy-preserving location- and context-based platform aims at being generic enough to support a number of diverse applications. Some examples are described below.

3.1 Highly Personalized Navigation

Mobility and navigation are important for a modern lifestyle. However, current navigation applications are typically limited to a few transportation modes and miss complex context and user related data. Exploiting data on user preferences, transportation modes, and the environment, allows a more effective user-oriented navigation and recommender system. The data may include real-time traffic data, public transportation, rental vehicles, air quality, weather conditions, safety ratings and user habits. The system suggests places to visit, transportation modes, as well as important traffic and environmental data to city officials. Users benefit by improved social interactions, handling mobility more sustainably and efficiently. Security and privacy issues may arise, such as untruthful users, non-trusted local infrastructure (easy to tamper with), and protecting user privacy.

3.2 User-Optimized Coupon Dispensing

An empirical study presented in [**?**] found that *proximity* drives coupon redemption. In particular, the study considered the behavior of people, while moving in the proximity of Subway restaurants: the authors showed that the distance to a restaurant is inversely proportional to the amount of monetary incentive needed to prompt people to redeem the restaurant coupons. However, the physical distance to a shop is not the only driving factor for an optimized coupon distribution. In fact, a better insight into potential customers' profile would allow a more effective optimization of coupon dispensing. Along with proximity, other user-related information may be important driving factors, for example: *personal preferences* (i.e., a user who likes Italian food is most likely to visit nearby Italian restaurants) and *social network* (i.e., a user tends to go where their friends have already been). Consequently, a coupon distribution service could optimize the process of customer selection and coupon distribution, by exploiting the synergistic platform proposed in this paper, for retrieving location- and context-related user information.

3.3 Dissemination of Localized and User-centric Recommendations

With respect to data dissemination, synergistic analytics could be exploited to optimize the propagation of information to users, according to their location and social profile similarity. For example, in the crowded environment of a very touristic city, the dissemination of localized recommendations (i.e., interesting events and places in the city) among users would be a more effective solution than a static, provider-based data distribution, both in terms of resource usage (allocation of the downlink communication resources) and in terms of time for the recommendations to reach the target users [**?**]. Such an environment is usually populated by people (e.g., tourists, workers, students, unemployed, etc) with various social profiles and interests. The availability of rich information about users (e.g., location, context, activity, social interest, etc.) may improve the dissemination of localized recommendations by identifying the people and/or communities with similar profiles and interests. On the other hand, this also raises the question of whether users should trust a local municipal network with their information.

4. EARLY EXPERIENCES IN SYNERGISTIC USER ↔ CONTEXT ANALYTICS

Considering the above use case examples, as well as the discussed challenges, we present in the following our ongoing efforts on three main aspects of the proposed privacy-preserving location- and context-based platform: a testbed with two units that we aim at integrating (a crowdsensing unit with smartphones and an Internet of Things (IoT) unit with sensors/actuators); a data model and storage solution, for efficiently representing and processing the highly heterogeneous information collected from the smartphones and from the sensors/actuators; and a predictive analytics engine.

4.1 Data Collection

As a first step towards a generic platform for location- and context-based services, we need access to real(istic) data and to be able to easily develop, deploy and debug software on real(istic) end devices. To this end, we are building VIVO, a novel human- and sensor-based testbed with volunteers.

4.1.1 The VIVO volunteer testbed

The VIVO testbed is based on the concept of *enrolled crowdsourcing*, that allows the deployment of several experiments, as opposed to the traditional usage of crowdsourcing for a single experiment. VIVO provides a secure and privacy-respecting platform for *testbed users*, to collect social, physical and environmental information. The information can be accessed remotely, as in traditional testbeds. However, VIVO differs from traditional testbeds in that it allows to test algorithms and solutions by scheduling and running them *in real time*, on real mobile phones of people participating in the testbed (also called *volunteers*, not to be confused with *users*). Further, VIVO also provides an emulation environment for *users* to run and test experiments on already existing data, stored in the VIVO database.

