Enhanced Routing Algorithm for Opportunistic Networking

On the Improvement of the Basic Opportunistic Networking Routing Algorithm by the Application of Machine Learning

Ladislava Smítková Janků and Kateřina Hyniová

Institute of Department of Digital Design, Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00 Praha 6, Czech Republic

Keywords: Opportunistic Networking, Information Retrieval, Probabilistic Models, HMM, Mobility Models, Autonomous Mobile Robots, Machine Learning based Routing Method.

Abstract: The opportunistic communication networks are special communication networks where no assumption is made on the existence of a complete path between two nodes wishing to communicate; the source and destination nodes needn't be connected to the same network at the same time. This assumption makes the routing in these networks extremely difficult. We proposed the novel opportunistic networking routing algorithm, which improves the basic opportunistic networking routing algorithm by application of machine learning. The HMM Autonomous Robot Mobility Models and Node Reachability Model are constructed from the observed data and used in a proposed routing scheme in order to compute the combined probabilities of message delivery to the destination node. In the proposed routing scheme, the messages are coppied between two nodes only if the combined probability of the message delivery to the destination node is higher than the preliminary defined limit value. The routing scheme was developed for the networks of autonomous mobile robots. The improvement about 70% in a network load is reported.

1 INTRODUCTION

The opportunistic communication networks are special communication networks where no assumption is made on the existence of a complete path between two nodes wishing to communicate (Pelusi, 2006); the source and destination nodes needn't be connected to the same network at the same time. This assumption makes the routing in these networks extremely difficult. In contrast to the common ad-hoc networks, source and destination nodes needn't be connected to the same network at the same time, but they are allowed to exchange messages between them due to techniques of the opportunistic networking. These techniques allow nodes to connect and disconnect anytime. In opportunistic networking, no limitations are also set on the nodes to keep their positions; the nodes can move. This opportunity networking paradigm opens a space for a number of novel application scenarios.

This paper describes the improvement of the basic opportunistic network routing algorithm by application of machine learning. Machine learning is applied to construct robot mobility models and node reachability model. These models are furthermore used in routing to make decisions about messages.

1.1 Basic Opportunistic Networking Routing Schema

The basic opportunistic networking schema is based on a flooding algorithm extended by sending of message delivery acknowledgement. There are several variants of flooding algorithm. Usually each node tries to forward every message to every one of its neighbours except the source node. Every message is delivered to all reachable parts of the network. One of the most important disadvantages of this routing schema is the fact, that messages can become duplicated in the network. It increases the network load and can cause the network overloading.

1.2 Previous Work

There are two main categories of mobility models: macroscopic models and microscopic models (Helbing, 2001); (Qiang, 2009). The macroscopic

Smítková Janků L. and Hyniová K..
Enhanced Routing Algorithm for Opportunistic Networking - On the Improvement of the Basic Opportunistic Networking Routing Algorithm by the Application of Machine Learning.
DOI: 10.5220/0004923507710776
In *Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods* (ICPRAM-2014), pages 771-776
ISBN: 978-989-758-018-5

models operate with the global parameters, e.g. an average node velocity or a traffic density. The microscopic models model each node of the network particularly. There are many subcategories of these models.

The Stochastic Models are based on the assumption of random movements of the nodes (Timo, 2007). Random Walk or Random Walk or Random Direction models belong to this category of mobility models. Traffic Stream Models explore the mathematical and physical techniques developed in hydrodynamics. In Car Following Models, the behavior of each driver is modeled according to vehicles.

Statistical models, also called Survey-based Models (Harri, 2009), constitute a second important group of mobility models. These models are based on probability functions describing properties of the node behavior.

Both the synthetic and statistical models described above are the dynamic models; all of them can generate mobility of patterns, but they can't operate on pattern data only. A different approach can be followed. Instead of developing complex models and then calibrating them using mobility traces or surveys, it is possible to extract generic mobility patterns from movement traces. This approach is called Trace-based Modeling (Harri, 2009).

2 METHOD DESCRIPTION

This section describes the whole process of the creation of the new routing method.

2.1 Mobility Model

The mobility model is constructed for each of the autonomous mobile robots from the finite set of its positions in time. Mobility model has to cover both the changes in a space and the changes in time.

Hidden Markov Model (HMM) is defined by the set of the observation symbols, initial state probablity vector and the state transition probability matrix. Since the HMMs were successfully applied in speech recognition, they become a standard tool for recognition and modeling spatio-temporal data.

