Reinventing Mobile Community Computing and Communication — **Source link** ⧉

Yann Busnel, Nuno Cruz, Denis Gillet, Adrian Holzer ...+1 more authors

Institutions: University of Nantes, University of Lisbon, École Polytechnique Fédérale de Lausanne

Topics: Mobile computing, Mobile search, Mobile community, Mobile device and Context awareness

Related papers:

- Context-awareness in wireless and mobile computing revisited to embrace social networking

- Exploring the logic of mobile search

- How are you feeling? A social network model to monitor the health of post-operative patients

- Role of mobile handhelds in redefining how we work, live and experience the world around us: some challenges and opportunities

- The mobile future

Reinventing Mobile Community Computing and Communication

Yann Busnel, Nuno Cruz, Denis Gillet, Adrian Holzer, Hugo Miranda

Reinventing Mobile Community Computing and Communication

(Position Paper)

Yann Busnel*, Nuno Cruz†§, Denis Gillet‡, Adrian Holzer‡, Hugo Miranda†

*University of Nantes
Nantes, France

†University of Lisbon
Lisbon, Portugal

‡EPFL
Lausanne, Switzerland

§IPL-ISEL
Lisbon, Portugal

Abstract—In less than a decade, smartphones and mobile applications spread like wildfire and dramatically improved aspects of our professional and private lives, from efficiency to safety. However, these applications are still in their infancy and mostly provide mobile versions of online Internet services or arcade games. With the exception of simple location-based query applications, context-awareness is largely ignored. However, it is not hard to imagine advanced mobile social networking applications – SNAPPs for short – that could proactively assist users in everyday tasks, improving their quality of life. Such services would require massive data collection, processing and communication between mobile devices. Unfortunately, the current centralised communication paradigm represents a major barrier to such intense networking. In this paper, we claim that a fundamental paradigm shift in communication is required to allow such application to see the light of day. The paper claims that such a shift is possible and that it resides in moving towards decentralised communication by taking advantage of the largely untapped network, storage and processing power capabilities offered by idle mobile devices. The paper presents and discusses a number of research questions that must be addressed in order to achieve this paradigm shift.

I. Motivation

Application development for mobile devices has been following a sub-optimal and predictable pattern that consists in replicating, for the wireless setting, the services that have for long been available on wired desktops. Examples are distributed games and ubiquitous interaction with social networks. A few remarkable exceptions include travel assistants, which use information made available from mobile device sensors (in particular, from GPS receivers) to facilitate user mobility [1]. Although this approach improves user productivity and satisfaction, it falls short of the expanding range of possibilities that the increasing computing power, networking capabilities and sensors integrated in mobile devices make available.

A. Advanced Social Networking Apps (SNAPPs)

In our vision, mobile devices should become effective human assistants, reacting to the events observed in the environment that are relevant to the user by offering advanced mobile context-dependent Social Networking APPlications, henceforth called SNAPPs. In order to motivate our vision, we present below two scenarios that make use of SNAPPs.

Scenario 1: Bob is visiting a shopping mall in a foreign city. While walking, Bob's smartphone receives a number of messages from other smartphones, indicating their user's

movements. Bob's smartphone is not familiar with the location and simply registers that the location is crowded and users tend to move. However, smartphones more familiar with the location notice that the movement pattern is unusual, given that all users in some region are moving in the same direction at an above average speed. As a result, they start sending alert signals both to their users and to other smartphones using their radio interfaces. Due to the large number of alert signals, Bob's smartphone suggests him to move away from the location, a few minutes before the fire alarm is triggered. Unfortunately, Charlie – an independent senior citizen – is currently in the mall and his health monitor notices a change in his condition, probably due to the panic. As a result, Charlie's smartphone places an automatic call to the emergency services, providing all the details about Charlie's health condition and location. After confirming the need for assistance, the smartphone mediates the communication between the emergency services, guiding them to Charlie's precise location. Moreover, the smartphone also finds a doctor currently in the flow of evacuating people, which provides assistance until the arrival of the emergency services.

Scenario 2: Alice and her son Tommy are at the amusement park when all of a sudden, Tommy disappears. After calling for help, Alice uses the emergency search app on her cell phone to spread a picture of Tommy that she had taken earlier. An alarm with the picture is gossiped in the background by the smartphones at the park, who start to look for Tommy in photographs shot by the devices recently. However, because face recognition is CPU intensive, smartphones extract faces from the photographs and delegate the facial recognition process to other smartphones in the neighbourhood with spare CPU time, effectively sharing the load and speeding up the process. In addition, some of the smartphones alert their users, thus creating a surveillance net set to increase every 5min. The picture is equally delivered to autonomous robots in the park who use real time face recognition software to try to find Tommy. Chloe, an autonomous robot who has received the picture, recognises Tommy and notifies Alice through the emergency app, telling her where to pick up Tommy.

