Position Paper for W3C/PrimeLife Federated Social Web Europe Conference 2011

"Do we know each other or is it just our Devices?": A Federated Context Model for Describing Social Activity Across Devices

George Gionis[1], Heiko Desruelle[2], Dieter Blomme[2], John Lyle[3], Shamal Faily[3] and Louay Bassbouss[4]

[1] National Technical University of Athens - DSS *lab*; {gionis}@epu.ntua.gr
[2] Ghent University - IBBT; {heiko.desruelle, dieter.blomme}@intec.ugent.be
[3]Oxford University Computing Laboratory; {john.lyle,shamal.faily}@comlab.ox.ac.uk
[4] Fraunhofer Institute for Open Communication Systems, FOKUS; {louay.bassbouss}@fokus.fraunhofer.de

1. Introduction

The availability of connected devices is rapidly growing. In our everyday life, we already use a multitude of personal devices that are connected to the Internet. The number of shipped smart-phones at the end of 2010 even surpassed the traditional computer segments for the first time in the US [IDC]. From PC, to mobile, to home entertainment and even in-car units, consumers should be preparing for a connected experience.

At the same time rise of social media also had a major influence on how people communicate and collaborate. Applications such as Twitter, Facebook and YouTube have become the first place to go when there is breaking news. As devices turn web enabled, this evolution should allow us to connect the day-to-day activities of millions of users regardless of their location and regardless of the type of device they are using. The number of possible activities are endless and range from sending instant messages to friends, taking photographs, looking for deals online, navigating to our destination, reading newspapers and magazines, sharing content and sometimes even talking to others.

However, as the web turns ubiquitous, all the more it becomes painfully apparent that we completely rely on dedicated applications in support of our needs. The web is everywhere, but we just can't live without our apps.

2. A Use Case: "*which device to use to view my photo album?*"

Up to now, connections over social networking platforms can be described across two broad areas:

1. Connecting with people through some social media platform, i.e. friending or un-friending someone, following and un-following, asking and receiving answers, subscribing, or un-subscribing to his stream, retransmitting or replying to someone's posts, etc.
2. Connecting with content-items on some social media platforms, i.e. liking a page, posting a status update, commenting on a post, sharing a link, tagging a photo, rating a review or a video, rsvp-ing to an event, etc.

The collection of all connections for a profile constitutes the profile's social graph. In the majority of online social networking sites, the underlying social graph does not contain the concept of a "device" as an endpoint in a social connection. This is expected since the device is the technological medium through which any user - when authorized - can create social connections with other users or items (as shown above). As a result we have very little knowledge on the "connection" a user may have with a particular device.

Yet, in the (near) future, we will be able to access and use applications that work across devices allowing us to have an uninterrupted usage experience. Capabilities such as migrating the session of an application from one device to another or seamlessly accessing and sharing content among applications running on different devices that (could) belong to different users are in the core of several research initiatives today [2]. Having the capability to use applications across devices allows the connections that users maintain among themselves to be propagated to their set of owned devices too.

Let's try and see a potential use case of this situation. It is Friday afternoon, Lia with a couple of her colleagues step into a bar across their office building to have some drinks before the weekend starts. George, who works in the same area, is also there with some colleagues for drinks. **Lia and George have never met**

before. That afternoon the two companies sit side by side. Lia and George meet, they start talking, have some drinks, and quickly the two companies mix. Lia uses her smart-phone to take some pictures of the whole company having fun together. She creates a new album "Friday Evening Happy Hour". Lia has this new application installed on her device called "CrazyHats" which uses photos from an album and puts funny looking hats on the persons in it. Using the application she puts some funny looking hats on a photo of her and George. After modifying the picture, she clicks the "share" button to share it with George. She gets several options, i.e. share in Facebook, share in Twitter, send as email, share with another device. **As they have just met, Lia does not have any connection with George through her social networks.** So she clicks the “share with another device” option. Her phone discovers George's tablet among a number of other devices of their friends who are nearby and have enabled bluetooth discovery. She selects it, the handshake is completed and she passes him the photo. After a couple of hours, the evening at the bar comes to an end and George invites all the company to continue the evening at his house which is nearby. The company heads over to George's house. In the meanwhile Lia has processed all the photos in the "Friday Evening Happy Hour" album using “CrazyHats” and now all the members of the company have funny hats on them in the photos. George thinks it will be fun for everyone to see the pictures together. Since Lia's smart-phone has a very small screen, so they decide to use George's HD TV to view the photos. Lia selects the album on her device and again she hits the "share" button. This time George's TV comes right on top of the list with a small label "suggested device" next to it. She selects this option, George approves it and soon the whole company watches their photos on high definition.

