International Journal of Information Technology and Knowledge Management
January-June 2011, Volume 4, No. 1, pp. 315-318

A GENETIC ALGORITHM FOR MANAGEMENT DATA STREAM IN VANET

S. Raghay[1], S. El Hadaj[2] & T. Noel[3]

Vehicular Ad-Hoc Networks (VANET) consists of moving vehicles exchanging data wirelessly without an existing infrastructure. With such communication mechanism a network participant can receive information directly from its close neighbours, as well as indirectly from any other vehicles in the network by using the rest of its peers as intermediate relays.

Different types of information can be exchanged in the context of VANET, especially to alert drivers of the occurrences of significant, possibly life-threatening, events. As a result, during their travel, the vehicles are flooded by information provided by others. The relevant data are then used locally to inform or alert the driver, or stored to be further disseminated afterwards. In the recent years, many works, such as [1], [2], [3] and [4], were aimed at studying some of the aspects of inter-vehicular information transmission and evaluation.

These studies have used the gathered data only to alert the drivers, considering it obsolete afterwards. In [4], the authors focused on the dynamic generation of knowledge from data collected by the vehicles to provide information to drivers, even if no neighbour is currently communicating.

In this paper, we propose a model for managing the data flow and we use genetic algorithms (GA) to determine the sufficient number of vehicles to generate knowledge from the data gathered by the population.

Keywords: Data Stream, VANET, Genetic Algorithms, Inter-vehicular Information Transmission

1. Introduction

Information has become, in our era, very important in the business, in the home and on the road in the car or vehicle in general. In addition to being a significant information generator, the car is also used as a means of disseminating this information. One of the problems yet to be addressed is how to determine the potential aptitude of vehicles most likely to convey in shortest time without relying on an existing infrastructure (e.g. in an ad-hoc manner).

Many algorithms have been designed in attempt to solve this problem, each with different number and type of parameters and varying complexity. We site a few in follows

- Single sampling algorithms [5] based on random selection of data from vehicles or other sources depending on the application domains.
- Reservoir sampling algorithms: an element of a given buffer is randomly selected, depending on the life time of each element in the buffer. The buffer is initially set to the first N elements of the data source [6].
- The algorithms with priority reservations: As in the previous algorithms when an item expires from the main sample, it is replaced by the item corresponding to highest priority, above those generated randomly for each element of the reservoir [7].
- The Vector algorithm [8]: In this paper, the authors have presented the description of a new method of forming a representative segment using the reservoir sampling algorithm and based on vector calculations.

These techniques select a data segment [9], which may probably be the most representative, from sources which can be as varied, generating a huge flow of data. However in our paper in addition to the more representative segment we treat the most appropriate sources.

We introduce in this paper genetic algorithms to choose the best generation (in this set of vehicles) that can circulate information as soon latencies, more accurately and for a large of vehicle having received the information and ensuring the best performance of diffusion.

1.1. Problematic

Today in order to optimize the traffic flow and to safeguard human lives, the means of transport need to exchange information on road and traffic conditions. The necessary condition for the dissemination of this information is the choice of the vehicles most likely to deliver it to the destination as soon as possible while minimizing the consumption of bandwidth of the wireless network.

[1]LAMAI, FSTG, B.P 549, Marrakech
[2]ENCG, Marrakech
[3]LSIIT, Université de Strasbourg
Email: said.raghay@hotmail.com[1], elhadajs@yahoo.fr[2], noel@unistra.fr[3]

To do this, in a network of vehicles (population), we are looking for the best group of elements (generation) which will generate and transport data. Thus the selection of these vehicles will follow a research process using genetic algorithms.

Genetic Algorithms

Genetic algorithms were originally developed by Bremermann [10] in 1958 and popularized in the early 1970s by John Holland [11]. These are processes that have the probability to find an approximate solution to an optimization problem based on the concept of natural selection.

The core principle of GA is to not exhaustively cover all points of the solution space, but from a set of solutions probably best known as initial population we apply transformations to just to keep improving the best one generation after a number of iterations.

The execution of a GA imitates biological processes, by simulating mutation and crossing of individuals of the population in question.

Mutation: A minor transformation of a person.

Crossing: A combination of two individuals, which generates a new individual.

Selection: As the GA are based on the principle of natural selection, an evaluation stage of the individuals by mutation or crossover is needed to proceed with the selection of the best elements of the population.

Thus the various stages of execution of a GA can be summarized in the following algorithm:

1. Choose the initial population of individuals.
2. Evaluate the fitness of each individual in that population.
3. Repeat on this generation until termination: (Stabilization of population size)
   a. Select the best-fit individuals for reproduction.
   b. Breed new individuals through crossover and mutation operations to give birth to offspring.
   c. Evaluate the individual fitness of new individuals.
   d. Replace least-fit population with new individuals.