The applications above have in common the emergence of a spontaneous collaborative system, composed by the mobile devices co-located in some region of interest. To be useful, these applications require a significant amount of computing power, operate in the background monitoring the environment (in Scenario 1) or waiting for some event (Scenario 2), but in any case without disturbing the users' other activities. In addition, applications will be required to share a large amounts of raw or preprocessed data, that is either stored or produced at the

---

Authors are listed in alphabetic order

mobile devices, whose relevance is constrained to the location of the devices and whose exchange either between the devices or with a wired third-party is incompatible with the state-of-the-art infrastructured networks. The following section makes a more in-depth discussion of the limitations that the current trend on application development and infrastructured wireless networks pose to the successful deployment of SNAPPs.

B. Limits of Current Systems

The development of a computing environment capable of supporting these applications is challenging due to a combination of factors that include the limited computing power and energy available at each mobile device, the lack of support to infrastructureless (ad hoc) networking on mobile operating systems, the large amount of context data and privacy constraints. Furthermore, the direction in which the state of the art on application development for mobile devices is moving is at odds with some fundamental principles needed to develop SNAPPs, given that, for example, it has been preferring point-to-point communication between the devices and servers on the wired infrastructure over device to device communication. Unfortunately, this raises a number of barriers as for cost, privacy and scalability issues, devices rarely communicate their sensor value in the wild.

Research that aims to extend the computing power of mobile devices can be broadly arranged in four classes, differentiated by the location where it is made available. The most popular approach makes the delegation visible for applications, which are required to explicitly perform remote calls to cloud hosted servers. This is the case of, for example, travel assistants and voice recognition applications. A second class uses the available infrastructure to delegate the computing power on resourceful servers executing virtualised versions of the mobile devices (*e.g.*, [2], [3], [4]).

Cloudlets, on the other hand, avoid the latency involved in transferring state to and from the virtualised servers by placing the computing power on specialised devices in the vicinity of the wireless network [5]. The fourth class attempts to distribute the computing power among the mobile devices in proximity. In this paper, we claim that this is the preferable approach as it alleviates bandwidth on long range wireless links and mitigates privacy issues. Unfortunately, instantiations of this approach rely on mechanisms that do not take into account the particularities of the wireless setting (by using resource hungry frameworks like MapReduce [6], or others [7], [8]) and do not propose useful mechanisms to cope with the large amounts of data that must be transferred between the devices.

Therefore, unless there is a dramatic paradigm shift, SNAPPs will not be part of our daily life anytime soon.

C. Towards a paradigm shift

In this paper, we present a research agenda divided in four distinct, yet complementary, research streams that aims at achieving a paradigm shift to produce an environment where SNAPPs can thrive. As depicted in Fig. 1, streams can be mapped on building blocks, each providing one class of services to the others and to SNAPPs and that put together different research areas. Each stream is presented in a distinct section as follows:
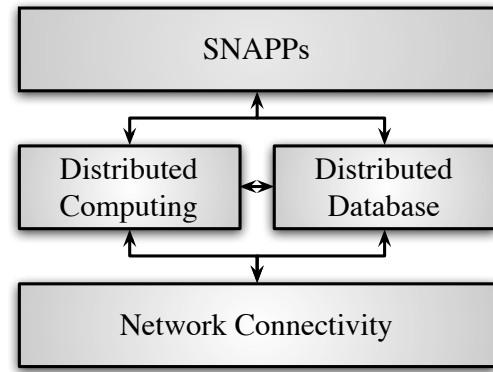

Fig. 1.   Research Streams Alignment

Connectivity: Section II gives an overview of how Mobile Ad Hoc Networks (MANET) connectivity needs to be rethought in order to provide the adequate building blocks for message exchange between mobile devices. The section makes the case for the promotion to first class citizens of communication models whose properties are in general neglected like anycast, geocast and content-based publish-subscribe.

Databases: Section III gives an overview of distributed data storage issues and shows how database systems could be rethought in order to accommodate the huge amount of decentralised data that needs to be stored for SNAPPs. We aim for the implementation of a location dependent distributed database model, providing the insertion, retrieval, update and interpretation of information. Ideally, the database should support efficient retrieval and inference of information using data streams.

Computing: Section IV gives an overview of the issues related with the limited computing power available on mobile devices and proposes to rethink distributed computing in order to facilitate delegation of computing tasks to idle mobile devices nearby.

Applications: Section V reflects on application design and how incentive systems must be rethought to encourage adoption of SNAPPs. Namely, the use of bartering schemes, reward mechanisms and privacy enforcement to foster collaboration is discussed.

For each stream, the paper formulates a few research questions that raise light on the open issues that must be addressed in order to achieve this paradigm shift.

