

Participatory sensing in policy modelling: a complex systems view.

Vittorio Loreto^{1,2}, Andrea Capocci¹ and Vito D.P. Servedio¹

¹Dipartimento di Fisica, Sapienza Università di Roma,
Piazzale A. Moro 5, 00185 Roma, Italy

²ISI Foundation, Viale S. Severo 65, 10133 Torino, Italy

There is now overwhelming evidence that the current organisation of our economies and societies is seriously damaging biological ecosystems, social structures and human living conditions in the very short term, with potentially catastrophic effects in the long term. A lot can and must be done from the technology and policy point of view. However, it is only when people become fully aware of their particular situation and its future consequences that the needed behavioural changes will truly happen. With a growing realisation that only through bottom-up actions we can deal with today's challenges, there is an urgent need to create an ICT fabric that can support the local actions of citizens by supporting collaborative monitoring, exposing actionable local information, and enabling an evidence-based dialogue among stakeholders. The idea is that the availability of locally-relevant digital data, together with their analysis, processing and visualization should trigger a bottom-up improvement of social strategies. It is important to combine digital technologies to gather data and opinions with established and novel theoretical methods to analyze them, with the ultimate goal of providing real-time, user-centered results through standard and largely available communication networks. The integration of participatory sensing with the monitoring of subjective opinions is novel and crucial, as it exposes the mechanisms by which the local perception of, say, an environmental issue, corroborated by quantitative data, evolves into socially shared opinions, and how the latter, eventually, drive behavioural change. Enabling this level of transparency critically allows an effective communication of desirable strategies to the general public and to institutional agencies.

Introduction

The latest evolution of Information and Communication Technologies (ICTs) has increasingly concerned the inclusion of users in the production of information. Nowadays, users are not only able to exchange messages, images and sounds with other individual peers, but also with whole communities whose size can be tuned by the

user him/herself. Moreover, the digital paradigm has allowed the integration of multiple information and communication sources, including connected PCs, phones and cameras. Accordingly, the distinction between consumers and producers of information, typical of a past era dominated by newspapers and television, is vanishing.

Such an interconnected communication network has dramatically enlarged the access to information sources with an undeniable advantage for the citizenship. At the same time, it is presenting new challenges. Information broadcast by uncontrolled sources could overload the network with noisy signals, preventing the meaningful ones from being received by the desirable recipients. As a consequence, users' attention could be exhausted by useless information.

To overcome the obstacles to the usability of the increased amount of data, a number of technologies have been developed. More sophisticated communication platforms have emerged - up to current Web2.0 social networks accessible from PCs and cell phones - where users have been given the opportunity of collectively categorizing and evaluating the content they browse, providing the community with an efficient information filter. The classification of digital resources is typically performed by assigning labels (called *tags*) or scores to resources. This collaborative categorization has given birth to several web-based *folksonomies* (from "folks" and "taxonomies"). Consequently, the most popular websites now incorporate some sort of collaborative categorization tools. These socio-semantic systems have also attracted much attention from the scientific community, to investigate quantitatively how cooperative phenomena arise and can be harnessed to improve the performance of such collective tasks (see [Mathes, 2004], [Quintarelli, 2005], [Golder, 2006] and the work accomplished by the authors, among others, in the framework of the EU project TAGora [TAGora]).

This ICT infrastructure has been applied not only to favour data exchange among people, but also to outsource productive tasks. Companies and institutions are increasingly relying on the recruitment of networking volunteers through the Internet to perform tasks. The main difference between this "crowdsourcing" and traditional labour markets lies in the absence of prearranged duties that workers deal with an idiosyncratic effort, while the infrastructure takes care of summing up all contributions despite the heterogeneity and number of users. First examples have concerned highly specialized tasks, such as open source software development or scientific programs [Benkler, 2002] that could be broken into smaller operations performed by uncoordinated volunteers. More recently, the range of activities being crowdsourced has expanded and forms a world-wide labour market facing tasks proposed by various agencies and companies asking, e.g., for technological and marketing solutions, or by research groups looking for volunteers for test and data mining activities [Brabham, 2008; Kittur, 2008]. These kinds of infrastructures, therefore, are particularly appropriate for the involvement of citizens in distributed sensing experiments.

Pervasive computing and participatory sensing

Devices employed to get connected to communication networks have converged in size and technological standards. Cell phones have integrated many functions traditionally accomplished by personal computers. In turn, computer manufacturers have privileged products designed for an easy mobile usage, featuring low-weight and low-cost, albeit with limited computing power. Moreover, cell phones and PCs incorporate sensors of increasing accuracy: GPS sensors, cameras, microphones, accelerometers, thermometers are already a default equipment in most of the mentioned devices. Networks have also accompanied this process, by expanding the availability of an Internet connection throughout daily life.

