Procedia Computer Science

Volume 51, 2015, Pages 542–551

ICCS 2015 International Conference On Computational Science

Collaborative Knowledge Fusion by Ad-Hoc Information Distribution in Crowds

George Kampis[1,2] and Paul Lukowicz[3]

[1] DFKI, German Research Insititute for Artificial Intelligence, Kaiserslautern, Germany
george.kampis@dfki.de
[2] ITMO University, St. Petersburg, Russia
[3] DFKI, German Research Insititute for Artificial Intelligence, Kaiserslautern, Germany
paul.lukowicz@dfki.de

Abstract

We study situations where (such as in a city festival) in the case of a phone signal outage cell phones can communicate opportunistically (for instance, using WiFi or Bluetooth) and we want to understand and control information spreading. A particular question is, how to prevent false information from spreading, and how to facilitate the spreading of useful (true) information? We introduce collaborative knowledge fusion as the operation by which individual knowledge claims are "merged". Such fusion events are necessarily local, e.g. happen upon the physical meetings of knowledge providers. We study and evaluate different methods for collaborative knowledge fusion and study the conditions for and tradeoffs of the convergence to a global true knowledge state under various conditions.

Keywords: knowledge fusion, agent based simulation, crowd control, embedded intelligence, large-scale systems

1 Introduction

Managing large crowds of people at public events and in urban areas is a complex and highly dynamic problem. This involves an early detection of potential events of common relevance (including security events) and the means to communicate these events with the crowd in an efficient and timely manner. Over the past years, we have developed a smart-phone based crowd management system, which analyzes smartphone sensor data voluntarily contributed by visitors of public events and creates a real time overview about crowd conditions [5, 3]. The system was tested in various real life situations such as the Zurich festival of 2013 [3]. A key concern that has emerged during the above deployments was the ability to deal with network outages.

In recent work [1] we thus follow a basic opportunistic networking approach by making use of the smartphones' built-in WiFi hotspot functionality which in combination with the

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.319

devices switching between access point and client modes facilitates the propagation of messages on a multi-hop basis in such a case. We presented a large-scale simulation [4] based on a dataset consisting of movement traces from 28.000 people (recorded at the Zurich city festival mentioned above). Using our simulation, the influence of various parameters on the system was investigated and the random mode switching strategy could be optimized.

In this paper we go an important step further and ask, by using further large-scale simulations, what happens if information has no single entry point but individual persons start spreading their own messages. We want to study the possibility of self-organization of information in such systems into coherent global knowledge states but also the possibility of information control. Messages may contain wrong or false-alarm information, and it can be vital to make sure that such messages do not spread into the entire system, whereas in cases where information is useful we want to facilitate spreading. We suggest to grasp an essential factor in the "quality of information" (such as its truth content) and we will examine cases where information with different quality is locally introduced in the system. We study how that information can spread but also how the spreading depends on the rules of a local interaction.

The concept that makes the idea work is that of *collaborative knowledge fusion*. This will be understood as the operation by which individual knowledge claims are "merged" so that the fusion events are local, e.g. they happen upon the physical meetings of knowledge providers. We study and evaluate different methods for collaborative knowledge fusion and analyze their consequences, and study the conditions for convergence to a global true knowledge state in the function of various conditions. We introduce and study 3 different local algorithms for collaborative knowledge fusion and compare them for adequacy and efficiency.

Note that the notion of "quality of information" as used here is quite general and goes significantly beyond crowd safety and control, opening a way towards advanced *smart city* applications where crowdsourced information is used to build and maintain a global distributed knowledge base of any kind, such a movie or restaurant database etc. [6, 2].

2 Information Spreading Without a Signal

Modern smartphones have a built-in WiFi-hotspot functionality that can be used if signals are out. The underlying principle is that, if a certain amount of devices take over the role as hotspots and the others are trying to connect to those hotspots, information can be exchanged between the connected devices via the micro-network of the hotspot. If these modes are switched periodically, then information can be propagated to all participating devices (a recent implementation is FireChat [8]).