How could this work?

3. Device Discovery using Social Proximity: “*Lia knows George ...or is it just their devices that know each other?*”

Whenever Lia hits the “share” button, this invokes a number of discovery protocols on her mobile phone. Some of the best known examples are multicast DNS (mDNS) and DNS Service Discovery (DNS-SD), the Simple Service Discovery Protocol (SSDP) for UPnP, and the Service Location Protocol (SLP) as IETF effort towards a standard protocol. In the use case above, such service discovery protocols are used in George’s house over his local WiFi network, or over bluetooth in the bar to discover devices. In the two cases though there are different devices that get discovered by Lia’s smart-phone through different protocols - in the bar George’s tablet using bluetooth and in his house his HD TV over his WiFi. **Therefore there is no prior knowledge, or any device discovery history, for Lia’s smart-phone to characterize George’s discovered HD TV as “a suggested device” among possibly other discovered devices of their friends or of the people next door.**

A potential solution to this problem that receives a lot of attention lately is to use the underlying social graph that people maintain in various social networks in order to ascertain whether the discovered devices belong to a (socially) known, and therefore most probably trusted, person - for example in the use case above if George was in one of Lia’s networks. **However, in our use case Lia and George do not have any connections in any online social network level.** Furthermore, they are among friends, with whom most probably they have already established relationships in online networks. Therefore, in such a case this approach could yield conflicting results in identifying a discovered device as the preferred one. On the other hand, a deeper analysis of their social graphs could show that George and Lia share indeed some degree of social proximity, i.e. they connect through the people they know. The research domain of graph analysis in order to define the level of proximity among two randomly picked nodes is currently a vibrant one, working on a number of techniques such as the shortest graph distance or the maximum information flow between two nodes, the number of common neighbours between two nodes and even more sophisticated proximity measures that involve infinite sums over the ensemble of paths between the nodes (Katz measure [Katz]), rooted PageRank [Liben-Nowell] algorithms, and escape probability [Tong]. Yet many such measure techniques that are highly effective in relatively small social networks become almost computationally prohibitive in large online social networks with millions of nodes [Sarkar] and a dynamic behaviour.

In addition, the numerous ad-hoc connections we encounter in our daily lives are often impossible to detect this way (e.g. delivery people, people you start talking with on the train, etc.). For example in our use case above, George and Lia spent an entire evening in close physical and social proximity and this has not yet been depicted in their online social network graphs. **In such a case further information is needed in order to ascertain any existing social proximity.** Research shows that a wealth of such information is related with our mobile devices. Data such as call logs, Bluetooth devices in proximity application usage, phone status etc. can provide the social fabric to infer on underlying social relationships [Raento]. Lia and George do not have any prior connection in an online social network, yet in the bar Lia shared a photo with him through an application on her smart-phone. This activity properly captured and structured could later - i.e. when they are in George's house - provide the necessary context for her smart-phone to discover George's HD TV as the "suggested device" among others in order to view the photo album (a content item which they have already shared between them before).

The solution we advocate for can be considered as an extension to the approach of validating the discovered devices based on whether or not their owners share a "social" connection in one of their online social networks. We propose that this approach can be supplemented by discovering also ephemeral or dormant connections among device owners through the activities they perform by their devices. Daily we use a number of different applications on our devices to perform a multitude of activities, access information, schedule our life, find our way around, connect with friends and share information with them. What is more, in the future our capabilities through such smart-enabled devices will be further augmented - i.e. sharing the photo camera of my smart-phone with an application on the HD TV of a friend or watching the movie I just paid for in my smart-phone on an application on the HD TV at my girlfriends house are just a few of such (near) future scenarios [webinos UCs].

So why not try and structure these activities that are performed through our smart devices in a way that could "make sense" and make this information available in a privacy preserving way get a better user experience!

4. A Federated Context Model for Describing Social Activity Across Devices

The idea of acquiring context data locally in a device and structuring them in a meaningful way has been developed to support context awareness capabilities within the EU FP7 webinos project.

Webinos Project and Social Proximity

Webinos is a "Service Platform" project under the EU FP7 ICT Programme [webinos]. The project aims to deliver a platform for web applications across mobile, PC, home media and in-car devices. From this perspective, Webinos represents a leap forward as it is a federated web runtime that offers a common set of APIs to allow applications to easily access cross-user, cross-service, and cross-device functionality in an open yet secure manner. Within the project, we are working on realizing the above described concept of ubiquitous detection of social proximity among the users' devices.