2.1.1 HMM Background

HMM is the spatio-temporal model defined as (Rabiner, 1989):

$$\Theta = (A, B, \pi) \tag{1}$$

where A is a state transition probability matrix, B is a observation probability matrix and π is an initial state probability vector.

$$a_{ij} = P\left(q_t + 1 = s_j \mid q_t = s_i\right), 1 \le i, j \le n \tag{2}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \tag{3}$$

$$b_j\left(v_t\right) = P\left(o_t = v_t \mid q_t = s_j\right) \quad i \le j \le n, i \le k \le m \tag{4}$$

$$\sum_{i=1}^{m} b_{ij}\left(v_t\right) \tag{5}$$

$$\pi_i = P\left(q_1 = s_i\right), 1 \le i \le n \tag{6}$$

$$\sum_{i=1}^{n} \pi_i = 1 \tag{7}$$

2.1.2 The Set of Mobile Robot Positions

The set of mobile robot positions consists of triples containing three elements: x-coordinate, y-coordinate and time. The set of mobile robot positions is finite.

In simulations, the mobile robot positions are generated by the simulator. Otherwise, the data set can be collected by any device, which makes possible to use GPS for position localization.

The finite data set of mobile robot positions is collected independently fo each mobile robot.

2.1.3 Observation Symbols in a Mobility Model

The set of mobile robot positions is too large to be used directly as the set of observation symbols in HMM based robot mobility model. We need to find a smaller set of symbols.

We decided to use cluster analysis algorithm, which was was proposed in order to construct Human Mobility Models (Hyunuk, 2012). The clustering vonsists from the following operations: cluster initialization

2.1.4 Autonomous Robot Mobility Model Identification

The well-known algorithms for the estimation of the Hidden Markov Model are EM algorithm and Viterbi training and their derivations. We used EM algorithm.

2.2 Network Scanning

The data on how the messages are boradcasted in the network are collected during the scanning. Each mobile robot generates a scanning message. The mobile robot moves and copies the scanning messages to the memory of each robot he meets during the scanning. The time of message copying is preserved as a part of the message in a form of a time stamp.

When a mobile robot meets another mobile robot, it copies its message memory to it. Message memory contains the messages, which the mobile robot received from the other robots and which are not expired. Expired messages are removed from the message memory. Network scanning algorithm is described below.

There are two important parameters of network scanning: maximum delivery time *tmax* and the maximum number of message copying *cmax.* The output of the network scanning is a log file, which contains a) messages, which were delivered to the destination node in time less or equal to *tmax* and simultaneously were not copied more times than c*max* times b) messages, which were copied *cmax* times and cannot be copied anymore.

2.2.1 Maximum Delivery Time

Maximum delivery time is defined as a maximum time between the scanning message transmission by the source mobile robot (source node) and its reception by the destination mobile robot (destination node).

During the scanning process, only the scanning messages delivered to the destination before the maximum delivery time is reached, are logged.

In our experiments, mobile robots are synchronised using the same internal clock in the simulator, e.g. we do not need solve the problems which occur when each mobile robot has its own clock.

2.2.2 Maximum Number of Message Copying

The maximum number of message copying *cmax* defines the depth of the network scanning. If the number of copying of a scanning message is equal to the *cmax,* the scanning message is removed from the node memory and it is logged.

2.2.3 Scanning Message Structure

The scanning message has the following structure:

Number of the Source Node … unique number of the source mobile robot (source node),

Number of the Destination Node … unique number of the destination mobile robot (destination node),

Message Id – a number of scanning message sent by the the source mobile robot (source node),

Source Node Cluster Id … a unique number of the mobility model cluster, in which the source mobile robot starts its movement,

Copying Counter … it identifies how many times the message was copyied,

Copy Info Section: the copy info section is repated in the message cmax times to preserve information of each copying of the message.

i-th Copy Info Node Number …. number of the node the message is copyied into, i=1 to cmax.

i-th Time Stamp ... time, when the message copy was done, i=1 to cmax

i-th “To” Node Cluster … unique number of the mobile robot (node), which received a copy of a message, i=1 to cmax.

2.2.4 Network Scanning Algorithm

This section describes the network scanning algorithm. The Network Scanning Process has the following phases:

1. Initialization. The message buffer of each mobile robot is cleaned in this step. The mobile robots are in their initial positions.
2. Scanning. The mobile robots move in some meaningful ways, e.g. solve some tasks. Every time the mobile robot meets another mobile robot, the followng operations are performed:
    1. Each mobile robot generates a scanning message; the another mobile robot makes a copy of this message with the time-stamp.
    2. Each mobile robot makes a copy of the message buffer of another mobile robot. During the copying process, the mobile robot reads each scanning message. If the mobile robot is the destination node for some message, it put the time stamp to this message and wrote the message copy in the log file; the message is cleaned. Otherweise, it modified the messages by adding the information on copying to each message. This information contains time and mobile robot ID. Also, the Copying Counter in each message is incremented.

