A Preliminary Study on SVM based Analysis of Underwater Magnetic Signals for Port Protection

Davide Leoncini, Sergio Decherchi, Osvaldo Faggioni, Paolo Gastaldo, Maurizio Soldani, Rodolfo Zunino

Dept. Biophysical and Electronic Engineering, University of Genoa,
16145 Genoa, Italy
{davide.leoncini, sergio.decherchi, paolo.gastaldo, rodolfo.zunino}@unige.it

Istituto Nazionale di Geofisica e Vulcanologia,
19025 Portovenere (SP), Italy
{osvaldo.faggioni, maurizio.soldani}@ingv.it

Abstract. People who attend to the problem of underwater port protection usually use sonar based systems. Recently it has been shown that integrating a sonar system with an auxiliary array of magnetic sensors can improve the effectiveness of the intruder detection system. One of the major issues that arise from the integrated magnetic and acoustic system is the interpretation of the magnetic signals coming from the sensors. In this paper a machine learning approach is proposed for the detection of divers or, in general, of underwater magnetic sources. The research proposed here, by means of a windowing of the signals, uses Support Vector Machines for classification, as tool for the detection problem. Empirical results show the effectiveness of the method.

Keywords: underwater detection systems, port protection, magnetic signal processing, Support Vector Machine.

1 Introduction

For many years security has not been perceived by people as a necessity. Today, after some dramatic events such as September 11 2001, security issue has become a serious concern not only for governments. In this scenario the importance of physical security has increased; in particular, during the last five years, the research concerning underwater port protection has made some substantial achievements [1, 3-6].
First of all the target of underwater intruder detection systems has been extended from a military one, such as an enemy nation navy submarine, to a terrorist one, such as a diver intruder. This produced a secondary effect concerning the up to date of the technology used to detect underwater sources:

traditional sonar systems resulted to be insufficient to solve this task, bringing back importance to magnetic based systems [4, 5].

The analysis and comparison of the performances of the two different approaches point out their peculiarities: acoustic arrays guarantee optimum volumetric control but lack in peripheral surveillance; vice versa magnetic subsystems achieve high peripheral security performances but partially fail in volumetric control. These considerations suggest the integration of both detection approaches into a dual system [6].

This integration guarantees a good effectiveness to the complete system: overlapping of the acoustic and magnetic subsystems supplies shadow areas avoidance and consequently prevents possible intrusions. Moreover in the zone of maximum uncertainty of each method the lack in performance of one approach is counterbalanced by the co-occurring presence of the other cooperating subsystem. While acoustic systems today are a commercial reality, magnetic underwater surveillance is still an open research field.

These introductions lead to the demand of proper tools able to analyze the magnetic subsystem output. Beside classical analysis techniques [1, 3] the purpose of this paper is introducing a machine learning tool, Support Vector Machine for classification, as a possible approach for the detection of diver intrusion patterns on the supplied data. In particular, machine learning techniques have been already successfully used when involving with sonar signals [7]; here the purpose is showing that an analogous approach can be also carried when dealing with magnetic signals. Section 2 introduces the magnetic subsystem architecture while Section 3 exposes SVM theory, data extraction and experimental results.

2 The "MACmag" Magnetic Subsystem

Nowadays magnetic sensors have extremely high sensitivities and are able, in theory, to detect signals generated by divers without any problem. This capability is strongly compromised in practice by the spectral content of the Earth's magnetic field in high noise environments, such as port areas, characterized by an extremely wide band and high amplitude components, which often hide the target signal. Given M spectral components of the magnetic field, if we call E_i the energy associated with the i-th component, the information content Q is given by [2]:

$$Q = \sum_{i=1}^{M} E_i \quad \mathbf{(1)}$$

Whereas the information capacity C_i, that is the capacity associate to the i-th elementary spectral component with its physical generator, is given by the ratio between the energy E_i and the total energy in which it is contained:

$$C_i = \frac{E_i}{Q} \quad \textbf{(2)}$$

The range of value of C_i is between 1 (monochromatic signal) and 0 (white noise or insufficient target signal amplitude):

$$\lim_{Q \to E_i} C_i = 1 \qquad \lim_{Q \to \infty} C_i = 0 \qquad \lim_{E_i \to 0} C_i = 0 \quad \textbf{(3)}$$

Given two magnetometers, one as sentinel and the other as reference, to protect a critical area, one indicates with N the noise measured by both the magnetometers. By T is indicated the target signal acquired only by the sentinel magnetometer. As shown in [3] it can be stated that the sentinel listen to N+T and the reference measures the environmental noise N. If measured signals are acquired with the same clock, then the filtering operation is a simple subtraction in the time domain; in the case of uncertainties (even minimal) in the clock timing, the filtering operation is performed in the frequency domain with an increase in numerical inaccuracy.