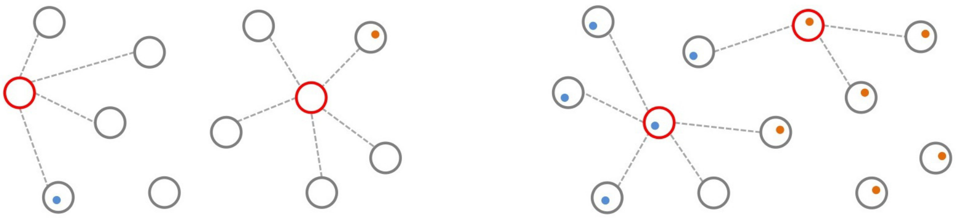


Figure 1: Spreading without a signal: opportunistic communication among mobile devices. Left: initial position (dots denote messages). Right: position after one iteration.

In [1] we introduced a mode switching strategy based on node mobility. We showed that using this strategy, information is spread with a low amount of devices acting as access points, and at a greater speed compared to a random mode switching. Using our simulation model, various crowd situations could be studied and different communication algorithms tested.

However, such mode switching opportunistic communication heavily relies on a *trust* about the functioning of the nodes. Every node is assumed to be an idealized agent that communicates information faithfully and does not start to spread own messages of unknown quality (for the quality of information, see below). Hence, a critical next question is to study similar opportunistic communication systems but where these overly optimistic assumptions are relaxed. That will be the goal of the subsequent sections.

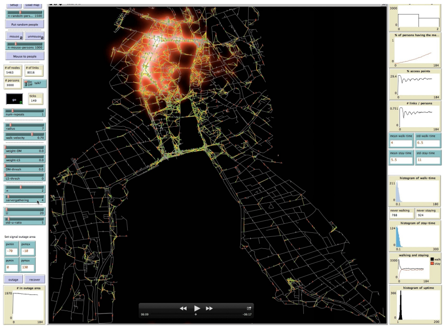


Figure 2: Simulation using a GIS shape file for structuring pedestrian motion (Zurich example).

3 Knowledge Fusion to Collaborative Knowledge Fusion

Knowledge fusion (KF) is an operation by which different knowledge claims from multiple information sources are merged to obtain a single knowledge claim with optimal properties. The best known example is based on using binary classifiers (where knowledge is encoded as the probability p of a $0-1$ classification). Binary or binomial classification is the problem of classifying (partitioning) the elements of a given set into two groups on the basis of a classification rule. Classification decisions are thus the simplest case for representing knowledge, and making yes/no decisions is a ground case.

The KF problem for binary classifiers, simplified, sounds as this: given n different binary classifiers, define a new one which is at least "as good" as the best among the n. In other words, KF is a (usually variable) mapping from $(p_1, p_2, \ldots p_n)$ to p' where p' marks the classification by the new synthetic classifier.

Traditional KF assumes informational completeness (i.e. that the complete set of alternative knowledge claims is permanently available and accessible). Also, no *a priori* limits on computational resources are set – although KF methods are usually learning methods with a carefully selected computational complexity, yet this does not exclude that in practice they can be computationally intensive. However, for this kind of KF off-the-shelf methods exist [7] that ensure fast and efficient computation in the face of problem size n as well as optimality in terms of the quality of obtained fusion knowledge.

By contrast, we suggest studying collaborative knowledge fusion (CKF) here. CFK is understood [1] as a version of KF where fusion events are local, e.g. happen upon the meetings of individual knowledge providers, and global fusion happens due to the collective (hence "collaborative") interaction dynamics. As there can be many individuals in a population, we are also facing the problem of information propagation under the ad-hoc interactions. CKF is incremental in that it proceeds forwards in a step-by-step fashion and the entire information set is never simultaneously available. In our context we will furthermore assume that only the simplest local one-step computations are permitted (such as addition, rewriting, etc.) as is typical for embedded systems of low power and compute resources. Because individual interactions form its basis, a local interaction rule replaces the global learning rule.

It should be upfront clear that in general there can be no question of optimality for CKF as understood here. Interaction rules typically use simple heuristics, and not optimality: thus when analysing CFK systems, we are in fact looking for a good (or best) heuristics for the local rules.