Unlike recent similar efforts LiveLab [**?**] and SmartLab [**?**] (where a single specific and static application is installed on each smartphone to constantly save the data collected from the sensors), VIVO aims to offer more flexibility. More precisely, VIVO *testbed users* can dynamically deploy their own application on each *volunteer*'s device, as in PhoneLab[**?**]. However, while PhoneLab requires *volunteers* to run a modified version of the Android OS on their mobile (thus limiting the set of potential participants), VIVO applications run on normal Android versions, with no hardware extra requirements. In addition, VIVO also promotes reproducibility of the experiments, via its emulation environment.

Fig. **??** represents the architecture of the VIVO testbed, including Syndesi 2.0 [**?**], our existing IoT unit that can interact with smartphones. Besides the IoT unit (described in

Sec. ??), the VIVO testbed has three main components: the *VIVOTestbed-UserInterface*, the *VIVOServer* and the *VIVO-Client*; and two main actors *VIVOTestbed-Users* and *VIVO-Volunteers*, which we describe below.

- The **VIVOTestbed-UserInterface** enables the communication between *VIVOTestbed-Users* and the *VIVOServer*. Through this interface, users reserve time slots and upload experimental code to be run on the VIVO testbed. Users specify the environment where they want the experiments to be run (i.e., the real or the emulated environment).
- The **VIVOServer** is the main component of the VIVO testbed and it consists of the following elements:
  1. The *VIVOManager* handles incoming requests from *VIVOTestbed-Users* and, based on their preferences, forwards experiments to be run either on *RealVIVO* or *EmulVIVO*. Once an experiment is completed, it sends back the secured and anonymized results to the requesting *VIVOTestbed-User*.
  2. *RealVIVO* manages and schedules the experiments to be executed by available *VIVO-Volunteers*, with the support of *RealVIVO Manager*. While the experiments are running, all data collected through *VIVO-Volunteers* is temporarily stored into the *RealVIVO-DB*. Additionally, part of the collected data is dumped to the *EmulVIVO-DB* by the *RealVIVO Manager*, in order to make it available for the emulation environment. Once an experiment is completed, the *RealVIVO Manager* transmits the secured collected data to the *VIVOManager*.
  3. *EmulVIVO* offers an environment to run experiments on existing data, available in the *EmulVIVO-DB*. The reasoning component of this module is the *EmulVIVO Manager*, which is in charge of receiving requests from *VIVOManager*, retrieving the corresponding data from the *EmulVIVO-DB* and allocating the emulation-running environment for the requested experiments.
  4. The *EmulVIVO-DB* stores anonymized data collected by *VIVO-Volunteers* while running experiments. This data can be used to emulate and reproduce a running environment through *EmulVIVO*.
- The **VIVO-Volunteers** are people equipped with smartphones (nodes), who accept to run VIVO experiments. Volunteers are enrolled for each experiment, based on their characteristics and availability.
- The **VIVOClient** is a VIVO dynamic service, which is in charge of hosting the scheduled experiments, on the volunteer's smartphone.

4.1.2 A Crowd-augmented Experimenting Facility

Syndesi 2.0 [?] is an IoT testbed architecture for smart buildings, that enables the seamless and scalable integration of crowdsourced resources, provided by the end-users[1] of the facility. This integration increases the awareness of the facility both in terms of sensory capabilities as well as in terms of end-user preferences and experienced comfort.

[1]The end-users of the smart building are equivalent to VIVO *volunteers*, and different from *testbed users*. End-users of the building are not necessarily also VIVO volunteers and vice versa.

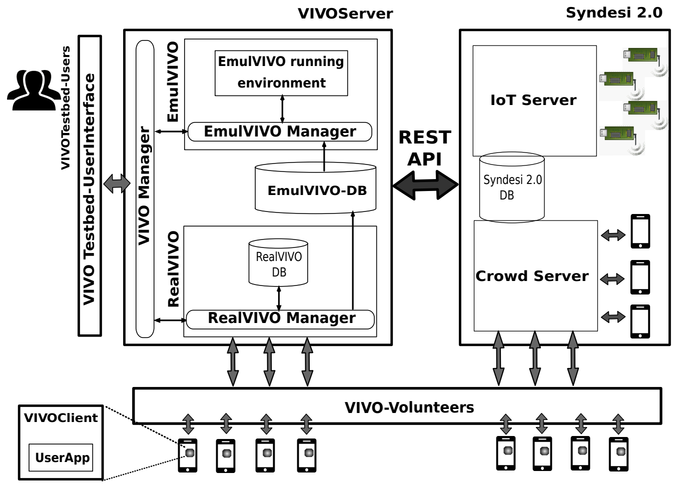

Figure 2: VIVO Testbed Architecture

Combining smart actuation, IoT communication and networking technologies, the testbed user is provided with an agile experimenting platform.