This result can be obtained using two different architectures of the magnetic subsystem: the first employs the magnetic field acquired from the previous or next sensor in the array as noise reference (so that each instrument in the array operates both as sentinel and as reference) and is known as SIMAN-type network (Self-referred Integrated MAgnetic Network); the second is based on a sensor array and another external device used to obtain noise reference values (so that all the instruments in the array operate only as sentinel) and is called RIMAN-type network (Referred Integrated MAgnetic Network) [4, 6]. The system employed in the present work consists of two magnetometers in a SIMAN configuration. However, this configuration does not represent a full operational unit of the SIMAN network; a diver crossing halfway between the two sentinel magnetometers induces an analogous signal in both the devices and, consequently, this produces the target signal removal in the filtering process. Therefore, a full operational unit needs a third magnetometer which allows a comparison $\Delta(1,2)$, between the first pair of sensors, and $\Delta(1,3)$, between the second pair, such that the removal of the target can occur for at most one pair only (see Fig. 1) but not for the whole system.

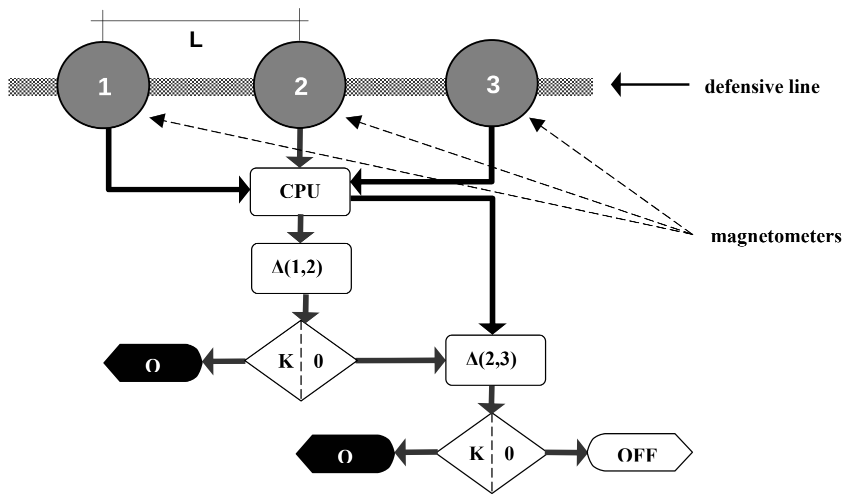

Fig. 1 - Operative structure of the elementary cell of the MAC mag subsystem.

Nevertheless the experimental configuration employed, including the two magnetometers, is clearly suitable for experimental validation of the MACmag component, with the exclusion of target crossings halfway between the two sensors. The magnetic signal used in ours experiments has been grabbed in this way from the sentinel and reference sensors in noisy environmental conditions and considering a civil diver as target.

3 Support Vector Machines for Classification

Support Vector Machines (SVM) constitutes a robust and well known classification algorithm [8]. The good classification performance of SVMs is due to the concept of margin maximization, whose roots are deeply connected with Statistical Learning Theory [8]. As usual in learning machines, SVM has a learning phase and a prediction phase. In the learning stage the machine sees the training patterns and learns a rule (an hyperplane) able to separate data in two groups according to data labeling. Conversely in the forward (prediction) phase the machine is asked to predict labels of new and unseen patterns.
From the formal point of view the following notation will be used:

- n_p is the number of patterns used as training set
- $\mathbf{X}$ is the training set
- $\mathbf{x} \in R^{n_i}$ is a pattern belonging to $\mathbf{X}$ where n_i is the data dimensionality

- $f(\mathbf{x}) = sign(\mathbf{wx} + b)$ is the prediction function based on the hyperplane defined by the normal $\mathbf{w}$ and the bias b
- $\mathbf{y}$ is the vector of labels of the training set, with $\mathbf{y} \in \{-1,1\}$

Given these definitions the cost function to be minimized for obtaining optimal $\mathbf{w}$ and b is:

$$C\sum_{i=1}^{n_p}(1 - y_i(\mathbf{wx} + b))_+ + \frac{1}{2}\|\mathbf{w}\|^2 \tag{4}$$

Where the positive constant C controls the tradeoff between data fitting (the first term) and regularization (the second term that represents margin maximization), and where $(k)_+$ indicates $\max(0,k)$.

