A Language Geographical Distribution Algorithm Based on Ant Colony Optimization

Hong Chen[1, a, *], Wensheng Zhang[1, b], Zhijian Zhao[2, c] and Haozhe Liu[2, d]

[1]Xi'an Eurasia University, No. 8 Dongyi Road, Xi'an, Shaanxi, China.

[2]Shaanxi University of Science and Technology, Shaanxi, China.

[a, *] messhiro@gmail.com, [b]7646524@qq.com, [c]496690080@qq.com, [d]liuhaozhest@gmail.com

Keywords: Ant Colony; Pheromones; Migration model.

Abstract. This article uses the ant colony algorithm to study the distribution change of languages in various regions under the same time period. We combined the neural network algorithm and made the population flow gradient as the main basis of pheromones to establish a world language distribution dynamic model.

1. Background

Globalization is an irreversible trend in current society. Under this circumstance, the multilingual promotion stimulates the market exploration and has become the necessary method in enterprises' internationalization strategy. And the multilingual speakers have become the practitioner of human civilization with global vision and humanistic quality.

2. Introduction

The ant colony algorithm is based on the bionic algorithm designed to simulate the shortest path behavior of ants looking for food. Therefore, in general, ant colony algorithm is used to solve the shortest path problem, and really in the traveling salesman problem (TSP, a problem of finding the shortest path) have achieved good results. At present, it has also been gradually applied to other fields and has applications in graph coloring problems, vehicle scheduling problems, integrated circuit design, communication networks and data clustering analysis.

3. Our Work

3.1 Algorithm Chosen

Ant System or Ant Colony System was first proposed by the Italian scholar Dorigo, Maniezzo et al. in the 1990s. In their study of ants foraging, they found that the behavior of a single ant is relatively simple, but the ant colony can reflect some intelligent behaviors. For example, colony can find the shortest path to food source in different environment. This is because the ants within the ant colony can pass the information through some information mechanism. After further study found that ants in the path of its release of what can be called a "pheromone" of the material. The ants within the colony have the ability to sense "pheromones," they follow a path of higher concentration of "pheromones," and every passing ant leaves a "pheromone" on the road, creating a similar Positive feedback mechanism, so that after a period of time, the entire ants will be along the shortest path to reach the food source.

3.2 Task Details

We are required to build a mathematical model to investigate trends of global languages and provides optimal location options for new offices opened by a transnational corporation. The problem is mainly analyzed into following part:

Demonstrate the process that geographic distributions of these languages change with the global population and human migration patterns over this same period of time.

4. ANN Based on Ant Colony Algorithm

4.1 The Model Description

We created a hypothesis assuming that the language distribution changes along with the flow of language speakers and predicted that the change of language distribution tend to be stabilized. Due to the expansion of internationalization trend, it's essential to put the population migration as the dominant factor and build the related model independently, so as to study the change of language distribution area.
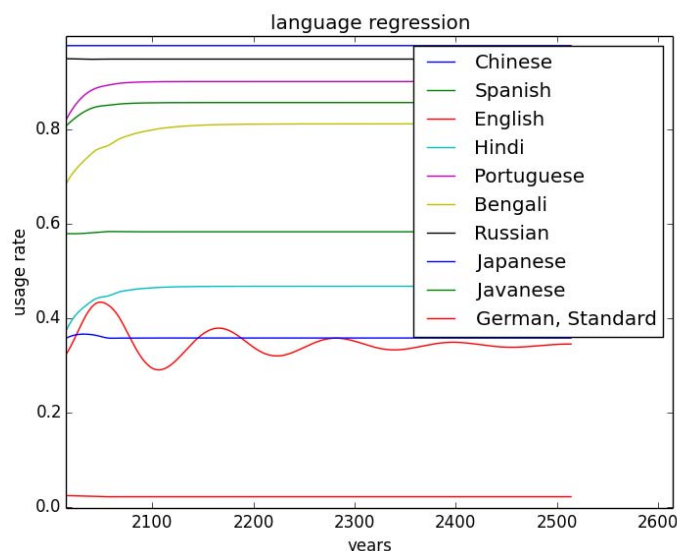

FIG. 1. Regression of Languages

Then we used ant colony algorithm. Specifically, the main analyzing object is language factor which is the mathematical abstract representation of each kind of language, like the ant in an ant colony in model. The place where language factors having the characteristic of finding the optimal solution are uninterested is likely to appeal the most pheromones (The population mobility gradient). When confronting obstacles, people would move randomly. But if there were no pheromone guides, people would move habitually towards the original direction, and they would also remember the passing points to avoid the loop. In the process of language flow, once finding the area of high mobility gradient (mobility gradient is a partial derivative of population flow to time), the language factor will release more pheromones. Conversely, the pheromones will become less when the language factors go far away.

The quantization of pheromones is the experience point of the whole algorithm. On the basis of the accuracy of algorithm, we utilized the neural network to definite the pheromones.

We used the historical information value in our model, set the population mobility gradient in an area as input, set the language rising gradient as output. And the corresponding rising gradient is the positive correlation factor of pheromones. The rising gradient of language is a partial derivative of the language to time.

Our hypothesis is not groundless. Based on our LDN model, we predict that the distribution of language tends to converge, and changes along with the flow of language speakers.

