

Mobilité d'un réseau de capteurs sans fil basée sur les méthodes à noyaux

Nisrine GHADBAN^{1,2}, Paul HONEINE¹, Clovis FRANCIS², Farah MOURAD-CHEHADE¹, Joumana FARAH³, Maya KALLAS¹

¹Institut Carles Delaunay (UMR CNRS 6279), Université de Technologie de Troyes, Troyes, France

²Laboratoire d'Analyse et de Surveillance de Systèmes, Faculté de Génie, Université Libanaise, Liban

³Département de télécommunications, Faculté d'ingénierie, Université Saint-Esprit de Kaslik, Kaslik, Liban

nisrine.ghadban@utt.fr, paul.honeine@utt.fr, cfrancis@ul.edu.lb farah.mourad@utt.fr,
joumanafarah@usek.edu.lb, maya.kallas@utt.fr

Résumé – Les réseaux de capteurs sans fil ont reçu une attention considérable au cours de la dernière décennie en raison de leur faible coût, de la facilité de leur déploiement et de leur aptitude à être employés pour des techniques de surveillance efficaces. Le présent article propose des schémas de mobilité qui permettent d'améliorer le suivi d'un phénomène physique. A cette fin, les méthodes à noyaux permettent de construire un modèle local pour chaque capteur, dans lequel l'entrée du modèle est la position géographique du capteur et la sortie est l'estimation de la quantité physique à mesurer. Nous montrons que les méthodes à noyaux fournissent un cadre élégant à ce problème d'estimation. Cela nous permet de définir des schémas de mobilité des capteurs pour améliorer l'efficacité des modèles, de façon à minimiser l'erreur d'approximation. Ainsi, plusieurs techniques d'optimisation sont-elles proposées en tenant compte de la première et de la seconde dérivée de l'erreur d'approximation. Les expérimentations réalisées dans ce cadre, visant à estimer la diffusion d'un gaz, montrent la pertinence de la méthode proposée.

Abstract – Wireless sensor networks have received considerable attention during the last decade for efficient monitoring, due to their low cost, their easy deployment and their capacity to locally process information. This paper derives original mobility schemes that allow improving the tracking of a physical phenomenon. To this end, we use kernel-based methods to construct a local model for each sensor, using the learning process where the input is the position of the sensor and the output is the estimation of the physical phenomenon. We show that kernel-based methods provide an elegant way to optimize the model. This allows us to derive mobility schemes for sensors in such a way to improve the efficiency of the models, by minimizing the estimation error. Sensors are moved according to several optimization techniques, by considering the first and second derivatives of the approximation error. Experimentations aim at estimating a gas diffusion in the space at any location devoid of sensor.

1 Introduction

Les réseaux de capteurs sans fil constituent un champ de recherche multidisciplinaire, à la croisée de communautés telles que la microélectronique, les télécommunications sans fil et le traitement du signal. Initialement développés pour des applications militaires, leur domaine d'application ne cesse de s'élargir pour inclure la surveillance dans les domaines civils et industriels [11]. L'autonomie de ces capteurs intelligents impose des contraintes de limitation de ressources, d'énergie en particulier, qui sont au fondement d'un nouveau paradigme : une optimisation collaborative, où chaque capteur résout localement un sous-problème d'optimisation, en collaborant de proche-en-proche avec ses voisins [6]. Des algorithmes à faible complexité de calcul doivent y être implémentés. Un nouveau défi clé concerne l'émergence des réseaux de capteurs mobiles, qui attirent beaucoup d'attention de nos jours, comme c'est le cas d'une flotte de drones par exemple. Un capteur est alors apte à se déplacer vers une nouvelle position pour y extraire une information pertinente. Ainsi sommes-nous à la recherche de l'information pertinente dans la région sous surveillance, tout en respectant des contraintes de couverture, d'énergie de déplacement et de communication.

Cet article étudie la surveillance et le suivi de l'évolution d'un champ de diffusion [1, 7, 3]. Il peut s'agir de la diffusion d'un gaz, d'une espèce biochimique ou encore de la pollution, pour ne citer que ceux-ci. Une première approche a été étudiée dans [4, 5], en considérant une estimation globale du champ étudié. Dans le présent travail, nous proposons d'effectuer une estimation locale pour chaque capteur, à partir d'informations collectées des capteurs voisins. Le cadre étudié concerne une mise en œuvre de modèles locaux des méthodes à noyaux en apprentissage statistique. Cela nous permet de définir des schémas de mobilité des capteurs pour améliorer l'efficacité des modèles locaux, en améliorant l'erreur d'approximation. Ainsi, le socle proposé permet-il de développer plusieurs techniques d'optimisation en tenant compte de la première et de la seconde dérivée de l'erreur d'approximation. Les expérimentations réalisées sur l'estimation de la diffusion d'un gaz montrent la pertinence de la méthode proposée.

