

A flexible infinite HMM model for accurate characterization and segmentation of RTT timeseries

Maxime Mouchet, Sandrine Vaton
IMT Atlantique, IRISA, UBL
Brest, France

Thierry Chonavel
IMT Atlantique, Lab-STICC, UBL
Brest, France

Abstract—The study of round-trip time (RTT) measurements on the Internet is of particular importance for improving real-time applications, enforcing QoS with traffic engineering, or detecting unexpected network conditions. On large timescales, from 1 hour to several days, RTT measurements exhibit characteristic patterns due to inter and intra-AS routing changes and traffic engineering, in addition to link congestion. We propose the use of a nonparametric Bayesian method to fully estimate HMM parameters from delay observations, including the number of states. We validate the model through three applications: the clustering of RIPE Atlas measurements, the detection of significant delay changes, and the reduction of the monitoring cost in routing overlays using Markov decision processes.

I. INTRODUCTION

The Internet is the subject of continuous measurements of its performance, whether the goal is to detect attacks such as denial-of-service, to route packets on better-performing paths (traffic engineering), or to improve protocols design. Among the various quality of service metrics that can be obtained, the round-trip time (RTT) is of particular importance as it directly impacts real-time services such as audio and video communications, as well as user experience (e.g. time to first byte). While various statistical models for the RTT have been proposed in the literature, they either focus on short-time delay prediction with autoregressive models and recurrent neural networks, or on modelling delay distributions on longer timescales but without accounting for temporal dependencies. However on timescales from 1 hour to several days, the delay exhibits characteristic patterns due to changes in the routing and the traffic level, as shown in Figure 2. We propose the use of a Bayesian nonparametric timeseries model, the hierarchical Dirichlet process hidden Markov model (HDP-HMM) to capture these patterns in a statistical model. Nonparametric Bayesian methods are concerned with the estimation of models where the number of parameters is allowed to change with the observations. Such a method has been recently applied to Internet measurements in [1]. In contrast to this work, where a model is proposed to characterize the distribution of the delay of several hosts observed at a major Internet backbone link, we propose a model for the delay of a single origin-destination pair over time, with exact inference of the transition matrix.

The model is validated through three applications. 1) We show on RIPE Atlas measurements that the hidden states inferred by the model match well with the IP path observed in the traceroutes. 2) We show on a labelled changepoint dataset

that the HDP-HMM performs at least as well as state of the art changepoint detection methods. 3) We exploit the Markovian property of the model to solve a parsimonious monitoring problem and reduce by 90% the number of measurements.

II. MODELLING RTTs USING THE HDP-HMM

We seek to learn RTT models from active measurements performed at regular intervals. A common model for discrete-time timeseries is the hidden Markov model (HMM). An HMM is defined by hidden states (z_t) and transition probabilities among them such that $\pi_{kl} := \mathbb{P}(z_t = l | z_{t-1} = k)$, and by observation parameters θ_k (e.g. $\theta_k = (\mu_k, \sigma_k)$ for a Gaussian HMM). The generative model is defined by $z_t | \{\pi_{k\bullet}\}_{k=1}^K, z_{t-1} \sim \pi_{z_{t-1}\bullet}$ and $y_t | \{\theta_k\}_{k=1}^K, z_t \sim p_{\theta_{z_t}}$. However the number of states K is unknown when observing the delay on the Internet. The HDP-HMM [2] is a Bayesian nonparametric extension that allows the number of hidden states to be inferred from the data. Furthermore the distribution of the delay is prone to change over time. For example in some states the delay might follow a normal distribution, while in some states the delay might follow an exponential distribution. To allow for a wide range of distributions we propose to use another Bayesian nonparametric model, the Dirichlet process Gaussian mixture model as the observation model.

III. VALIDATION AND APPLICATIONS

There exists no RTT dataset with ground-truth on the network "state". We propose to validate the model through three applications: clustering of RIPE Atlas measurements, detection of delay changes, and reduction of the monitoring cost in routing overlays using Markov decision processes.

A. Statistical Analysis of RIPE Atlas Measurements

In order to study the correlation between learned hidden states and the IP path we rely on Atlas anchoring mesh measurements. These measurements are setup between every anchors pairs in a full-mesh topology and provide delay measurements every four minutes as well as traceroute measurements every fifteen minutes. Measurements are performed using the ICMP protocol and traceroutes are available for the forward and reverse paths. Considering the subset of anchors that were online over a 7 days period, we learned a model for the 301 RTT series of 2520 data points each. The median number of states learned was 2 and the highest 11. Looking

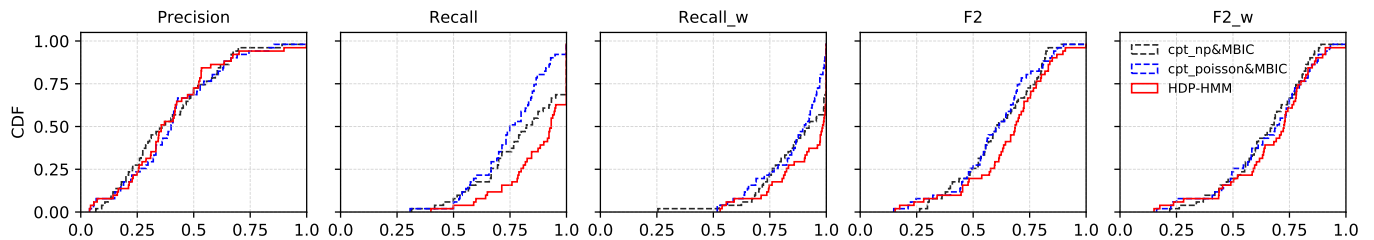


Fig. 1. Comparison of the HDP-HMM with the changepoint detection methods evaluated in [3].

