

Chapter 15

Immune system inspired computing

Matti Pöllä and Timo Honkela

15.1 Introduction

Artificial immune systems, motivated by the natural immune principles, are emerging to overcome the limitations of conventional model-based techniques in combating with the uncertainties, vagueness, and high-dimensions of real-world problems.

In a collaboration project, lead by Professor Seppo Ovaska from the Institute of Intelligent Power Electronics of TKK, we have been investigating the fusion of artificial immune systems and soft computing [4] with an application in text data mining [5].

15.2 Anomaly detection

The task of detecting anomalies in a collection of data is one of the primary research topics of immunology-inspired engineering. Biological immune systems have evolved into various successful mechanisms for detecting bacteria and viruses while having no prior information on them. Correspondingly, immunology-inspired anomaly detection attempts to mimic these mechanism to develop classification algorithms for anomaly detection.

In [3] we have introduced a method for applying a negative selection algorithm to anomaly detection in textual data. The sparsity of discrete sequential data, such as written language, is decreased by analyzing individual character frequencies in a subset of the corpus. We use a collection of Wikipedia articles to show how little information on the original article is needed to detect and locate the changed parts.

The use of generative statistical models has also been the topic of recent research on text anomaly detection. In [1] we have augmented the mixture model scheme by Stibor [2] for arbitrary strings into a mixture-of-multinomials model. Aside from text mining, the presented method for anomaly detection is applicable for other types of symbolic sequential data such as gene and protein sequence analysis. We also compare the use of such generative models with the one-class support vector machine.

References

- [1] Matti Pöllä. A Generative Model for Self/Non-Self Discrimination in Strings. *Proceedings of ICANNGA'09: International Conference on Adaptive and Natural Computing Algorithms*, pages 293-302. Springer-Verlag, April 2009.
- [2] Thomas Stibor. Discriminating Self from Non-Self with Finite Mixtures of Multivariate Bernoulli Distributions. *Proceedings of Genetic and Evolutionary Computation Conference*, pages 127-134. ACM Press, 2008.
- [3] Matti Pöllä and Timo Honkela. Change detection of text documents using negative first-order statistics, *Proceedings of AKRR'08, The Second International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning*, pages 48-55. Porvoo, Finland, September 2008.
- [4] Matti Pöllä, Timo Honkela, and Xiao-Zhi Gao. Biologically inspired clustering: Comparing the neural and immune paradigms. In *Proceedings of NISCO 2007 Workshop on Nature Inspired Cooperative Strategies for Optimization*, pages 179-188, Acireale, 2008. Springer-Verlag. Other paper.
- [5] Matti Pöllä and Timo Honkela. Change detection of text documents using negative first-order statistics. In *Proceedings of AKRR'08*, pages 48-55, Porvoo, 2008. Other paper.