Traffic Prediction in Wireless Mesh Networks Using Process Mining Algorithms

Kirill Krinkin
Open Source and Linux lab
Saint Petersburg, Russia
kirill.krinkin@fruct.org

Eugene Kalishenko
Saint Petersburg Electrotechnical University "LETI"
Saint Petersburg, Russia
ydginster@gmail.com

Abstract

Prediction of the traffic flow in particular systems will expedite discovering of an optimal path for packet transmitting in dynamic wireless networks. The main goal is to predict traffic overload while changing a network topology. Machine learning techniques and process mining enables prediction of the traffic produced by several moving nodes. Several related approaches are observed. The idea of process mining approach is proposed.

Index Terms: Wireless mesh networks, routing, process mining, traffic overload.

I. INTRODUCTION

Nowadays, representative class of program systems for intellectual data processing allows to extract data structures and analyze regularities in them. Usually such systems consist of a raw application environment model, experimental information sources, data mining methods and instruments. Timely task now is to develop such methods for process analysis system, which allows an accumulation and systematization of system's behavior templates and use this information in order to correct an application environment model and predict system state in the near future.

Traffic flow prediction is aimed to escape situations presented on Fig. 1. For instance, intensive nodes moving resulted in traffic flow overload on several nodes. Theoretically, such situations in mesh networks can be predicted and traffic redistribution can be made in advance.

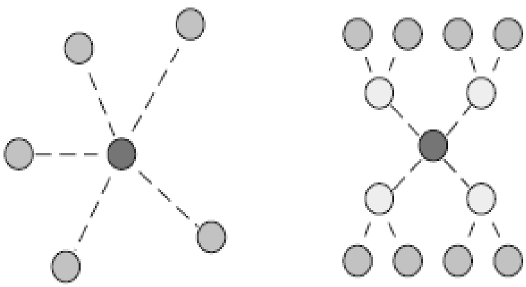

Fig. 1. Overload situations. Center - stressed node

Network traffic prediction plays an important role in guaranteeing network QoS. It is important to note that algorithms should be constructed to work in real time and be based on a minimal amount of historical information. Let's observe several main traffic prediction methods.

II. Relative works

A. Prediction using wavelet neural networks

The traditional prediction methods such as time series analysis, regression method etc. are difficult to put in practice because the WMN's traffic is complicated and lacks effective mathematical model. Moreover most of traditional prediction algorithms are resource-intensive. Thus the paper [1] presents a multipath routing algorithm based on wavelet neural networks.

Wavelet neural network prediction model adopts three-layer structure. There are Pinput for input layer, namely one input containing Pvalues of the time series. There are Nneurons on the Hidden layer. And there is one neuron on the Output layer. The offered equation shows that a function can be approximated by orthogonal function sets and thus can be implemented using a neural network.

As shown in Fig. 2 in some cases such an approach leads to a perfect results and can be used for traffic forecasting.

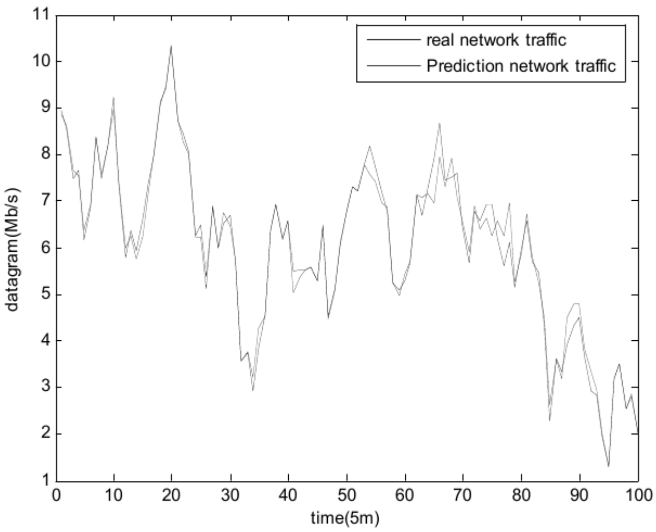

Fig. 2. Five-step prediction of wavelet-neural network

B. Clastering Approach

Besides communication between nodes there is a problem of load balancing between base stations in WMNs which provides access to external services like internet or telephony. A given geographical area consists of hexagonal cells each served by a base station (BS), as on Fig. 3. The BS is a part of the wireless infrastructure that controls one or multiple cell sites and radio signals [2].