Modeling

In this work, the goal is to find the best population of vehicles that generate and distribute a packet of information in a VANET network. Global methods, such as those based on genetic algorithms are particularly useful in treatment of new problems including the nature of the solution space is unknown.

To use a genetic algorithm, we need a selection criterion that depends on a fitness function. This criterion should allow the selection of a good population. In this work, we considered the fitness function:

$$F(£, P, \zeta) = \alpha_1 £ + \alpha_2 P + \alpha_3 \zeta$$

Where:

α_k: effect coefficient with $\alpha_1 + \alpha_2 + \alpha_3 = 1$,

£: Latency,

P : Percent of vehicles that received the message,

ζ : Intelligence of a vehicle (calculated according to the vehicle's ability to evaluate information before dissemination, the performance of a broadcast, the accuracy in estimating the location of other vehicles on the road, ...).

The performance of genetic algorithms depends on the right choice of mutation and crossover operators. To use these operators in a VANET, we modeled an individual as an information packet consisting of two parts: one containing only the parameters relative to the vehicle (ID, location, speed, etc.) and a second part containing the set of data to be disseminated.

It is said that a vehicle has mutated if one component of the first part of the package has changed.

We say that the individual X was crossed with an individual Y when X sends a packet of information to Y, which in turn amends the second part of this package. We create a virtual element that is injected into the population to apply the actions of selection, mutation and crossover.

An individual X is called efficient if its fitness exceeds a minimum calculated as an average performance of the initial population.

NUMERICAL RESULTS

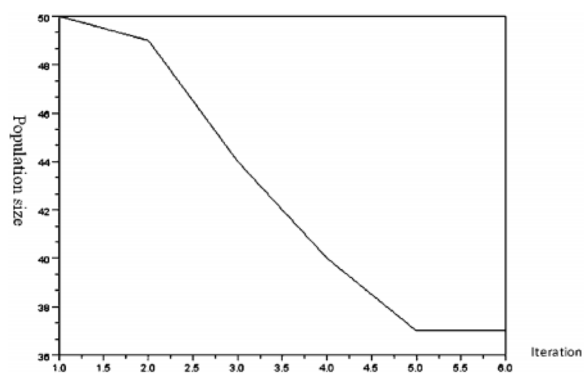

Fig.1: Latency's Coefficient = 1/3, Initial Population Size = 50, Length Space Segment = 1Km

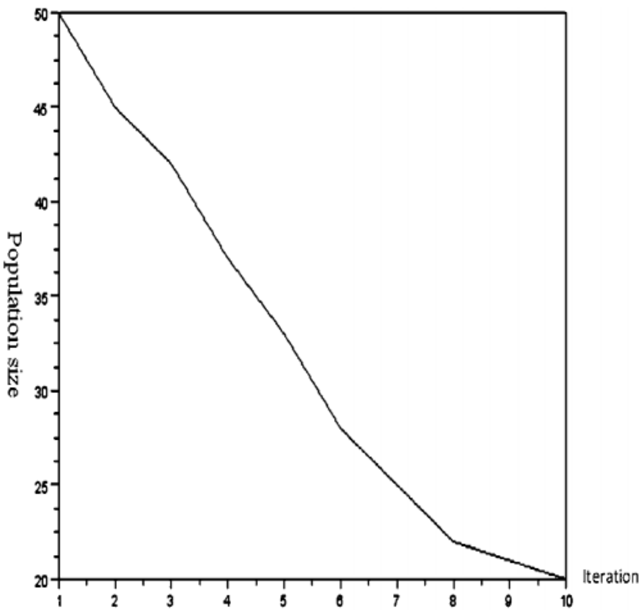

Fig.2: Latency's Coefficient = 1/2, Initial Population = 50, Length Space Segment = 1Km

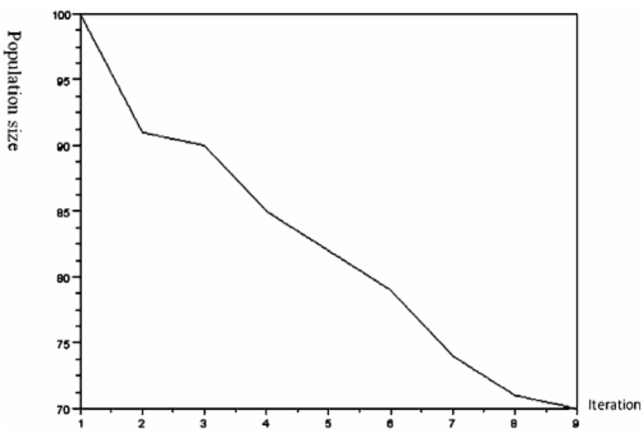

Fig.3: Latency's Coefficient = 1/2, Initial Population = 100, Length Space Segment = 1Km

