Availability Estimation of External IP-Optical Network Connections Using Bayesian Modeling

Filippos Christou, Andreas Kirstädter

Institute of Communication Networks and Computer Engineering (IKR), University of Stuttgart, Germany
filippos.christou@ikr.uni-stuttgart.de

Abstract *We present a Bayesian model for multi-domain networking that estimates external domain connection availabilities. The model uses informed priors from an internal domain model and the novel notion of "hyperlinks". The performance of the probabilistic estimations proves superior compared to interval availability. ©2025 The Author(s)*

Introduction

Since the introduction of optical fibers in the late twentieth century, optical networks have been established as the backbone of today's communication. As critical services using this infrastructure grow, the importance of reliability in optical networks also increases. There are two ways of improving the reliability of the system. The first is strictly connected to Capital Expenditures (CapEx), where, for example, the hardware equipment is updated. The second is to focus on efficient network control and sophisticated algorithms, e.g., involving optical protection or restoration. However, when dealing with multi-domain networks, hardware updates in an external domain are impossible, and any influence on control mechanisms is limited. This paper offers a model towards a digital twin to probabilistically estimate the availability of an external connection. The model outperforms the observed interval availability in terms of accuracy, which is defined as the experienced fraction of time during which a system is operational over a finite period. Operators can build algorithms utilizing this model to achieve the desired end-to-end connection availability.

This work has been conducted as part of the CELTIC-NEXT SUSTAINET project and relates to efforts within the network tomography field[1], with recent advances also for optical networks[2]. Although there is much work on estimating availability in a single domain like [3] and the literature mentioned therein, no work has been found to bring up this topic for multi-domain networks. In [4], a hierarchical Bayesian model was introduced, utilizing considerations of [5] to estimate link and path availabilities. The current work develops a Bayesian model that uses data to estimate the availability of external connections by introducing the idea of *hyperlinks* and employing an internal domain model[4] as informed prior knowledge. The following section presents the model before we move on to the evaluation in order to show its superiority against the commonly used interval availability, which is sometimes wrongly assumed even to be equivalent to the real availability[3].

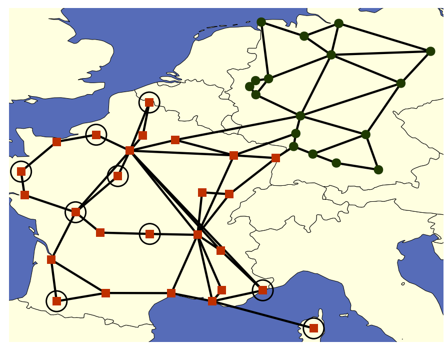

Fig. 1: Multi-domain topology

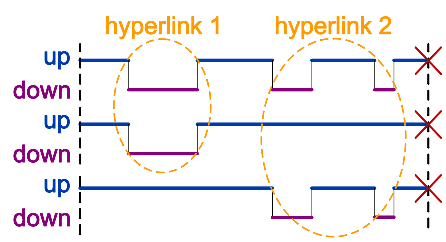

Fig. 2: Hyperlinks in connections

Modeling

We consider an IP-optical multi-domain network, like Fig. 1, where the operator's network (internal) is connected to another (external) network, with the border nodes being the only shared information. The internal domain issues external connectivity requests, with the border nodes being the sources and some external IP routers being the destinations. IP protocols can report the connection distance. We assume path deployment remains unchanged and failures originate only from links. For every connection, we observe a series of downtimes and uptimes in the external network. When a connection is disrupted simultaneously with another, it is evident that a common underlying equipment, i.e., links, failed. We attribute the minimum identifiable joint disruption of the connections to the failure of a so-called *hyperlink (HL)*, as we cannot be sure of the number of failing links. Fig. 2 shows three connections failing. The simultaneous failure and repair times of the first and second connections suggest the existence of one HL. Similarly, the simultaneous last two downtimes between the first and third connection indicate the existence of a second HL. Connections $c \in C$ are modeled to consist of a series of HLs $h \in H$. A compensation hyperlink (CHL) $h' \in H'$ is added to each connection to account for the remaining links, where no failure was observed.

