Clustering and Dynamic Data Visualization with Artificial Flying Insect

S. Aupetit[1], N. Monmarché[1], M. Slimane[1], C. Guinot[2], and G. Venturini[1]

[1] Laboratoire d'Informatique de l'Université de Tours,
École Polytechnique de l'Université de Tours - Département Informatique
64, Avenue Jean Portalis, 37200 Tours, France.
{monmarche,oliver,venturini}@univ-tours.fr
sebastien.aupetit@etu.univ-tours.fr
[2] CE.R.I.E.S.,
20 rue Victor Noir, 92521 Neuilly sur Seine Cédex.
christiane.guinot@ceries-lab.com

Abstract. We present in this paper a new bio-inspired algorithm that dynamically creates and visualizes groups of data. This algorithm uses the concepts of flying insects that move together in complex manner with simple local rules. Each insect represents one datum. The insect moves aim at creating homogeneous groups of data that evolve together in a 2D environment in order to help the domain expert to understand the underlying class structure of the data set.

1 Introduction

Many clustering algorithms are inspired from biology like genetic algorithms [1, 2] or artificial ant algorithms [3,4] for instance. The main advantages of these algorithms are that they are distributed and they generally do not need an initial partition of data as it can be often needed. This study takes its inspiration from different kinds of animals that use social behavior for their movement (clouds of insects, schooling fishes or bird flocks) that have not been applied and extensively tested on clustering problems yet. Models of these behaviors that can be found in literature are characterized by a "swarm intelligence" which consists in the appearance of macroscopic patterns obtained with simple entities obeying to simple local coordination rules [6,5].

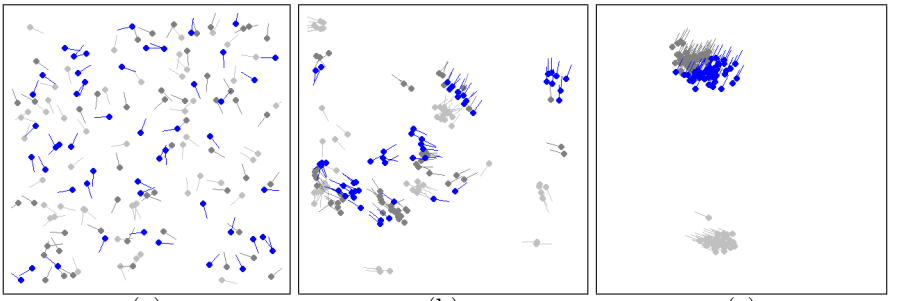
2 Principle

In this work, we use the notion of flying insect/entity in order to treat dynamic visualization and data clustering problems. The main idea is to consider that insects represent data to cluster and that they move following local behavior rule in a way that, after few movements, homogeneous insect clusters appear and move together. Cluster visualization allow the domain expert to perceive

E. Cantú-Paz et al. (Eds.): GECCO 2003, LNCS 2723, pp. 140–141, 2003.

the partitioning of the data. Another algorithm can analyze these clusters and give precise classification as output.

An example can be observed in the following pictures :



(a) (b) (c)

where (a) corresponds to the initial step for 150 objects (Iris dataset), (b) and (c) are screen shots showing the dynamic formation of clusters.

3 Conclusion

This work has demonstrated that flying animals can be used to visualize data structure in a dynamic way. Future work will concerns an application of these principles to present results obtained by a search engine.

References

1. R. Cucchiara. Analysis and comparison of different genetic models for the clustering problem in image analysis. In R.F. Albrecht, C.R. Reeves, and N.C. Steele, editors, *International Conference on Artificial Neural Networks and Genetic Algorithms*, pages 423–427. Springer-Verlag, 1993.
2. D.R. Jones and M.A. Beltrano. Solving partitioning problems with genetic algorithms. In Belew and Booker, editors. *Fourth International Conference on Genetic Algorithms*. Morgan Kaufmann, San Mateo, CA, 1991., pages 442–449.
3. E.D. Lumer and B. Faieta. Diversity and adaptation in populations of clustering ants. In D. Cliff, P. Husbands, J.A. Meyer, and Stewart W., editors, *Proceedings of the Third International Conference on Simulation of Adaptive Behavior*, pages 501–508. MIT Press, Cambridge, Massachusetts, 1994.
4. N. Monmarché, M. Slimane, and G. Venturini. On improving clustering in numerical databases with artificial ants. In D. Floreano, J.D. Nicoud, and F. Mondala, editors, *5th European Conference on Artificial Life (ECAL'99), Lecture Notes in Artificial Intelligence*, volume 1674, pages 626–635, Swiss Federal Institute of Technology, Lausanne, Switzerland, 13-17 September 1999. Springer-Verlag.
5. G. Proctor and C. Winter. Information flocking: Data visualisation in virtual worlds using emergent behaviours. In J.-C. Heudin, editor, *Proc. 1st Int. Conf. Virtual Worlds, VW*, volume 1434, pages 168–176. Springer-Verlag, 1998.
6. C. W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics (SIGGRAPH '87 Conference Proceedings)*, 21(4):25–34, 1987.