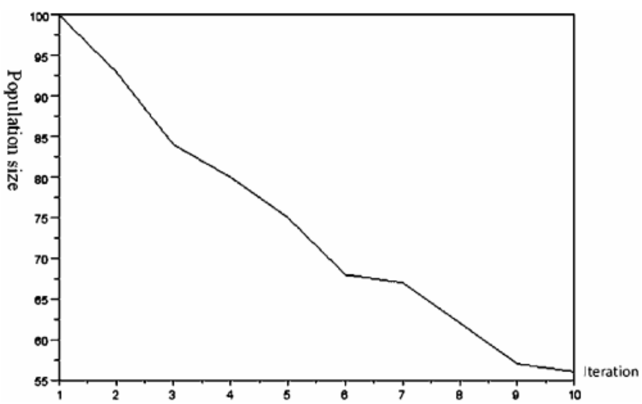

Fig.4: Latency's Coefficient = 1/3, Initial Population = 100, Length Space Segment = 1Km

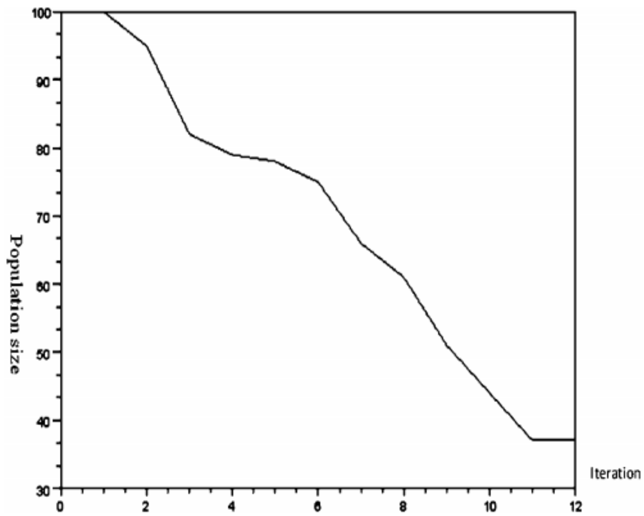

Fig.5: Latency's Coefficient = 1/2, Initial Population = 100, Length Space Segment = 3Km

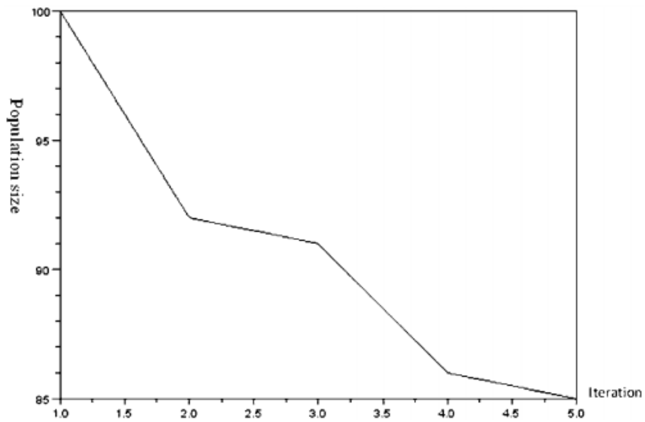

Fig.6: Latency's Coefficient = 1/3, Initial Population = 100, Length Space Segment = 3Km

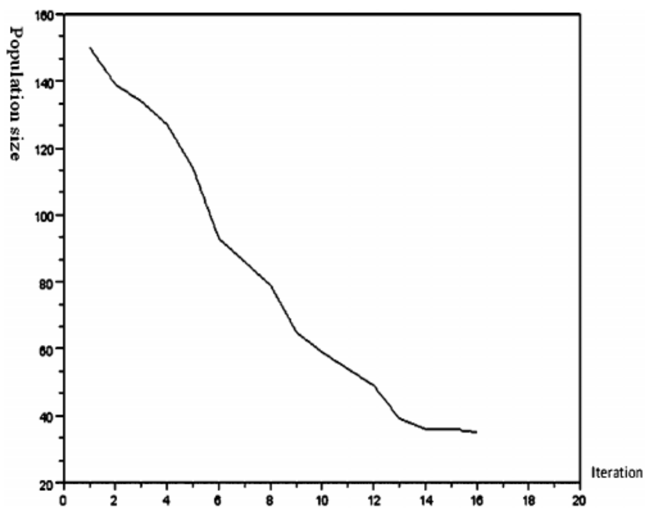

Fig.7: Latency's Coefficient = 1/2, Initial Population = 150, Length space segment = 3Km

