Simulating Collaborative Sensor Calibration: Convergence and Cost

Léo Le Taro, Hervé Rivano

Simulating Collaborative Sensor Calibration: Convergence and Cost

Léo Le Taro and Hervé Rivano

Univ. Lyon, INSA Lyon, INRIA, CITI
F-69621 Villeurbanne, France
{fname.lname}@inria.fr

***Abstract*—Air pollution is an increasingly concerning issue in urban areas because of its impact on citizens' health. To tackle pollution effectively, accurate monitoring is a must. Precise stations managed by governmental or specialised institutions do exist, but they are both costly and bulky, which limits the potential to deploy them densely. However, recent progress in micro, connected sensors brings new alternative deployment schemes for dense monitoring by low-cost, imprecise sensors. For such a deployment to be relevant relative to urban air quality monitoring aspects, we are concerned with maintaining the system's properties over time. Indeed, one of the major drawbacks of cheap sensors is their drift: chemical properties degrade over time and alter the measurement accuracy. We challenge this issue by designing distributed, online recalibration procedures. We present a simulation framework modelling a mobile wireless sensor network (WSN) and we assess the system's measurement confidence using trust propagation paradigms. As WSN calibrations translate to information exchange between sensors, we also study means of limiting the number of such transmissions by skipping the calibrations deemed least profitable to the system.**

Keywords—Wireless Sensor Networks; Air Pollution Monitoring; Distributed Algorithms; Sensor Calibration.

I. Introduction

Urban pollution monitoring is traditionally carried out using governmental air quality stations, precise but costly and sparsely deployed. In large cities, this limits monitoring resolution to neighbourhood level, actual estimation being computed by numerical models. Recent advances in nanotechnologies is giving birth to small, affordable electrochemical sensors. One of the major research challenges is to achieve a higher spatial resolution with dense and mobile deployments of such low-cost sensors. To keep an exploitable measurement accuracy, sensor calibration must be considered [1]. However, low-cost electrochemical sensors such as ones we use to measure NO_2 concentrations degrade over time [2], hence recalibration is required if the system is to remain usable over a long deployment. Recalibrating nodes in-place is called non-blind calibration [3]. Ground truths, i.e., reference measurements, are needed to adjust the gain and offset of the low-cost nodes. However, such an approach does not scale because it is infeasible to move a high-quality reference to periodically visit hundreds of sensors.

Another paradigm, denoted blind calibration, assumes unknown ground truths and develops techniques to predict and compensate errors in measurements. Blind calibration methods rely on the underlying signal of interest being either band-limited (i.e., varying smoothly in time and space) [4][5][6] or sparse [7]. Yet NO_2 fields prove to be neither, exhibiting large spatiotemporal variations [8], thus negating the possibility to exploit such properties.

Mobile sensing is gaining more attention as recent studies found that a few mobile nodes on well-selected routes can reflect data as accurately as many static ones [9]. On top of it, mobile sensor networks offer opportunities for multihop calibration, i.e., freshly calibrated sensors may in turn calibrate others. Work has been conducted to minimise error propagation in multihop calibration [10].

II. Problem Statement

Given a heterogeneous set of sensing units, comprised of precise base stations and low quality sensors, assuming all low quality sensors are initially uncalibrated, we are wondering whether the system converges to an exploitable state, if so, we are concerned with the time required to reach a permanent regime.

Secondly, we are interested in limiting the number of energy-hungry transmissions between sensors, and wonder how saving data exchanges would affect the system's accuracy.

In Section III, we present the model upon which we based our simulation framework. In Section IV, we detail the algorithm that the framework implements. Section V presents mathematical analyses of binary calibration in our model. Simulation results are laid out in Section VI. We then conclude the paper in Section VII, suggesting future work ideas as well.

III. Model

We model our system as a discrete time and space process in which sensors follow a stochastic mobility pattern. There are $S = \{1..s\}$ initially uncalibrated mobile sensors randomly moving within a space of $P = \{1..p\}$ positions.
For convenience, mobile sensors move following a uniform distribution: at any given time, a sensor moves to $x \in P$ with probability $1/p$. Several sensors may share the same position.

To simulate the presence of precise institutional or governmental air quality stations in a real-life urban scenario, let $R \subset P$ be the subset of r positions featuring a static reference station, assumed perfectly calibrated and reliable.