Le reste de cet article est organisé comme suit : Dans la section suivante, nous présentons les modèles locaux dans le cadre des méthodes à noyaux. Dans la Section 3, nous traitons le problème d'optimisation. Plusieurs schémas d'optimisation sont décrits dans la Section 4. La Section 5 fournit des résultats expérimentaux, tandis que la Section 6 conclut l'article.

2 Méthode à noyaux

Considérons un réseau composé de N capteurs déployés dans une région \mathbb{X} , avec $\mathbb{X} \subset \mathbb{R}^2$ ou $\mathbb{X} \subset \mathbb{R}^3$. Soit un capteur i de position $\mathbf{x}_i \in \mathbb{X}$, et soit ϑ_i l'ensemble des indices des ℓ plus proches capteurs voisins au capteur i , où par convention $i \notin \vartheta_i$. A tout instant, tout capteur i collecte de chacun de ses voisins, $j \in \vartheta_i$, sa position \mathbf{x}_j et la valeur de la grandeur physique mesurée y_j . Chaque capteur a deux tâches : la première consiste à modéliser le champ, tandis que la seconde est de déterminer son déplacement qui permet de réduire l'erreur d'estimation. Lors de la première tâche, le capteur i , $1 \leq i \leq N$, recueille des mesures à partir de ses capteurs voisins j , $j \in \vartheta_i$, et estime le champ local représenté par la fonction $\psi_i(\cdot)$ définie sur \mathbb{X} , telle que

$$\psi_i(\mathbf{x}_j) \approx y_j \quad \text{pour tout } j \in \vartheta_i.$$

Pour définir cette fonction, nous considérons le formalisme de noyau reproduisant [10, 9, 12]. Soit \mathbb{H} un espace de Hilbert à noyau reproduisant constitué de fonctions de \mathbb{X} dans \mathbb{R} . Soient $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ le produit scalaire dans \mathbb{H} et $\|\cdot\|_{\mathbb{H}}$ la norme correspondante. Un noyau (défini positif) $\kappa(\cdot, \cdot)$ désigne une fonction définie sur $\mathbb{X} \times \mathbb{X}$, et vérifiant $\psi(\mathbf{x}_j) = \langle \psi(\cdot), \kappa(\mathbf{x}_j, \cdot) \rangle_{\mathbb{H}}$ pour tout $\mathbf{x}_j \in \mathbb{X}$ et $\psi(\cdot) \in \mathbb{H}$.

Le problème d'optimisation est posé en minimisant l'erreur quadratique moyenne entre la sortie du modèle en \mathbf{x}_i et la sortie désirée y_i , et en utilisant la régularisation de Tikhonov, selon :

$$\psi_i(\cdot) = \arg \min_{\psi \in \mathbb{H}} \sum_{j \in \vartheta_i} |y_j - \psi(\mathbf{x}_j)|^2 + \nu \|\psi\|_{\mathbb{H}}^2, \quad (1)$$

où ν est un paramètre de régularisation qui contrôle le compromis entre le raccord aux données disponibles et la douceur de la solution. On note que l'utilisation d'un voisinage permet aussi un tel contrôle, réduisant ainsi la contribution de ν dans le problème d'optimisation. D'après le Théorème de Représentation [8], la solution de ce problème est donnée sous la forme d'une combinaison linéaire de noyaux $\kappa(\mathbf{x}_j, \cdot)$ associés à l'espace \mathbb{H} :

$$\psi_i(\cdot) = \sum_{j \in \vartheta_i} \alpha_{i,j} \kappa(\mathbf{x}_j, \cdot).$$

En regroupant les coefficients à déterminer, $\alpha_{i,j}$ pour $j \in \vartheta_i$, dans un vecteur $\boldsymbol{\alpha}_i$, et dans \mathbf{y}_i les mesures y_j pour $j \in \vartheta_i$, on peut facilement montrer que

$$\boldsymbol{\alpha}_i = (\mathbf{K}_i + \nu \mathbf{I})^{-1} \mathbf{y}_i, \quad (2)$$

où \mathbf{K}_i est la matrice de Gram de taille $\ell \times \ell$, d'éléments $\kappa(\mathbf{x}_m, \mathbf{x}_n)$ pour $m, n \in \vartheta_i$, et \mathbf{I} est la matrice identité de taille $\ell \times \ell$.