Elements of the Context Model: Context Sources, Social Activities and User Profiles

A context source can be seen as a virtual object (context object) with a set of properties which represent the state of this object in the real world. Consumers of context sources are in general applications or services which are usually interested to get a snapshot of the state of a context source or to subscribe to a context source to receive real-time notifications each time the state of the associated context source is changed. In both cases a concept for modelling context sources is needed. Several approach exist in literature. Chen and Kotz [ChenKotz] identified six different categories for modelling context information (location model, key-value pairs, tagged encoding, object-oriented model, logic-based model, and others). Strang [Strang] also evaluated six similar categories, ranging Key-Value pairs where all context information gets stored in a table-like data structure, Markup Schema modelling based on Standard Generic Markup Language (SGML), Graphical context modelling that creates an intuitive model that can be quickly grasped by an observer, Object-Oriented modelling allowing capabilities such as encapsulation, reusability and inheritance, Logic based context modelling that can enable a highly formalized means to create expressive descriptions of factual circumstances and Ontology based context modelling that can describe concepts as a hierarchy that also defines relationships and properties.

In our case it is very important to have a clear identification of the context objects and their properties in the context model but also a clear view of the activities that take place around these context object - for example the photo album that gets shared in the use case above. Therefore a context modelling approach that fits the Ontology based context modelling characteristics is more desired. In this direction a set of emerging initiatives from the social web domain, such as the Activity Streams [Activity] and Portable Contacts [POCO] already meet a high adoption level and they can be used to provide a methodological background to describe context objects and also the set of social activities that could be performed on them across connected devices. Towards this direction a solution that we currently work on is to enrich the two initiatives with the required activity verbs, context object and Contact attributes in order to be able and describe activities such as "Lia *shared a photo album with George"* or several other similar ones without losing information such as Lia's source and/or applications and George's target device and/or application

Addressing and Accessing Context Data from Multiple Devices in the Cloud

Users today own an increasing number of devices. Additionally, some devices could even be owned (or used) by many users such as the family's HD TV or some of the in-car applications. Therefore it is inevitable that contextual information will be captured at device level, since users own and use many devices. Yet in order to have a collective view, it is necessary that this contextual information is transferred and synchronized with the User's profile in the cloud.

Within webinos this is addressed using the concepts of a Zone and a ZoneHub. A Zone is a user-defined area where the user maintains his devices along with a number of other items such as licences for services he has purchased, trusted relationships, etc. The zone provides a number of supporting services such as single sign on, discovery of the services provided by the devices in the zone, data replication and synchronisation. External access to the Zone is provided via the ZoneHub which is a representation of the Zone in the cloud and also a part of Zone itself. The ZoneHub supports the discovery of other ZoneHubs using the User name (or any other identifier) of the owner and is the host of all context data in the Zone. Therefore upon connecting to the Zone each device synchronizes its context data with the ZoneHub.

To address a specific context object in the ZoneHub it is necessary to be able and singularly identify it. Each context object is always associated to a specific user e.g. location holds the current position of a user. The state of the location context object will be updated each time the GPS-unit of the user´s device reports a new position to be used by the user or some application that requires it. To address context objects in this category an identifier of the context object *coid* (e.g. location) and an identifier of the user uid (e.g. alice) are needed. Finally, the address location@alice can be used to locate the location context source on Alice´s device.

An indicative specification for addressing context objects in general could be as following <co-id>[@<uid>][?<key>=<value>(<key>=<value>)*].

The address of a context object as described above is needed to access the content of this object. There are two approaches for accessing the data of a context source:

- Request/Response: In this case consumer applications access the context object on demand and get a snapshot of the object's state (attributes).
- Publish/Subscribe: Here, consumer applications subscribe one time for a context object and receive real-time notifications each time the state of the object is changed.

For accessing context objects it is necessary to address ZoneHub of the associated user. From that point on, provided that the accessing entity has the proper level of authorisation, all context objects can be accessed from this single point.

Protecting Privacy

Acquiring ad-hoc social context data raises privacy concerns, particularly when the controls usually provided by social networks may be absent on personal devices. Users may be concerned that their smart device will betray their activities by recording too much information. This information might have unforeseen implications and consequences, perhaps making someone appear to have a closer relationship with another person than they really do, or encouraging unwanted attention from advertisers. There is also potential for abuse by authorities, who might try to track associations between people to counteract activist groups, thereby threatening freedom of assembly. As a result, the potential for making new connections must be mediated by privacy controls which allow users to give their informed consent to data disclosures.