Each mobile sensor a is assigned a trust value $C_a \in [0..1]$, 0 meaning completely inaccurate and 1 meaning completely accurate. Unless recalibrated, mobile sensors degrade and lower their trust following an exponential decay of rate $\gamma \in [0..1]$: $C_a(t) = C_a(0) \cdot e^{-\gamma t}$.
Reference stations retain a constant trust of 1. Initially, all mobile sensors start with a trust of 0.

Mobile sensors are said to have a *rendez-vous* with a reference station or another mobile sensor when, at the same time instant, they stand at the same position as, respectively, a reference station or another mobile sensor.

Finally, we introduce the recalibration threshold q which conditions calibrations between mobile sensors. The rationale behind this parameter is detailed in Section IV.

IV. Implementation

Each sensor executes the following pseudo-code algorithm at every time step:

Data: γ, q
$C \leftarrow C \cdot e^{-\gamma}$;
foreach n **in** find_neighbours(p, t) **do**
  if n *is a reference station* **then**
    $C \leftarrow 1$;
  else if n *is a mobile sensor* **and** $C + q < n.C$ **then**
    $C \leftarrow n.C$;
  end
end

Figure 1. Sensor trust updating process

We can observe that mobile sensors unconditionally recalibrate to reference stations, however calibration to a peer is conditioned by the variable q. In the real world, each calibration means wirelessly exchanging information. We wish to carry out only meaningful calibrations, with a high benefit. This threshold q allows skipping the calibrations deemed not worthwhile, as they would not lead to a noticeable increase of the trust.

As a corollary, a q of 1 means mobile sensors do not cooperate between each other and only recalibrate to reference stations. Conversely, a q of 0 means that a sensor will always upgrade its trust to the maximum of its neighbors'.

V. Binary Calibration: Analytic Evaluation

Binary calibration is the concept of considering either fully uncalibrated or fully calibrated sensors, with trust values being respectively 0 or 1. In our framework, binary calibration is equivalent to letting γ be zero. Let the random variable T represent the date of a sensor's first calibration.

A. Sensors do not cooperate

Each sensor has probability r/p at each time step to encounter a reference station, independently of other sensors. The probability for such a sensor to remain uncalibrated at t follows a geometric progression of rate $1 - r/p$: $P(T > t) = (1 - r/p)^t$.

The system's global average trust can be derived: $\overline{C}(t) = 1 - P(T > t) = 1 - (1 - r/p)^t$

B. Sensors cooperate

Being calibrated at t means:

- either having calibrated between 1 and $t - 1$ (noted $T \leq t - 1$);
- or *not* having calibrated between 1 and $t - 1$ *but* having a rendez-vous at exactly t (noted $T = t$).

Both cases are mutually exclusive, let us sum their probabilities, keeping in mind sensors start the simulation uncalibrated (time step 0):

$$P(T = 0) = 0$$
$$P(T = t) = 1 - (1 - \frac{1}{p})^{r + sP(T \leq t-1)}$$
$$P(T \leq t) = P(T \leq t - 1) + (1 - P(T \leq t - 1))P(T = t)$$

The rest of this paper presents results of our simulation framework whose purposes are to validate our binary calibration analysis, as well as study cases of non-binary calibration, with $\gamma > 0$.

VI. Non-binary Calibration: Performance Evaluation

In our system, individual sensors' trust occasionally jump when they recalibrate, then continuously variate to lower values because of decay. This behaviour is illustrated by the results of a simple simulation plotted in Figure 2 with a reduced number of sensors and grid positions to preserve readability. This figure hints at the possibility of a permanent regime, because after an initial period of 250 steps no sensor seems to fall below a trust of 0.8. How does the average trust of all sensors behave over a longer period of time and with different parameters?

Figure 3 shows that after a short hysteresis-shaped transient phase, the system's average trust converges to a permanent regime where the trust remains stable. Almost all sensors' trusts evolve between 0.5 and 0.95 despite a pessimistic decay rate $\gamma = 10^{-3}$. Cooperative recalibration is therefore able to maintain the accuracy of the system.