Deux principaux types de noyaux sont disponibles : projectif et radial. Dans cet article, nous utilisons les noyaux radiaux de la forme :

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = g(\|\mathbf{x}_i - \mathbf{x}_j\|^2), \quad (3)$$

où $g(\cdot)$ est une fonction réelle positive. Les noyaux radiaux s'adaptent bien à l'étude d'un problème de diffusion, car ils ne dépendent que de la distance et sont invariants à la translation. Sans perte de généralité, nous dérivons des expressions générales pour ce type de noyaux, avant d'envisager le cas typique du noyau Gaussien où $g(\cdot) = \exp^{-\frac{1}{2\sigma^2}(\cdot)}$, avec σ le paramètre de largeur de bande.

3 Mobilité

L'objectif de la méthode proposée consiste à adapter de manière itérative la position des capteurs dans le but d'améliorer le modèle de régression. Soit $\vartheta_{i,t}$ l'ensemble de voisinage du capteur i à l'itération t , où $\vartheta_{i,0}$ est l'ensemble initial des voisinages du capteur i , et soit $\psi_{i,t}(\cdot)$ le modèle calculé conformément à la Section 2 en utilisant $\vartheta_{i,t}$, soit :

$$\psi_{i,t}(\cdot) = \sum_{j \in \vartheta_{i,t}} \alpha_{i,j,t} \kappa(\mathbf{x}_{j,t}, \cdot), \quad (4)$$

où les coefficients $\alpha_{i,j,t}$ sont obtenus en utilisant l'expression (2) avec l'ensemble $\vartheta_{i,t}$. L'erreur de validation du modèle à tout couple (\mathbf{x}_*, y_*) dans le voisinage du capteur i est donnée par l'erreur quadratique entre la sortie désirée y_* et la sortie du modèle estimé $\psi_{i,t}(\mathbf{x}_*)$:

$$\varepsilon_{i,t}(\mathbf{x}_*) = |y_* - \psi_{i,t}(\mathbf{x}_*)|^2. \quad (5)$$

A noter qu'une optimisation simultanée de la position $\mathbf{x}_{j,t}$ de tous les capteurs est un problème difficile, voire irréaliste dans le cadre distribué des réseaux de capteurs sans fil, même avec un centre de calcul. Pour ces raisons, nous proposons d'effectuer une optimisation locale de chaque capteur, à raison d'un capteur à chaque itération. A ce but, nous choisissons, parmi tous les capteurs, le capteur i dont le modèle donne la plus grande erreur d'estimation $\varepsilon_{i,t}(\mathbf{x}_{i,t-1})$. Nous adaptons ensuite sa position $\mathbf{x}_{i,t-1}$ en $\mathbf{x}_{i,t}$ de façon à réduire l'erreur de validation du modèle, notamment en minimisant la fonction $\varepsilon_{i,t}(\cdot)$:

$$\mathbf{x}_{i,t} = \arg \min_{\mathbf{x} \in \mathbb{X}} \varepsilon_{i,t}(\mathbf{x}). \quad (6)$$

Les positions de tous les autres capteurs restent inchangées, à savoir $\mathbf{x}_{j,t} = \mathbf{x}_{j,t-1}$ pour tout $j \neq i$. Il est à noter que, dans cette stratégie, le sur-apprentissage est évité, puisque $i \notin \vartheta_i$ par construction, c'est à dire que les informations $(\mathbf{x}_{i,t-1}, y_{i,t-1})$ du capteur i n'interviennent pas dans l'estimation du modèle associé $\psi_{i,t-1}$, mais a posteriori dans son ajustement. Le problème de minimisation de $\varepsilon_{i,t}(\mathbf{x})$ est un problème difficile, car il s'agit d'un problème d'optimisation non linéaire et non convexe, principalement à cause de la nature du noyau considéré. Dans ce qui suit, nous proposons trois solutions possibles à ce problème.

