

# Calibration en ligne d'un réseau de capteurs mobiles par factorisation matricielle

YAHAYA Farouk<sup>(1)</sup>, DORFFER Clément<sup>(2)</sup>, PUIGT Matthieu<sup>(1)</sup>,  
DELMAIRE Gilles<sup>(1)</sup>, ROUSSEL Gilles<sup>(1)</sup>

[farouk.yahaya@univ-littoral.fr](mailto:farouk.yahaya@univ-littoral.fr)  
[clement.dorffer@ensta-bretagne.fr](mailto:clement.dorffer@ensta-bretagne.fr)  
[matthieu.puigt@univ-littoral.fr](mailto:matthieu.puigt@univ-littoral.fr)  
[gilles.delmaire@univ-littoral.fr](mailto:gilles.delmaire@univ-littoral.fr)  
[gilles.rousseau@univ-littoral.fr](mailto:gilles.rousseau@univ-littoral.fr)

(1) Univ. du Littoral Côte d'Opale, LISIC – EA 4491, F-62228 Calais France

(2) Lab-STICC, UMR 6285, CNRS & ENSTA-Bretagne, France

La surveillance de la qualité de l'air d'un milieu urbain est généralement réalisée via des mesures automatisées, qui nécessitent la maintenance d'un réseau de capteurs utilisés dans des conditions contraignantes et onéreuses. Ceci limite un déploiement de masse de la mesure fixe, et par conséquent réduit les possibilités d'une cartographie fine sans une modélisation très robuste, et à ce jour encore non-atteinte. En effet, à l'échelle urbaine, la modélisation reste relativement imprécise. Plusieurs projets ont alors vu le jour pour réaliser une intensification de la mesure (en utilisant une foule de capteurs à coût plus modéré) et pour réaliser une cartographie plus proche de la réalité terrain. Pour favoriser le déploiement bas coût et l'engagement citoyen dans une collecte participative des mesures de la qualité de l'air, une approche prometteuse consiste à faire l'usage de micro-capteurs, et parfois même de proposer une version libre et bas coût de ces capteurs, réalisable par auto-construction à la portée de tous.

Dès 2014, le projet **OSCAR** (Observation et Sensibilisation Citoyenne de la Qualité de l'air en Région) [7] a cherché à mettre en place une observation répartie de la qualité de l'air sur une zone d'intérêt, en l'occurrence Dunkerque, à travers une approche par mesure participative (le vocable *crowd-sensing* passe dans le langage courant) via une série de capteurs mobiles connectés à un smartphone. En incitant la population à mesurer la qualité de l'air dans son environnement et en utilisant la géolocalisation fournie par les données GPS, nous obtenons rapidement une base de données de grande dimension, permettant la cartographie des données mesurées avec une finesse inconnue aujourd'hui. Un des aspects importants du problème de la mesure citoyenne est aussi le respect de l'anonymat des porteurs, ce qui a été le principe mis en place dans la plate-forme de *crowdsensing* APISENSE® [6]. Cependant, la masse d'informations à traiter doit faire face (i) à l'hétérogénéité de la qualité des données recueillies due à la disparité des composants de caractéristiques différentes, (ii) au non-étalonnage *a priori* des capteurs recueillant les données et (iii) à l'échantillonnage irrégulier tant spatialement que temporellement.

Le *mobile crowdsensing* est donc un moyen efficace d'acquérir des données datées et géolocalisées [1], en utilisant une foule de capteurs mobiles. L'exploitation des données provenant de tels capteurs nécessite le développement de méthodes spécifiques de traitement de l'information. En particulier, l'étalonnage (ou calibration) de capteurs ne peut être effectué rapidement et économiquement en laboratoire et doit donc être réalisé à partir de données mesurées directement, via des approches aveugles. Nous proposons de revisiter le problème d'étalonnage aveugle de capteurs de pollutions comme un problème informé de factorisation matricielle à données manquantes [2], où les facteurs contiennent

