Frequency-based inversion of a single wireless link for indoor passive target detection

Frequency-based inversion of a single wireless link for indoor passive target detection

Giorgio Gottardi, Mohammad Abdul Hannan, Alessandro Polo, Marco Salucci, and Federico Viani

ELEDIA Research Center (ELEDIA@UniTN - University of Trento), Via Sommarive 9, I-38123 Trento, Italy

E-mail: giorgio.gottardi@unitn.it

Abstract. In this work, the wireless detection of transceiver-free targets is formulated as an inverse problem and solved using the received signal strength of existing wireless devices. The main goal of the proposed solution is to opportunistically exploit a single wireless link established between two commercial wireless access points for the real-time detection of humans entering the monitored domain. The proposed computational method exploits both an empirical mode decomposition for denoising and a discrete Fourier transform for filtering the received signal strength. A reduced set of highly informative features are used for the automatic learning of a support vector machine classifier and the prediction of target absence/presence probability. A preliminary experiment carried out in a real indoor test site is presented to assess the advantages and limitations of the proposed inversion method.

1. Introduction

In the last years, the passive wireless localization, also called device-free localization (*DFL*), using the electromagnetic (*EM*) propagation of radio-frequency (*RF*) wireless networks has revealed strong potentialities in many applications related to security, surveillance, health-care monitoring, building management, and many other fields where location-based services need the detection and localization of non-cooperative people in indoor scenarios [1]-[5]. The main advantage of *DFL* strategies is that the target needs neither to carry active devices nor to cooperate to the localization procedure. The basic principle relies on the fact that the *RF* propagation is affected by the changes in the environment, including the time-varying presence and position of people acting as *EM* scatterers, especially if wireless systems operate at the microwave frequencies of the common wireless standards, such as the *IEEE802.11* (e.g., WiFi). Among the first solutions available in the state of the art for passive wireless localization, the localization problem has been modeled as a machine learning problem for the real-time estimation of target positions in wireless environments, through the real-time processing and inversion of the received signal strength (*RSS*) [5]. Many other reference works have been proposed to solve the numerous challenges of *RSS*-based passive localization, introducing more and more complex algorithms and techniques for performance improvement and for multi-target localization [6], including compressive sensing [7], Bayesian grid [8], support vector machines (*SVM*) [9].

This work addresses the challenge of detecting non-cooperative and device-free targets through the *opportunistic* exploitation of a *single* wireless link, without any kind of hardware and/or driver customization. The proposed method is aimed at maximizing the reliability of the passive

target detection by applying a customized inversion algorithm based on the analysis of the *RSS* time series in the frequency domain. More in detail, in order to enhance the effects of the target presence on the *RSS* data stream, the empirical mode decomposition (*EMD*) has been applied for the denoising of the raw input series. Successively, the denoised sequences have been transformed into the frequency domain by means of a discrete Fourier transform (*DFT*) for the extraction and selection of a reduced set of *DFT* coefficients. Finally, the small set of denoised and filtered coefficients have been used as training features of a learning by examples (*LBE*) method based on an *SVM* binary classifier [9]-[11]. The output of the *SVM*-based detector is the probability that a target occupies the area in proximity of the wireless link. The mathematical formulation of the proposed method is presented in Sect. 2, where the EMD-based technique, the feature extraction strategy, and the SVM-based detector are described in Sect. 2.1, 2.2, 2.3, respectively. The results of the preliminary experimental validation are presented in Sect. 3, whereas some conclusions and final remarks are summarized in Sect. 4.

2. Mathematical Formulation

Let us consider a single wireless link between the transmitter and the receiver located in the domain Ω in known positions $\underline{r}_{TX}$ and $\underline{r}_{RX}$, respectively, $\underline{r} = (x_\Omega, y_\Omega, z_\Omega)$ being the position vector. Heterogeneous obstacles like furniture and walls occupy the domain, as well as a variable number of $p = 1, ..., P$ human targets. Let $\mathbf{S}(t) = \{s(t - k\Delta t)\,;\, k = 0, ..., K-1\}$ denote the *RSS* vector available at the receiver at the time instant t and composed of K readings stored with a constant sampling rate Δt. The *RSS* time series is analyzed by considering the subset of data stored in the time window $\underline{\omega}(t) = [s(t - U\Delta t + u\Delta t)\,;\, u = 0, ..., U-1]$ of length $U \leq K$ and ending at the time instant t. The data samples contained in $\underline{\omega}(t)$ are directly processed by the proposed EMD technique (Sect. 2.1) without other pre-processing and/or filtering procedures.