4 Schémas d'optimisation

Dans cette section, nous proposons des schémas itératifs d'optimisation du premier et second ordre pour la mobilité. Bien que les méthodes proposées peuvent être appliquées à tout type de noyau, nous considérons dans cet article les noyaux radiaux de la forme (3). Dans ce cas, le gradient de la fonction d'erreur $\varepsilon_{i,t}(\mathbf{x})$ est donné par

$$\begin{aligned} \nabla \varepsilon_{i,t}(\mathbf{x}) &= -4 \left(y_{i,t-1} - \sum_{j \in \vartheta_{i,t-1}} \alpha_{i,j,t} \kappa(\mathbf{x}_{j,t-1}, \mathbf{x}) \right) \\ &\quad \times \sum_{j \in \vartheta_{i,t-1}} \alpha_{i,j,t} g'(\|\mathbf{x} - \mathbf{x}_{j,t-1}\|^2) (\mathbf{x} - \mathbf{x}_{j,t-1}), \end{aligned}$$

où $g'(z)$ est la première dérivée de $g(z)$ par rapport à z . Nous supposons que \mathbf{x} reste dans une région autour de $\mathbf{x}_{i,t-1}$

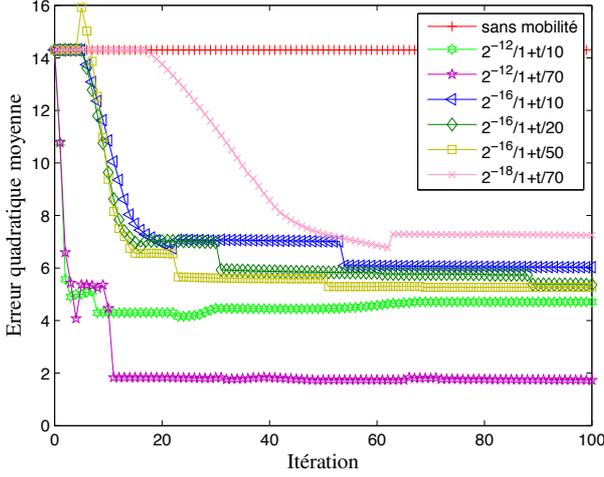


FIGURE 1 – Evolution de l’erreur quadratique moyenne de la méthode de descente de gradient pour différentes valeurs du pas η_t .

où la mesure $y_{i,t-1}$ varie très peu. Nous proposons plusieurs stratégies de mobilité des capteurs, en minimisant la fonction coût (5), c’est à dire en annulant l’expression de son gradient.

Méthode du point-fixe

Le minimum de $\varepsilon_{i,t}(\mathbf{x})$ est obtenu lorsque son gradient s’annule, c’est à dire $\nabla \varepsilon_{i,t}(\mathbf{x}) = 0$. De cette équation, on peut déduire l’expression de la nouvelle position $\mathbf{x}_{i,t}$ à partir de la précédente position $\mathbf{x}_{i,t-1}$, selon :

$$\mathbf{x}_{i,t} = \frac{\sum_{j \in \mathcal{D}_{i,t-1}} \alpha_{i,j} g'(\|\mathbf{x}_{i,t-1} - \mathbf{x}_{j,t-1}\|^2) \mathbf{x}_{j,t-1}}{\sum_{j \in \mathcal{D}_{i,t-1}} \alpha_{i,j} g'(\|\mathbf{x}_{i,t-1} - \mathbf{x}_{j,t-1}\|^2)}.$$

La méthode du point-fixe peut être instable et risque parfois de ne pas converger, notamment lorsque le dénominateur ci-dessus se rapproche de la valeur nulle. Les problèmes liés à cette méthode sont probablement dus à l’absence d’un paramètre de pas d’adaptation pour permettre un contrôle de la convergence de l’algorithme.

Méthode de descente de gradient

Le capteur i se déplace suivant la direction opposée au gradient, selon

$$\mathbf{x}_{i,t} = \mathbf{x}_{i,t-1} - \eta_t \nabla \varepsilon_{i,t}(\mathbf{x}_{i,t-1}),$$

où η_t est un pas qui dépend du temps. Dans le présent article, et comme préconisé en optimisation itérative (voir par exemple [2]), on choisit le pas selon l’expression $\eta_t = \frac{\eta_0}{1+t/\tau}$. Le paramètre de réglage τ détermine la durée de la phase de recherche initiale, avec $\eta_t \simeq \eta_0$ (lorsque $t \ll \tau$), avant une phase de convergence où η_t diminue selon η_0/t (lorsque $t \gg \tau$). En pratique, ce choix du pas permet de contrôler les déplacements des capteurs, et ainsi conserver l’énergie.