II.   RETHINKING MANETs CONNECTIVITY

Network connectivity and low level messaging are the cornerstones of any distributed architecture. It is therefore essential to assess its potential limits and opportunities for SNAPPs.

Network support to SNAPPs is expected to exhibit two fundamental characteristics. One is the capability to detect other devices in proximity. The other is broadband, to make possible the exchange of the large amounts of data required by the applications. Unfortunately, none is provided by current

infrastructure-dependent wireless technology. Infrastructure, in particular 3G and 4G, prevents bandwidth reuse in short distances, what reduces the number of simultaneous broadband connections and increases their cost.

Infrastructure-less networks, usually named ad hoc networks, are self-organised networks exclusively composed by the participating devices and could play a relevant role in short range communication. However, most of the research on ad hoc networks was concentrated on solving the problems raised by mirroring the TCP/IP communication stack on mobile devices. The large amounts of broadcast (e.g. [9], [10]), point-to-point routing (e.g. [11], [12], [13]) and unique addressing protocols (e.g. [14], [15]) investigated during the last decade, aiming to a number of arguably usable scenarios, supports this claim. In contrast with MANETs, this issue was addressed and resolved in sensor networks and vehicular ad hoc networks (VANETs). In the former by having communication protocols to trade off a full blown stack by energy performance (see [16] for a survey). In the later by inserting parallel communication stacks along with TCP/IP to support particular applications such as short messaging [17].

Point-to-point is only one of the communication models that should ideally support SNAPPs. Locating devices with idle CPU cycles, listening to movement direction announcements or delivering an image to some of the devices in proximity are all examples of services not adequately supported by point-to-point reliable delivery. In practice, this implies the definition of new communication paradigms that extend and refocus research on ad hoc networks, challenging the role of the TCP/IP communication stack in these networks and making other self-organised paradigms like epidemic dissemination, anycast and publish-subscribe, first class citizens of ad hoc networks. Removing the TCP/IP communication stack limits application, however it paves the way into a rupture with known communication paradigms. Most of the past attempts to remove or reinvent TCP/IP for ad hoc wireless environments have failed to be accepted by the community (e.g. [18], [19]), however if the dismissal of TCP/IP is driven by the application, a good outcome should be expected. A precise identification of the paradigms to be supported will mostly depend of the computing and database requirements and therefore, is left as an open research question.

RQ 1 – What communication paradigms should be provided to support SNAPPs?

Early implementations of these novel communication paradigms for MANETs made extensive use of flooding, an algorithm that consists in having all mobile devices to retransmit the message once. Unfortunately, network interfaces are known to be one of the most power demanding components of mobile devices [20] and the use of flooding further amplifies the problem. Luckily, work is underway to devise more efficient implementations (see for example [9], [21]) of these paradigms. However, the performance of these implementations depends on several variables like network topology, node density and number of messages. A crucial aspect to foster user adherence to SNAPPs, and which is raised in RQ 2, is that SNAPPs must be perceived as providing a benefit that compensates the cost of the decreased availability of the mobile devices due to the additional battery consumption.

RQ 2 – Will energy consumption prevent a wide adoption of SNAPPs? Should SNAPPs be available only in the less energy demanding scenarios?

The effort to improve power consumption on MANETs contrasts with the simplicity (from a mobile device perspective) of its implementation when the support of an infrastructure is available. In the later, mobile devices would be required to send a single point-to-point message with the data and the communication paradigm headers to the infrastructure, who would be responsible for enforcing the required communication properties and deliver it to the destinations. However, this model is not free of costs. Elsewhere [22], it was shown that transmissions using the cellular networks interface consume more energy than WiFi and Bluetooth. In addition, implementing these communication paradigms in infrastructure supported networks would require a stronger knowledge of the network, what could not be achieved without frequent message exchanges between the infrastructure and the mobile devices. Question 3 addresses these issues by seeking the borderline cases where the use of an infrastructure would be preferable.

RQ 3 – In which application scenarios and topologies will the use of infrastructure-based implementations of novel communication paradigms provide a more efficient usage of the device's batteries?

III. RETHINKING DISTRIBUTED DATABASES

To further increase the interoperability of SNAPPs, smartphones would be required to understand and learn from the messages circulating on the wireless broadcast media. Message interpretation can be supported by the definition of an ontology. However, in spite of existing efforts, it is hard to define an ontology encompassing all possible events and knowledge relevant in the scope of users' everyday activities. Therefore, message interpretation must be complemented with the ability to infer information simply from the message patterns and by the presence or absence of the messages. A similar behaviour can be found in Nature, where a limited set of sounds, smells or pheromones is used to implement a basic language that permits to communities of bees, wolf packs, *etc.* to alert all its members of eminent danger and organise the community.

