Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 123 (2018) 122–127

Procedia Computer Science

www.elsevier.com/locate/procedia

8th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2017

Study of Efficiency of Dividing the Problem Space as a Means to Improve Solution of Multi-parameter Inverse Problem by Adaptive Methods*

Alexander Efitorov[1], Tatiana Dolenko[1,2], Sergey Burikov[1,2], Kirill Laptinskiy[1,2], and Sergey Dolenko[1]

[1] D.V.Skobeltsyn Institute of Nuclear Physics, M.V.Lomonosov Moscow State University, Moscow Russia

[2]Physical Department, M.V.Lomonosov Moscow State University, Moscow, Russia

a.efitorov@sinp.msu.ru, dolenko@sinp.msu.ru

Abstract

The considered multi-parameter inverse problem is determination of concentrations of salts or ions in multi-component water solutions of inorganic salts by Raman spectroscopy with subsequent spectra analysis by a non-linear adaptive method (multilayer perceptron type artificial neural networks (ANN) and by a linear adaptive method (partial least squares (PLS) method based on principal component analysis). Dividing the problem space into parts by data clustering simplifies the problem within each cluster but reduces the number of samples. This study compares efficiency of application of this approach for problems with different complexity (determination of concentrations of five salts, or ten salts, or ten ions) and with various distributions of samples over concentration range of the components. Based on experimental results, limitations and areas of application of the approach are discussed.

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures

Keywords: inverse problems, artificial neural networks, partial least squares, clustering, Raman spectroscopy

1 Introduction

Development of methods for rapid diagnostics of water media, in particular, for determination of their salt composition, is an extremely important issue at present time. Its relevance is due to the fact that water occupies an exceptional position in nature, and plays a special role in human life. To a large extent the quality of water is determined by its ionic composition – by inorganic compounds present in it in various concentrations and aggregate states. Violation of concentration balance of ions can have a dramatic negative impact on both human health and ecology state of the environment. Currently, extremely urgent tasks are monitoring of technical and waste waters; diagnostics and control of water of reservoirs used for irrigation of agricultural land; control of composition of mineral waters; determination of the salt composition of sea, river and reservoir waters etc.

Currently, there are many methods for determining ion composition of water media. They can be divided into two classes: 1) contact methods – chemical (titration, chromatography); 2) non-contact methods – radiometry and optical (spectrophotometry, spectroscopy of infrared (FTIR) absorption and Raman spectroscopy) (Crompton, 2002).

Chemical (analytical) methods provide high enough accuracy of determination of the concentration of ions – from units to a few hundredths of µg/l (Crompton, 2002). These methods are individual for

* This study has been performed at the expense of the grant of Russian Science Foundation (project no. 14-11-00579).

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures
10.1016/j.procs.2018.01.020

each ion; their implementation requires a long time, proper sample preparation and consumption of expensive reagents (Michalski, et al., 2011) (American Society for Testing and Materials, 1999). In addition, chemical methods require operative acquisition of preliminary information about the type and concentration of the ions present in the aqueous medium.

At present, to obtain operational information about the ionic composition of aqueous media, more and more involved are non-contact methods of diagnostics, providing information in real time – methods of vibrational spectroscopy (Raman and IR absorption) (Li, et al., 2011) (Ning, et al., 2012).

As it is well known (Chaplin, 2016) (Rudolph & Irmer, 2017) (Burikov, et al., 2005) (Rudolph & G., 2007) various ions and their concentrations significantly affect the shape and position of Raman and IR bands of water. In addition, if the salt contains complex cations or anions (e.g., NH_4^+, NO_3^-, PO_3^{2-}, SO_4^{2-}), with the change in the concentration of ions, proper bands of Raman vibrational groups of complex ions also change (Rudolph & G., 2007). This sensitivity of the Raman spectrum of aqueous solutions to the type and concentration of dissolved ions provides background to the development of an express method for remote determination of the type and concentration of ions in natural waters by Raman spectra. In (Burikov, et al., 2005) (Rudolph & G., 2007), methods for determining the concentration of specific ions in single-component aqueous solutions either by the change of the valence band (Burikov, et al., 2005), or by the change of proper vibrational bands of complex ions (Rudolph & G., 2007), have been proposed.

The authors of this study were the first who developed non-contact methods of identifying simultaneously several salts and determining concentrations of each of them in multi-component solutions by their Raman spectra. Such multi-parameter inverse problems (IP) were successfully solved by the use of artificial neural networks (ANN) (Burikov, et al., 2005). First, the problem was solved for three-component aqueous solutions of salts (Burikov, et al., 2005), and then for five-component solutions (Burikov, et al., 2010) (Dolenko, et al., 2012) (Dolenko, et al., 2014) (Efitorov, et al., 2015a) (Efitorov, et al., 2016) (Efitorov, et al., 2016a).

In their most recent studies, the authors increased to 10 the number of salts that may be present in the solution, and considered the IP of determination of concentration of separate ions (Dolenko, et al., 2015) (Efitorov, et al., 2015), which turned out to be much more difficult. For each of the salts, the spectrum shape is affected by presence of both the cation and the anion, and separate determination of concentration of each ion increases the non-linearity of the problem and the error of its solution.

2 Dividing the Problem Space: General Approach

The general approach to ANN solution of the described type of IP is the following. The ANN (usually a multi-layer perceptron, MLP) has as many outputs as there are components whose concentrations should be determined (salts or ions), and the desired value at each output equals the concentration of the corresponding component. Thus, a single MLP provides processing for the whole data array and for all components possibly present in the studied sample solution (simultaneous determination of IP parameters).

A commonly used approach to the solution of a complex problem is to divide it into several simpler ones. One type of simplification is creation of a separate single-output ANN to determine the concentration of each component (autonomous determination of IP parameters). However, such simplification succeeds only for not very complicated IPs (determination of concentrations of 5 salts). Increase in the problem complexity (determination of concentrations of 10 ions) makes autonomous determination inefficient (Efitorov, et al., 2016a).

Another way of simplification is to divide the problem space into several smaller areas, either by classification or by clusterization, and to solve the IP within each area by a separate ANN trained on the samples that fell into the corresponding class/cluster. This study compares the efficiency of this approach for three problems with various levels and types of complexity.