The purpose of integrating crowdsourced resources, such as smartphones and tablets, is dual. First, the embedded sensory capabilities of the resources provided by the crowd are combined with the sensing capabilities of the building for efficient smart actuations. Second, the system is able to interact with its end-users in a direct, personal way, both for incentivizing them to provide sensory data from their devices and to receive feedback on their preferences and experienced comfort. Supported testbed functionalities like the these will be abstracted (i.e. exposed) to the experimenters as services via RESTful APIs, thus enabling their usage in the context of webservices. In fact, given the implemented testbed APIs, an experimenter will be able to use them (or even develop) in custom made experimentation tools (e.g. for post-experiment data processing and visualisation) while being agnostic of the technical details regarding the provided functionalities. Such architectures, in which testbed functionalities are exposed as services, have lead to the notion of Testbed as a Service (TBaaS).

We have demonstrated the use of Syndesi 2.0 via a crowd-enabled, smart luminance scenario. In this scenario, the embedded sensory infrastructure of the building (based on an IPv6-enabled wireless sensor network (WSN)) is opportunistically augmented, by integrating any available smart devices located in its area of operation. Since such devices are carried by people, the system initially provides incentives to their owners in exchange for providing access to their devices. The system then pulls data on the ambient luminance conditions from the devices of the end-users that have accepted the incentives (volunteers) and combines them with the data collected from the WSN infrastructure. This way live luminance maps of the building are composed and then used to optimize indoor lighting towards achieving a tradeoff between energy efficiency and user comfort.

Thanks to its modular, service-oriented architecture, Syndesi 2.0 can be further leveraged by integrating it into the VIVO testbed presented above. All testbed resources of Syndesi 2.0, along with the accompanying mechanisms (e.g. defining the incentivization strategies towards the end-users) are exposed as services over RESTful APIs. These services can be consumed by the VIVO testbed, thus leading to the integration of the two.

4.1.3 VIVO Privacy and Security

The potential security issues facing the VIVO testbed (including the IoT unit) can be clearly defined through specifying different trust models. Firstly, we can assume that the user trusts the application with their data, but may not trust the central VIVO Server. The user definitely does not trust the intervening network. The server, on the other hand, cannot be sure that the application (or users) are providing truthful information.

A number of security components are available to make sure that the system is functioning property. Mechanism design can be used to give incentives to users to provide truthful information. Differentially private statistical models can be used so as to optimally trade off user privacy requirements with utility of the service, in a task-dependent manner. Finally, cryptographic methods can be used for the secure communication between the server and the users.

The particular provisioning of the VIVO testbed for trust and privacy preserving issues along with the capability of supporting a heterogeneous set of information (social, physical, environmental) will enable the facility to be used in more diverse experiments, by a higher number of end-users. For example, for monitoring and collecting data on environmental conditions in out-door settings (via sensors for ambient noise and luminance levels, barometric pressure, etc) and their correlation to user preferences. The extracted data can then be utilized in order to emulate and study more populous crowds in the EmulVIVO running environment.

4.2 Tackling the Heterogeneous Data Challenge

In addition to the challenges of collecting and unifying the data, our proposed platform also needs an appropriate data model, that allows easy and efficient querying, processing and analytics. As already shown in previous work [**?**, **?**, **?**], efficiently storing, processing and analyzing continuous streams of heterogeneous, dynamic data is a complex task.

