Projective Adaptive Resonance Theory Revisited with Applications to Clustering Influence Spread in Online Social Networks

Jianhong Wu
LIAM, Department of Mathematics and Statistics, York University,
Toronto, Canada, M3J1P3.
Email: wujh@yorku.ca

Abstract— **We revisit the theory and applications of the Projective Adaptive Resonance Theory (PART) neural network architecture for clustering high dimensional data in low dimensional subspaces and nonlinear manifolds. We put a number of PhD theses, research publications and projects of the York University's Laboratory for Industrial and Applied Mathematics (LIAM) in a coherent framework about information processing delay, high dimension data clustering, and nonlinear neural dynamics. The objective is to develop both mathematical foundation and effective techniques/tools for pattern recognition in high dimensional data.**

Keywords-projective clustering; nonlinear dynamics in processing high dimensional data; influence soread in online social network.

I. INTRODUCTION

In a series of studies starting in the papers [1] [2] and the thesis [3], the Laboratory for Industrial and Applied Mathematics (LIAM) at York University (Toronto, Canada) has been developing a comprehensive framework about information processing delay, high dimension data clustering, and nonlinear neural dynamics. The objective is to develop both mathematical foundation and effective techniques/tools for pattern recognition in high dimensional data. Some of the earlier developments have been reported in the monograph [4] and the survey [5]. The survey also provided a heuristic description of the philosophy about how the modern nonlinear dynamic systems theory (invariant manifolds, domain attractions, global convergence, Lyapunov functions etc.) provides some theoretical principles based on recent biological evidences for novel neural network based clustering architectures to speed up information processing to assist decision making. Here, we briefly describe the current status (Section II) and then summarize the interdisciplinary nature (Section III) of a high dimensional data projective-clustering driven academic-industrial collaboration based on nonlinear dynamics and neural networks.

II. CURRENT STATUS

Under the framework "Projective Adaptive Resonance Theory" (PART), we *developed a novel neural network architecture and algorithm to detect low dimensional patterns in a high dimensional data set.* These are the patterns characterized by the so-called projective clusters, in nonlinear subspaces or nonlinear manifolds. PART *has received much attention by data science researcher and end-user community,* and has formed the core data analytics tools of three Collaborative Research Development projects (CRD), funded by the Natural Science and Engineering Research Council of Canada (NSERC). In particular, the NSERC CRD project Enterprize Software for Data Analytics in collaboration with InferSystems Inc. is based on the application of PART to analyzing odd bidding behaviors in a real-time bidding auction; while the project An Online Integrated Health Risk Assessment Tool brings together a team of investigators with expertise for an interdisciplinary and multi-institutional collaboration to develop systematic analyses converging on a single number, modeled after the single-number Credit Score, to inform chronic disease decision making, both at the population and individual levels. These are also part of a newly funded NSERC Collaborative Research and Training Experience Program in Data Analytics & Visualization.

The PART algorithm has since been used in a number of applications. It was used to develop a powerful gene filtering and cancer diagnosis method in [6][7][8], which shows that *"the results have proven that PART was superior for gene screening".* PART was also used for clustering neural spiking trains [9], ontology construction [10], stock associations [11], and information propagation in online social networks [12][13]. The PART algorithm has also been extended to deal with categorical data in the thesis [14].

The PART architecture is based on the well known ART developed by Carpenter and Grossberg, with a selective output signaling (SOS) mechanism to deal with the inherent sparsity in the full space of the data points in order to focus on dimensions where information can be found. The key feature of the PART network is a hidden layer of neurons which incorporates SOS to calculate the dissimilarity between the output of a given input neuron with the corresponding component of the template (statistical mean) of a candidate cluster neuron and to allow the signal to be transmitted to the cluster neuron only when the similarity measure is sufficiently large. *Recently discovered physiological properties* of the nervous system, the adaptability of transmission time delays and the signal losses that necessarily arises in the presence of transmission delay, enabled us to interpret SOS as a plausible mechanism from the self-organized adaptation of transmission delays driven by the aforementioned dissimilarity. The result is a

 ISBN: 978-1-61208-423-7

novel clustering network, termed PART–D, *with physiological evidence from living neural network and rigorous mathematical proof of exceptional computational performance* [15].

Such an adaptation can be regarded as a consequence of the Hebbian learning law, and the dynamic adaptation can be modeled by a nonlinear differential equation using dissimilarity driven delay in signal processing. This links to the PhD thesis [16], which proposed an alternative neural network formulation of the Fitts' law for the speed-accuracy trade-off of information processing, and its subsequent publications including [17][18][19][20]. When the delay adaption rates are in certain ranges, we observe nonlinear oscillatory behaviors (clustering switching) and this oscillation slows down the convergence of the clustering algorithm. How to detect and prevent these oscillations is the focus of the thesis [21] the studies [22][23].

III. SUMMARY

In summary, there have been increasing physiological evidences to support the idea of projective clustering using neural networks with delay adaption, there has been some theoretical analysis to show why such a network architecture works well for high dimensional data, and there have been sufficient applications to illustrate PART clustering algorithms are efficient. An interdisciplinary approach for high dimensional data clustering clearly shows the potential to develop a dynamical system framework for pattern recognition in high dimensional data.