Authors [2] use the clustering approach to form clusters based on the traffic load. Clustering is used to classify the base stations into heavily loaded, moderately loaded and idle base stations. In case of heavy loaded BS group detection, radial stations with lesser load can intercept some part of the traffic.

ISBN number will go here (do not remove this line)

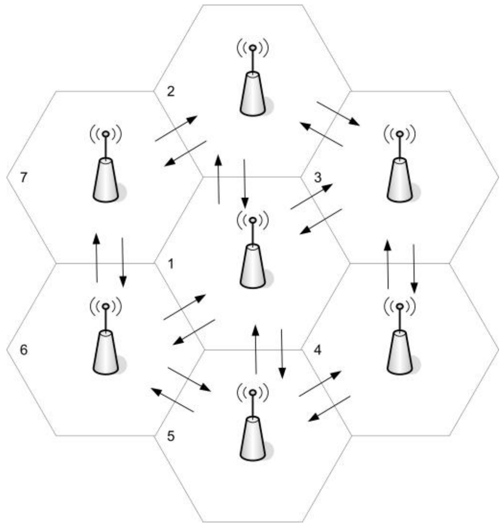

Fig. 3. Basic stations network

C. Graph Mining

One more approach comes from road traffic load prediction. As wireless network load forecast is aimed to optimize routes, road traffic analysis is aimed to choose an optimal way considering arterial load.

The basic model in [3] algorithm is a graph where vertexes are cross-road and edges are streets. The main metric of each edge is speed. It is needed to mention that in general case of transport it is impossible to determine the end point of particular unit. But in case of networks destination node is always known. The main idea of presented in [3] algorithm is to determine edges weights on some moment in near future.

D. Time series analysis

Time series analysis is widely used in wired networks, but it needed to be modified to use in client wireless networks to decrease amount of information being analyzed. On way proposed in [4] is to consider only internal mesh-network traffic between nodes.

The input is sequence of network load values (in bytes per hour for instance). After several filter steps and prediction algorithm authors achieved results which can be used in traffic prediction (Fig. 4).

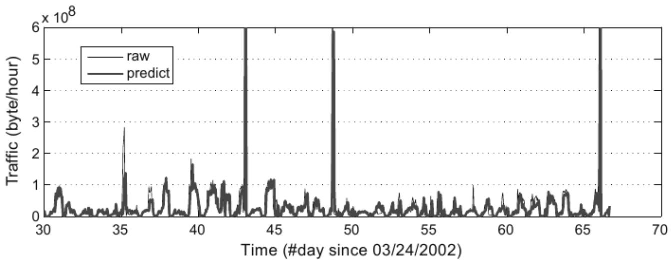

Fig. 4. Time series analysis

III. Process mining approach

Each of algorithms briefly described above has its limitations:

1. *Wavelet neural networks:* can be less efficient with lack of representative traffic statistics

2. *Clastering:* as was mentions, algorithm is designed for a static basic stations and can't be effectively applied to the dynamic mesh topology
3. *Graph Mining:* resource-intensive algorithm
4. *Time series analysis:* only internal traffic with periodic-type traffic

The other proposed approach concerned with extraction of template topologies using process mining algorithms [5]. The main idea is looking for topology changes and redistribute traffic in case of overload nodes prediction (Fig. 5).