Defining the data model is the first step towards solving this task. In our connected world, data is heavily interrelated and the main goal of analytics is to identify and exploit both the obvious and the less obvious relationships among data. Thus, a graph-based model is the natural choice. Thankfully, this fact has been increasingly recognized in the past few years, catapulting graph models and graph databases to the forefront of the analytics world. Google's knowledge graph, Facebook's social graph and Twitter's interest graph are the best known examples of this trend. Other growing commercial uses include cloud management, bioinformatics, content management, and security and access control.

In the case of synergistic analytics, we are dealing with multiple node types (users, locations, activities etc) and multiple link types ("knows", "is interested in", "is currently at" etc). In addition, both nodes and links may have attributes, such a demographic information for users, usage for locations or statistical information for links. Finally, while graphs normally only support edges between two nodes, it would be clearly beneficial to be able to represent links among several nodes, forming hypergraphs. For example, as shown in Fig. **??**, an *interest* in art is connected both to the interested *user* and to a *gallery*.

Storing this type of information in an efficient, yet easy to handle manner is challenging. On the one hand, the new generation (hyper)graph databases that have become popular in the past few years seem like the natural choice, since we are dealing with a graph. On the other hand, the competing RDF (Resource Description Framework) databases offer a simple, uniform data model and a powerful declarative query language, that have already proven their worth. Choosing between the two (or potentially additional options) will highly depend on the type of processing to be done on the graph, which we discuss in the next section.

4.3 Prediction Tasks

The prediction engine of our synergistic analytics platform enables different types of predictions, such a user mobility, behavior or service use predictions, as shown in Fig. **??**.

The prediction engine uses i) location, ii) activity, iii) social profile and activity, iv) physical environment, v) infrastructure and vi) application-specific data to predict the users' next place and their behavior, the users' service usage, as well as any required application-specific predictions (e.g. shopping behavior). Social data mainly consists of the user's social profile (e.g., habits, interests, etc.) and social contacts/activity information. Physical data contains the user's mobility history, activity, sensed data from different embedded sensors of mobile phone and physical contacts with other people. Environmental and infrastructure information may include ambient noise levels, ambient luminance, barometric pressure, public transportation schedules, road traffic data etc. Finally, the application-specific data should be provided by the contracting entity (e.g. retailer, hotel owner etc).

The heterogeneity of the collected data gives high potential to the prediction engine, which is then able to perform a deeper analysis of the user and context related data. In terms of mobility, it predicts the user's next-visited physical location together with its semantic meaning (i.e., where the user is willing to go), and additionally it predicts the user's next physical contact. The behavioral prediction includes user activity (i.e., what the user is willing to do), mood (i.e., how the user feels), social contact and activity (i.e., which people the user is willing to meet).

The prediction methodology is based on both historical and current data. The historical data is analyzed to create a user-dynamic mobility and behavioral model. This allows a characterization of the user in terms of mobility aspects (i.e., more sportive and active or sedentary person) and the identification of the locations which are relevant both for the user itself and for the social community he belongs to (according to users' social profile similarities). The model changes dynamically with time, in order to keep track of the changes