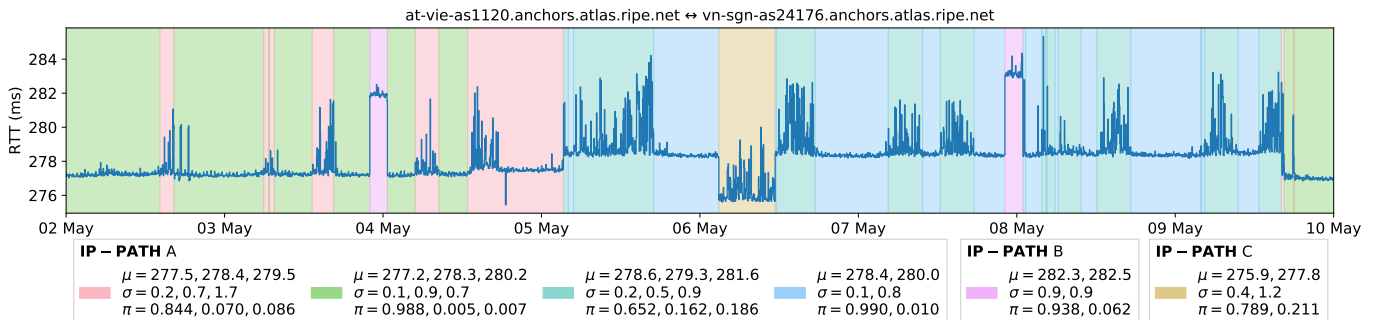


Fig. 2. Segmentation of RTT observations between `at-vie-as1120` and `vn-sgn-as24176` using an HDP-HMM with DP-GMM emissions. Each color identifies a state.

at the correlation with traceroutes, we found that 80% of the states match with exactly one AS path, and 67% with exactly one IP path. States associated with more than one path can be explained by delay differences too small to be separated. Note that HDP-HMMs allow to summarize RTT timeseries with few parameters. Figure 2 displays one of the model learned.

B. Changepoint Detection

For traffic engineering purpose one may want to detect significant changes in RTTs over time. In [3] a labelled changepoint dataset and an evaluation framework is proposed to compare detection methods. In Figure 1 we reproduce the results of [3] and compare the results obtained with the HDP-HMM described in Section II to the two best-performing methods of the original paper. The weighted recall gives more importance to large RTT changes, we refer the reader to [3] for more information on the evaluation methodology. We remark that the HDP-HMM performs at-least as well as the other methods based on the recent PELT method.

C. Parsimonious Monitoring

Current routing overlays approaches use all-pair probing which makes the probing cost prohibitive ($O(n^2)$ with n , the number of nodes). Assuming a Markov model of delay on Internet paths, we propose a method to balance the number of measurements with the routing error based on Markov decision processes. Given the current knowledge of paths delays, our method gives which path to monitor, and when, thus it is sparse in space and in time. While related works studied the problem of predicting path changes from traceroute measurements, our

approach is based uniquely on delay observations. We simulated a 30-nodes topology from RIPE Atlas measurements by randomly choosing five anchors on each of the six continents. We managed to reduce the number of measurements by 90% while having on average an optimized path only 0.65ms longer than the optimal path. We refer the reader to [4] for more information.

IV. CONCLUSION

We’ve shown that the HDP-HMM, a nonparametric Bayesian model, can be used to estimate HMM parameters, including the number of states, from RTT observations. We have checked on real Internet measurements that the states learned map well to AS paths. Such a flexible model enabled us to achieve precise clustering of delay series. It can also be used to forecast delays and limit the monitoring cost in a network management context.

REFERENCES

- [1] R. Fontugne, J. Mazel, and K. Fukuda, “An empirical mixture model for large-scale RTT measurements,” in *2015 IEEE Conference on Computer Communications (INFOCOM)*, April 2015, pp. 2470–2478.
- [2] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, “A sticky HDP-HMM with application to speaker diarization,” *The Annals of Applied Statistics*, pp. 1020–1056, 2011.
- [3] W. Shao, J. L. Rougier, A. Paris, F. Devienne, and M. Viste, “One-to-One Matching of RTT and Path Changes,” in *2017 29th International Teletraffic Congress (ITC 29)*, vol. 1, Sept 2017, pp. 196–204.
- [4] S. Vaton, O. Brun, M. Mouchet, P. Belzarena, I. Amigo, B. J. Prabhu, and T. Chonavel, “Joint minimization of monitoring cost and delay in overlay networks: optimal policies with a Markovian approach,” *Journal of Network and Systems Management*, 2018.