The existence of a permanent regime raises the question of how much time is necessary to reach it. We define the time to converge $t_{0.9}$ as the time by which the average trust of the system $\overline{C}(t_{0.9})$ exceeds 0.9. $t_{0.9}$ is plotted against r in Figure 4, which validates our framework against the analytic evaluations conducted in in Section V for binary calibration schemes ($\gamma = 0$). It shows that cooperation between sensors

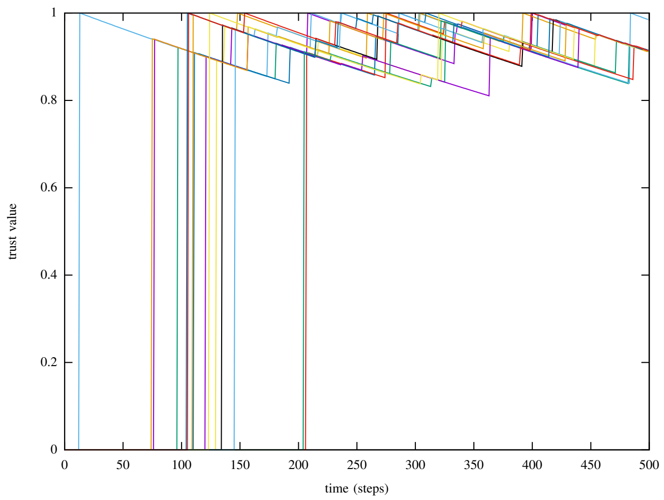

Figure 2. Sensor trust overview
($p = 400, s = 15, r = 1, \gamma = 10^{-3}, q = 0$)

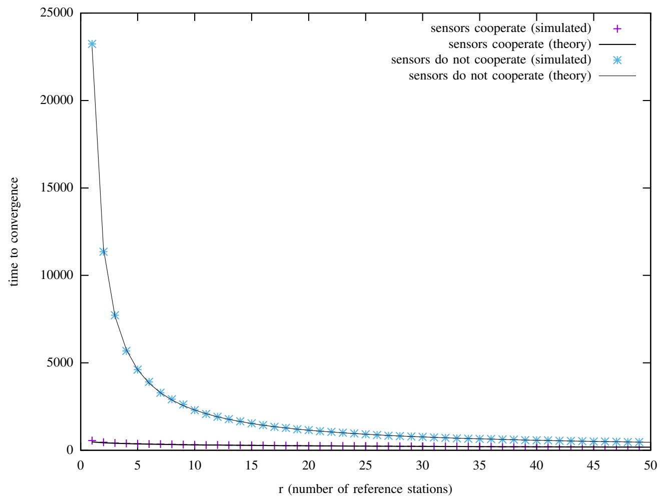

Figure 4. Time to converge
($p = 10000, s = 150, \gamma = 0$)

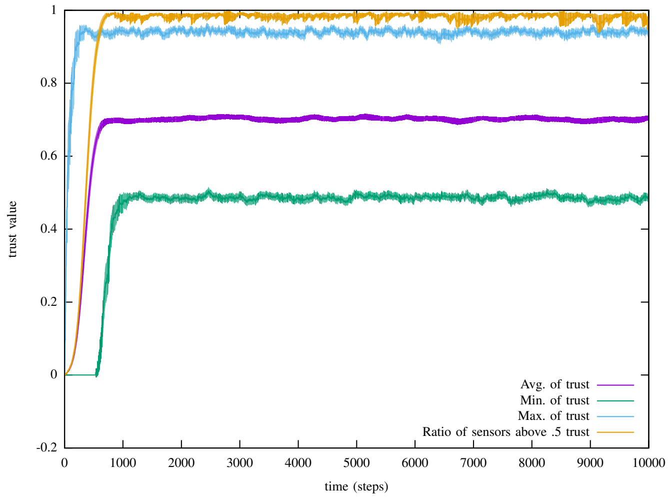

Figure 3. System state over time
($p = 10000, s = 150, r = 1, \gamma = 10^{-3}, q = 0$)
Curves represent a 95% confidence interval over 100 simulations.

have a much stronger impact on the convergence speed than the number of reference stations.

We illustrate the impact of the decay γ on the system's average trust, as well as the maximal and minimal trusts achieved by non-reference sensors in Figure 5. We also consider the ratio of sensors above a trust of 0.5.