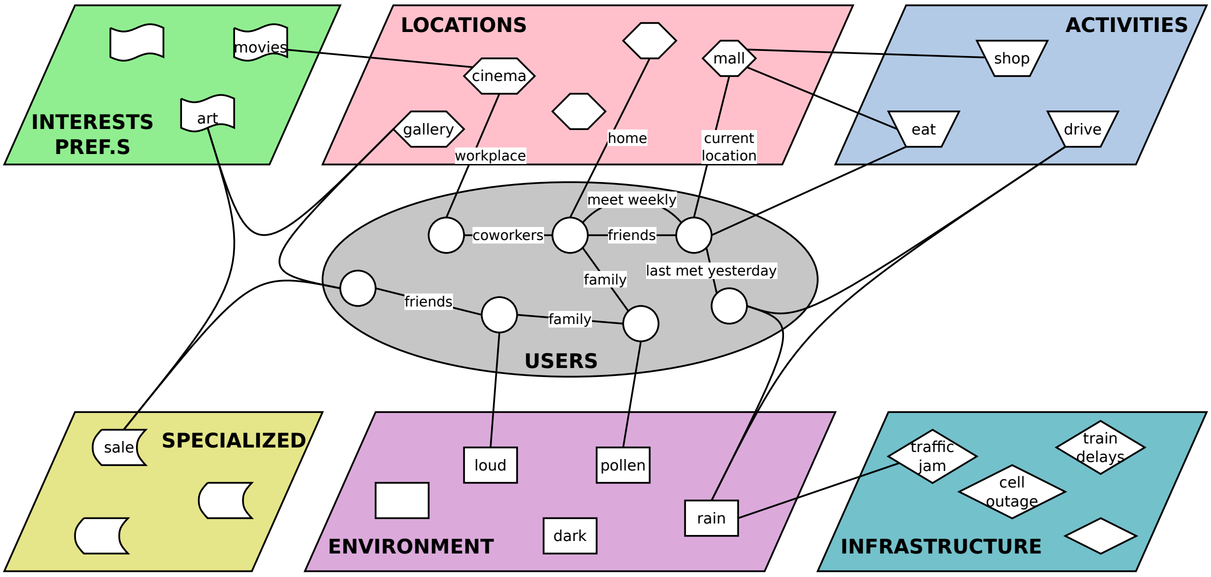

Figure 3: Graph model for synergistic user ↔ context analytics

in the user mobility and behavior. The system exploits also currently collected data in order to be adaptive to the current user's context, therefore allowing more accurate predictions.

Some preliminary analysis on the potentiality of including social aspects to location prediction confirmed the validity of the approach presented above. We showed that analysing the user's mobility history only, already enables a classification of the visited locations according to their relevance to the user. This classification is then used to retrieve the user's mobility and behavioral *characteristics*: a user is identified as a *wanderer*, if she spends most of her time visiting new places; she is instead a *creature-of-habits*, if she spends most of her time in very well-known locations (home, work, etc). Even this simple information about the user profile already improves the next-visited location prediction [**?**, **?**]. The synergistic platform presented above will, among others, combine our initial results with personality and social behavior information, in order to further refine the location prediction.

5. CONCLUSIONS

In this paper, we introduced synergistic user ↔ context analytics, a concept that extends the recent proposals of (socio-)physical and/or personal analytics, by including more comprehensive data sources. Specifically, we argued that, in addition to smartphone sensors and (online) social interactions, also the environment (e.g. weather, infrastructure) and application-sepcific information is valuable for gaining insights into the interactions between users and their context, and among users. After discussing some example use cases of synergistic analytics, we presented our ongoing work on a synergistic analytics platform: a a testbed, based on mobile crowdsensing and the Internet of Things (IoT), a data model for representing the different sources of data and their connections, and a prediction engine, for analyzing the data and producing the insights.

6. ACKNOWLEDGMENTS

We thank Steven Mudda and Alan Ferrari from SUPSI for their contributions. This work is supported by the Swiss National Science Foundation (SNSF) via the SwissSenseSynergy project: http://p3.snf.ch/project-154458.

7. REFERENCES

[1] Beach A, Gartrell M et al. *Fusing mobile, sensor, and social data to fully enable context-aware computing*. ACM SIGMOBILE HotMobile 2010.
[2] Nandakumar R, Rallapalli S et al. *Physical analytics: A new frontier for (indoor) location research*. Tech. Rep. MSR-TR-2013-107, 2013.
[3] Misra A, Jayarajah K et al. *Socio-physical analytics: Challenges & opportunities*. ACM MobiSys WPA 2014.
[4] Lee Y and Balan RK. *The case for human-centric personal analytics*. ACM MobiSys WPA 2014.
[5] Microsoft-Research. *Phytics: Physical analytics*. http://research.microsoft.com/en-us/projects/phytics/, 2015.
[6] IBM-Research. *Physical analytics*. http://researcher.watson.ibm.com/researcher/view_group.php?id=1836, 2015.
[7] Ruckus-Wireless. *Location: The new battlefield for business & digital innovation*. Tech. rep., 2014.
[8] Lund J. *Google analytics for physical environments*. http://popupcity.net/google-analytics-for-physical-environments/, 2013.
[9] Pham T. *CTY: Analytics for physical spaces*. http://cty.io, 2014.
[10] Ganti RK, Ye F et al. *Mobile crowdsensing: current state and future challenges*. IEEE Communications Magazine, 49:32–39, 2011.
[11] Angelopoulos CM, Nikoletseas SE et al. *Characteristic utilities, join policies and efficient incentives in mobile crowdsensing systems*. IFIP Wireless Days, Rio de Janeiro, Brazil, November 12-14, 2014, 1–6.
[12] Krause A, Horvitz E et al. *Toward community sensing*. International Conference on Information Processing in Sensor Networks (IPSN). 2008.
[13] Kremer I, Mansour Y et al. *Implementing the "wisdom of the crowd"*. Journal of Political Economy, 122(5):988–1012, 2014.
[14] Chen L and Pu P. *Survey of preference elicitation methods*. Tech. rep., 2004.
[15] Boutilier C. *A pomdp formulation of preference elicitation problems*. AAAI/IAAI, 239–246. 2002.
[16] Zhou Y, Wilkinson D et al. *Large-scale parallel collaborative filtering for the netflix prize*. Algorithmic Aspects in Information and Management, 337–348. Springer, 2008.
[17] Dwork C. *Differential privacy*. Encyclopedia of Cryptography and Security, 338–340. Springer, 2011.