However, access control on mobile devices is difficult, due in part to the dynamic nature of the environment - resulting in the need to make decisions quickly - as well as the limitations of mobile user interfaces. It is therefore necessary to compromise in favour of usability and simplicity rather than providing fine-grained controls to match every circumstance. Quick settings which enable anonymity, prevent device discovery and clear saved context data are required. Furthermore, users must be able to query their devices to find out what information their devices have collected on other users, as well as what other users may have collected from them. We are also unlikely to escape the need for runtime authorisation, particularly when encountering devices for the first time.

The webinos project is researching many potential improvements in mobile privacy. Removing the reliance on passwords through techniques developed in the PrimeLife project [PrimeLife] would make a big difference, reducing password re-use and therefore account link-ability, limiting the impact of a privacy breach. Providing better information to users when mobile applications request authorisation would also enable users to make more informed decisions. For example, if an application requests access to a user's contact list, it should state why it needs this information and what it's privacy policy is on data usage and disclosure. This can then be compared to pre-defined user privacy preferences, as well as decisions made by trusted friends within a social network. Similarly, choosing whether to trust another user's device can potentially be enhanced through attestation [Nauman] - knowing that the other user is running trusted software which will maintain control of our data. Finally, the impact of a standardised cross-device privacy and security framework would be significant: allowing users to make one decision about an application across all devices would increase usability while maintaining consistent security policies.

5. Summary

As the availability of connected devices is rapidly increasing, users gain the ability to access their applications anywhere, anytime and on any device. The advent of the social web over the last few years has clearly shown some interesting new possibilities for combining the ubiquitous character of the web and the social connections we encounter in our day-to-day activities. We can extract numerous scenarios in which applications could benefit from an accurate detection of newly created or ad-hoc social connections. For example, detecting the ad-hoc connection and activities we set up with a new person we only just met. It is important to realize that people are not only connected with their closest friends. Interacting with persons to whom we share a much vaguer relationship is also a significant part of our daily life, as well as our business life.

Acknowledgments. The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7-ICT-2009-5) under grant agreement number 257103.

References

[Activity] Atom Activity Extensions, M. Atkins, D. Reordon, C. Messina, M. Keller, A. Steinberg, and R. Dolin., Draft specification, March 9th 2010, available at http://activitystrea.ms/schema/1.0/activity-schema-01.html

[Chen] Chen G., Kotz D., A survey of context-aware mobile computing research, 2000.

[IDC] International Data Corporation (IDC), 2011, available at http://www.idc.com/research

[Katz] Katz.L., A new status index derised from sociometric analysis. Psychometrika, 1953.
[Liben-Nowell] Liben-Nowell D., Kleinberg. J. The link-prediction problem for social networks. J. Am. Soc. Inf. Sci. Technol., 2007.

[Nauman] Nauman et al. Beyond kernel-level integrity measurement: enabling remote attestation for the android platform, Proceedings of TRUST'10: The 3rd International Conference on Trust and Trustworthy Computing, pp 1-15, Springer LNCS, June 2010.

[PoCo] Portable Contacts 1.0, J. Smarr. Draft specification, August 5th 2008. available at http://portablecontacts.net/draft-spec.html

[PrimeLife] PrimeLife project and the Identity Mixer. May 2011. available at http://www.primelife.eu/results/opensource/55-identity-mixer

[Raento] Raento M, Oulasvirta A, Petit R, Toivonen H (2005) ContextPhone: A prototyping platform for context-aware mobile applications. IEEE Pervasive Computing 4:51–59.

[Sarkar] Sarkar P., Moore A. W., Prakash A.,. Fast incremental proximity search in large graphs. In Proc. of ICML, 2008.

[Tong] Tong H., Faloutsos C., Koren Y., . Fast direction-aware proximity for graph mining. In Proc. of KDD, 2007.

[Strang] Strang, T., Linnhoff-Popien, C., A context modeling survey, Workshop on Advanced Context Modeling, Reasoning and Management as part of UbiComp, 2004.

[webinos] Webinos project, 2011, available at http://www.webinos.org

[webinos UCs] Webinos Project: Use Cases and Scenarios, March 22nd 2011, available at http://webinos.org/archives/744