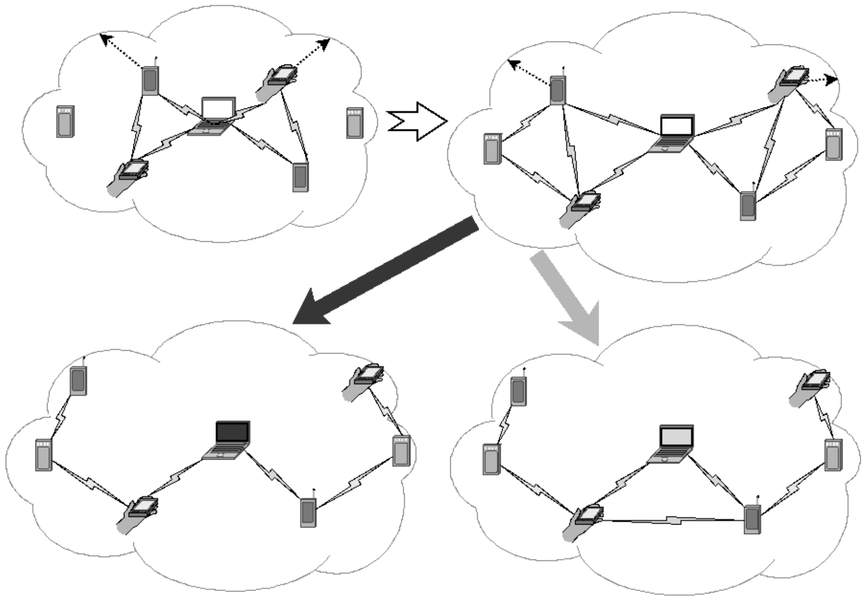

Fig. 5. Traffic redistribution

A. NS-3

NS-3 is a discrete-event network simulator used to model different kinds of networks both wired and wireless. It has a flexible object-oriented and plug-in architecture for creating particular network topologies and extend it with new traces output formats [6]. NS-3 is used to build and model a set of dynamic mesh networks (like Fig. 5) in order to get traces for analysis using process mining techniques.

To model dynamic mesh network in NS-3 it necessary to do the following in model program (in C++):

1. Create a set of network nodes
2. Install WiFi devices on created nodes: NS-3 provides a set of "Helpers" for some kinds of networks. It is enough to use such helper for mesh network to make a topology with given set of nodes
3. Set node mobility properties: initial position in a grid and vector of velocity
4. Install IP addresses
5. Install network applications on each node
6. Start modeling: as a result of modelling process NS-3 gives a trace, for instance in tcpdump format, which can be analyzed in, for example, wireshark (an application for network traffic visualization)

B. ProM

ProM is a generic framework for implementing process mining tools in a standard environment. The ProM framework receives as input logs in the XES or MXML (Mining XML) format. Currently, this framework has plug-ins for process mining, analysis, monitoring and conversion [8].

NS-3 architecture provides a plug-in mechanism for extending a set of formats for trace generation. Thus it is possible to implement a MXML plug-in for NS-3 and pass modeling results directly to ProM for analysis.

The main goal of using ProM is to find a set of effective process mining algorithms for determination of topology changes, which can lead to traffic overload in some points of the network.

D. Algorithm

The basic research approach consists of several steps:

1. *Dynamic network modeling:* use NS-3 to model dynamic mesh network with several nodes witch form spatial configurations with potentially hotspot nodes.
2. *Trace converting:* convert NS-3 trace into MXML format for ProM.
3. *Process extraction:* process mining techniques are aimed to extract, analyze, extend or optimize different kinds of processes. Process extraction algorithms are used to extract information about network topology and traffic distribution changes.
4. *Template search:* if current extracted process is similar to someone in database of dangerous process templates, a potential hotspot node has to be informed about traffic overload in near future. If overload already occurred then new template added to the database to prevent such routes distribution in future.
5. *Recommendations for redistribution:* some metrics of potentially overloaded node are changed to prevent new routes through it. If process is still similar to some of dangerous templates, active routes redistribution may be done between overloaded node neighbors.

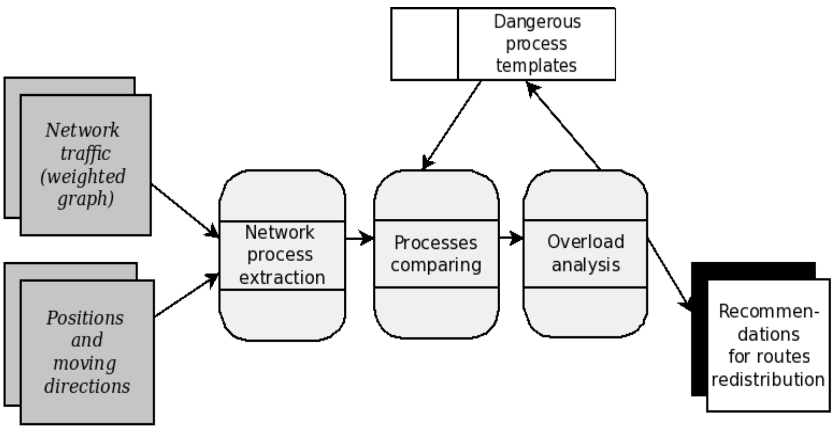

Fig. 6. Basic scheme

The main advantage is a minimal information about traffic history. Taking into account nodes position leads to an ability of extraction different kinds of topologies and traffic load portraits. Now research is in process of defining the best process mining algorithms for topology analysis.