2.1. Empirical Mode Decomposition for RSS *Denoising*

The *RSS* stream is decomposed in a finite number of intrinsic mode functions (*IMFs*), which represent the oscillatory modes embedded in the signal, adopting the so-called *sifting* process [12]. The reconstruction of the denoised signal is obtained combining only a subset of *IMFs* in order to remove the undesired frequency components. The proper selection of the significant *IMFs* is performed considering the Hurst exponent $H \in]0, 1[$, which is a measure of long-term memory of a time series [13]. Higher values indicate a persistent behavior and a long-term auto-correlation of the time series, while lower values refer to fast changes between adjacent pairs of the series (i.e., lower auto-correlation). According to this rule, high Hurst exponents are more related to the signal components while lower values are representative of the noise. The reconstructed time window obtained with the selected *IMFs* is

$$\underline{\tilde{\omega}}(t) = \sum_{g=1}^{G} \underline{\varsigma}_g(t) \tag{1}$$

where $\underline{\varsigma}_g(t)$, $g = 1, ..., G$, $G \leq I$, are the *IMFs* selected according to the following rule

$$\underline{\varsigma}_g(t) = \begin{cases} \underline{c}_i(t) & if\ \mathcal{H}(\underline{c}_i(t)) \geq H_{th} \\ 0 & if\ \mathcal{H}(\underline{c}_i(t)) < H_{th} \end{cases} \quad g = 1, ..., G \tag{2}$$

$\mathcal{H}(\cdot)$ being the function devoted to the computation of the Hurst exponent [12], and $H_{th} \in]0, 1[$ a user-defined threshold.

2.2. RSS *Features Extraction in the Frequency Domain*

The Euclidean distance between two *RSS* time series acquired in absence and in presence of targets has been adopted as a similarity measure in order to analyze the effects on the *RSS* stream caused by the target presence. In particular, the distance has been formulated as

$$\delta\left(\underline{\tilde{\omega}}^{(A)}(t), \underline{\tilde{\omega}}^{(P)}(t)\right) = \sqrt{\sum_{u=0}^{U-1}\left(\tilde{\omega}_u^{(A)}(t) - \tilde{\omega}_u^{(P)}(t)\right)^2} \tag{3}$$

where $\underline{\tilde{\omega}}^{(A)}(t)$ and $\underline{\tilde{\omega}}^{(P)}(t)$ are the denoised time windows acquired in absence and in presence of targets, respectively. In order to extract few and representative features of such difference, the time sequences are mapped in the frequency domain to isolate the frequencies of interest for the detection problem at hand. The *DFT* provides a mathematical tool to represent the denoised time window $\underline{\tilde{\omega}}(t)$ into periodic Fourier series. The sequence of coefficients $\underline{\xi} = [\xi_f;\ f = 0, ..., U-1]$ is given by

$$\xi_f = \frac{1}{U}\sum_{u=0}^{U-1} \underline{\tilde{\omega}}_u(t)\ e^{-\frac{j2\pi f u}{U}} \quad f = 0, ..., U-1 \tag{4}$$

where $j = \sqrt{-1}$ is the imaginary unit, and ξ_f, $f = 1, ..., U-1$, are complex numbers (ξ_0 is real since the input signal $\underline{\tilde{\omega}}(t)$ is real). According to the Parseval theorem, the energy of the signal in the time domain is the same as the energy in the frequency domain, and this implies that the Euclidean distance between two sequences in the time domain is the same as their distance in the frequency domain, formulated as

$$\hat{\delta}\left(\underline{\xi}^{(A)}, \underline{\xi}^{(P)}\right) = \sqrt{\sum_{f=0}^{U-1}\left(\xi_f^{(A)} - \xi_f^{(P)}\right)^2} \tag{5}$$

where $\underline{\xi}^{(A)}$ and $\underline{\xi}^{(P)}$ are the *DFT* coefficients computed in absence and in presence of targets, respectively. In order to further enhance the effects of the target presence on the Euclidean distance computed in the frequency domain, a band-pass filtering procedure is applied to the *DFT* coefficients. The bandwidth of the filter is calibrated to discard the frequency contributions not affected by the target absence/presence. The Euclidean distance computed with the filtered coefficients is formulated as

