**"Vasile Alecsandri" University of Bacău**
**Faculty of Sciences**
**Scientific Studies and Research**
**Series Mathematics and Informatics**
**Vol. 23 (2013), No. 1, 23 – 28**

# ON THE ROBUSTNESS OF ANT COLONY OPTIMIZATION

GLORIA CERASELA CRIŞAN

**Abstract.** The real life faces many situations when people have to handle inconsistent data (i.e. contradictory facts, opposite opinions, etc.), managing them using personal and social experience and beliefs. Generally speaking, social life forms successfully and gracefully treat these conflicting cases – for example, by identifying the false facts and dropping them. In Computer Science and Information Technology, the problem of error identification and correction is a central one, approached in Code Theory, Software Engineering, Databases, Computer Architecture, etc.

This paper studies the behavior of a multi-agent software application when some of the agents have correct information about the solved problem, but the others do not. The well-known Traveling Salesman Problem (TSP) is solved using a biologically-inspired algorithm that models the way real ants manage to find the shortest path from nest to food source. The results are encouraging, showing that the artificial ants manifest the same robustness as real ants.

## 1. INTRODUCTION

We currently face situations when a decision has to be made based on uncertain or conflicting premises. From the childhood, we are trained to reason using probable, possible, or dynamic facts; as adults, our social experience often helps in successfully solving this type of situations. When computing systems have to solve complex real-life problems, they have to use large amount of (imperfect) data in order to provide good solutions.

---

**Keywords and phrases:** Big data, Information quality, Inconsistency, Ant algorithms, Combinatorial Optimization.
**(2010) Mathematics Subject Classification:** 68T20, 68P99, 90B80, 49N30.

The didactic problems, having well-known solving methods (for example the exact methods for the P problems, or the approximation methods for the NP problems) have to be re - defined, in order to tackle the uncertainty in data. This supplementary level of complexity brings new challenges to the researchers: the formalization of the problem is modified; the uncertain attributes could complicate beyond tractability the search process; the current models could no longer be appropriate.

This paper uses a social model from nature (millions of years of evolutionary processes can be seen as successful, as they brought the complex and beautiful world we live in) in order to solve some difficult variants of optimization problems, with uncertain features. Section 2 explains the starting idea of this investigation. Section 3 introduces the problem approached and Section 4 describes the method used. Section 5 presents the results and concludes the paper.

## 2. THE MOTIVATION: UNCERTAINTY, A COMMON FEATURE OF CURRENT DATA

These days, a huge amount of data is produced and can be stored, interpreted, analyzed, and visualized. At the time this paper was prepared, Facebook announced the March 2013 operational results: 665 million daily active users (DAUs) on average, 1.11 billion monthly active users (MAUs), and 751 million Mobile MAUs [4]. The Large Hadron Collider (LHC) from Geneva produces data in the order of one million gigabytes per second. Due to sophisticated and instantaneous selection procedures, only 0.00001% of these data are actually stored. After such a drastic data reduction, 25 petabytes per year is collected and investigated by the Worldwide LHC Computing Grid (WLCG) [8].

The Big Data – data that are inappropriate to manage using the RDBMSs, as we understand them today – brings to the scientists a new way of approaching the Science: the Data-Intensive Discovery paradigm. We are now facing new fundamental changes, affecting the world at different scales: the Information governance at global scale, the NoSQL [7] technology developments, and the new job of chief data officers at micro scale are just three of their effects. Such complex, dynamic and huge volume of data with some degree of inconsistencies, uncertainties, and approximations is now available to the scientists. The exploration and exploitation of data that is not perfect are now at their beginnings [2, 5].

As a model of the Big Data imperfection, this paper studies the modification of a didactic optimization problem – the well known Traveler Salesman Problem – by injecting some degree of inconsistency in data, as defined in [1].

## 3. THE PROBLEM: TRAVELER SALESMAN PROBLEM WITH INCOMPLETE INFORMATION

One of the oldest combinatorial optimization problems is the Traveler Salesman Problem (TSP) which basically seeks for a shortest tour connecting all the cities in a map. With unclear origins, the formal TSP was stated in the past '30s and since then it is constantly receiving attention from the scientists. TSP has obvious representation in real life; with a simple formulation and multiple variants, it is suited to multiple generalizations and interpretations.

The current developments of computing and communication systems allow solving high-dimension TSP instances, with supplementary restrictions, as the real-world situations request. Such TSP variants are: asymmetric TSP, bottleneck TSP, generalized TSP, etc. Other important optimization problems can be seen as TSP generalization; one such example, the Vehicle Routing Problem (VRP), has important economic implications, as the price of virtually any traded good has transportation costs included.

Some recent preoccupations concerning TSP focus on considering imperfect data: instead of complete knowledge at the solving moment, the computer application has to provide the solution using only incomplete information. Such situations are: stochastic TSP, fuzzy TSP, dynamic TSP, interval TSP, TSP with errors. All these cases have enhanced difficulty, and sometimes even the theoretical notions have to be re-defined (for example, the *solution* concept in the interval TSP case).

This paper focuses on the TSP with errors (TSPE) [1]. It supposes that a degree of data incoherence is manifested, due for example, to errors in land measurements. When processing incoherent data, the algorithm is expected to provide a weaker solution than in the perfect data processing case. The scientists are interested in finding fault-tolerant algorithms, supplying reliable solutions. The TSP variant we approach here is solved using a nature-inspired method, as we hope that modeling the robustness expressed by life forms developed along millions of years of evolution can be useful in solving our modern problems.

Specifically, we attached two sets of distances between the cities: the correct (consistent) ones and the altered ones. Data alteration is made by introducing a degree of incoherence, measuring how much data is altered, how much far (in average) is the altered value from the correct one [1], and how many artificial ants use the altered dataset. This supplementary metric is used in order to study the robustness of the solving method, presented in the following.