Méthode de Newton

Cette méthode permet de suivre la courbure de l’erreur quadratique $\varepsilon_{i,t}(\cdot)$, en considérant ses dérivées partielles secondes.

En regroupant ces dernières dans la matrice Hessienne \mathbf{H} (dont l’expression n’est pas donnée ici par manque de place), on obtient :

$$\mathbf{x}_{i,t} = \mathbf{x}_{i,t-1} - \mathbf{H}^{-1} \nabla \varepsilon_{i,t}(\mathbf{x}_{i,t-1}).$$

Ainsi le pas est-il adapté à chaque itération. Le prix à payer est, d’une part le calcul de la matrice Hessienne à chaque itération, et d’autre part le problème d’une bonne estimation de cette matrice à partir de peu d’information disponible.

5 Simulation

Afin d’illustrer les résultats, nous considérons la diffusion d’un gaz à partir d’une source, dans une région carrée $\mathbb{X} = [-0.5; 0.5] \times [-0.5; 0.5]$. Cette diffusion est régie par l’équation différentielle suivante :

$$\frac{\partial G(\mathbf{x}, t)}{\partial t} - c \nabla_{\mathbf{x}}^2 G(\mathbf{x}, t) = Q(\mathbf{x}, t),$$

où $G(\mathbf{x}, t)$ est la densité du gaz dépendant de la position et du temps, $\nabla_{\mathbf{x}}^2 G(\mathbf{x}, t)$ est l’opérateur spatial de Laplace, $Q(\mathbf{x}, t)$ correspond à la quantité de gaz ajoutée et c est la conductivité de la surface. Une source de gaz placée au centre de la région est activée. La distribution de gaz dans l’espace est considérée stationnaire ; elle est supposée être invariante par rapport au temps d’estimation. On utilise 30 capteurs déployés aléatoirement dans l’espace considéré. Les simulations sont faites sur le logiciel Matlab. Un bruit blanc Gaussien de moyenne nulle et d’écart type 0.01 est ajouté aux mesures de chaque capteur. La largeur de bande du noyau Gaussien est fixée à $\sigma = 0.1$. Dans ce qui suit, 20 réalisations indépendantes de Monte Carlo sont effectuées, chaque réalisation composée de 100 itérations, i.e., $t = 1, 2, \dots, 100$. A chaque itération, un seul capteur est déplacé par l’une des méthodes de mobilité décrites dans la Section 4.

La pertinence du modèle est évaluée à l’aide de l’erreur quadratique moyenne estimée en utilisant 10 capteurs de test. Ces capteurs de test sont déployés aléatoirement dans la région \mathbb{X} et mesurent, à leurs positions, la densité de gaz correspondante. Les capteurs de test ont des positions fixes et les informations correspondantes ne sont pas utilisées dans le processus d’apprentissage.

Contrairement aux autres méthodes où aucun choix de pas n’est nécessaire, les performances de la méthode de descente de gradient dépendent du choix du pas. La FIGURE 1 montre l’évolution, à chaque mouvement, de l’erreur quadratique moyenne sur l’ensemble des capteurs de test, pour plusieurs combinaisons des valeurs du paramètre de η_0 et τ dans l’expression du pas $\eta_t = \eta_0/(1+t/\tau)$. En se basant sur ces

TABLE 1 – Distance moyenne et distance maximale, parcourues pour chacune des méthodes de mobilité proposées dans cet article.

Stratégie de mobilité :	Point-fixe	Gradient	Newton
Distance moyenne :	0.0759	0.0443	0.0021
Distance maximale :	0.3234	0.2542	0.0642

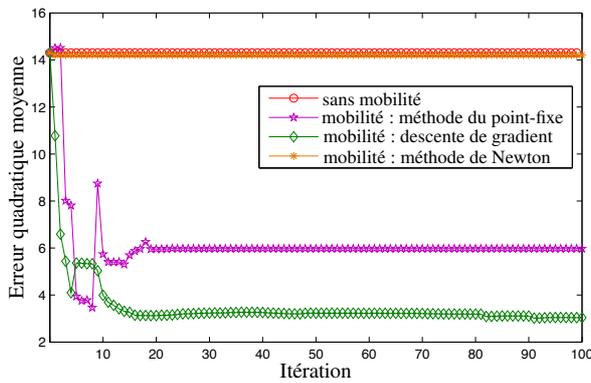


FIGURE 2 – Performances des méthodes proposées.

observations, nous adoptons les valeurs $\eta_0 = 2^{-12}$ et $\tau = 70$ pour le schéma d'optimisation de descente de gradient.