$$\Delta\left(\underline{\tilde{\xi}}^{(A)}, \underline{\tilde{\xi}}^{(P)}\right) = \sqrt{\sum_{f=f_{min}}^{f_{max}}\left(\xi_f^{(A)} - \xi_f^{(P)}\right)^2} \tag{6}$$

where $\underline{\tilde{\xi}}^{(A)} \subset \underline{\xi}^{(A)}$ and $\underline{\tilde{\xi}}^{(P)} \subset \underline{\xi}^{(P)}$ are the subset of filtered *DFT* coefficients, and f_{min} and f_{max} are the corresponding lower and upper coefficients indexes, respectively. The number of filtered *DFT* coefficients is $F = f_{max} - f_{min}$.

2.3. SVM*-based Target Detector*

The automatic estimation of the target absence/presence from the *RSS* data stream is based on a *SVM*-based classifier trained with a finite set of $m = 1, ..., M$ known input-output samples $\{\underline{x}_m, y_m;\ m = 1, ..., M\}$, where $\underline{x}_m$ is the m-th vector of input features, and $y_m \in \{-1, +1\}$ indicates the target absence (i.e., $y_m = -1$) and presence (i.e., $y_m = +1$). More in details, the input vectors include the Euclidean distance Δ_m, $m = 1, ..., M$, calculated between a

reference configuration in absence of targets (the corresponding DFT coefficients are $\underline{\tilde{\xi}}_m^{(A)}$) and a training configuration $\underline{\tilde{\xi}}_m^{(train)}$, $m = 1, ..., M$, which includes both the absence of targets (for the training samples $m = 1, ..., M/2$) and the presence of targets (for the remaining samples $m = M/2 + 1, ..., M$). Moreover, in order to also consider the trend of the DFT coefficients during the learning process, the input vector includes the statistical indicators of the coefficients, as follows

$$\underline{x}_m = \left[\Delta_m, \overline{\xi}_m^{(A)}, var\left(\underline{\xi}_m^{(A)}\right), \overline{\xi}_m^{(train)}, var\left(\underline{\xi}_m^{(train)}\right)\right], \tag{7}$$

where $\overline{\xi}$ and $var\left(\underline{\xi}\right)$ are the mean and the variance of the considered DFT coefficients. A linear SVM-based learning method is devoted to separate the training data set with the function $y = sign\left[w^T \varphi\left(\underline{x}\right) + w_0\right]$, where $sign\left(\cdot\right)$ is the binary sign function, w and w_0 are weight parameters computed during the training phase [14], and $\varphi\left(\cdot\right)$ is the linear function mapping the data from the input space to the feature space. However, since the RSS data acquired in the two absence/presence status are not linearly separable, the non-linear radial basis function (RBF) has been adopted for the features mapping (i.e., $k\left(\underline{x}, \underline{x}_m\right) = \varphi\left(\underline{x}\right)^T \cdot \varphi\left(\underline{x}_m\right) = e^{-\gamma \|\underline{x} - \underline{x}_m\|^2}$, $\gamma > 0$ being the RBF parameter [9][15]). Once the training phase is done and the model parameters are estimated, the SVM prediction computed during the online test phase is given by

$$y^{(test)} = sign\left[\Phi\left(\underline{x}\right)\right] = sign\left[\sum_{m=1}^{M} \lambda_m y_m k\left(\underline{x}, \underline{x}_m\right) + w_0\right] \tag{8}$$

λ_m, $m = 1, ..., M$, being the Lagrange multipliers computed by the optimization strategy during the SVM-based minimization procedure subject to constraints [15] and $\Phi\left(\cdot\right)$ the decision function. Moreover, instead of only predicting the binary labels of the test data, the proposed approach approximates the a-posterior class probability