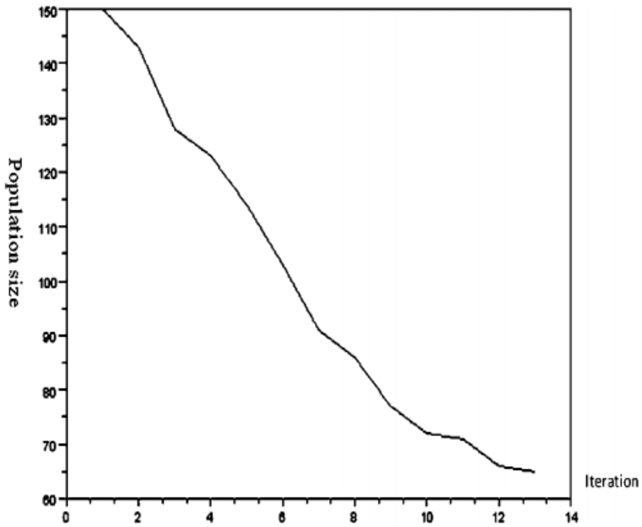

Fig. 8: Latency's Coefficient = 1/3, Initial Population = 150, Length Space Segment = 3Km

In these figures we present the change in population size based on iterations of the genetic algorithm by varying the parameters needed for the fitness function.

We find the speed of the algorithm and the decrease in the number of generations produced after application of mutation operators, crossover and selection. At convergence, we get a generation of reduced size and comprised of well-performing individuals. This result saves bandwidth, avoids network congestion and ensures the relevance of information.

Table 1

Initial population size	α_1	α_2	α_3	Length space	Final population size
50	1/3	1/3	1/3	1 km	34
50	1/2	1/4	1/4	1 km	26
100	1/3	1/3	1/3	1 km	50
100	1/2	1/4	¼	1 km	39
100	1/3	1/3	1/3	3 km	36
100	1/2	1/4	1/4	3 km	24
150	1/3	1/3	1/3	3 km	39
150	1/2	1/4	1/4	3 km	21

CONCLUSION

In this paper we gave a new application of genetic algorithm in a VANET. We have used the GA to find the best generation of vehicles which can generate and manage a data flow while minimizing the consumption of bandwidth of the wireless network. Using a random values of £, P and ζ we find that the initial population size is reduced depending on corresponding effect coefficients (α_1, α_2, α_3).

REFERENCES

[1] C. Adler, "Information Dissemination in Vehicular Adhoc Networks", Master's Thesis, Univ. of Munich, 2006.

[2] N. Cenerario and T. Delot, "Evaluation Continue De requêtes dans les réseaux de communication inter-véhicules", In Atelier INFORSID sur la Gestion de Données dans les Systµemes d'Information Pervasifs (GEDSIP'07), 2007.

[3] N. Cenerario, T. Delot, and S. Ilarri, "Dissemination of Information in Inter-vehicle ad hoc Networks", In Intelligent Vehicles Symposium (IV'08), 2008.

[4] B. Defude, T. Delot, Jose-Luis Zechinelli, "Extraction de connaissances dans les réseaux ad hoc inter-véhicules", ACM International Conference Proceeding Series, 277, ISBN:978-1-59593-980-7, 2008.

[5] Kawarasaki, J., and M. Sibuya, "Random Numbers for Simple Random Sampling without Replacement", Keio Math. Sem. Rep. No. 7, I-9. 1982.

[6] J. Vitter, "Random Sampling with a Reservoir", ACM Trans. Math. Softw., 11(1), 37–57. 1985.

[7] Babcock B., M. Datar, R. Motwano, "Sampling From a Moving Window Over Streaming Data", Proc. of the 2002 Annual ACM-SIAM Symp. on Discrete Algorithms, 633–634.

[8] S. El Hadaj, S. Raghay, T. Noel, "A Vector Algorithm for Managing Data Stream", 3rd IEEE International Conference on Web and Information Technologies (ICWIT'10), Marrakech, Morocco-june 2010.

[9] Chikly R., "Résumé de flux de données distribues", Rapport de thèse, à l'ENST de ParisTech. 2009.

[10] Bremermann, H.J., "The Evolution of Intelligence, The Nervous System as a Model of its Environment", Technical Report, No.1, Contract No. 477(17), Dept. Mathematics, Univ. Washington, Seattle, July, 1958.

[11] J. H. Holland, "Adaptation in Natural and Artificial Systems", Ann Arbor: The University of Michigan Press, 1975.