Thus, users can now easily form dedicated networks providing data that monitor particular issues. Such sensing networks can be of opportunistic or participatory type. In the first type, data are provided by monitoring devices collecting data autonomously, with no personal commitment from the user [Campbell, 2006]. For examples, GPS sensors continuously track the position and the displacement of users or vehicles to collect traffic information. Participatory sensing, instead, focuses on motivated groups willing to engage themselves as digital pressure groups, e.g. in monitoring the quality of a metropolitan environment [Steels, 2008; Paulos, 2007] or proposing urban plans for redeveloping areas [Burke, 2006].

The participation of users in the monitoring affects both the resolution and the quality of the data collected. Traditional sensing generally involves a small number of highly controlled observation points. The low spatial resolution of the data gathered in this way is compensated by the high data quality certified by the controlling agency. On the other hand, distributed sensing relies on the possibility of gathering large amounts of data from many uncontrolled sources, which cannot ensure high data quality standards; however, by means of statistical methods together with the possibility of storing and post-processing large datasets, this quality gap with respect to traditional sensing can be overcome.

Reasonably, users provide larger quantities of data if the observed phenomenon and its management directly concern the community involved in participatory sensing experiments. For example, people are interested in reporting meteorological observations in order to improve existing models and receive more accurate weather forecasts, and this, as a virtuous feedback, could be a reason for a citizen to provide more data to meteorological centres (see, for an example, the website www.ilmeteo.it). Moreover, large communities allow the monitoring of a wider range of situations. As a general rule, the larger the number of participants, the better the monitoring. However, the number of users involved in a participatory sensing experiment is often unknown in advance, as participants are free to involve in, or discontinue, the sensing activity. A low required effort and an efficient feedback mechanism are crucial in encouraging the participation. Users should benefit from the participation even when their number is not large enough; otherwise, a too large critical mass, i.e. the mini-

mum number of participants needed to self sustain the feedback mechanism, would be required and could never be reached so to make the experiment fail [Lane, 2008].

Data of dubious quality may not only be the result of inaccurate sensors. Data can be affected by involuntary or malicious sensor misuse, and data flows may vary strongly according to the individual effort provided by users. Therefore, infrastructures assisting the social interaction should be able to detect biases, filter data, aggregate them and extract meaningful information even from a very noisy data set. As we explain in the following, the knowledge of the underlying social interaction is crucial for such a task.

The environmental monitoring represents a very interesting area to be explored by participatory sensing. For such an issue, the involvement of individual citizens is crucial for a number of reasons: the environment quality is strongly affected by the behaviour of individuals in their most ordinary daily situations; citizens' behaviour, in turn, depends on their awareness; many bad environmental practices arise when citizens do not coordinate in order to attain a global optimal usage of collective resources, but rather pursue their own profit selfishly - resulting in an even worse long term individual performance. For these reasons, the application of a novel ICT-based sensing framework may have a stronger impact here than in other areas, and we will refer to this one in most of the examples cited in the following.

Empirical data and subjective opinions

Along with sensors, human themselves can act as a probe to monitor many phenomena, especially in the environmental area. In fact, the overall assessment of a situation should often take into account many parameters resulting in a very complex quality landscape. The human perception, synthesized by opinions, can be considered as a probe of such landscape. Thus, the opinion of a user often conveys relevant information, although it is influenced by subjective biases.

The comparison of sensor data and opinions has a twofold importance. On one hand, it allows to understand how users perceive combinations of multidimensional observations: which of the environmental characteristics (temperature, air quality, noise pollution etc.) has the stronger impact on their perception? On the other hand, the knowledge of both the environmental conditions and the social network a user has been exposed to, allows estimating how much social biases affect his/her perception of the quality of the environment and individual behaviour. Detecting the opinion leader in social networks, spotting the imitation mechanisms at work and the inertial effects as opposed to opinion shifts, is crucial if one seeks not only to monitor the existing practices, but also to induce better ones. At this aim, so called "sociophysics" has developed many tools and models to study the opinion dynamics taking place on social networks. This interdisciplinary field employs concepts borrowed from the

theory of complex systems in statistical physics. Statistical physics has proven to be a very fruitful framework to describe phenomena outside the realm of traditional physics [Castellano, 2009]. The last years have witnessed the attempt by physicists to study collective phenomena emerging from the interactions of individuals considered as elementary units in social structures: from opinion, cultural and language dynamics to crowd behaviour, hierarchy formation, human dynamics, social spreading. In all these social phenomena the basic constituents are not particles but humans and every individual interacts with a limited number of peers, usually negligible compared to the total number of people in the system. In spite of that, human societies are characterized by stunning global regularities [Buchanan, 2007]. There are transitions from disorder to order, like the spontaneous formation of a common language/culture or the emergence of consensus about a specific issue. It may be surprising, but the idea of a physical modelling of social phenomena is in some sense older than the idea of statistical modelling of physical phenomena. The discovery of quantitative laws in the collective properties of a large number of people, as revealed for example by birth and death rates or crime statistics, was one of the factors pushing for the development of statistics and led many scientists and philosophers to call for some quantitative understanding (in the sense of physics) on how such precise regularities arise out of the apparently erratic behaviour of single individuals. Hobbes, Laplace, Comte, Stuart Mill and many others shared, to a different extent, this line of thought [Ball, 2004].