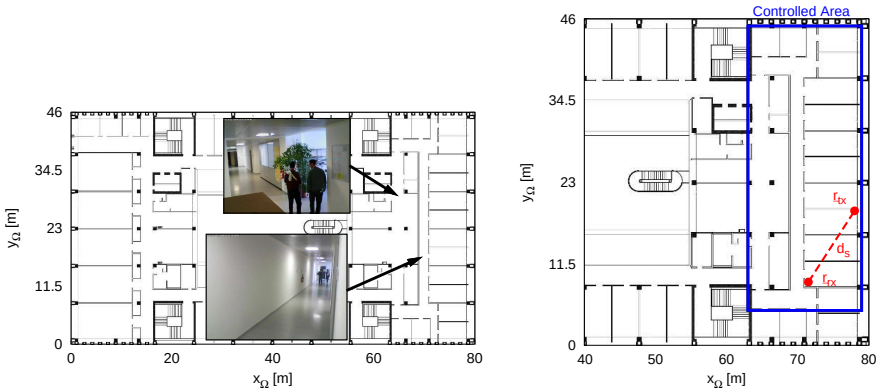
$$\Pi = Pr\left(y^{(test)} = 1 \middle| \underline{x}\right) = \frac{1}{1 + \exp\left(\alpha \Phi\left(\underline{x}\right) + \beta\right)} \tag{9}$$

where α and β are the parameters of the Sigmoid function estimated according to the Platt algorithm [15].

3. Preliminary Experimental Validation

The effectiveness and the robustness of the proposed method have been preliminarily assessed considering a real indoor test field located in the laboratories of the ELEDIA Research Center at the University of Trento, Italy (ELEDIA@UniTN). The whole building floor covers an area of $X_\Omega = 80$ $[m]$, $Y_\Omega = 46$ $[m]$, and $Z_\Omega = 2.8$ $[m]$ organized in offices and laboratories, with different and heterogeneous obstacles such as furniture and lab equipment [Fig. $1(a)$]. A single wireless link between two *IEEE 802.11bg* WiFi access points has been considered. The access points belong to the existing WiFi network, they are installed at the ceiling level and configured to operate at the working frequency $\phi = 2.4$ $[GHz]$ with transmitting power $P_{TX} = 18$ $[dBm]$. The proposed system stores the RSS readings in the vector $\mathbf{S}\left(t\right)$ with a constant sampling rate $\Delta t = 0.6$ $[sec]$ [1]. Figure $1(a)$ shows two sample snapshots of the camera views, while Fig. $1(b)$ reports the boundary of the area controlled by the surveillance system. Two surveillance cameras has been used to obtain the occupancy ground truth. The occupancy information (i.e., the number of occupants inside the detection area and their position) has been inferred from the video recordings and used for the manual labeling of the dataset.

[1] The sampling rate depends by the adopted hardware platform. The well-known Linksys WRT54GL has been adopted for validation. The standard scanning procedure provided by the OpenWRT operating system (v. 8.09, r14511, Kamikaze) has been used for the acquisition of the *RSS*.



(a) (b)

Figure 1. Indoor test site for experimental validation, (a) blueprint of the building floor with snapshots of ground truth acquisition, (b) test area with the wireless link of *Experiment 1*, $d_S = 11.9$ $[m]$.

Let us define the status of "target presence" when $P \geq 1$ humans occupy such sub-domain of the considered test site. The focus of this work is to estimate the binary status of target presence and target absence, while the impact of the number of humans on the detection probability will be investigated in future studies. The detection performance of the proposed strategy has been evaluated in terms of false positive rate ρ_{FP} [%], which refers to the wrong detections generated in absence of targets, false negative rate ρ_{FN} [%] pointing out how much the method fails to detect the target presence, and finally of failure rate ρ [%], which is the total error rate considering both false positives and false negatives. Consequently, the detection rate can be also computed as $\rho_D = (100 - \rho)$ [%]. More in detail, the performance metrics have been formulated as follows

$$\rho_{FP} = \frac{100}{K-U} \sum_{q=1}^{K-U} \left(\chi_q^{(+)} \Big|_{y_q^{(test)}=-1} \right) \tag{10}$$

$$\rho_{FN} = \frac{100}{K-U} \sum_{q=1}^{K-U} \left(\chi_q^{(-)} \Big|_{y_q^{(test)}=+1} \right) \tag{11}$$

$$\rho = \frac{100}{K-U} \sum_{q=1}^{K-U} \left(\chi_q^{(-)} \Big|_{y_q^{(test)}=+1} + \chi_q^{(+)} \Big|_{y_q^{(test)}=-1} \right) \tag{12}$$

where

$$\chi_q^{(+)} = \begin{cases} 1 & if\ Pr\left(y_q^{(test)} = 1 \middle| \underline{x} \right) > 0.5 \\ 0 & otherwise \end{cases} \tag{13}$$

and

$$\chi_q^{(-)} = \begin{cases} 1 & if\ Pr\left(y_q^{(test)} = 1 \middle| \underline{x} \right) \leq 0.5 \\ 0 & otherwise\,. \end{cases} \tag{14}$$