Bayesian modeling is a statistical approach that uses expert knowledge to define a stochastic process's behavior and prior knowledge[6]. The priors are initial beliefs in the form of probability distributions about parameters used in the statistical model. During inference, the model is fed data, and the corresponding final posterior distributions are deducted. The eqs. (1) to (14) describe the statistical model. The left-hand-side (LHS) of the equations is either a random variable (RV) if it has a hat ˆ or data otherwise.

$$\hat{l_h} \sim \text{Truncated-Normal}(l_h^p, \frac{l_h^m - l_h^p}{2}, 0, l_h^m) \; \forall \, h \in H \tag{1}$$

$$\hat{l_H^c} = \sum_{h \in H_c} \hat{l_h} \; \forall \, c \in C \tag{2}$$

$$\hat{l_{h'}} = e^{|l_c - \hat{l_H^c}|} \; \forall \, c \in C \tag{3}$$

$$l_c \sim \text{Normal}(\hat{l_H^c}, \frac{l_c}{2(|H_c|+1)}) \; \forall \, c \in C \tag{4}$$

$$\hat{f_h^s} = \text{IDM}(\hat{l_h}) \; \forall \, h \in H \tag{5}$$

$$\hat{f_h} \sim \text{Gamma-MLE}(\hat{f_h^s}) \; \forall \, h \in H \tag{6}$$

$$u_h \sim \text{Exponential}(\hat{f_h}) \; \forall \, u_h \in U_h \; \forall \, h \in H \tag{7}$$

$$n_h^f \sim \text{Poisson}(\frac{t_h^u}{\hat{f_h}}) \; \forall \, h \in H \tag{8}$$

$$\hat{f_{h'}^s} = \text{IDM}(\hat{l_{h'}}) \; \forall \, h' \in H' \tag{9}$$

$$\hat{f_{h'}} \sim \text{Gamma-MLE}(\hat{f_{h'}^s}) \; \forall \, h \in H \tag{10}$$

$$0 \sim \text{Poisson}(\frac{T}{\hat{f_{h'}}}) \tag{11}$$

$$\hat{r_H} \sim \text{Inverse-Gamma}(2, 40) \tag{12}$$

$$d \sim \text{Exponential}(\hat{r_H}) \; \forall \, d \in D \tag{13}$$

$$n_H^r \sim \text{Poisson}(\frac{t_H^d}{\hat{r_H}}) \tag{14}$$

$\hat{l_h}$ in (1) denotes the estimated total distance of HL h. It is modeled via a truncated normal distribution between 0 and l_h^m. l_h^m is the maximum possible HL distance, which is calculated as the minimum connection distance that this HL is involved in. l_h^p is the prior of the Normal mean and is calculated from the connection distance being equally divided across the contained HLs and CHL. $\hat{l_H^c}$ from (2) is the sum of all HL distances in a connection c. H_c is the set of HLs contained in connection c. (3) calculates the distance of the CHL h', $\hat{l_{h'}}$, as the difference between the total HL distance and the connection distance l_c and then passes it through the exponential function to render it strictly positive and amplify its effect. The modeling of the CHL distance to depend on other RVs neatly avoids the introduction of further RVs. (4) is the likelihood model of the connection distance l_c as a Normal distribution. (5) uses the internal domain model (IDM) from[4], which receives a (HL) distance as an input and gives a chain of mean time to failure (MTTF) samples $\hat{f_h^s}$ as an output. These samples are then used in (6) to fit a Gamma distribution using maximum likelihood estimation (MLE), which makes the prior for the HL MTTF $\hat{f_h}$. (7) assumes that the uninterrupted uptimes $u_h \in U_h$ of HL h follow an Exponential distribution. As a HL is composed of a potential series of links, a Hyperexponential distribution would theoretically be more appropriate. However, the Exponential distribution was preferred due to the lack of data, the reduction of parameters, and the computational complexity. (8) utilizes the number of failure events n_h^f of HL h and models them with a Poisson process, with the average rate of occurrence determined by the total uptime of HL h, t_h^u, divided by the mean of the underlying distribution $\hat{f_h}$. Similar to eqs. (5), (6) and (8), the eqs. (9) to (11) serve the same purposes but for CHLs, with the distinction that CHLs have no failure events. T is the total measurement time, where downtime and uptime measurements are interrupted. (12) defines the prior for the mean time to repair (MTTR) $\hat{r_H}$ common for all HLs. (13) models the downtimes $d \in D$ of all the HLs using an Exponential distribution. (14) uses again a Poisson process to model the number of the repair events n_H^r during the aggregated uninterrupted downtimes of all HLs t_H^d. The model from eqs. (1) to (14) is implemented using the Turing.jl[7] probabilistic programming language (PPL). We use the No-U-Turn Sampler (NUTS)[8], a Monte Carlo Markov Chain (MCMC) inference technique, to fit the model to the data and get the posterior distributions. MCMC techniques are powerful because they enable the inference of posterior distributions without analytical solutions. Therefore, no formula for the posterior distribution is found, but rather a chain of samples representing the posterior distribution. The posterior samples can be used to calculate moments of the posterior distributions or to compute new quantities by performing operations on them.

