Classification of Ion Mobility Data Using the Neural Network Approach. T. A. Duong and I. Kanik[+],
Jet Propulsion Laboratory, California Institute of Technology 4800 Oak Grove Dr. Pasadena, CA 91109

Determination of atmospheric and surface elemental and molecular composition of various solar system bodies is essential to the development of a firm understanding of the origin and evolution of the solar system. Furthermore, such data is needed to address the intriguing question of whether or not life exists or once existed elsewhere in the Solar System. As such, these measurements are among the primary scientific goals of NASA's current and future planetary missions. In recent years, significant progress toward both miniaturization and field portability of *in situ* analytical separation and detection devices have been made with future planetary explorations in mind. However, despite all these advances, accurate *in situ* identification of atmospheric and surface compounds remains a big challenge. In response to that we are developing various hardware and software tools which would enable us to uniquely identify species of interest in a complex chemical environment. As a part of this effort, we developed a novel Cascade Error Projection (CEP) Neural Network (NN) approach for accurate classification of target compounds based on their ion mobility spectra. CEP is known as one of the most hardware-friendly algorithms based on the simplicity of the architecture, less bit resolution requirement for synapse, and fast convergence. Our goal is to demonstrate the feasibility of utilizing neural network approach to establish that ion mobility spectral data contain sufficient information to permit the development of a novel "*automated identification system*" applicable to NASA's future *in situ* planetary missions.

Since 1990, neural networks have been widely used in chemistry for classification of spectral data into chemical classes or functional groups [Bell et al., 1999; Zheng, et al., 1996; Boger and Karpas, 1994]. This approach has proven its capability to capture the nonlinear mapping between given input samples to target samples by parameterizing the strength of the interconnection matrix (called weight space) in the optimal way. This technique is called supervised learning technique. If the sufficient samples are given for the network to learn, when done, it will form an effective non-

linear transformation to predict the new and untrained samples in the most optimal approximation (non linear interpolation).

We developed a novel hardware learning algorithm for neural network namely "Cascade Error Projection (CEP)" [Duong, 1995] which is used to serve as a tool for identification and classification of the chemicals. This approach has been demonstrated for learning convergence in weight limited weigh space (hardware constraints) [Duong and Stubberud, 2000] and verified in hardware approach [Duong and Blaes, 2001]. The CEP has also been successfully demonstrated for its performance for color segmentation [Duong, and Daud, 2000], target recognition [Duong and Duong, 2003], and prediction [Duong and Blaes, 2001]. The CEP algorithm is implemented in a chip ($<5x5mm^2$) shown in Fig 1.
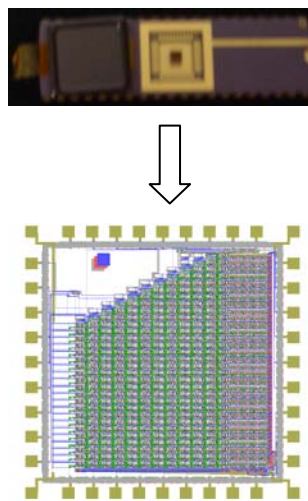

Fig. 1: 21 x 10 x 5 Cascading Neural Network chip developed by Lincoln Lab/MIT is shown.

Most supervised learning neural networks focus on a software-based approach, in which the learning trajectory is often smooth. When the weight component and the weight update are of 8-bit value (which is typically based upon hardware implementation constraints), learning convergence is often very difficult [Hollis et al., 1990].

On the other hand, hardware based neural network approaches are power hungry and not compact and is not suitable for mobile and/or remote application. Due to the nature of our hardware, it was necessary that we use a neural network learning algorithm which can be used on a miniaturized platform suitable for mobile and remote applications.

In this work, the ion mobility spectral data, obtained from electrospray ionization ion mobility spectrometer system, are used as input data for CEP neural network code to learn and validate. The details of the IMS system can be found in our earlier publications [Johnson et al., 2003]. In brief, *IMS* is a high-pressure time-of-flight (TOF) method based on elastic collisions of an ion with a buffer gas. It is advantageous over the other detection methods because it can operate at the same pressure as the ESI resulting in no differential pumping when the ions go from ionizer to analyzer. When an ion is placed in the electric field of the IMS it migrates in the direction of the field until it collides with another molecule. At that point it begins to accelerate again until it suffers another collision and so forth. This results in each species having an average drift velocity, which is proportional to the applied electric field and the ion mobility (K_m). The mobility itself is related to the size and shape of the ion imparting a second dimension of selectivity which enables IMS to separate isomers such as leucine and isoleucine [Beegle et al., 2001]. IMS has been demonstrated to be very sensitive in detecting organic compounds [Beegle et al., 2002] and is currently the instrument of choice for field detection of explosives and chemical/biological warfare agents [Baumbach et al., 1999].

Conclusions: It is demonstrated that the high-resolution ESI/IMS technique, equipped with a novel hardware learning algorithm (CEP neural network) has great potential to fulfill NASA requirements for detecting and accurately analyzing chemical compounds *in-situ*. Such an instrument would be to able to quickly detect and accurately identify organic compounds (such as biotic amino acids, abiotic amino acids, carboxylic acids, peptides, etc.) as part of an *in-situ* experiment on the surface of a planetary body such as Mars, Europa and Titan. Furthermore, it possesses ppb detection sensitivity and is largely free from any fragmentation problems owing to the soft ionization method (ESI).

References:

Baumbach J I and Eiceman G A, App Spec 53, 338A (1999)

Beegle L W et al., Anal Chem 73, 3028 (2001)

Beegle L W et al., J. Mass Spec. 216, 257 (2002)

Bell, S., Nazarov, E, Wang, Y. F. and Eiceman, G. A., Analytica Chimica Acta, 394 (1999) 121-133 (1999).

Boger, Z. and Karpas, Z. Anal. Chim. Acta 292 (1994) 243 (1994).

Duong T. A. and V.A. Duong, V. A., Journal of Advanced Computational Intelligence and Intelligent Informatics, pp. 289-293, Vol.7 No.3, 2003

Duong, T. A. and B. Blaes, B., Proc. Of the 5th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2001) and the 7th International Conference on Information Systems Analysis and Synthesis (ISAS 2001), pp.447-451, July 22-25, Orlando, 2001.

Duong, T. A. and Stubberud, A. L., International Journal of Neural System, Vol. 10, No. 3, pp. 199-210, June 2000.

Duong, T. A., Daud, T.,Proceeding of IASTED Intelligent System and Control, pp. 305-308, Honolulu, Hawai, August 14-17, 2000.

Duong, T. A., Proceeding Int'l IEEE/ICNN in Perth, Western Australia, vol. 1, pp. 175-178, Oct. 27-Dec 1, 1995.

Hollis, P., W., Harper, J., S., Paulos, J.,J. "The effects of Precision Constraints in a Back propagation learning Network", Neural Computation, vol 2, pp. 363-373 (1990).

Johnson, P. V. et al., J. Phys. Chem.. A, 108, 5785 (2004).

Zheng, P. et al. , Chemom. Intell. Lab. Syst. 33 121 (1996).

+: Corresponding author