respectivement le modèle d'étalonnage fonction du phénomène physique observé (ce modèle peut-être notamment affine [2], ou non linéaire [3]) et les paramètres d'étalonnage de chaque capteur. Ces approches sont dites informées car (i) les facteurs matriciels sont structurés par la nature du problème [2,3], (ii) le phénomène observé peut-être décomposé sous forme parcimonieuse dans un dictionnaire connu [4] et (iii) nous connaissons la fonction d'étalonnage moyenne des capteurs à étalonner [4]. Dans l'application de surveillance de la qualité de l'air que nous considérons, nous supposons avoir à notre disposition des mesures très précises mais distribuées de manière très parcimonieuse dans l'espace et le temps, que nous couplons aux multiples mesures issues de capteurs de mobile crowdsensing. Ces approches faisant appel aux rendez-vous entre les différents capteurs mobiles et capteurs conventionnels et une décomposition parcimonieuse sont plus performantes que des méthodes basées sur la complétion de la matrice de données observées ou des techniques multi-sauts faisant appel à des approches robustes de régression [5]. Les travaux en cours portent sur l'accélération de ces techniques [8], pour pouvoir gérer le déluge de données dans l'application considérée.

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing : current state and future challenges," IEEE Communications Magazine, vol. 49, no. 11, pp. 32–39, Nov. 2011

[2] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Blind calibration of mobile sensors using informed nonnegative matrix factorization," in Proc. of LVA/ICA, LNCS 9237, pp. 497-505, 2015.

[3] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Nonlinear mobile sensor calibration using informed seminonnegative matrix factorization with a Vandermonde factor," in Proc. of SAM, 2016.

[4] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration," IEEE Transactions on Signal and Information Processing over Networks, to appear. <https://hal.archives-ouvertes.fr/hal-01580604/>

[5] O. Saukh, D. Hasenfratz, and L. Thiele, "Reducing multi-hop calibration errors in large-scale mobile sensor networks," Proc. of IPSN, 2015.

[6] R. Sommenard, R. Rouvoy, "Towards Privacy-Preserving Data Dissemination in Crowd-Sensing Middleware Platform". In Proc. of UbiMob'16, July 2016, Lorient, France. <https://hal.inria.fr/hal-01332588>

[7] OSCAR, projet du Programme Chercheurs Citoyens (PCC) de la région Hauts de France. Label Repère (Réseau d'échange et de Projets pour le Pilotage de Recherche et L'Expertise), colloque du 27 juin 2018, Paris Nord, Matthieu Puigt, Gilles Roussel. Romain Rouvoy, Isabelle Sagnier.

[8] F. Yahaya, M. Puigt, G. Delmaire, and G. Roussel, "Faster-than-fast NMF using random projections and Nesterov iterations". In Proc. of iTwist'18, Nov. 2018, Marseille, France.

Remerciements : Le projet OSCAR a été financé par la Région Hauts-de-France dans le cadre de son programme « chercheurs citoyens ». F. Yahaya remercie vivement la Région Hauts-de-France qui finance en partie sa thèse de doctorat.

# Online calibration of a mobile sensor network by matrix factorization

YAHAYA Farouk<sup>(1)</sup>, DORFFER Clément<sup>(2)</sup>, PUIGT Matthieu<sup>(1)</sup>,  
DELMAIRE Gilles<sup>(1)</sup>, ROUSSEL Gilles<sup>(1)</sup>

[gilles.rousseau@univ-littoral.fr](mailto:gilles.rousseau@univ-littoral.fr)

(1) Univ. du Littoral Côte d'Opale, LISIC – EA 4491, F-62228 Calais France

(2) Lab-STICC, UMR 6285, CNRS & ENSTA-Bretagne, France

Monitoring the air quality of an urban environment is generally carried out via automated measurements, which require the maintenance of a sensors network used under constraining and expensive conditions. This limits a mass deployment of the fixed measurement, and consequently reduces the possibilities of a fine mapping without a very robust modeling, which is still not reached today. Indeed, at the urban scale, modelling remains relatively imprecise. Several projects thus emerged to provide a finer measurement coverage (by using a crowd of low-cost sensors) and to realize a finer cartography which would be closer to the ground truth than state-of-the-art techniques. To promote low-cost deployment and citizen involvement in participatory air-quality data acquisition, a promising approach consists of using micro-sensors or even proposing an open-source and low-cost version of these sensors, which can be easily built, within the Do-It-Yourself trend.