Research should be focused in the support of "ether" knowledge bases by providing interfaces for insertion, retrieval, update and inference (using data streams) of information on location dependent databases. A possible solution focuses on continuous queries and streams, which rely on large-scale distributed information systems with decentralised management and avoid single points of control and failure. Within the system, each participant contributes with some physical resources (storage, computation, communication) and information composed of both data and queries. Data as well as queries, which are submitted to the system, are revealed to authorised participants who can then analyse them. This clearly improves information symmetry. Furthermore, participants interested in similar queries could take advantage of this knowledge to organise themselves to acquire necessary information, compute the query (as volunteer computing [23]), store and disseminate the results. Hence, the more interested users there are for a query, the more resources are devoted to

the query, the cheaper it is for a given participant to obtain results and the more the results are replicated and available.

A. Query and indexation

The organisation of participants may be leveraged through a flexible analysis of relations between the queries they issue. This relies on the definition of an algebraic formalisation as conducted in databases. For example, a simple "equivalence relation" permits to share the cost of a query evaluation through relevant reuse of results. Other more complex relations involving queries and sub-queries have to be studied to avoid repeated evaluations. Each considered relation leads to specific organisation, but they all share the general principle of organising participants accordingly to their queries. To sketch the general picture, a potential solution should be composed of two layers. The first one enables accessing and querying data with a simple language and offers basic data indexing capabilities. The second layer takes advantage of both distribution and algebraic analysis of queries to bring back the full power of querying to the users. They are freely developed using any kind of query language and tools without any limitation as far as they have access to basic information (data or queries). Moreover, many users may share enough resources to compute very complex queries which are currently out of their individual scope. Recent research has demonstrated the effectiveness of epidemic protocols in a variety of settings [24], [25], [26]. These include not only traditional peer-to-peer applications such as data dissemination or overlay maintenance, but also user-oriented tasks such as identifying communities with similar interests, or recommendation systems. Still, an open research question is:

RQ 4– How can group communication paradigms contribute to perform efficient queries on such alternative infrastructure? How can data and users be efficiently organised?

B. Stream analysis

In the applicative contexts of SNAPPs, it may be necessary to quickly and precisely process a huge amount of data. The problem of extracting pertinent information in a data stream is similar to the problem of identifying patterns that do not conform to the expected behaviour, which has been an active area of research for many decades. Unfortunately, a common feature of these techniques is their large space complexity and their computational cost, as they rely on full space algorithms for analysing their data. Given our settings — the real time monitoring of ad hoc network traffic — relying on full space algorithms for analysing input data is not feasible. In contrast, two main approaches exist to monitor in real time massive data streams. The first one consists in regularly sampling the input streams so that only a limited amount of data items is locally kept [27], [28], [29]. This allows to exactly compute functions (metrics, pattern matching, *etc.*) on these samples. However, accuracy of this computation, with respect to the stream in its entirety, fully depends on the volume of data that has been sampled and their locations in the stream. On the other hand, the streaming approach consists in scanning each piece of data of the input stream on the fly, and in locally keeping only compact synopses or sketches that contain the most important information about data items. This approach permits to derive some data streams analyses with guaranteed

error bounds without making any assumptions on the order in which data items are received at nodes (*i.e.*, data items ordering can be delayed by the ad hoc routing scheme [30]). Most of the research done so far with this approach has focused on computing functions or statistic measures with error ε using $\text{poly}(1/\varepsilon, \log n)$ space where n is the domain size of the data items. These include the computation of the number of different data items in a given stream [31], [32], the frequency moments [33], the most frequent data items [33], [34], the entropy of the stream [35], or the relative entropy between one data stream and the uniform one [36], [37]. A comprehensive survey of these techniques, their advantages and their drawbacks is given in [38].

Unfortunately, most of the work proposed for a single stream analysis is clearly not adaptable to the distributed functional monitoring model [39]. However, a combination of sampling techniques and information-theoretic methods can be used to extract pertinent information from collections of streams (metrics, summaries, pattern matching, *etc.*). In this combination, the input data is read on the fly (as in the streaming model), and processed with a minimum workspace and time what makes it suitable for the combination of limited resources and broadcast medium of the SNAPPs environment. In the communication complexity model, each node receives an input data stream, performs some local computation, and communicates only with a coordinator who wishes to continuously compute or estimate a given function of the union of all the input streams. The challenging issue in this model is for the coordinator to compute the given function by minimizing the number of communicated bits [39], [40], [41]. Cormode *et al.* [39] pioneer the formal study of functions in this model by focusing on the estimation of the first three frequency moments F_0, F_1 and F_2 [33]. Arackaparambil *et al.* [40] consider the empirical entropy estimation [33] and improve the work of Cormode by providing lower bounds on the frequency moments, and finally distributed algorithms for counting at any time t the number of items that have been received by a set of nodes from the inception of their streams have been proposed in [42], [43]. However, an open research issue is:

RQ 5– How can coordination be distributed among all participants of SNAPPs? How can on-line stream analysis provide inferred pertinent information dedicated to SNAPPs?