The experiment has been performed using test dataset not belonging to the training dataset. More in detail, the testing periods are temporally shifted respect to the training periods (the reference configuration assumes the test period with one week delay after the training). The considered APs are located in $\underline{r}_{TX} = (78.0,\ 19.0,\ 2.6)$ $[m]$ and $\underline{r}_{RX} = (71.5,\ 9.0,\ 2.6)$ $[m]$ and

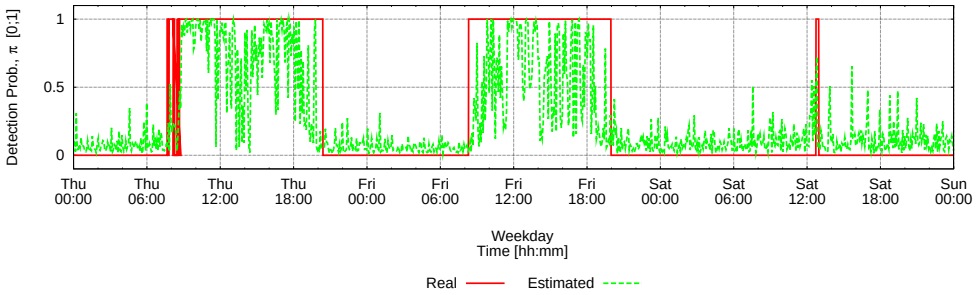


Figure 2. *Experimental Validation* ($U = 500$, $H_{th} = 0.6$, $f_{min} = 1$, $f_{max} = 10$, $M = 200$) - Time evolution of the real and estimated detection probability Π.

Table 1. *Experimental Validation* (short link, $U = 500$, $H_{th} = 0.6$, $f_{min} = 1$, $f_{max} = 10$, $M = 200$) - Performance metrics of the *3*-days measurement campaign.

Test Period	Performance Metrics		
	ρ_{FP} [%]	ρ_{FN} [%]	ρ [%]
Day 1 (*Thursday*)	1.02	7.09	6.73
Day 2 (*Friday*)	0.00	12.53	8.37
Day 3 (*Saturday*)	1.04	0.01	1.16
Total	0.69	6.54	5.42

the arising wireless link with line-of-sight (LOS) length $d_S = 11.9$ $[m]$ shown in Fig. $1(b)$ has been taken into account for the acquisition of the RSS time series. A 3-days measurement campaign has been carried out during working days and during the weekend in order to consider different user presence patterns. The area has been occupied by a variable number of targets $0 \leq p \leq 25$ during the measurements. A total number of $K = 442 \times 10^3$ RSS readings have been stored in $\mathbf{S}(t)$ and successively processed applying a sliding window $\underline{\omega}(t)$ of size $U = 500$. It has to be noticed that the detection algorithm is executed every $\Delta t = 0.6$ $[sec]$ and that each estimation has a "memory" of $U \times \Delta t = 300$ $[sec]$. The EMD denoising has been performed with a threshold $H_{th} = 0.6$ applied to the Hurst exponent for the selection of the most informative $IMFs$. The band-pass filtering procedure has been applied and the DFT coefficients between the lower index $f_{min} = 1$ and the upper index $f_{max} = 10$ have been selected, pointing out a significant reduction of the most important features, from $U/2 = 250$ (only half coefficients are considered by the method due to the symmetry of the Fourier spectrum) to $F = 9$. The SVM-based classifier has been trained with $M = 200$ samples well balanced between the absence and presence classes.

