

# Analysis of Infectious Disease Data Based on Evolutionary Computation

Dong-bin Zhao and Jian-qiang Yi

Institute of Automation, Chinese Academy of Sciences  
{dongbin.zhao, jianqiang.yi}@ia.ac.cn

## 1 Project Background

An international cooperative research project on intelligence and security informatics (ISI) was initialized in 2005 with funding support by the Chinese Academy of Sciences. This abstract summarizes an ongoing study on ISI data analysis using evolutionary computation methods, conducted by the Chinese team of this international project.

Infectious diseases outbreaks, no matter by nature or terrorism, are critical threats to public health and national security. Analyzing infectious disease data will provide valuable information to prevent, detect, respond to, and manage infectious diseases. A critical challenge facing infectious disease data analysis is the clustering of disease cases or related observations in the spatial and temporal context, also known as spatial-temporal hotspot analysis [1]. Lots of researchers have invested their efforts on such an analysis. As an example of recent attempts, a risk-adjusted support vector clustering approach was successfully developed to cluster dead bird sightings as early warning of West Nile Virus outbreaks [2].

Evolutionary computation has become a major branch of computational intelligence in the past twenty years. Evolutionary computation algorithms are inspired and abstracted from natural evolution of species and have been applied to solve all kinds of optimization, modeling, and control problems. Examples of evolutionary computation algorithms include genetic algorithms, the recently proposed particle swarm optimization algorithm, the ant colony optimization algorithm, and the artificial immune algorithm.

Our research is aimed at developing data analysis approaches based on evolutionary computational algorithms and applying these approaches in the domain of infectious disease informatics.

## 2 Major Research Issues

Our research is currently focused on spatial-temporal hotspot analysis. Given the center and radius of candidate clusters, a clustering function can be formulated. This function can be optimized using evolutionary computation approaches to (near) optimally decide the center and area of infectious disease hotspots. Take the genetic algorithm as an example. The center and radius can be encoded as

an individual and the number of the infectious disease cases divided by the radius can be adopted as the fitness measure. Through a genetic learning process, an optimal or sub-optimal solution could be found, representing the infectious disease hotspots. Other kinds of evolutionary computation methods can also be applied to solve such optimization problems. We are currently experimenting with a selected set of such methods to perform spatial-temporal hotspot analysis. These methods are being evaluated based on their run-time performance, scalability, accuracy, flexibility, and extensibility in the context of ISI data analysis. One particular challenge is to identify multiple hotspots from the data at the same time. Possible solutions include (a) incorporating multiple cluster centers/radii into one fitness function, and (b) using immune algorithms with intrinsic multiple-modal considerations.

We are currently collecting data on Severe Acute Respiratory Syndrome (SARS) and avian influenza outbreaks. Such datasets will be used to test the proposed algorithms. For evaluation purposes, we will carry out a comparative study using both simulated and real-world datasets to compare our approach with existing ones (e.g., scan statistic-based approaches and the risk-adjusted support vector clustering method).

**Acknowledgement.** This work is supported by the National Science Foundation of China Grants 60573078 and 60334020.

## References

1. Chen, H.C., Wang, F.Y., Zeng, D.: Intelligence and Security Informatics for Homeland Security Information, Communication, and Transportation. *IEEE Transactions on Intelligent Transportation Systems* 5(4) (2004) 329-341
2. Chen, H.C., Wang, F.Y.: Artificial Intelligence for Homeland Security. *IEEE Intelligent Systems* 20(5) (2005) 12-16