A key notion that will be helpful is the *quality of knowledge state*. This notion measures collaborative fusion. For the purposes of our analysis, it will be assumed that there always exists a "correct decision", known to the experimenter yet not known to the agents - a useful assumption in the training and analysis phase to monitor information spreading (to be abandoned in the test phase where a chosen interaction rule is exploited). Without restriction of generality we will assume that the correct decision is always 0 and that $p = 1$ implies the correct decision - for symmetry reasons, choices are equivalent. To characterize the quality of the knowledge state of an individual agent, we simply compare its decision (made on the basis of its p value) with the correct decision (i.e. 0). To characterize the quality of the knowledge state of an entire population, we take each agent's decision and compare their average with the right decision. Here we assume that an agent's decision will be 0 if its p value is strictly larger than 0.5 - remember that by construction, p was the probability of the zero decision. The quality of knowledge Q of the population is thus a real number between 0 and 1. The *average decision error* of a population will be defined as $1 - Q$.

4 Well-Informed Agents

As mentioned above, the rationale of our approach is to study situations where (such as in a city festival) in a signal outage cell phones can communicate opportunistically by WiFi or Bluetooth, and we want to understand and control the information spreading. A particular question is, how to prevent false information from becoming ubiquitous, and how to facilitate the distribution of true and useful information?

We will assume in the following that some (zero or nonzero proportion) of the agents can be "well-informed" (well-informed agents, WIAs), and remain so all along in a given simulation run. A well-informed agent is one that makes the right decision at $p = 1$ and also "knows" (represents) that that is the right information[1]. While WIAs also undergo meetings and interaction, and thus enter the exchange of information, we will nevertheless assume that the WIAs' knowledge is never updated (based on the fact that the WIAs' knowledge is already known to be perfect, and cannot get any better). WIAs are, therefore, reliable and permanent sources of information - introduced by hand at the beginning or in an emergent fashion dynamically.

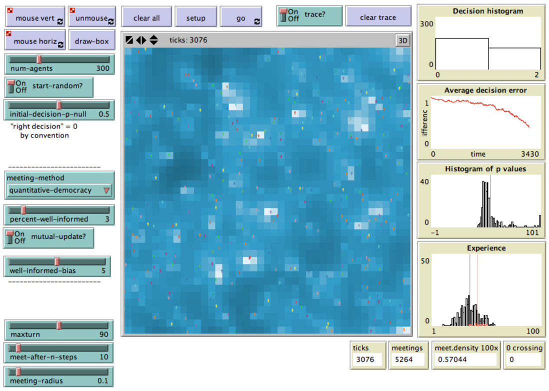
Intuitively, WIAs play an important role in the knowledge dynamics and influence convergence towards high quality knowledge states in CKF – an intuition to be tested in the subsequent simulations. Can WIAs stop the spreading of false alarms in a crowd to avoid panic and catastrophy? If so, how many WIAs will be needed, and where lie the tradeoffs, if any?

[1] This is similar to the famous KK rule in the modal logic of knowledge by Hintikka.

5 Incremental Fusion: Three Model Variants

We have developed a large-scale agent based simulation framework that forms the basis of the following experiments (referred to hereinafter as "the model"). The model uses a population of agents (of size N) that navigate in 2-space. Each agent possesses a position (known to the observer) and performs random Brownian motion allowed by the topology (torus, square, or a walled structure of streets or offices - our interactive model framework offers tools for hand-drawing the latter, or for importing a maze map or a GIS-based street shapefile). When an agent moves, its position is updated accordingly. The agents' meeting is understood as an event when agents are found in the vicinity of each other within a meeting radius d. (Note that at $d = 0$ the agents can no longer meet, positions being represented as real numbers that have a negligible chance to match up exactly). Upon meeting, agents exchange information.