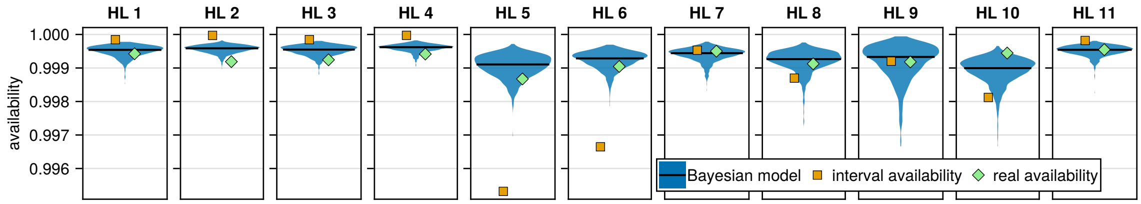

Fig. 3: Hyperlink availability estimations

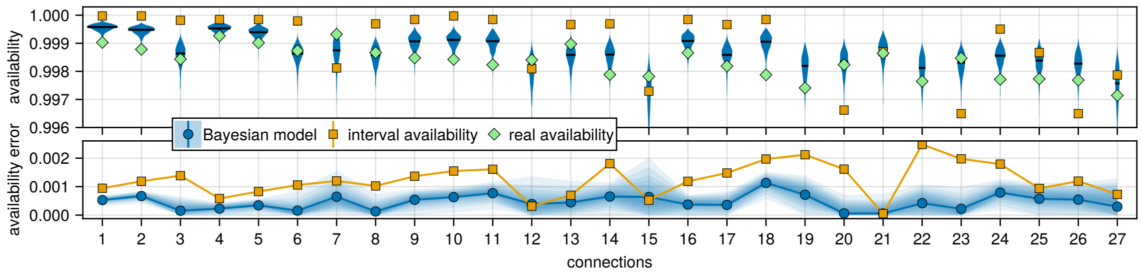

Fig. 4: Connection availability estimations

Evaluation

This section evaluates the developed model on the simulated multi-domain topology of Fig. 1. We assume to operate the German network[9] and issue connectivity requests from its border nodes to the circled nodes of the French network[10]. The data and the real link MTTFs and MTTRs are generated similarly to[4]. It involves exponential link uptimes and downtimes using a reciprocal relationship between the link distance and failure events and an added layer of Gaussian noise. The real availability can then be deducted using the availability formula eq. (15).

$$\text{availability} = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} \tag{15}$$

Combining the posterior chain samples from eqs. (6) and (12), we compute the posterior HL availability $\hat{a_h}$ with eq. (16) and plot the result in Fig. 3.

$$\hat{a_h} = \frac{\hat{f}_h}{\hat{f}_h + \hat{r_H}} \tag{16}$$

Fig. 3 shows the posterior availability distribution for all HLs. A Gaussian kernel estimate is used to visualize the posterior samples, and a black line is drawn for the median value. As a baseline, we use the interval availability, i.e., the experienced availability of the assumed HL, given by eq. (17) and denoted with a rectangle.

$$\text{interval availability} = \frac{\text{uptime}}{\text{uptime} + \text{downtime}} \tag{17}$$

The real availability of the HL is noted in Fig. 3 with a rhombus. We observe the probabilistic nature of our model, thanks to which it can incorporate uncertainty. Furthermore, taking the mean point estimates of the posterior availability distribution across all HLs approaches the real availability on average $47\,\%$ better than the interval availability.