Since 2014, the OSCAR project (Observation and Citizen sensitizing of Air Quality in the Region) [7] has sought to set up a distributed air-quality observation in an area of interest, i.e., Dunquerque, through a participatory measurement approach (commonly named crowd-sensing) which uses a set of mobile sensors connected to a smartphone. By promoting the collect of geolocated air quality measurements in the people environment—using the geolocation provided by GPS data of volunteers' smartphones—we quickly obtain a large database which allows the sensed data mapping with an unknown accuracy. One of the important aspects of the problem of citizen measurement is also the respect of the anonymity of sensor holders, which has been the principle implemented in the APISENSE® crowdsensing platform [6]. However, the mass of information to be processed must cope with (i) the heterogeneity of the quality of the data collected due to the disparity of the components with different characteristics, (ii) the *a priori* non-calibration of the sensors collecting the data, and (iii) the irregular sampling, both spatially and temporally.

Mobile crowdsensing is therefore an efficient way to acquire time-stamped and geo-located data [1], using a crowd of mobile sensors. The exploitation of data from such sensors requires the development of specific information processing methods. In particular, sensor calibration cannot be performed quickly and economically in the laboratory and must therefore be directly performed from sensor readings, using blind approaches. We propose to revisit the problem of blind calibration of mobile pollution sensors as an informed matrix factorization problem with missing data [2], where the factor matrices respectively contain the calibration model as a function of the observed physical phenomenon (affine [2] or nonlinear model [3]) and the calibration parameters of each sensor. These approaches are said to be informed because (i) the matrix factors are structured by the nature of the problem [2,3], (ii) the observed phenomenon can be broken down into parsimonious form in a known dictionary [4], and (iii) we know the average calibration function of the sensors to be calibrated [4]. In the air quality monitoring application we are considering, we assume that we are provided

sensor readings which are very precise but very sparsely distributed in both space and time, that we couple to the multiple measurements from mobile crowdsensing sensors. These approaches using *rendezvous* between the different mobile sensors and conventional sensors and sparse decompositions are more efficient than methods based on the completion of the observed data matrix or multi-hop techniques using robust regression approaches [5]. Current work concerns the acceleration of these techniques [8], in order to process the data deluge due to the considered application.

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing : current state and future challenges," IEEE Communications Magazine, vol. 49, no. 11, pp. 32–39, Nov. 2011.

[2] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Blind calibration of mobile sensors using informed nonnegative matrix factorization," in Proc. of LVA/ICA, LNCS 9237, pp. 497-505, 2015.

[3] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Nonlinear mobile sensor calibration using informed seminonnegative matrix factorization with a Vandermonde factor," in Proc. of SAM, 2016.

[4] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration," IEEE Transactions on Signal and Information Processing over Networks, to appear. <https://hal.archives-ouvertes.fr/hal-01580604/>

[5] O. Saukh, D. Hasenfratz, and L. Thiele, "Reducing multi-hop calibration errors in large-scale mobile sensor networks," Proc. of IPSN, 2015.

[6] R. Sommenard, R. Rouvoy, "*Towards Privacy-Preserving Data Dissemination in Crowd-Sensing Middleware Platform*". In Proc. of UbiMob'16, July 2016, Lorient, France. <https://hal.inria.fr/hal-01332588>

[7] OSCAR, projet du Programme Chercheurs Citoyens (PCC) de la région Hauts de France. Label Repère (Réseau d'échange et de Projets pour le Pilotage de Recherche et L'Expertise), colloque du 27 juin 2018, Paris Nord, Matthieu Puigt, Gilles Roussel. Romain Rouvoy, Isabelle Sagnier.

[8] F. Yahaya, M. Puigt, G. Delmaire, and G. Roussel, "Faster-than-fast NMF using random projections and Nesterov iterations". In Proc. of iTwist'18, Nov. 2018, Marseille, France.

Acknowledgements: The OSCAR project was funded by the Région Hauts-de-France within the « chercheurs-citoyens » program. F. Yahaya greatly acknowledges the Région Hauts-de-France to partly fund his Ph.D. fellowship.