3. Each mobile robot scans its message buffer for the "garbage" messages. Garbage Messages are:
   a) the messages, which are too old, e.g. messages which are not valid, because they exist longer than the Maximum Delivery Time parameter dwfines,
   b) the messages which cannot be copied anymore; the message cannot be copied anymore if the Copy Counter in the message is equal to the Maximum Number of Message Copying parameter.

All the garbage messages are moved to the log file and removed from the mobile robot message buffer.

3. End of Logging. The network scanning is ended by the human user on demand.

2.2.5 Node Reachability Model

Node Reachability Model is constructed from the data collected during the network scanning. Node Reachability Model is defined on the MxMxCl, where M is the set of the autonomous network nodes (autonomous robots) and Cl is the union of all the observation symbols used in HMM Robot Mobility Models. The elements of the Node Reachability Model are values of the probability function. Each value represents the probability than the message sent from the i-th source node in k-th cluster reaches the j-th destination node in time less than tmax and the number of message copying will not be higher than cmax.

2.3 Routing Algorithm Alpha 09

We proposed the novel opportunistic networking routing algorithm called ALPHA09. ALPHA09 improves the basic opportunistic networking routing algorithm by application of HMM Autonomous Robot Mobility Models described in section 2.2. and Node Reachability Model described in section 2.3.5. The detailed description of the algorithm needs more space than is available in this paper. The next section presents the most important parts of the ALPHA09.

2.3.1 How the ALPHA09 Works

Let A be the source node and Z be the destination node and *m* be a message generated by the node A. Let tmax be a maximum delivery time. Let cmax be a maximal number of a message copying, e.g. message m sent from the source node A can be most highly cmax times copyied on its way through the network until it reaches the destination node Z.

If the source node A meets an unknown node X, the following steps are done:

1. Node A asks unknown node X for its Id Number;
   1a. if IdX IS IdZ, the message m is copied directly from the node A to the node Z. The message m was delivered. The acceptation message is copied from Z to A and m is removed from the node A memory.
   1b. if IdX IS NOT IdZ, then
2. Node A asks node X if it already carries the message m.
   2a if node X carries a message m, the Copy Count of message m in node X is set to 1.
   2b if node X does not carry a message, node A asks node X for its HMM mobility model. Node A computes a combined probability of message m delivery on assumption the message is copied to the node X. This combined probability PAX is computed using the Node Reachability Model and the HMM Mobility Model of the node X as a sum of the all probabilities over the set of output symbols of the HMM Mobility Model of the node X.
3. Node A compares PAX to Pe;
   3a. if PAX >= Pe, the message is copied from node A to node X.
   3b. if PAX < Pe, the message is not copied from node A to node X. The continue with the step 4.
4. Node A stops the active communication.

Let M be a node, which carries a k-th copy of the message m generated by the source node A.

If the node M meets an unknown node X, the following steps are done:

1. Node M asks unknown node X for its Id Number;
   1a if IdX IS IdZ, the message m is copied directly from the node M to the node Z. The message m is delivered. The acceptation message is copied from Z to A and m is removed from the node A memory.
   1b if IdX IS NOT IdZ,
2. Node M compares k and the cmax.
   2a If k=cmax-1, the message is not copied and M stops the active communication.
   2b k<cmax-1; the node M asks node X for its HMM mobility model. Node M computes a combined probability of message m delivery on assumption the message is to the node X. This combined probability PMX is computed using the Node Reachability Model and the HMM Mobility Model of the Node X as a

weightened sum of the all probabilities of reaching the destination node computed over the all output symbols of the HMM Robot Mobility System. The weightening is done by the paramter 1/k.

3. Node M compares PMX to Pe;
   3a. if PMX >= Pe, the message is copied from node M to node X.
   3b. if PMX < Pe, the message is not copied from node M to node X. The continue with the step 4.
4. Node M stops the active communication. Pe is the threshold probability. Pe is defined by the human user. The value of Pe describes the netowrk senstivity. The efficiency of the proposed method is strongly influenced by the value of Pe.

3 DATA COLLECTION

The machine learning enhanced routing algorithm for opportunistic networking ALPHA9 was tested on artificially created data generated by the simulating software NESCUAR 1.0. (Natural Environment Simulator for Cooperative Unmanned Aerial Robots, version 1.0).

3.1 The Simulation Survey

NESCUAR 1.0. is a simulation environment designed and developed at the Faculty of Information Technology at Czech Technical University for the purposes of research, development and testing of a communication platform for mobile robots.

The random 2D model of a homogeneous landscape was generated in the NESCUAR 1.0. Then, the traffic of mobile robots was modeled. The trajectory of each mobile robot was simulated separately using probabilistic models and task-oriented models, e.g. it is supposed the robots do some meaningful movements. Mobile robots are simulated as autonomous devices. Some random noise was added to the data during the data collection process to simulate measurement inaccuracies.