Information is represented in the agents as decision probability p (as discussed above). The model is static: there always exists a fixed "correct decision", and this decision stays unchanged within a run. Agents further possess a parameter called *experience* which is the sum of prior meetings (i.e. the amount of exposure to other agents and their knowledge claims).



num-agents 300
On Off start-random?
initial-decision-p-null 0.5
"right decision" = 0 by convention
meeting-method quantitative-democracy
percent-well-informed 3
On Off mutual-update?
well-informed-bias 5
maxturn 90
meet-after-n-steps 10
meeting-radius 0.1

ticks	meetings	meet.density 100x	0 crossing
3076	5264	0.57044	0

Figure 3: The model GUI for interactive exploration. The model is written in NetLogo and available from `http://ccl.northwestern.edu/netlogo/models/`

The model comes with a user interface for interactive exploration, selecting model versions from a roll-down menu. We tested the model using 3 different local interaction rules (meeting methods) called (i) *quantitative democracy*, where knowledge is averaged upon a meeting (ii) *experience takes all*, where a more experienced agent (such as a teacher) overwrites the prior knowledge of the less experienced one (the "student"), and (iii) *transitive experience* where not only the knowledge but also experience itself is handed over to the meeting partner (imagine this such that a student of a guru becomes a guru). In what follows we describe and analyze each of these model versions with tools that include extended parameter studies. For the algorithm specification we refer to the self-documenting code available with the model.

Simulations were tested using the following parameters and intervals (Table 1). "Mutual update" here means whether only the active agent (which is a well-defined entity in an agent based simulation) is updated at a meeting or both meeting parters are: the model permits either, and here we assume the second. In terms of calibration, we can conveniently think of a time step as 0.1s of real time.

All simulation runs were performed from $t = 0$ to $t_{max} = 10,000$ steps and each parameter combination was tested using 10 different runs. When zero crossing did not happen, time was

parameter	**value**	**range**
population N	300	fixed
meeting radius r	0.1	fixed
mutual update	yes	fixed
initial decision $p(0)$	0.1	0.1-1.0
percent of WIAs	1	1-10
WIAs' experience bias	1	1-10

Table 1: Simulation parameters tested.

set to t_{max} to help visualization. Population number N was kept low for analysis: the system can be run on hundreds of thousands of agents representing real people.

5.1 Quantitative Democracy (QD)

Under this rule, when two agents meet, they compute and share the average of their knowledge values: $p_{new} = (p_1 + p_2)/2$.

Clearly, without any WIAs the population using this rule undergoes a random process towards a convergent end state. The naive expected value of the population's decision error is 0.5, but this exact value is never realized in an individual run. Instead, every individual run ends with 0 or 1 as the entire population's knowledge state flips after a certain time - these two values are equally probable in the absence of WIAs and the exact value of 0.5 has zero probability. (At the other extreme, if all agents are WIAs, the population obviously starts from a perfect knowledge state and remains there.)

Starting from a random population with some WIAs, how many WIAs does it take to get a "directed process", where the end value of the population's decision error is a guaranteed zero, independently from the initial knowledge quality? Again, it is easy to see that for this one a single WIA is enough, as there can be no equilibrium state until all agents converge to the WIA values. Agents no doubt meet sometime, sooner or later, with the available WIAs (whether there are many or just a few) and from this meeting they gain knowledge (i.e. improve their knowledge quality) whereas the WIAs are never compromised. Then the agents meet again with further WIAs and with further updated agents and so on – so the agents continually improve more and more, in an endless process that necessarily converges to the right decision at $p = 1$ (the WIA's value).

The characteristic zero crossing time (to use a statistical physics term, i.e. the time to a zero error), grows, however, dramatically with the decreasing values of WIA proportions. We studied this using parameter sweeps, to be discussed next. But before that note that there is a long-standing memory effect of the initial knowledge state of the population. If the initial knowledge state is entirely random (i.e. the expected value of p is 0.5 at t_0) then the convergence process is fast in the presence of WIAs. However, if the initial population is extremely badly informed ($p = 0$ for the 0 decision) then the convergence to the right decision can be significantly slower.

Figure 4 now shows the effect of WIAs (in the function of their increasing proportion) as well as the effect of the quality of initial knowledge state of the population. What we can observe is that that in combination they can lead to a fast convergence in QD to a high quality emergent knowledge state, and the increasing WIA proportions reduce the standard error of the convergence time. We further observe that the histogram of p values in the population (Figure 5) follows a unimodal distribution, i.e. the individual values gather around a mean value marked as a vertical grey line. The right column corresponds to WIAs at $p = 1$.