## 4. THE METHOD: ANT COLONY OPTIMIZATION

The biologists discovered that the ants (virtually blind, social insects) manage to find the shortest path from their nest to the source food. During their walks, they deposit pheromones, and the following ants prefer to use the high pheromone-intensity paths. This cooperation was modeled in Computer Science by Ant System (AS) - a multi-agent system made by artificial ants that concurrently solve TSP and place artificial pheromone of the components of their solutions [3]. By reinforcing the "good" components (appearing more often in "good" solutions), the following ants tend to often use them, and so exploiting the previous knowledge. Although many of the solutions are weak, when repeatedly searching for a solution, the artificial ants tend to more and more use the shorter paths, as ants do in real life. AS was a success and many other ant-based algorithms were since defined for many other problems.

Ant Colony Optimization (ACO) is the meta-heuristic that globally refers to the Ant algorithms devised for combinatorial optimization problems [3]. These days, ACO algorithms are used for solving transportation, logistic, communication, allocation problems [6]. Specific variants of ACO are used for incomplete information problems, providing very good results in dynamic or stochastic environments.

The TSPE (described in Section 3) is solved using two matrices of distances. The matrix $D$ holds the correct (consistent) distances between each pair of cities, and the matrix $E$ holds the incorrect distances. The matrix $E$ is constructed by initially copying the matrix $D$, and then $p$% stochastically chosen values (and their symmetric elements) are altered within an interval of length $l$, centered in the correct value. From the total artificial ants, $e$% use the $E$ distance matrix. Three new parameters are added to the classic AS, in order to adapt it to the TSPE. Of course, if $p = 0$, we have the classic TSP. So, we can consider that these new parameters ($p$, $l$, and $e$) measure the TSPE's *degree of incoherence*.

## 5. RESULTS AND CONCLUSIONS

In order to test the ACO robustness on the modified TSP already described in Section 3, we used the instance *lu980* (Luxembourg) having 634 nodes from the National [9] benchmark collection. The optimum integer value for this instance is *11340*.

We implemented the AS in Java, and ran it on a common PC (Pentium dual-core at 2.5 GHz, with 2GB RAM). The application was executed 5 times for each parameter value, and the integer average for the results was taken. All other AS parameters are maintained constant. The solution provided by the same algorithm for the classic, consistent TSP is *13320*. The results are presented in Table 1.

| p | l | e | TSPE solution | TSPE % error |
|---|---|---|---|---|
| 0.10 | 5 | 2 | 13298 | -0.17% |
| 0.10 | 5 | 5 | 13337 | 0.13% |
| 0.25 | 5 | 2 | 13316 | -0.03% |
| 0.25 | 5 | 5 | 13304 | -0.12% |
| 0.50 | 5 | 2 | 13284 | -0.27% |
| 0.50 | 5 | 5 | 13368 | 0.36% |
| 0.75 | 5 | 2 | 13298 | -0.17% |
| 0.75 | 5 | 5 | 13357 | 0.28% |
| 1.00 | 5 | 2 | 13416 | 0.72% |
| 1.00 | 5 | 5 | 13386 | 0.50% |

**Table 1. AS results for different TSPE parameter values**

The results show a very good fault tolerance of the AS: the application solving TSPE always finds solutions within a 1% error interval, manifesting the same robustness as a colony of real ants. Even if a part of the artificial ants operate with erroneous data, they are successfully balanced by the majority that uses correct information. More investigation has to be made in finding the characteristics of the triplets (*p*, *l*, *e*) that trigger bad behavior, meaning that the ant colony can not cope with the severity of data errors.

This paper has a two-fold contribution. Firstly, it offers a model of the newest evolutions in data: with an affordable computing system it approaches an imperfect collection of data, with a controllable degree of incoherence. Secondly, it shows that nature-inspired methods designed for perfect data, didactic problems can be successfully generalized for solving complex, real-life situations, manifesting some data imperfection.

## REFERENCES

[1] G. C. Crişan, **Ant Algorithms in Artificial Intelligence**, PhD Thesis, Faculty of Informatics, Al. I. Cuza Unversity of Iaşi, Romania, 2007.

[2] G. C. Crişan, C. Pintea, C. Chira, **Risk assessment for incoherent data**, Environmental Engineering and Management Journal 11 (12) (2012), 2167-2172.

[3] M. Dorigo, T. Stützle, **Ant colony optimization**, MIT Press, 2004.

[4] Facebook Reports First Quarter 2013 Results, available at http://investor.fb.com/releasedetail.cfm?ReleaseID=761090 accessed June 8$^{th}$, 2013.

[5] E. Nechita, C. Muraru, M. Talmaciu, **A Bayesian approach for the assessment of risk probability. Case study for digital risk probability**, Environmental Engineering and Management Journal 11 (12) (2012), 2249-2256.
[6] C. M. Pintea, D. Dumitrescu, P. C. Pop, **Combining Heuristics and Modifying Local Information to Guide Ant-based Search,** Carpathian J. Math. 24 (1) (2008), 94-103.
[7] P. J. Sadalage, M. Fowler, **NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence**, Addison-Wesley Professional, 2012.
[8] The Grid at CERN, available at http://cds.cern.ch/record/1421296/files/CERN-Brochure-2012-001-Eng.pdf accessed June 8th, 2013.
[9] Traveling Salesman Problem benchmarks, available at http://www.tsp.gatech.edu/ accessed June 8th, 2013.

"Vasile Alecsandri" University of Bacău
Faculty of Sciences
Department of Mathematics, Informatics and Education Sciences
157 Calea Mărăşeşti, Bacău, 600115, ROMANIA
e-mail: ceraselacrisan@ub.ro