We can observe than for $\gamma = 0.002$, the average trust is close to 0.5, which is also the median trust with 50 percent of sensors lying above that value.

Like all others, this simulation was run 100 times and results were aggregated to mitigate the stochastic nature of the results. Nevertheless, we notice that the strongest sensor's trust plot (in light blue) has a quite random aspect, explained by the fact that it depends a lot on its "luck" with rendez-vous rather than exclusively the values of simulation parameters. We deduce that while the average trust we can expect from our system is quite predictable, we cannot guarantee the reliability of a given sensor. Hence, in a real-world urban situation, areas where monitoring is critical (e.g., school or retirement home neighbourhoods) should be covered by many fixed sensors, and/or reference stations.

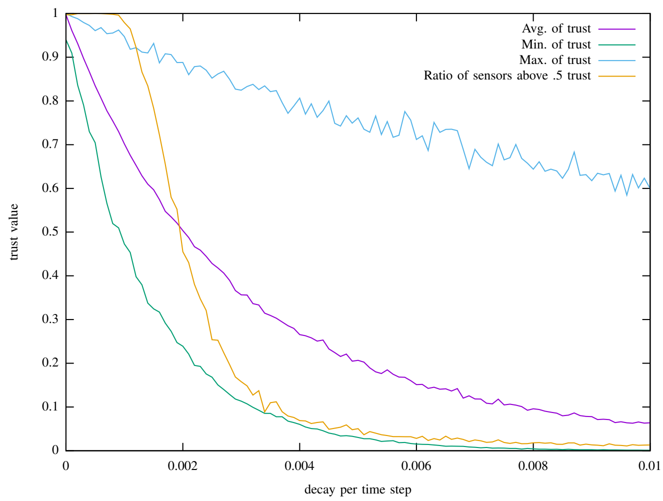

Figure 5. System state versus Gamma
($p = 10000, s = 150, r = 1, q = 0$)

Assuming sensors exchange information once for each recalibration, neglecting the cost of polling one's neighbour's trust, Figures 6 and 7 depict the trade-off between the number of wireless transmissions and the accuracy of the system, i.e., the average trust. The recalibration threshold q is the parameter that restricts transmissions to the case of "*useful*" recalibrations. Numbers were gathered during the permanent regime. An increase of q degrades the system's trust as expected. However, using even a low q dramatically limits the number of wireless transmissions while sacrificing very little trust.

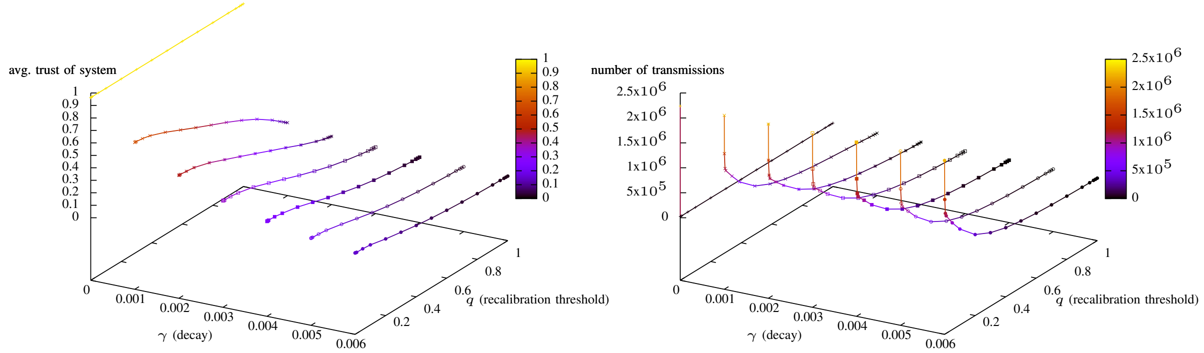

Figure 6. Tradeoff between trust and transmission cost
$(p = 10000, s = 150, r = 1)$