We get the posterior connection availabilities by multiplying the HLs and CHL posterior availabilities of a connection, i.e., their posterior samples. These are similarly plotted in the upper part of Fig. 4, with the interval connection availability and real connection availability values. We notice that the real availability almost always falls within the plotted range of the Bayesian estimations. The lower part of Fig. 4 plots the absolute errors of the Bayesian model and the interval availabilities. The mean of the posterior availability distribution samples was used for the circular markers. The error using the full posterior distribution is depicted starting from a $90\,\%$ highest density interval (HDI) with lighter shades to a $10\,\%$ HDI with darker shades and increments of $10\,\%$. The HDI is the narrowest interval containing the specified probability mass[6]. On average, a $54\,\%$ improvement over the interval availability is achieved.

Conclusions

This paper introduced the first model to estimate external IP-optical connection availabilities using Bayesian techniques. Modeling external domains is vital for multi-domain decision-making but challenging because information is missing. The developed model only needs the external connection uptimes and downtimes, based on which the hyperlinks and their availabilities are derived. It uses priors from an informed model built based on the internal domain. The model improved accuracy at the order of $50\,\%$ compared to the de facto approach of interval availability. The probabilistic nature of the estimations enables the incorporation of uncertainty in operation strategies. Along these lines, Bayesian modeling emerges as a powerful tool and an exciting research area for optical networks.

Acknowledgements

This work has been performed in the framework of the CELTIC-NEXT EUREKA project SUSTAINET inNOvAte (Project ID C2024/3-4), and it is partly funded by the German BMBF (Project ID 16KIS2260).

References

[1] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, "Network Tomography: Recent Developments", *Statistical Science*, vol. 19, no. 3, pp. 499–517, 2004. DOI: 10.1214/088342304000000422. [Online]. Available: https://doi.org/10.1214/088342304000000422.

[2] R. Shinzaki, M. Eto, K. Tajima, *et al.*, "Optical network tomography: Demonstration of anomaly loss monitoring and fiber-type identification", in *2023 Optical Fiber Communications Conference and Exhibition (OFC)*, 2023, pp. 1–3. DOI: 10.1364/OFC.2023.M3Z.2.

[3] D. A. Mello, H. Waldman, and G. S. Quitério, "Interval availability estimation for protected connections in optical networks", *Computer Networks*, vol. 55, no. 1, pp. 193–204, 2011, ISSN: 1389-1286. DOI: https://doi.org/10.1016/j.comnet.2010.07.018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128610002483.

[4] F. Christou, "Availability estimation of optical network links using a bayesian model", *J. Opt. Commun. Netw.*, vol. 16, no. 5, B1–B15, May 2024. DOI: 10.1364/JOCN.510726. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-16-5-B1.

[5] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks", in *DRCN 2005). Proceedings.5th International Workshop on Design of Reliable Communication Networks, 2005.*, 2005, 8 pp.-. DOI: 10.1109/DRCN.2005.1563848.

[6] R. McElreath, *Statistical Rethinking A Bayesian Course with Examples in R and STAN*, 2nd. Chapman and Hall/CRC, 2020, ch. 3, Chapter 3 https://doi.org/10.1201/9780429029608. DOI: https://doi.org/10.1201/9780429029608.

[7] H. Ge, K. Xu, and Z. Ghahramani, "Turing: A language for flexible probabilistic inference", in *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*, 2018, pp. 1682–1690. [Online]. Available: http://proceedings.mlr.press/v84/ge18b.html.

[8] M. D. Hoffman and A. Gelman, "The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo", *Journal of Machine Learning Research*, vol. 15, no. 47, pp. 1593–1623, 2014. [Online]. Available: http://jmlr.org/papers/v15/hoffman14a.html.

[9] S. Orlowski, R. Wessäly, M. Pióro, and A. Tomaszewski, "Sndlib 1.0—survivable network design library", *Networks*, vol. 55, no. 3, pp. 276–286, 2010. DOI: https://doi.org/10.1002/net.20371. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/net.20371. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/net.20371.

[10] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The internet topology zoo", *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 9, pp. 1765–1775, 2011. DOI: 10.1109/JSAC.2011.111002.