4.2 The Model Results

Type The preliminary guess is that the language distribution may fluctuate due to the influence of internationalization trend. In the next 50 years, the flow of population may be the main factor affecting the distribution of language. A large number of flows may cause the gradual disappearance of small languages and promote the integration of languages. Along with the passing of time, the geographical distribution of the top language is likely to be stable, and the distribution rate in possible areas of high density of low-ranking language may abruptly fall, even disappear.

4.3 The Model Assessment

The model fully acknowledges the random events and gives an event model big enough to the solution domain.

The positive feedback mechanism is used to converge the searching process and finally the result approaches the optimal solution.
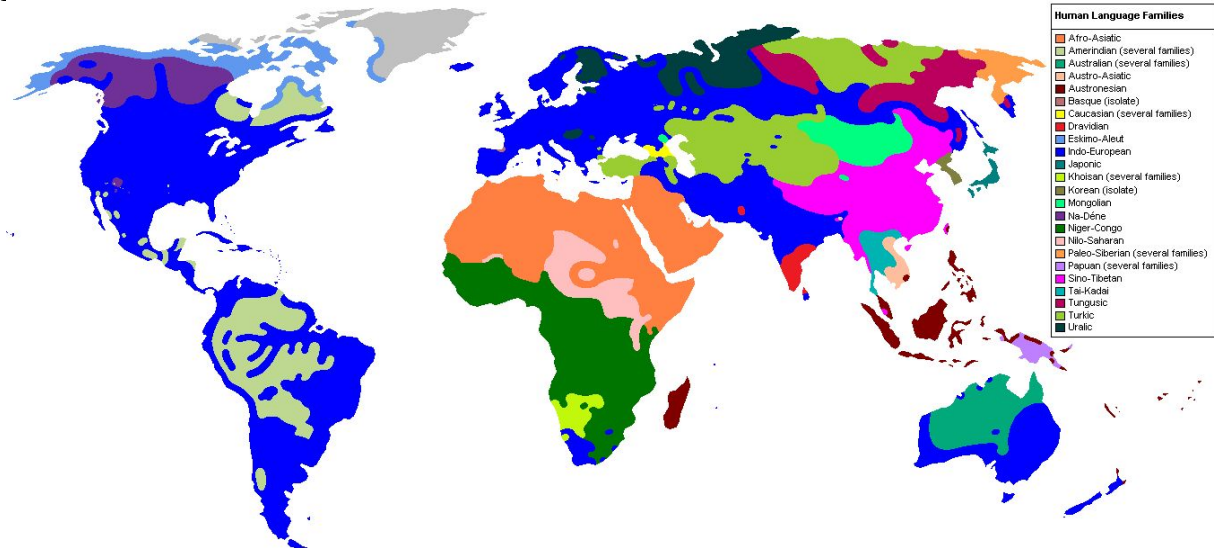

FIG. 2. Distribution of Languages

Each individual can change the environment and perceive the changes in the surroundings.

Distributed parallel computing is used to improve the computing power and operational efficiency.

Do not worry to be obsessed with the local optimal solution, and the global optimal solution is easy to obtain.

5. Strength and Weakness

Strengths:

Results processed by the ant colony algorithm objectively reflect the population flow around the world.

The model performs well when adequate data is provided.

Different from the other algorithm, the model itself does not need a large amount of population flow data. It only depends on the current distribution of population and PPP.

Weakness:

Since the huge amount of data comes from all over the world, we need many computer resources to calculating, so as to obtain the referential results.

6. Conclusion

The ant colony algorithm proposed in this paper is a good prediction of the future geographical distribution of language. We suspect that the language distribution may fluctuate under the influence of the trend of internationalization. Future population movements are likely to be the main factors affecting the distribution of languages. A large number of movements may result in the gradual disappearance of small languages and promote the integration of languages. As time progresses, the geographical distribution of the top-ranked languages is likely to gradually stabilize while that of the top-ranked languages where there is a possibility of high distribution regions may drop sharply or even die out.

Acknowledgements

References

[1]. S. Hochreiter and J. Schmidhuber. Long short-term memory. Neural Computation, 9(8):1735–1780, 1997.

[2]. A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, J. Schmidhuber. A Novel Connectionist System for Improved Unconstrained Handwriting Recognition. IEEETransactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 5, 2009.

[3]. Neural Machine Translation by Jointly Learning to Align and Translate.

[4]. M. Zlochin, M. Birattari, N. Meuleau, ET M. Dorigo, Model-based search for combinatorial optimization: A critical survey, Annals of Operations Research, vol. 131, pp. 373-395, 2004.

[5]. R. Hadji, M. Rahoual, E. Talbi and V. Bachelet "Ant colonies for the set covering problem," Abstract proceedings of ANTS2000, pp.63-66, 2000.

[6]. M. Ebling, M. Di Loreto, M. Presley, F. Wieland, et D. Jefferson, An Ant Foraging Model Implemented on the Time Warp Operating System, Proceedings of the SCS Multiconference on Distributed Simulation, 1989.

[7]. Clive Granger, (1981) "Some Properties of Time Series Data and Their Use in Econometric Model Specification", Journal of Econometrics 16: 121-130.

[8]. Fotheringham, AS; Wong, DWS (1 January 1991). "The modifiable areal unit problem in multivariate statistical analysis". Environment and Planning A. 23 (7): 1025–1044.