En procédant à 20 simulations de Monte Carlo, la TABLE 1 montre les distances (moyenne et maximale) parcourues par les capteurs, pour les différentes stratégies de mobilité. Ces distances sont directement liées à la contrainte énergétique des capteurs et elles sont minimales avec la méthode de Newton. La FIGURE 2 illustre l'erreur quadratique moyenne sur l'ensemble des capteurs de test. Cette dernière montre la pertinence de la mobilité pour diminuer l'erreur, et illustre l'efficacité de la méthode de descente de gradient au prix du choix du pas η_t . Contrairement à ce qui a été prévu, la méthode de Newton n'a pas de bonnes performances comparée aux méthodes du premier ordre. Ceci est principalement dû à de la non-linéarité de la fonction coût en fonction de la position et à la variation de la mesure du capteur mobile qui nécessite d'être négligée dans le calcul des dérivées.

6 Conclusion

Dans cet article, nous modélisons un champ de gaz mesuré par un réseau de capteurs sans fil mobiles. Nous optimisons ce modèle en déplaçant une trentaine de capteurs déployés aléatoirement dans la zone à surveiller. Les modèles construits avec les deux cas statique et mobile sont comparés. Dans les deux cas, l'objectif est de définir, pour chaque capteur, le modèle local qui correspond le mieux à la quantité mesurée de gaz dans le voisinage du capteur. Ces modèles sont déterminés au moyen d'un processus d'apprentissage non linéaire en utilisant le noyau Gaussien. D'après les simulations réalisées, nous avons démontré une bonne réduction des erreurs par les méthodes proposées, par rapport au cas statique.

Remerciements

Ce travail est soutenu par l'Université Libanaise, et par la Région Champagne-Ardennes dans le cadre du Projet "WiDiD" ("Wireless Diffusion Detection").

Références

- [1] J. Chen, C. Richard, P. Honeine, and J. C. M. Bermudez. Non-negative distributed regression for data inference in wireless sensor networks. In *Proc. 44th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove (CA), USA, 7–10 November 2010.
- [2] P. Honeine. Online kernel principal component analysis : a reduced-order model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9) :1814–1826, September 2012.
- [3] P. Honeine, M. Essoloh, C. Richard, and H. Snoussi. Distributed regression in sensor networks with a reduced-order kernel model. In *Proc. 51st IEEE GLOBECOM Global Communications Conference*, New Orleans, LA, USA, 2008.
- [4] P. Honeine, C. Richard, J. C. M. Bermudez, H. Snoussi, M. Essoloh, and F. Vincent. Functional estimation in hilbert space for distributed learning in wireless sensor networks. In *Proc. 34th IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 2009.
- [5] P. Honeine, C. Richard, H. Snoussi, J. C. M. Bermudez, and J. Chen. A decentralized approach for non-linear prediction of time series data in sensor networks. *Journal on Wireless Communications and Networking*, 2010.
- [6] D. Puccinelli and M. Haenggi. Wireless sensor networks : Applications and challenges of ubiquitous sensing. *IEEE Circuits and Systems Magazine*, 5 :19–31, 2005.
- [7] C. Richard, P. Honeine, H. Snoussi, A. Ferrari, and C. Theys. Distributed learning with kernels in wireless sensor networks for physical phenomena modeling and tracking. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Honolulu (Hawaii), USA, 25–30 July 2010.
- [8] B. Schölkopf, R. Herbrich, and A. J. Smola. A Generalized Representer Theorem. In *COLT '01/EuroCOLT '01 : Proceedings of the 14th Annual Conference on Computational Learning Theory and and 5th European Conference on Computational Learning Theory*, pages 416–426, London, UK, 2001. Springer-Verlag.
- [9] B. Schölkopf, S. Mika, C. J. C. Burges, P. Knirsch, K.-R. Müller, G. Rätsch, and A. J. Smola. Input space versus feature space in kernel-based methods. *IEEE Transactions on Neural Networks*, 10 :1000–1017, 1999.
- [10] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, June 2004.
- [11] D. C. Steere, A. Baptista, D. Mcnamee, C. Pu, and J. Walpole. Research challenges in environmental observation and forecasting systems. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom'00)*, pages 292–299, 2000.
- [12] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.