[18] Gentry C. A fully homomorphic encryption scheme. Ph.D. thesis, Stanford University, 2009.

[19] Goldreich O. *Secure multi-party computation*. Manuscript Preliminary version, 1998.

[20] To H, Ghinita G et al. *A framework for protecting worker location privacy in spatial crowdsourcing*. PVLDB, 7(10):919–930, 2014. URL http://www.vldb.org/pvldb/vol7/p919-to.pdf.

[21] Yuan M, Chen L et al. *Protect you more than blank: Anti-learning sensitive user information in the social networks*. Journal of Computer Science and Technology, 29(5):762–776, 2014. ISSN 1000-9000. doi:10.1007/s11390-014-1466-1. URL http://dx.doi.org/10.1007/s11390-014-1466-1.

[22] Spiekermann S, Rothensee M et al. *Street marketing: how proximity and context drive coupon redemption*. Journal of Consumer Marketing, 28(4):280–289, 2011.

[23] Garg K and Giordano S. *Towards developing a generalized modeling framework for data dissemination*. ewsn 2015, 9, 2015.

[24] Werner-Allen G, Swieskowski P et al. *Motelab: A wireless sensor network testbed*. Proceedings of the 4th international symposium on Information processing in sensor networks, 68. IEEE Press, 2005.

[25] Nandugudi A, Maiti A et al. *Phonelab: A large programmable smartphone testbed*. Proceedings of First International Workshop on Sensing and Big Data Mining, 1–6. ACM, 2013.

[26] Balan RK, Misra A et al. *Livelabs: Building an in-situ real-time mobile experimentation testbed*. Proceedings of the 15th Workshop on Mobile Computing Systems and Applications, 14. ACM, 2014.

[27] Larkou G, Mintzis M et al. *Demonstration abstract: Sensor mockup experiments with smartlab*. Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on, 339–340. IEEE, 2014.

[28] Gluhak A, Krco S et al. *A survey on facilities for experimental internet of things research*. Communications Magazine, IEEE, 49(11):58–67, 2011. ISSN 0163-6804. doi:10.1109/MCOM.2011.6069710.

[29] Atzori L, Iera A et al. *The internet of things: A survey*. Comput Netw, 54(15):2787–2805, 2010. ISSN 1389-1286. doi:10.1016/j.comnet.2010.05.010.

[30] Angelopoulos CM, Evangelatos O et al. *A user-enabled testbed architecture with mobile crowdsensing support for smart, green buildings*. IEEE International Conference on Communications, ICC 2015, London, U.K., June 8-12, 2015.

[31] Evangelatos O, Samarasinghe K et al. *Syndesi: A framework for creating personalized smart environments using wireless sensor networks*. Distributed Computing in Sensor Systems (DCOSS), 2013 IEEE International Conference on, 325–330. 2013. doi:10.1109/DCOSS.2013.35.

[32] Papandrea M, Zignani M et al. *How many places do you visit a day?* Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on, 218–223. IEEE, 2013.