ProM presents results of its work as Petri Net. This Net is supposed to aggregate common and anomalous network behavior. From the other side the same network should be suitable for separation of common and anomalous process instances.

IV. CURRENT WORK

In general, process mining approach for dynamic mesh network analysis is proposed. Current work is concerned with active development of the research framework based on NS-3 and ProM in order to define methods and algorithms suitable for network analysis. Thus, the next main research steps are:

1) Extend a set of overloaded mesh networks: extend a set of networks being analyzed. Now several simple dynamic networks with up to ten moving nodes are implemented using NS-3, for instance "dispersion" model, where a set of nodes run out from the central one (Fig. 7). In order to check process mining algorithms stability it is needed to model several networks witch are closer to reality.

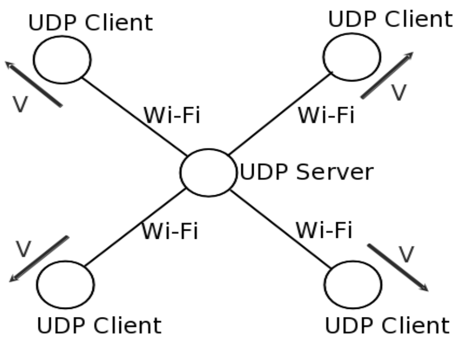

Fig. 7. Simple model

2) NS-3 MXML plug-in: develop an test a plug-in for converting NS-3 trace into MXML format for ProM. Now the basic version of a plug-in is implemented as a library, that can be used with an arbitrary NS-3 model. In order to use NS-3 logging framework, plug-in should implement a set of callbacks and register them on devices, applications etc... Callback receives an event, which parameters depends on its type. For example, common signature of the network type callback is: *PacketSentCallback (string context, Ptr<Packet const> packet).* Where context is an identifier of originator (for instance, "*/NodeList/0/DeviceList/1/$ns3::WifiNetDevice/PhyRxBegin*") and packet is a parameter, specific for the callback type. Library is extensible in terms of new event and callback types. In order to use ProM several event types were chosen to register callbacks:

- packet has begun transmitting (or being received) over the channel medium
- packet has been completely transmitted (or being received) over the channel
- packet has been dropped by the device during transmission
- packet has been received ...

C. Choose the best algorithms: after network traces analysis using ProM framework we need to choose a set of algorithms which can bring out network potentially dangerous, in terms of overload in a short-time period, topology templates. As MXML-plug-in is already implemented, now that stage is on focus of our research.

D. Implement correction metric: use process mining algorithms to implement a new routing metric in NS-3 for mesh networks. Then compare the average network overload in case of using developed metric and without it.

REFERENCES

[1] Li Zhiyua and Wang A. Ruchuan, “Multipath routing algorithm based on traffic prediction in wireless mesh networks” Proc. 5-th IEEE Conf. on Natural Computation. Tianjin, 2009.

[2] J. Usha, Ajay Kumar and A.D. Shaligram, “Clustering approach for congestion in mobile networks” International Journal of Computer Science and Network Security, 2010, vol. 2.

[3] Hans-Peter Kriegel, Matthias Renz, Matthias Schubert and Andreas Zuefle, “Statistical density prediction in traffic networks” Proc. 8-th SIAM International Conference on Data Mining, 2008.

[4] Liang Dai, Yuan Xue, Bin Chang, Yanchuan Cao and Yi Cui, “Optimal routing for wireless mesh networks with dynamic traffic demand” Mobile Networks and Applications, 2008, vol. 1.

[5] S. Elizarov, A. Bargesyan, M. Tess, M. Kupriyanov and I. Holod, Data and process analysis, BHV, *Saint-Petersburg*, 2009.

[6] NS-3 description http://www.nsnam.org/wiki.

[7] Yan Zhang, Jijun Luo and Honglin Hu, “Wireless Mesh Networking: Architectures, Protocols and Standards”, Auerbach Publications, 2006.

[8] ProM decsription http://www.processmining.org/.