IV. RETHINKING DISTRIBUTED COMPUTING

To mitigate computing power shortage, current data management paradigm is moving more and more away from storing and processing data directly on the end device and delegating it to remote computer farms. This solution however is not sustainable, due to the physical limitations of wireless bandwidth which will eventually become a bottleneck for data-intensive communication. There is however an unexplored alternative solution which consists in using the idle storage space and computing power of surrounding mobile devices to store and process information, creating an ad hoc distributed computing environment.

A. Alternative Models

To be useful, this ad hoc distributed computing environment must outperform, either in latency or battery consumption, two alternative approaches: strictly local computation

(LC) and delegation of the data on computer farms on the Internet (CF).

A comparison between the ad hoc computing environment and LC suggests that the parallelism that can be achieved by the former should improve performance. However, latency is affected by the need to coordinate the participants, in particular, to locate the peers, to deliver the data and integrate the results. In addition, ad hoc distributed computing is expected to be more power demanding (when measured by the integral over all the participating devices), as wireless bandwidth for message passing is also required.

Elsewhere [22], [44], it has been shown that long range wireless network interfaces like 3G or 4G consume significantly more power than short range like WiFi or Bluetooth. Therefore, the CF approach is expected to spend most battery on networking activities. The geographical distance between the smartphone and the servers as well as the limited bandwidth available on the long range wireless link (and which will be used to transfer data from the distributed database) are both expected to increase latency for CF.

Although the ad hoc distributed computing environment appears to be the most intuitive solution for obtaining computing power, it is not clear if ad hoc computing is capable to compensate the costs and latency of CF and LC. Therefore, a fundamental research question to be addressed is:

RQ 6– In which scenarios will ad hoc distributed computing outperform delegation on the cloud?

B. Mechanisms

Authors anticipate that SNAPPs would benefit from the increase in the concentration of smartphones in proximity, from judicious and careful definition of the computing units to be processed by each smartphone and from the gains provided by the novel communication paradigms for MANETs addressed in Sec. II. Two different research lines should be investigated. The first includes mechanisms to alleviate the burden of separating computing units from the rest of the software and aggregating their results, capitalising on the recent advances in programming languages and software engineering that have started to emerge with multi-core computing. The second line consists in investigating mechanisms for speeding up computing delegation, namely by using preloaded computing blocks, in order to approximate the syntax of the work delegation to the one found in remote procedure calls, complemented with innovative distributed approaches, of which pipelining is a fit-well example. Therefore, in the scope of distributed computing, research must be able to answer the following question:

RQ 7– What mechanisms should be put in place to improve the performance of ad hoc distributed computing?

V. RETHINKING MOBILE SOCIAL APPLICATIONS

SNAPPs can be described as being special cases of crowd-sourcing apps. Crowdsourcing captures systems that outsource tasks, data collection and aggregation to the crowd [45]. Examples of mainstream wired systems include YouTube, which collects and displays videos, Wikipedia, which aggregates and presents textual knowledge bases, Amazon, which outsources product reviews to customers, BitTorrent peer-to-peer file sharing or shared processing such as in SETI@home [23]. These systems allow to gather huge amounts of data from a large amount of users. In order to overcome information overload, they usually offer filter mechanisms such as recommendations and profile-based personalisation. Crowdsourcing apps generally leverage on the huge amount of users that can be tapped when geographical barriers are removed and on the user's spontaneous motivation to provide content. This contrasts with SNAPPs, which have to rely on the contributions of *nearby* users to store and process data and where content is mainly *inferred* and system-generated (a system with similar assumptions for wired networks is described in [46]). This poses several new challenges. Hereafter, we discuss three of them, namely challenges in business models, challenges in privacy and anonymity, and challenges in development models.

A. Business Models

SNAPPs strongly depends of its wide adoption as, to be useful, it requires a critical mass of co-located users. Building a successful business model for sharing resources is challenging, as many unsuccessful attempts illustrate. For instance an initiative that aimed at promoting hotspot sharing among neighbours failed to get enough users to gain traction. In this case, users felt that there was a lack of accountability [47] and thus they did not feel comfortable to share their connection. This problem illustrates an issue that was simply not taken into account when the business model was designed. New business models will have to be investigated to capture user assumptions and provide adequate value propositions [48] that will motivate users to adopt and use SNAPPs. These models might be based on innovative bartering schemes [49], decentralised monetary systems as well as contextual reputation and trust mechanism.