Problem **(4)** can be solved via quadratic optimization algorithms; despite this fact, problem (**4**) is usually solved using its Lagrange dual formulation. The dual formulation makes possible to use non-linear mapping functions called kernel functions [8] that lead to non linear separating surfaces (see Fig. 2). This operation is possible observing that the only operations in which data are directly involved are dot products. Calling $K(\mathbf{x}_l, \mathbf{x}_m)$ the kernel dot product between $\mathbf{x}_l$ and $\mathbf{x}_m$ it can be shown [8] that the dual problem of (**4**) is:

$$\min_{\alpha}\left\{\frac{1}{2}\sum_{l,m=1}^{n_p}\alpha_l\alpha_m y_l y_m K(\mathbf{x}_l, \mathbf{x}_m) - \sum_{l=1}^{n_p}\alpha_l\right\} \quad \text{subject to:} \quad \begin{cases} 0 \le \alpha_l \le C, \forall l \\ \sum_{l=1}^{n_p} y_l\alpha_l = 0 \end{cases} \tag{5}$$

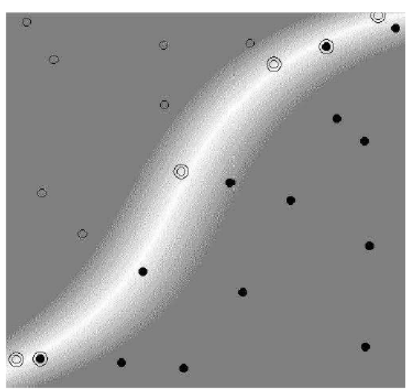

Fig. 2 - Non linear separating surface.

Where vector $\boldsymbol{\alpha}$ is of length n_p and represents the set of dual variables. Problem (**5**) poses the major problem of its optimization. To this regard fast optimization techniques has been developed [9]. One of these techniques, called Sequential Minimal Optimization [10], is the one that will be used for the following experimental section.

Once (5) has been optimized, as a final step, one has an efficient way to compute the bias b [10]. Finally, provided $\boldsymbol{\alpha}$ and b the non linear prediction function can be written as:

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{n_p} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b\right) \quad \textbf{(6)}$$

4 Experimental Results

The first addressed step is the definition of a suitable dataset for SVM based classification. Elaborated data refer to the problem of detecting the presence of a diver (class +1) or its absence (class -1).

Two quantities must be defined: the vector data $\mathbf{x}$ and its corresponding label y. The vector $\mathbf{x}$ can be created by windowing the signals coming from the magnetic subsystem: in particular given the original signal of length m, for each sample a window of width l is grabbed. This means that the total number of windows (superposition of windows is allowed) is m-l. Because the signals coming from the subsystem are two (reference and target), for each produced window the final pattern is built up by the concatenation of the two windows derived from the two signals. This translates in having m-l patterns $\mathbf{x}$ each of size $2l$. Using $l=100$ the number of produced patterns is considerable; for this reason a sub-sampling technique has been employed. To obtain a meaningful dataset the sections of the signal which are characterized by an intrusion have been more densely windowed than the sections in which no intrusion occurs (see Fig 3).

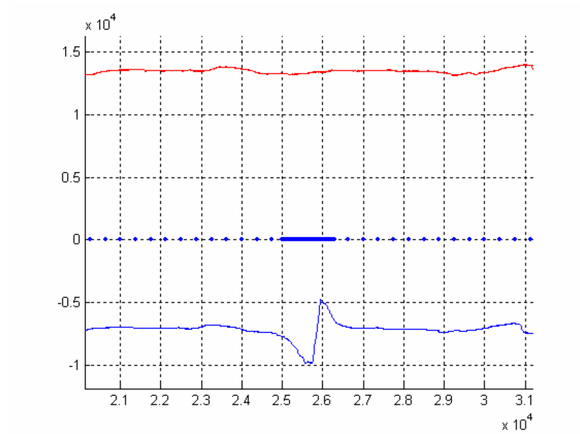

Fig. 3 - The upper signal is the reference signal; the lower signal is the target signal. Dotted line in the middle represents the windowing density.

Table 1 summarizes the statistics of training and test data after the above mentioned sub-sampling technique.