Data gathering, analysis and validation

Systems' modelling relies on large-scale data structures but these ones are often inaccessible or not envisaged as important until a main event occurs. Systems modelling will rely more in the future, on forms of data gathering involving individual agents moving across system domains. New ways of gathering and communicating data, enabled by ICT, produce new forms of involving the public. Sensor-based gathering of temperature and noise-level information, for example, allows collection of data on totally new scales. Use of mobile phones for this purpose seems a particularly powerful way of getting ordinary people involved, as it could integrate subjective data (moods, opinions) as well as scientific readings. The World Wide Web provides several tools, such as collaborative systems (e.g., del.icio.us), micro-blogs (e.g., Twitter), and other so-called Web 2.0 services to gather opinions in a user-friendly manner.

It is possible to make more sense of the collected data when it is displayed over a base map of the local streets either via GPS readings or by captures through a map interface. Data gathered in this way could, if socially accepted, induce widespread opinion dynamics leading to changes in behaviour. The idea is that the availability of locally relevant digital data, together with their analysis, processing and visualization should trigger a bottom-up improvement of social strategies. On the other hand, the augmented awareness could be a source of pressure on the relevant stakeholders and policy makers. Data are of course relevant also directly for policy makers. Every policy ought to be tested with data. While there is indeed an obsession in governments

with assessing their policies with data (impact assessment), there is a problem of gathering data on the right level and of the right type. Often there is a mismatch of scale and type. Here comes the issue of data validation and interpolation. Tools and techniques able to cope with huge sets of heterogeneous and often unreliable data to efficiently reconstruct dynamic system models at multiple levels are crucially needed. This includes data-rich probing technologies, protocols and experiments to gain realistic data on what goes now under the denomination of techno-social systems. A techno-social system, in this sense, acts like a “lens” that captures information from the environment: one has to explore the peculiarities of having human agents as sensing nodes, the role of noise sources at different scales, the effect of opinion bias, information spreading in the community supporting the techno-social system, network effects, and so forth. More generally, reliable data play a crucial role also in refinement of models as Science looks at the available data and stimulates model corrections (see for example the modelling of climate change at the beginning of the 1990s where a mismatch between models and data led to introducing aerosols into the equations that led to a far better match).

Modelling and predictability

The modelling activity is crucial to reach a deep theoretical and pragmatic understanding of social phenomena [Castellano, 2009]. When coupled with a serious data analysis activity devoted to the discovery of emergent features, it can result in a virtuous loop, where measures inspire models, model analysis suggests new measures and observations, which in turn allow the evaluation and refinement of models. Once a satisfactory level of agreement between theory and experiments is achieved, the theoretical description can suggest and inspire control strategies and directions for improving systems.

The modelling and the simulation of such multi-level systems, should take into account the relevant technological, psychological and social dimensions as well as the realistic diversity of behaviours, social and spatial structures and knowledge. The theoretical foundations for understanding and modelling the behaviour of such systems lie in uncovering the basic interactions between the user and the ICT system, as well as the interactions between users mediated by the ICT system. Realistic models of these interactions are still lacking in a validated form grounded on experimental data. Not only the technological aspects of the ICT platform, but also the psychological and cognitive factors come into play at this level, together with the social structure of the community and the spatial structure of the environment where users act. It is important to provide theoretical foundations for the dynamical aspects, grounding theoretical constructions on data from real systems and exploring the space of possible behaviour by means of computer simulations.

One of the main objectives of the modelling activity is that of coming up with a

notion of predictability for socio-technological systems. Several aspects are relevant here where the notion of predictability can be investigated. (i) **Inertia and critical mass:** An important aspect of the predictability in techno-social systems is related to the individual inertia, i.e., the resistance of an individual in changing his/her opinion and more generally his/her habits. The individual inertia, on its turn, will generally depend on the pressure exerted by the environment and by peers. It is thus important to investigate whether critical thresholds (critical mass) exist for triggering an opinion change and how these thresholds depend on the peer pressure or other social factors. (ii) **Response to a perturbation:** Another crucial aspect to assess the predictability of a generic system is its response function to external perturbations, e.g., a specific policy change; (iii) **Scale effects:** An additional possible perspective of the notion of predictability is to consider the role played by the system size. The question can be posed as follows. Suppose one has observed a given phenomenology in a small community, how much of the acquired knowledge can be transposed to a larger (sometimes much larger) community? This is a typical problem in statistical physics for which a lot of tools and methodologies are currently available.

Dissemination

A proper dissemination is a crucial aspect of the whole concept of participatory sensing. The goal is to raise awareness about the long-term benefits that can be reached with a self-sustained feedback mechanism, involving the public, the scientific and technological communities, and crucially, policy makers, relevant stakeholders and governmental organisations. In order to raise a larger public awareness into the capabilities of present day, particularly in young generations, so to reach the above-mentioned critical mass, it is necessary to develop demonstrators, to start up case studies, and to make a massive use of international popular and scientific press. Most importantly, the creation of on-line social communities, their interaction with the collected and processed information and a direct bridge with policy institutions can be largely favoured by the use of the Internet, at rather low costs.

2020 Applications: environment, healthcare, politics

The framework described above can be applied and have been indeed already applied to many areas besides the environmental one. For example, applications of distributed sensing are now being experimented in the health care systems, in order to gather real-time data to monitor patients' status (e.g., by measurement of health parameters) and send information to them (e.g. reminders and advises on suitable therapies) [DietSense]. Also, participatory sensing experiment are being set up to collect data from individuals monitoring the diffusion of potentially pandemic diseases [influeweb], providing earlier alerts about an imminent outbreak.

Policy makers, however, will be mostly interested in the application of participatory sensing techniques. Gathering in a distributed way the effects of simulated or actual policies will provide a more accurate picture of the concerned communities. But, most important, the knowledge of the underlying social networks will enlighten how opinions get formed and how social ties can be harnessed to induce changes in opinions and setting on the desired behavioural shifts. A wider use of participatory tools for environmental monitoring will generate a more demanding citizenship and evidence-based policies taking into greater account the social acceptance of proposed or actual plans, at the benefit of the whole community.

Bibliography

- [Mathes, 2004] A. Mathes, Folksonomies - Cooperative Classification and Communication Through Shared Metadata, <http://www.adammathes.com/academic/computer-mediatedcommunication/folksonomies.html>
- [Quintarelli, 2005] E. Quintarelli, Folksonomies: power to the people, ISKO Italy-UniMIB meeting
- [Golder, 2006] S.A. Golder and B. A. Huberman, Usage patterns of collaborative tagging systems, *Journal of Information Science* 32, 198 (2006)
- [TAGora] For a complete list of publications resulted from the EU project TAGora, see <http://www.bibsonomy.org/tag/tagorapub>
- [Benkler, 2002] Y. Benkler, Coase's penguin, or, Linux and The Nature of the Firm., *Yale Law Journal* 369, 112, 2002
- [Kittur, 2008] A. Kittur, E.H. Chi and B. Suh, Crowdsourcing user studies with Mechanical Turk, *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 2008
- [Brabham, 2008] D.C. Brabham, Crowdsourcing as a model for problem solving, *Convergence: The International Journal of Research into New Media Technologies* 14, 75 (2008)
- [Campbell, 2006] A.T. Campbell, S.B. Eisenman, N.D. Lane, E. Miluzzo, R.A. Peterson, People-centric human sensing, *WICON '06: Proceedings of the 2nd annual international workshop on Wireless internet*, 2006
- [Paulos, 2007] E. Paulos, R. Honicky, and E. Goodman. Sensing atmospheres. In *Sensing on Everyday Mobile Phones in support of participatory research at ACM SenSys 2007*, 2007.
- [Steels, 2008] L. Steels and E. Tisselli, Social Tagging in Community Memories. *Proceedings of the AAAI Spring Symposium on Social Information Processing*, AAAI
- [Burke, 2006] Burke, J., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., and Srivastava, N.B. Participatory sensing. In *Proceedings of the World Sensor Web Workshop, ACM SENSYS* (Boulder, CO, 2006)

[Lane, 2008] N.D. Lane, S.B. Eisenman, M. Musolesi, E. Miluzzo, Emiliano and A.T. Campbell, Urban sensing systems: opportunistic or participatory? HotMobile '08: Proceedings of the 9th workshop on Mobile computing systems and applications, 2008

[Castellano, 2009] C. Castellano S. Fortunato and V. Loreto, Statistical physics of social dynamics Rev. Mod. Phys 81, 591 (2009). [arXiv:0710.3256]

[Buchanan, 2007] Buchanan, M., 2007, The social atom (Bloomsbury, New York, NY, USA).

[Ball, 2004] Ball, P., 2004, Critical Mass: How One Thing Leads to Another (Farrar, Straus and Giroux, London, UK).

[DietSense] <http://urban.cens.ucla.edu/projects/dietsense/>

[influweb] <http://www.influweb.it/>