This leads to the following open research question:

RQ 8– What are the characteristics of business models that encourage the adoption and usage of SNAPPs?

B. Privacy & Anonymity

Privacy was identified early on as one of the key issues related to information sharing, as exemplified by Mason's 1986 seminal paper [50]. Twenty-five years later, and despite a tremendous change in technology, the core of these concerns still applies and became especially relevant with the advent of social networking [51]. With SNAPPs, privacy is especially important as data will mainly be system-generated, and user resources will be shared. They are key for ensuring the community acceptance of the model and the compliance with legislation that aims to protect users' right of protection of their personal information. According to the new European regulation about the principles relating to personal data processing, the notion of *purpose* and the *right to oblivion* are essential. Among others, these principles mention "personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes" and "data subjects should have the right that their personal data are erased and no longer processed, where the data are no longer necessary in relation to the purposes for which the data are collected [..]" [52]. Research contributions should be equally devoted to monitor, validate and seek alternatives to advances that compromise security or privacy of personal information.

Open research questions could investigate if it is possible to take advantage of the fact that decentralised distributed computing relies on devices positioned in a certain proximity at a certain time to create built-in privacy mechanisms. For instance, data could be made to exist only when a given number of devices are co-located at a given place and time. This departs from current practice where data centralisation makes it easy for someone to eavesdrop on one single entry point. Furthermore, anonymity can be a *sine qua non* condition in order to get users on board, especially when they are asked to share contextual information. However anonymity has some drawbacks that can lead to malicious use, resource hoarding, or spamming depending on the application scenario [53]. Restricting anonymous data to a collocated group can potentially decrease the feeling of anonymity and thus reduce its side effects. We believe that such parameters of anonymity should be further investigated and might provide a balance between engagement and potential drawbacks. Furthermore, we believe that privacy can also be achieved by investigating how access to individual user data should be ephemeral whereas aggregated information can be persistent. In short, we belief the following research question should be addressed:

RQ 9– What are the adequate models for privacy and anonymity to support SNAPPs in order to optimise user satisfaction?

C. Development Models

Finally, application development models should be rethought in order to involve more developers in the design of SNAPPs. New models such as the Metropolis model [54] can be used as starting point for a reflection. This model takes a community-driven approach where software is built in a decentralised fashion by open teams and there is little control on who contributes, using the concept of mashability, where software parts can be used and reused, and emergent behaviour, which implies that mechanisms that go beyond the vision of the originators may appear. This final point can be captured by the following research question:

RQ 10– What are the characteristics of development models that motivate developers to design SNAPPs?

VI. CONCLUSION

Developing advanced social networking apps (SNAPPs) is a challenging task. We believe that it will involve dramatic paradigm shifts in the way we think about mobile distributed systems. In this position paper, we presented a research agenda that aims at providing a path towards such paradigm shifts. This agenda is expected to open novel venues both at the societal and research level. At the research level, this agenda innovates in the experimentation of three novel concepts: the possibility of mobile devices in proximity to share CPU cycles, communicate using fuzzy messages, possibly learning on the fly the communication patterns and, finally, to cooperate in the storage of the knowledge that was previously acquired. These are interesting challenges that can open new research directions on the field of the mobile ad hoc networks and ubiquitous computing. At the societal level, it is expected that the public availability of software libraries and communication standards pave the way for the viral development of a novel class of applications that can contribute for an improved quality of life of worldwide citizens.

REFERENCES

[1] D. Raptis, N. Tselios, and N. Avouris, "Context-based design of mobile applications for museums: a survey of existing practices," in *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*. ACM, 2005, pp. 153–160.

[2] B.-G. Chun and P. Maniatis, "Augmented smartphone applications through clone cloud execution," in *Proceedings of the 12th Conference on Hot topics in operating systems (HotOS'09)*. Berkeley, CA, USA: USENIX Association, 2009, p. 8.

[3] G. Huerta-Canepa and D. Lee, "A virtual cloud computing provider for mobile devices," in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*, ser. MCS '10. New York, NY, USA: ACM, 2010, pp. 6:1–6:5.

[4] E. Cuervo, A. Balasubramanian, D.-k. Cho, A. Wolman, S. Saroiu, R. Chandra, and P. Bahl, "Maui: making smartphones last longer with code offload," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*, ser. MobiSys '10. New York, NY, USA: ACM, 2010, pp. 49–62.

[5] M. Satyanarayanan, V. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *Pervasive Computing, IEEE*, vol. 8, no. 4, pp. 14–23, 2009.

[6] E. E. Marinelli, "Hyrax: Cloud computing on mobile devices using MapReduce," Ph.D. dissertation, Carnegie-Mellon University, Sep. 2009.