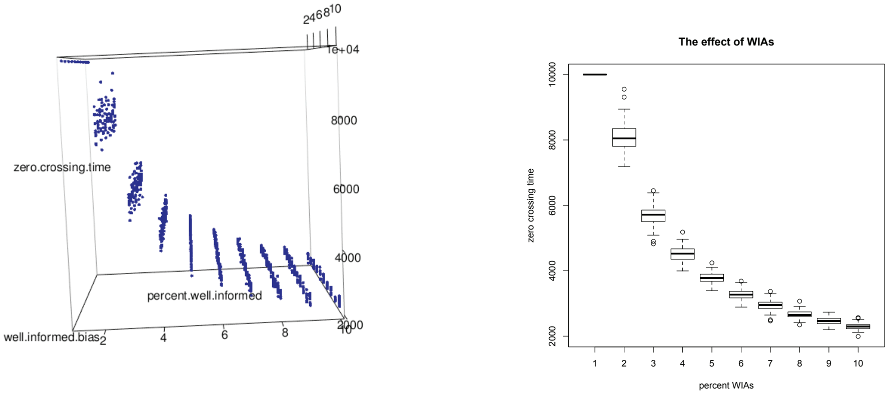


Figure 4: In QD, the introduction of more WIAs directs the process towards high quality knowledge states. Using the calibration, 4% WIAs yield a convergence in ca. 7 minutes.

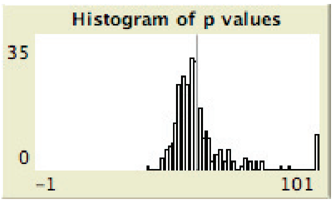


Figure 5: The dynamic histogram of binary classification probabilities under QD (to be compared with EXP and TRAN below).

5.2 Experience Takes All (EXP)

As a first step towards more realistic model versions, we tested an interaction rule "experience wins". A simple reputation system is thereby realized, where reputation is simply understood as the experience of prior meetings. Here the intuitive idea is that, when two agents meet, the naive agent takes its p value from the more experienced one, i.e. the one that had more prior meetings before and whose knowledge state thus accumulates information from all the prior meetings. We can imagine this as a teacher vs. student situation where the experienced partner (the teacher) hands out information to the less experienced (the student).

The teacher metaphor also helps us understand that, in order to distinguish WIAs from the other agents and to mark them as teachers, we need to add an initial bias or "experience" to them – i.e. an a priori trust or reputation at t_0 available to the WIAs to give them a kick-start and initial credibility.

Again, the first fews steps of the analysis are easy. If there are no WIAs or if the *WIAs have no bias* then "nothing interesting" can happen. Typically the system again converges to an average decision error around $p = 0.5$ – but not to a single coherent decision this time. The reason is the pullback effect of badly informed but highly experienced agents – informally we can call them *false prophets*. Unlike in QD where the p values form a distribution, here any agent with an arbitrary p value but a high experience can become a center for further spreading.

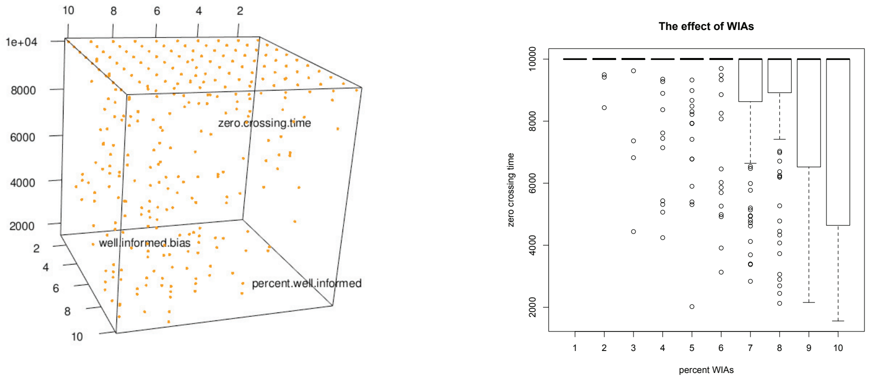


Figure 6: The effect of WIAs and "WIAs bias" in EXP. More WIAs reduce individual zero crossing time, but the average is unchanged: most runs do not converge to a singleton.