The detection probability Π has been computed for the whole *3*-days test period and the obtained results have been compared with the ground truth acquired by means of the video-surveillance system. The comparison shown in Fig. 2 points out a good matching of the estimated probability with the actual target presence, as also confirmed by the performance indicators reported in Tab. 1. The values ρ_{FP} [%], ρ_{FN} [%], and ρ [%] have been calculated for each day of the measurement campaign in order to analyze the detection performance during the different daily user activity patterns. As it can be seen, the higher false detection rate has been obtained during the second day ($\rho = 8.37$ [%]), mainly due to the high value of the false negative estimations ($\rho_{FN} = 12.53$ [%]) obtained during the Friday afternoon. It has to be noticed that, even if the ground truth declares the target presence during the whole afternoon, lower activity has been recorded by the surveillance system (less than 20 target movements have been detected from 2:30

PM to 7:00 PM). As expected, this result points out that the presence of targets that are standing within the domain in fixed positions without doing movements makes the passive detection from *RSS* perturbations very complex. However, even if the complexity to detect standing target is very high, the failure rate is still lower than 10 [%].

4. Conclusion

In this work, a *LBE* strategy for the detection of device-free targets in indoor scenarios has been presented. The *RSS* of a single wireless link has been analyzed in the frequency domain for the extraction of the frequency components more affected by the target presence. The proposed strategy has been designed to estimate the target detection for real-world location-based services where simple wireless architectures already exist for standard wireless communications. Further analyses are required to explore whether the proposed strategy can perform in different and more complex configurations, as well as to investigate the feasibility to localize and track the target when more wireless links are available.

Acknowledgments

This work benefited from the networking activities carried out within the Project "SMARTOUR - Piattaforma Intelligente per il Turismo" (Grant no. SCN_00166) funded by the Italian Ministry of Education, University, and Research within the Program "Smart cities and communities and Social Innovation", and the Project "CYBER-PHYSICAL ELECTROMAGNETIC VISION: Context-Aware Electromagnetic Sensing and Smart Reaction (EMvisioning)" funded by the Italian Ministry of Education, University, and Research within the PRIN2017 Program.

References

[1] Patwari N and Wilson J 2010 RF sensor networks for device-free localization: Measurements, models, and algorithms *Proc. IEEE* **98** 1961-1973
[2] Viani F, Migliore M D, Polo A, Salucci M and Massa A 2018 Iterative classification strategy for multi-resolution wireless sensing of passive targets *Electron. Lett.* **54** 101-103
[3] Ahmadi H, Polo A, Moriyama T, Salucci M and Viani F 2016 Semantic wireless localization of WiFi terminals in smart buildings *Radio Sci.* **51** 876-892
[4] Liu C et *al.* 2016 RSS Distribution-based passive localization and its application in sensor networks *IEEE Trans. Wireless Commun.* **15** 2883-2895
[5] Viani F, Robol F, Polo A, Rocca P, Oliveri G and Massa A 2013 Wireless architectures for heterogeneous sensing in smart home applications - Concepts and real implementations *Proc. IEEE* **101** 2381-2396
[6] Nannuru S, Li Y, Zeng Y, Coates M and Yang B 2013 Radio frequency tomography for passive indoor multi-target tracking *IEEE Trans. Mobile Comput.* **12** 2322 - 2333
[7] Wang J, Gao Q, Zhang X and Wang H 2012 Device-free localisation with wireless networks based on compressive sensing *IET Commun.* **6** 2395-2403
[8] Wang J, Gao Q, Cheng P, Yu Y, Xin K and Wang H 2014 Lightweight robust device-free localization in wireless networks *IEEE Trans. Ind. Electron.* **61** 5681-5689
[9] Massa A, Oliveri G, Salucci M, Anselmi N and Rocca P 2018 Learning-by-examples techniques as applied to electromagnetics *J. Electromagn. Waves Appl.* **32** 516-541
[10] Salucci M, Vrba J, Merunka I and Massa A 2017 Real-time brain stroke detection through a learning-by-examples technique - an experimental assessment *Microw. Opt. Technol. Lett.* **59** 2796-2799
[11] Salucci M and Oliveri G 2019 Robust real-time inversion of electrical impedance tomography data for human lung ventilation monitoring *Microw. Opt. Technol. Lett.* **61** 5-8
[12] Huang N E, Shen Z, Long S R, Wu M C and Shih H H 1998 The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis *Proc. of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **454** 903-995
[13] Hurst E 1951 Long-term storage capacity of reservoirs *Trans. Amer. Soc. Civil Eng.* **11** 770-799
[14] Cristianini N and Shawe-Taylor J 2000 *An Introduction to Support Vector Machines* (Cambridge, UK: Cambridge Univ. Press)
[15] Vapnik V N 1999 *The Nature of Statistical Learning Theory* (New York: Wiley)