[7] C. Borcea, D. Iyer, P. Kang, A. Saxena, and L. Iftode, "Cooperative computing for distributed embedded systems," in *Proceedings of the 22nd International Conference on Distributed Computing Systems (ICDCS'2002)*, 2002, pp. 227–236.

[8] D. G. Murray, E. Yoneki, J. Crowcroft, and S. Hand, "The case for crowd computing," in *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds (MobiHeld '10)*. New York, NY, USA: ACM, 2010, pp. 39–44.

[9] H. Miranda, S. Leggio, L. Rodrigues, and K. Raatikainen, "A power-aware broadcasting algorithm," in *Proceedings of The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)*. Helsinki, Finland: University of Oulu, Sep. 11–14 2006.

[10] B. Garbinato, A. Holzer, and F. Vessaz, "Context-aware broadcasting approaches in mobile ad hoc networks," *Computer Networks*, vol. 54, no. 7, pp. 1210–1228, 2010.

[11] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing," RFC 3561 (Experimental), Internet Engineering Task Force, Jul. 2003.

[12] D. B. Johnson, D. A. Maltz, J. Broch *et al.*, "DSR: the dynamic source routing protocol for multi-hop wireless ad hoc networks," *Ad hoc networking*, vol. 5, pp. 139–172, 2001.

[13] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)," RFC 3626 (Experimental), Internet Engineering Task Force, Oct. 2003.

[14] E. Baccelli and M. Townsley, "IP Addressing Model in Ad Hoc Networks," RFC 5889 (Informational), Internet Engineering Task Force, Sep. 2010.

[15] I. Jeong, H. Choi, and J. Ma, "Study on address allocation in ad-hoc networks," in *Computer and Information Science, 2005. Fourth Annual ACIS International Conference on*, 2005, pp. 604–609.

[16] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, 2009.

[17] H. Hartenstein and K. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *Communications Magazine, IEEE*, vol. 46, no. 6, pp. 164–171, june 2008.

[18] J. Liu and S. Singh, "ATCP: TCP for mobile ad hoc networks," *IEEE J.Sel. A. Commun.*, vol. 19, no. 7, pp. 1300–1315, Sep. 2006.

[19] K. Sundaresan, V. Anantharaman, H.-Y. Hsieh, and R. Sivakumar, "ATP: A reliable transport protocol for ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 6, pp. 588–603, Nov. 2005.

[20] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," in *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001)*, vol. 3. Anchorage, AK USA: IEEE, Apr. 22–26 2001, pp. 1548–1557.

[21] I. Pu and Y. Shen, "Enhanced blocking expanding ring search in mobile ad hoc networks," in *Proceeds of the 3rd International Conference on New Technologies, Mobility and Security (NTMS)*, Dec. 2009, pp. 1–5.

[22] R. Balani, "Energy consumption analysis for bluetooth, wifi and cellular networks," Networked & Embedded Systems Laboratory, NESL Technical Report TR-UCLA-NESL-200712-01, 2007.

[23] D. P. Anderson, "Boinc: A system for public-resource computing and storage," in *GRID '04: Proceedings of the Fifth IEEE/ACM International Workshop on Grid Computing*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 4–10.

[24] Z. J. Haas, J. Y. Halpern, and L. Li, "Gossip-based ad hoc routing," *IEEE/ACM Transactions on Networking (ToN)*, vol. 14, no. 3, pp. 479–491, 2006.

[25] M. Jelasity, S. Voulgaris, R. Guerraoui, A.-M. Kermarrec, and M. Van Steen, "Gossip-based peer sampling," *ACM Transactions on Computer Systems (TOCS)*, vol. 25, no. 3, p. 8, 2007.

[26] E. Riviere and S. Voulgaris, "Gossip-based networking for internet-scale distributed systems," *E-Technologies: Transformation in a Connected World*, pp. 253–284, 2011.

[27] B. K. Subhabrata, E. Krishnamurthy, S. Sen, Y. Zhang, and Y. Chen, "Sketch-based change detection: Methods, evaluation, and applications," in *Internet Measurement Conference*, 2003, pp. 234–247.

[28] V. Karamcheti, D. Geiger, Z. Kedem, and S. Muthuskrishnan, "Detecting malicious network traffic using inverse distribution of packet contents," in *Proceedings of the workshop on Mining Network Data (MineNet) co-located with ACM SICOMM*, 2005.

[29] A. Lakhina, M. Crovella, and C.Diot, "Mining anomalies using traffic feature distributions," in *Proceedings of the ACM SIGCOMM*, 2005.

[30] E. Anceaume, Y. Busnel, and S. Gambs, "Uniform and Ergodic Sampling in Unstructured Peer-to-Peer Systems with Malicious Nodes," in *Proceedings of the 14th international conference on Principles of distributed systems (OPODIS)*, vol. 6490, 2010, pp. 64–78.