The set of autonomous mobile robot positions consists of triples containing three elements: x-coordinate, y-coordinate and time. The finite data set of mobile robot positions was collected independently fo each mobile robot.

The simulation environment NESCUAR 1.0. enables integration of external routing algorithmes. Both the routing algorithmes, the were tested in this simulator. The number of mobile robots was set to 25.

4 RESULTS

The positions of each of the 25 mobile robots was clustered separately. Each cluster was described by the following parameters: center latitude, center longitude, maximal distance, mean distance, mean speed, number of data elements.

The number of clusters obtained from the mobile robot position data depends on the complexity of the mobile robot movements. The numbers of clusters generated for the different mobile robots changes from 5 (the least complex traces) to 21 (the most complex traces).

The mobility model based on HMM of each mobile robot was identified. We received 25 HMM based mobility models $\Omega_{MOBBIE01}$, $\Omega_{MOBBIE02}$, … $\Omega_{MOBBIE25}$. The numbers of internal states of HMMs generated for the different mobile robots changes from 2 to 7.

The Node Reachability Model was constructed from the data collected during the network scanning. It was implemented as a multi-dimensional matrix. Both the estimated Node Reachability Model and the HMM Autonomous Robot Mobility Models were used in the testing the ALPHA09 routing algorithm.

The number of mobile robots was set to 25. Maximum delivery time was set to 120 seconds. The maximal number of a message copying cmax was set to 10, e.g. each message sent from any source node could be most highly 10 times copyied on its way through the network until it reached the destination node. The pairs of communicating mobile robots were randomly generated before the simulation and were used in both simulations. The Pe delivery probability limit value was selected by a human operator; the automatic selection of this parameter is a challenge for the further research.

5 CONCLUSIONS

Thus paper deals with the proposal of the novel opportunistic networking routing algorithm, which improves the flooding routing algorithm by application of machine learning. The HMM Autonomous Robot Mobility Models and Node Reachability Model are constructed from the observed data and used in a proposed routing scheme in order to compute the combined probabilities of message delivery to the destination node. In comparision to the basic opportunistic networking routing algorithm, the significant improvement was observed in network load reduction about 70%. The future work will be focused on application of data preprocessing techniques, exploration of the algirithmes for the combined probability computation. We also intend to redesign the proposed method to enable fully distributed routing.

ACKNOWLEDGEMENTS

This work has been supported from the research grant SGS12/094/OHK3/1T/18 "Metody a postupy číslicového návrhu a jejich experimentální ověření" hold by Assoc. Prof. Hana Kubátová, CSc. at the Dept. of Digital Design at the Faculty of Information Technology of the Czech Technical University in Prague.

REFERENCES

Pelusi, L., Passarella, A., Conti, M., 2006.. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *In IEEE Communications Magazine, Vol. 44, No. 11, pp. 134-141*. IEEE.

Ravikiran, G.; Singh, S., 2004. Influence of mobility models on the performance of routing protocols in ad-hoc wireless networks. In *Vehicular Technology Conference, 2004. VTC 2004, vol. 4, p. 2185-2189.* IEEE.

Helbing, D., 2001. Traffic and Related Self-driven Many-particles Systems. In *Rev. Modern Physics, Vol. 73, 2001, pp. 1067-1141.*

Qiang, Z. et al. 2009. Formal taxonomy research on opportunistic networks. In *Proceedings of 2nd IEEE International Conference on Broadband Network & Multimedia Technology 2009, pp. 854–857.*

Ariyakhajorn, J. et al. 2006. A Comparative Study of Random Waypoint and Gauss-Markov Mobility Models in the Performance Evaluation of MANET. In *Proceedings of the International Symposium on Communications and Information Technologies 2006, pp. 894-899.*

Timo, R., Blackmore, K., Hanlen, L., 2007. Strong Stochastic Stability for MANET Mobility Models. In *Proceedings of the 15th IEEE International Conference on Networks 2007, pp. 13-18.* IEEE.

Harri, J., Filali, F., Bonnet, C., 2009. Mobility models for vehicular ad hoc networks: a survey and taxonomy. In *IEEE Communications Surveys & Tutorials. Vol. 11, No. 4, pp. 19-41.* IEEE.

Rabiner, L. R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition, In *Proceedings of the IEEE, Vol. 77, No. 2, 1989, pp. 257-286.* IEEE.

Hyunuk Kim, Ha Yoon Song, 2012 *Formulating Human Mobility Model in a Form of Continuous Time Markov Chain*, Procedia CS, Vol. 10, 2012.