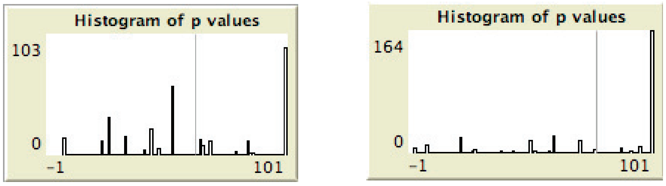


Figure 7: The dynamic histogram of binary classification probabilities under EXP (left) and TRAN (right) showing arbitrary centers of spreading. The right columns shows the effect of WIAs whose values are copied.

Similarly, then, in a less obvious case, where WIAs exist and an initial WIA bias is added, we experience that even at high values of this bias the head advantage of the WIAs can slowly evaporate due to the random effects of meetings and compensated by a higher number of meetings by some arbitrary, however badly informed agents. In other words, agents may for no systematic reason randomly accumulate higher experience up to a point that they continue to spread the wrong decision. As a result, the system can maintain a mixed knowledge state for a very long time. (Ultimately the entire system must collapse to a single knowledge state again for equilibrium reasons, since random effects drive it and $p = 0$ or 1 act as sinks or absorbing walls. But often the time needed to reach them is not on a practical scale.

Applying more WIAs and/or more bias may, of course, counterbalance this and drive the system to a right decision or drive it nearer: these quantitative effects are studied in parameter sweeps shown on Figure 6. It is seen that WIAs or their bias do not change the average process which shows no convergence in the studied time frame. Also on Figure 7 we see that instead of a continuous distribution, several single values dominate the system. These values can remain semi-permanent for a long time under EXP.

5.3 Transitive Experience (TRAN)

Finally, we developed a model variant informed by an idea borrowed from peer-to-peer (p2p) "gossip" systems [6, 2]. This new variant is the same as the experience-based model, with the important exception that reputation (i.e. experience) is transitive here: together with a knowledge claim, the experience accumulated in the lifetime of the agent is handed over to the less experienced partner upon a meeting. In other words, and to stay with the metaphor used above, by meeting with a teacher (and learning from her), we become teachers ourselves (and by meeting a grand master we become grand masters).

Parameter studies (Fig. 8) show TRAN to be significantly more depedable than EXP and converging to an end state in a rapid way. The system behaves much as in QD but even faster, however less predictably so: the cost is several outliers that render an individual application risky.

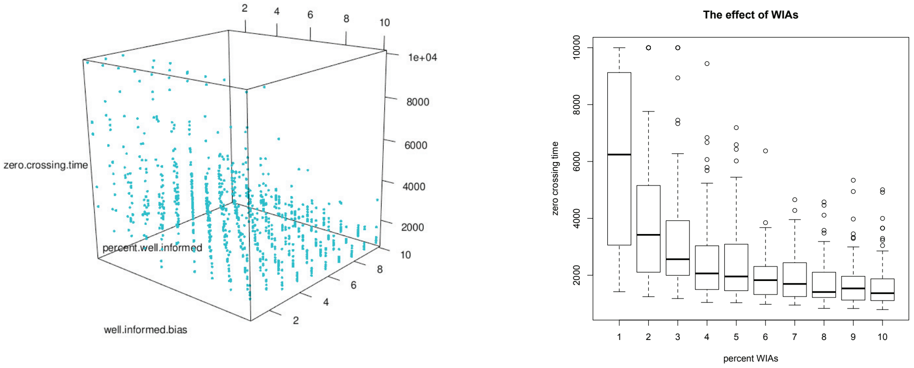


Figure 8: TRAN convergence vs. WIAs; at 4% zero crossing takes 3-6 minutes.