[31] P. Flajolet and G. N. Martin, "Probabilistic counting algorithms for data base applications," *Journal of Computer and System Sciences*, vol. 31, no. 2, pp. 182–209, 1985.

[32] D. M. Kane, J. Nelson, and D. P. Woodruff, "An optimal algorithm for the distinct element problem," in *Proceedings of the Symposium on Principles of Databases (PODS)*, 2010.

[33] N. Alon, Y. Matias, and M. Szegedy, "The space complexity of approximating the frequency moments," in *Proceedings of the 28th annual ACM symposium on Theory of computing (STOC)*, 1996, pp. 20–29.

[34] M. Charikar, K. Chen, and M. Farach-Colton, "Finding frequent items in data streams," *Theoretical Computer Science*, vol. 312, no. 1, pp. 3–15, 2004.

[35] A. Chakrabarti, G. Cormode, and A. McGregor, "A near-optimal algorithm for computing the entropy of a stream," in *ACM-SIAM Symposium on Discrete Algorithms*, 2007, pp. 328–335.

[36] E. Anceaume and Y. Busnel, "An information divergence estimation over data streams," in *Proceedings of the 11th IEEE International Symposium on Network Computing and Applications (NCA)*, 2012.

[37] E. Anceaume, Y. Busnel, and S. Gambs, "Ankle: Detecting attacks in large scale systems via information divergence," in *Proceedings of the 9th European Dependable Computing Conference (EDCC)*, 2012.

[38] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[39] G. Cormode, S. Muthukrishnan, and K. Yi, "Algorithms for distributed functional monitoring," in *Proceedings of the 19th annual ACM-SIAM Symposium On Discrete Algorithms (SODA)*, 2008.

[40] C. Arackaparambil, J. Brody, and A. Chakrabarti, "Functional monitoring without monotonicity," in *Proceedings of the 36th ACM International Colloquium on Automata, Languages and Programming (ICALP)*, 2009.

[41] P. B. Gibbons and S. Tirthapura, "Estimating simple functions on the union of data streams," in *Proceedings of the Thirteenth Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, 2001, pp. 281–291.

[42] Z. Haung, K. Yi, and Q. Zhang, "Randomized algorithms for tracking distributed count, frequencies and ranks," in *Proceedings of 31st ACM Symposium on Principles of Database Systems (PODS)*, 2012.

[43] Z. Liu, B. Radunovic, and M. Vojnovic, "Continuous distributed counting for non-monotonic streams," in *Proceedings of 31st ACM Symposium on Principles of Database Systems (PODS)*, 2012.

[44] G. Perrucci, F. Fitzek, G. Sasso, W. Kellerer, and J. Widmer, "On the impact of 2g and 3g network usage for mobile phones' battery life," in *Procs. of the 15th European Wireless Conference (EW 2009)*, 2009, pp. 255–259.

[45] A. Doan, R. Raghu Ramakrishnan, and A. Y. Halevy, "Crowdsourcing systems on the world-wide web," in *Commun. ACM 54, 4, 86-96*, 2011.

[46] A. Boutet, D. Frey, R. Guerraoui, and A.-M. Kermarrec, "Whatsup: News, from, for, through, everyone," in *Peer-to-Peer Computing (P2P), 2010 IEEE Tenth International Conference on*. IEEE, 2010, pp. 1–2.

[47] C. A. Middleton and A. Bryne, "An exploration of user-generated wireless broadband infrastructures in digital cities," *Telematics and Informatics*, vol. 28, no. 3, pp. 163–175, 2011.

[48] A. Osterwalder and Y. Pigneur, *Business Model Generation: A handbook for visionaries, game changers and challengers*. Wiley, 2011.

[49] A. Berlanga, P. Sloep, L. Kester, F. Brouns, P. Van Rosmalen, and P. Koper, "Ad hoc transient communities: towards fostering knowledge sharing in learning networks," *International Journal of Learning Technology*, vol. 3, pp. 443–458, 2008.

[50] R. O. Mason, "Four ethical issues of the information age," *MIS Quarterly*, vol. 10, no. 1, pp. 5–12, 1986.

[51] J. Parish, "Papa knows best: Principles for the ethical sharing of information on social networking sites," *Ethics and Information Technology*, vol. 12, pp. 187–193, 2010.

[52] E. Commission, "Regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation)," 2012.

[53] D. Davenport, "Anonymity on the internet: why the price may be too high," *Commun. ACM*, vol. 45, no. 4, pp. 33–35, Apr. 2002.

[54] "The metropolis model a new logic for development of crowdsourced systems," *Commun. ACM*, vol. 52, no. 7, pp. 76–84, Jul. 2009.