REFERENCES

[1] Y. Cao & J. Wu, "Projective ART for clustering data sets in high dimensional spaces", Neural Networks, 15, 105-120, 2002.

[2] Y. Cao & J. Wu, "Dynamics of projective resonance theory model: the foundation of PART algorithm", IEEE Trans Neural Networks, 15, 245-260, 2004.

[3] Y. Cao, "Neural networks for clustering: theory, architecture, algorithm, and neural dynamics", York University, 2002.

[4] G. Gan, C. Ma, and J. Wu, Data Clustering: Theory, Algorithms and Applications, ASA-SIAM Series on Statistics and Applied Probability, 466 pages, 2007.

[5] J. Wu, "High dimensional data clustering from a dynamical systems point of view," in Bifurcation Theory and Spatio-Temporal pattern Formation, The Fields Institute Communication, vol. 49, W. Nagata and N. Sri. Namachchivaya eds. Philadophia: American Mathematical Society, vol. 49, pp. 117-150, 2006.

[6] H. Takahashi, T. Kobayashi, and H. Honda, "Construction of robust progonostic predictors by using projective adaptive resonance theory as a gene filtering method", Bioinformatics, 21:179, 2005.

[7] H. Takahashi, Y. Murase, T. Kobayashi, and H. Honda, "New cancer diagnosis modeling using boosting and projective adaptive resonance theory with improved reliable index", Biochemical Engineering Journal, 33:100, 2007.

[8] H. Takahashi, T. Nemoto, T. Yoshida, H. Honda, and T. Hasegawa, "Cancer diagnosis marker extraction for soft tissue sarcomas based on gene expression profiling data by using projective adaptive resonance theory (PART) filtering method", BMC Bioinformatics, 7:1, 2006.

[9] J. Hunter, J. Wu, and J. Milton, "Clustering neural spike trains with transient", Proc.the 47th IEEE Conferenceon Decision and Control, (December 9-11, 2008), pp. 2000-2005. DOI: 10.1109/CDC.2008.47387292008.

[10] R.-C. Chen and C.-H. Chuang, "Automating construction of a domain ontoloogy using a projective adaptive resonance theory neural network and bayesian network", Expert Systems, 25(4):414-430, 2008.

[11] L. Liu, L. Huang, M. Lai, and C. Ma, "Projective ART with bu ers for the high dimensional space clustering and an application to discover stock associations", Neurocomputing, 72:1283-1295, 2009.

[12] M. Freeman, J. McVittie, I. Sivak, and J. Wu, "Viral information propagation in the Digg online social network", Physica A: Statistical Mechanics and its Applications. 415:1, 87-94, 2014.

[13] G. Gan, J. Yin, Y. Wang, and J. Wu, "Complex data clustering: from neural network architecture to theory and applications of nonlinear dynamics of pattern recognition", Proc. 13th International Symposium on Mathematical and Computational Biology (The Fields Institute, November 4-8, 2013), pp. 85-106.

[14] G. Gan, "Subspace clustering for high dimensional categorical data". Master's thesis, York University, Toronto, Canada, October 2003.

[15] J. Wu, H. ZivariPiran, J. Hunter, and J. Milton, "Projective clustering using neural networks with adaptive delay and signal transmission loss", Neural Computation, 23(6):1568-1604, 2011.

[16] D. Beamish, "50 years later: a neurodynamic explanation of Fitts' law", PhD thesis, York University, 2004.

[17] D. Beamish, M. Bhatti, S. MacKenzie, and J. Wu, "Fifty years later: a neurodynamic explanation of Ftts' law", J. Royal Society Interface, 3(10):649-654, 2006.

[18] D. Beamish, S.A. Bhatti, C.S. Chubbs, I.S. MacKenzie, J. Wu, and Z. Jing, "Estimation of psychomotor delay from the Ftts' law coefficients", Biological Cybernetics, 101(4):279{296, 2009.

[19] D. Beamish, S.A. Bhatti, J.Wu, and Z. Jing, "Performance limitation from delay in human and mechanical motor control", Biological Cybernetics, 99(1):43-61, 2008.

[20] D. Beamish, S. MacKenize, and J. Wu, "Speed-accuracy trade-o in planned arm movements with delayed feedback", Neural Networks, 19(5):582-599, 2006.

[21] Q. Hu, "Differential equations with state-dependent delay: global Hopf bifurcation and smoothness dependence on parameters", PhD thesis, York University, 2008.

[22] Q. Hu & J. Wu, "Global continua of rapidly oscillating periodic solutions of state-dependent delay differential equations", Journal of Dynamics and Differential Equations, 22(2):253-284, 2010.

[23] Q. Hu, J. Wu, and X. Zou, "Estimates of periods and global continua of periodic solutions for state-dependent delay equations", SIAM Journal on Mathematical Analysis, 44(4):2401-2427, 2012.