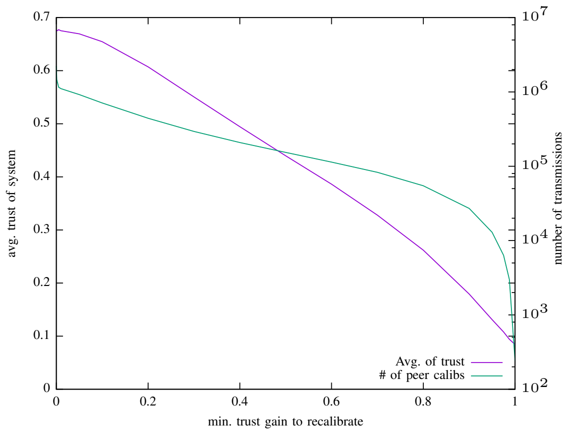

Figure 7. System trust and number of transmissions versus q
$(p = 10000, s = 150, r = 1, \gamma = 10^{-3})$

VII. Conclusion and Perspectives

In this paper, we have presented a simulation framework capable of providing insight on what we can expect from auto-calibrating low-cost electrochemical sensor networks. Cooperation between sensors make the system converge quickly and maintain a steady average accuracy despite the individual decay of the sensors, without requiring many reference stations. It is also shown that many wireless transmissions are spent on a very marginal improvement of the accuracy.

Adjusting certain parameters, we were able to make optimistic predictions of certain metrics like the measurement confidence of the system, its convergence time and the number of transmissions required to calibrate nodes.

Current and future work shall integrate more realistic mobility patterns into the model. One could expect that an urban mobility pattern induce correlation between subsets of sensors, hence a less stable “permanent” regime. The impact of the mobility model and the deployment of reference stations on the spatial mapping of trust values should then be investigated. Besides, realistic mobility models should take into account the distance between calibrating and calibrated nodes into the trust propagation function. Finally, calibration theory requires a non-correlated set of simultaneous calibrating and calibrated measurements. Such a set could be collected over a sequence of rendez-vous, at the cost of a more complex interpretation of the trust.

Acknowledgement

This work has been supported by the "LABEX IMU" (ANR-10-LABX-0088) and the "Programme Avenir Lyon Saint-Etienne" of Université de Lyon, within the program "Investissements d’Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

References

[1] G. Tolle *et al.*, “A macroscope in the redwoods,” in *Proceedings of the 3rd international conference on Embedded networked sensor systems*. ACM, 2005, pp. 51–63.

[2] P. Kumar *et al.*, “The rise of low-cost sensing for managing air pollution in cities,” *Environment international*, vol. 75, pp. 199–205, 2015.

[3] N. Ramanathan *et al.*, “Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks,” *Center for Embedded Network Sensing*, 2006.

[4] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak, “A collaborative approach to in-place sensor calibration,” in *Information Processing in Sensor Networks*. Springer, 2003, pp. 301–316.

[5] L. Balzano and R. Nowak, “Blind calibration of sensor networks,” in *Proceedings of the 6th international conference on Information processing in sensor networks*. ACM, 2007, pp. 79–88.

[6] J. Lipor and L. Balzano, “Robust blind calibration via total least squares,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 4244–4248.

[7] S. Ling and T. Strohmer, “Self-calibration and biconvex compressive sensing,” *Inverse Problems*, vol. 31, no. 11, p. 115002, 2015.

[8] S. Moltchanov *et al.*, “On the feasibility of measuring urban air pollution by wireless distributed sensor networks,” *Science of The Total Environment*, vol. 502, pp. 537–547, 2015.

[9] C. T. Calafate and B. Ducourthial, “On the use of mobile sensors for estimating city-wide pollution levels,” in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2015, pp. 262–267.

[10] O. Saukh, D. Hasenfratz, and L. Thiele, “Reducing multi-hop calibration errors in large-scale mobile sensor networks,” in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 2015, pp. 274–285.

[11] D. Hasenfratz, T. Arn, I. de Concini, O. Saukh, and L. Thiele, “Health-optimal routing in urban areas,” in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 2015, pp. 398–399.

[12] W. Tsujita, A. Yoshino, H. Ishida, and T. Moriizumi, “Gas sensor network for air-pollution monitoring,” *Sensors and Actuators B: Chemical*, vol. 110, no. 2, pp. 304–311, 2005.

[13] P. Buonadonna, D. Gay, J. M. Hellerstein, W. Hong, and S. Madden, “Task: Sensor network in a box,” in *Proceeedings of the Second European Workshop on Wireless Sensor Networks, 2005.* IEEE, 2005, pp. 133–144.