6 Discussion and Conclusions

Bringing results on the same plot (Fig. 9) summarises our findings. In the presence of WIAs, QD is fair and highly reliable. Using QD as comparison, EXP inhabits the upper and TRAN the lower region, the latter at the same time being significantly faster and less predictable (as visible on both Fig. 8 and 9).

We have demonstrated that simple local rules can facilitate the emergence of high quality global information in collaborative knowledge fusion (CKF) systems. Some of these rules, or yet others to be studied later, could be used in real-life deployments in the future. The question remains, however, which of them is superior, and what are the tradeoffs to be considered? Whereas this current first study cannot reply to all these questions, already some hints can be formulated. Of the 3 rules tested, TRAN is fastest but QD is most reliable, and both are superior to EXP. The last one however may preserve diversity. While the focus of the current paper was on CKF with a high quality emergent global state, other applications might focus on maintained diversity and heterogeneity.

Further research is needed to compare the tested methods with the p2p “gossip” algorithms that use an independent sampling of peers (i.e., where ”meetings” happen in virtual space as in a random graph). For example, independence is known to exclude local clusters as experienced in EXP and may improve on it. Exploration of these and other possibilities and are left to subsequent papers.

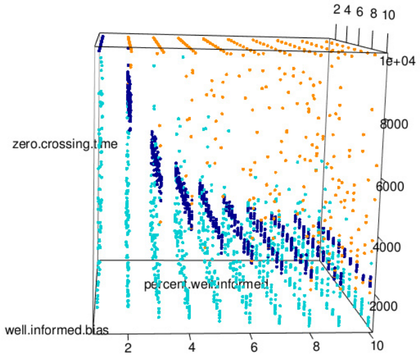


Figure 9: QD (blue), EXP (orange) and TRAN (turquoise) in one plot. Best results by TRAN yet at a high error and outliers. EXP is inferior in all parameter combinations.

7 Acknowledgements

This work is supported by the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement #600854 "Smart Society". G.K. gratefully acknowledges the partial support by the Russian Scientific Foundation, proposal #14-21-00137, "Supercomputer simulation of critical phenomena of complex social systems".

References

[1] T. Franke, G. Kampis, and P. Lukowicz. Leveraging human mobility in smartphone based ad-hoc information distribution in crowd management scenarios. Submitted to MobiSys 2015. `http://www.sigmobile.org/mobisys/2015/`.

[2] M. Jelasity, A. Montresor, and O. Babaoglu. Gossip-based aggregation in large dynamic networks. *ACM Transactions on Computer Systems*, 23(3):219–252, 2005.

[3] G. Kampis, J.W. Kantelhardt, K. Kloch, and P. Lukowicz. Analytical and simulation models for collaborative localization. *J. Computational Science*, 6(1):110, 2015.

[4] G. Kampis and P. Lukowicz. Collaborative localization as a paradigm for incremental knowledge fusion. 5th IEEE CogInfoCom 2014 Conference, 2014. `http://coginfocom.hu/conference/CogInfoCom14/downloads/Program_CogInfoCom_2014_final.pdf`.

[5] K. Kloch, P. Lukowicz, and C. Fischer. Collaborative PDR localisation with mobile phones. In *Proceedings of the 2011 15th Annual International Symposium on Wearable Computers, ISWC 11*, pages 37–40. IEEE Computer Society, Washington, DC, USA, 2011.

[6] R. Ormandi, I. Hegedus, , and M. Jelasity. Asynchronous peer-to-peer data mining with stochastic gradient descent. In *In Emmanuel Jeannot, Raymond Namyst, and Jean Roman, editors, Euro-Par 2011, volume 6852 of Lecture Notes in Computer Science,*, pages 363–368, New York, NY, USA, 2011. Springer-Verlag.

[7] D. Ruta and B. Gabrys. An overview of classifier fusion methods. *Computing and Information systems*, 7(1):1–10, 2000.

[8] T. Simonite. Firechat could be the first in a wave of mesh networking apps. MIT Technology Review. Retrieved 29 June 2014. `http://www.technologyreview.com/news/525921/the-latest-chat-app-for-iphone-needs-no-internet-connection/`.