Table 1. Dataset overview

DataSet	Class +1	Class -1
Training Set	142	145
Test Set	144	150

After this preliminary step all data were normalized for each attribute in the domain [-1, +1]. The experimental session was carried by using a SVM with standard linear kernel [8] and SMO [10] optimizer. In particular the accuracy on the optimality conditions was set to 1e-3, a typical value for SVM training convergence (Karesh Kuhn Tucker conditions [8, 10]). The model was selected according to the *C* regularization constant (as per (**4**)) that led to the lowest test set error.

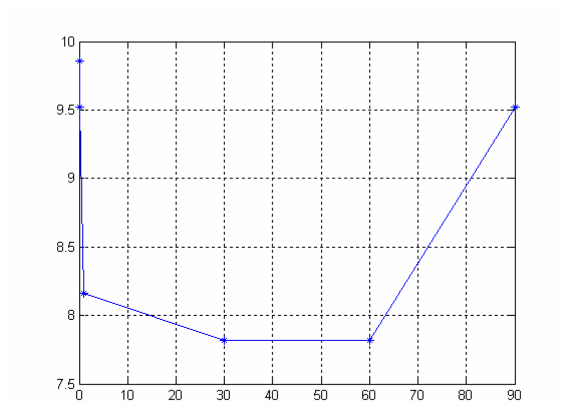

Fig. 4 – Model Selection curve: x axis are *C* values, y axis are percentage error values.

Figure 4 depicts the obtained curve for the *C* values {0.01, 0.1, 1, 30, 60, 90}; its shape is in accordance with theory [8], showing underfitting regions (small *C* values) and overfitting regions (big *C* values). The best performances are obtained with *C = 30* and *C = 60*; for both an error of 7.82% occurs. Recalling that an underfitting behavior is usually preferable to an overfitting one [8], the final selected parameter *C* was set to *C = 30.*

This preliminary experiment and the proposed approach seem to be promising although accuracy should be further improved to consider the system practically usable in on-field operations. Future research may investigate: extensions of the current preliminary study, other neural based techniques and alternative data preprocessing methods.

References

1. Faggioni O., Gabellone A., Hollett R., Kessel R.T., and Soldani M., Anti-intruder port protection “MAC (Magnetic ACoustic) System”: advances in the magnetic component performance, CD-ROM Proceedings of 1st WSS Conference, August 25-28, Copenhagen, Denmark, 2008.
2. Kanasewich E.R., Time Sequence Analysis in Geophysics, The University of Alberta Press, Edmonton, Canada, 1981.
3. Faggioni O., Soldani M., Gabellone A., Maggiani P., and Leoncini D., Development of anti intruders underwater systems: time domain evaluation of the self-informed magnetic networks performance, in Corchado E., Zunino R., Gastaldo P., and Herrero Á. (Eds.), Proceedings of the International Workshop on Computational Intelligence in Security for Information Systems CISIS 2008, Springer-Verlag, Advances in Soft Computing, vol. 53, pp. 100-107, Berlin Heidelberg, Germany, 2009.
4. Gabellone A., Faggioni O., Soldani M., and Guerrini P., CAIMAN Experiment, CD-ROM Proceedings of UDT Europe 2007 Conference, June 5-7, Naples, Italy, 2007.
5. Gabellone A., Faggioni O., Soldani M., and Guerrini P., CAIMAN (Coastal Anti Intruder MAgnetometers Network), CD-ROM Proceedings of RTO-MP-SET-130 Symposium on NATO Military Sensing, March 12-14, Orlando, Florida, USA, 2008, NATO classified.
6. Faggioni O., Soldani M., Zunino R., Leoncini D., Di Gennaro E., Gabellone A., Maggiani P.V., Falcucci V., and Michelizza E., Building the Synthetic “MAC System”: an Analytical Integration of Magnetic and Acoustic Subsystems for Port Protection Scenarios, CD-ROM Proceedings of UDT Europe 2009 Conference, June 9-11, Cannes, France, 2009, *in press.*
7. Hettich, S. and Bay, S. D. (1999). The UCI KDD Archive [http://kdd.ics.uci.edu]. Irvine, CA: University of California, Department of Information and Computer Science.
8. Vapnik V. Statistical Learning Theory, 1998, Wiley-Interscience Pub.
9. Schölkopf, B. and A.J. Smola: Learning with Kernels”. MIT Press, Cambridge, MA, USA (2002).
10. C.C.Chang, C.J. Lin., LibSVM: a library for Support Vector Machines [http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf].