Extreme Values in Optical Fiber Communication Systems

Seb J. Savory
University of Cambridge, Department of Engineering, Electrical Engineering Division, 9 J J Thomson Avenue, Cambridge, CB3 0FA, U.K
sjs1001@cam.ac.uk

Abstract: Extreme value theory provides a framework to assess rare but extreme events such as network outages or cycle slips. We present the theory of extreme value statistics and its application to optical fiber communication systems. © 2020 The Author

1. Overview

Optical fiber communication systems are often characterized by rare events whose extreme nature can have a devastating impact. Examples of these extreme events include cycle slips in the carrier recovery of a digital coherent receiver to a network outage caused by insufficient signal to noise being available at the receiver. In a well-designed system, the probability of these extreme events is reduced to an acceptably low value, requiring statistical techniques to estimate these rare events. Extreme value statistics provides a means of estimating these rare events[1]. The field of extreme value statistics began over ninety years ago with the pioneering work of Fréchet[2], Fisher and Tippett[3], Gumbel[4] and Weibull[5], and has been applied to numerous problems within the communications community, ranging from bit error rate estimation[6-15] to network traffic analysis[16-21]. While historically the primary use of extreme value statistics in optical fiber communication systems was the Weibull distribution for the analysis of the mechanical strength of optical fibers[22-25], system level issues are now being considered[26-32].

2. Mathematical preliminaries for extreme value statistics

2.1 Introduction to extreme value statistics

Extreme value statistics are concerned with the distribution of the maxima (or minima) of a set of samples. For example, if the worst-case daily bit error rate (BER) is recorded, what would be the distribution of these worst-case measurements over a period of a month? If having measured over a month, can it be estimated what the worst-case BER that might be expected to be measured over a year? Alternatively, if a hard decision FEC is employed – what is the probability that the BER might exceed the FEC threshold over a decade of operation?

In parallel with machine learning (ML), the approach of extreme value statistics is data driven, however in contrast to ML the data sets are often very small due to the rarity of extremes observed. To analyze these small data sets, order statistics are often used, e.g. if the worst-case BER is recorded over N days then an empirical distribution function is formed with $x_1, x_2, \ldots x_N$ being the measurements ordered such that $x_1 \leq x_2 \ldots \leq x_N$ and the corresponding probability is estimated as $p_i = i/(1+N)$. As we shall see in the subsequent, in general the distribution for the maxima (or minima) converge to a single generalized extreme value distribution, being akin to the central limit theorem for the average of independently and identically distributed (i.i.d.) random variables.

2.2 Fisher-Tippett Theorem

If we consider a set of n i.i.d. random variables $\mathrm{X}_1, \mathrm{X}_2, \ldots . \mathrm{X}_\mathrm{n}$ and define $M_n = \max\{X_1, X_2, X_3, \ldots . X_n\}$, it follows that if $F(x) = P(X \leq x)$ then $P(M_n \leq x) = [F(x)]^n$. Fisher and Tippett considered convergence to a limiting form $G(x)$ for large n such that for some parameters a_n and b_n we can write $[F(x)]^n = G(a_n x + b_n)$. Fisher and Tippett noted since we could write $n = n_1 n_2$ it followed $[G(x)]^n$ would also tend to the same limiting form thereby satisfying the functional relationship

$$[G(x)]^n = G(a_n x + b_n) \tag{1}$$

with solution for $1 + \xi x > 0$ given by

$$G(x) = \exp\left(-[1 + \xi x]^{-1/\xi}\right) \tag{2}$$

Which satisfies (3) with $a_n = n^{-\xi}$ and $b_n = (n^{-\xi} - 1)/\xi$ and hence we deduce the Generalized Extreme Value (GEV) distribution for the maxima from n samples is given by

$$P(M_n \leq x) = \exp\left(-\left[1 + \xi\left(\frac{x - \mu_n}{\sigma_n}\right)\right]^{-\frac{1}{\xi}}\right) \tag{4}$$

With three regimes existing according to whether $\xi = 0$, $\xi < 0$ or $\xi > 0$.

Type I $\xi = 0$ (Gumbel distribution) $\qquad F(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right)$

Type II $\xi > 0$ (Fréchet distribution)

$$F(x) = \begin{cases} 0 & x \leq \mu - \sigma/\xi \\ \exp\left(-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right) & x > \mu - \sigma/\xi \end{cases}$$

Type III $\xi < 0$ (Weibull distribution)

$$F(x) = \begin{cases} \exp\left(-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right) & x < \mu - \sigma/\xi \\ 1 & x \geq \mu - \sigma/\xi \end{cases}$$

2.3 Limiting distribution for the minima of a set of identically distributed variables

If we define the minima $m_n = \min\{X_1, X_2, X_3, \ldots. X_n\}$ then a similar distribution to that obtained for the maxima exists since

$$P(m_n \geq x) = [P(X \geq x)]^n = [1 - F(x)]^n \tag{5}$$

with the solution of $[1 - G(x)]^n = 1 - G(a_n x + b_n)$ giving

$$P(m_n \leq x) = 1 - \exp\left(-\left[1 - \xi\left(\frac{x - \mu_n}{\sigma_n}\right)\right]^{-\frac{1}{\xi}}\right) \tag{6}$$

and again three regimes existing according to whether $\xi = 0$ (Gumbel), $\xi < 0$ (Fréchet) or $\xi > 0$ (Weibull).

2.4 Peaks over threshold analysis

A modern alternative to the classical approach to consider the magnitude y of the peaks over some threshold u, with a good example of this being in [15]. If we define $y = (x - u) > 0$, the corresponding conditional probability is:

$$P(X > u + y | X > u) = \frac{1 - F(u + y)}{1 - F(u)} \tag{7}$$

If $F(x)$ satisfies the Tippett-Fisher theorem then $P(X > u + y|X > u)$ is a generalized Pareto distribution since

$$P(X > u + y|X > u) \approx \frac{\left[1 + \xi\left(\frac{u + y - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}}{\left[1 + \xi\left(\frac{u - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}} = \left[\frac{\sigma + \xi(u + y - \mu)}{\sigma + \xi(u - \mu)}\right]^{-\frac{1}{\xi}} = \left[1 + \frac{\xi y}{\tilde{\sigma}}\right]^{-1/\xi} \tag{8}$$

where $\tilde{\sigma} = \sigma + \xi(u - \mu)$ and for the approximation we note if $\exp(-f(x)) \approx 1$ then $|f(x)| \ll 1$ and hence $\exp(-f(x)) \approx 1 - f(x)$. We also note in the limit as $\xi \to 0$ that $P(X > u + y|X > u) = \exp(-y/\sigma)$

3. Application to Optical Fiber Communication Systems

For the case of optical fiber communication systems, extreme value methods lend themselves to problems for which no analytical model exists but for which measurement data is available. One such area within optical fiber communication systems relates to the temporal evolution of polarization within an optical fiber, with Gumbel distribution used to model the maximum of log(BER) due to polarization mode dispersion[26,28] and peaks over threshold analysis used for estimate outage due to polarization dependent loss for wavelength interleaved transmission[29]. Alternatively, it can be used to improve the simulation time compared to Monte Carlo techniques, with [27] indicating extreme value methods required fewer simulation runs than a Multicanonical Monte Carlo approach for the same accuracy. Likewise, when considering the network blocking probability, reprocessing of the data in [31] reveals excellent agreement between Monte Carlo (with 10,000 simulation runs) and extreme value analysis (with 100 simulation runs) for estimating the number of 100 GbE demands for a 1% network blocking probability.

4. Concluding remarks

As machine learning emerges, we are increasing moving towards data driven optical fiber communication systems, where models are based on measurement data rather than fundamental physics. We should however also consider alternative statistical data driven approaches such as extreme value methods where the focus is on modeling the tail of the distribution. This is of particular importance for an optical fiber communication system where link availability of 99.9999% or better might be required.

5. Acknowledgements

The author thanks BT, Ciena, Huawei, Microsoft and EPSRC through programme TRANSNET EP/R035342/1 for financial support. Takis Hadjifotiou is thanked for introducing the author to the field of extreme value statistics.

6. References

[1] S. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer, 2001).
[2] M. Fréchet, "Sur la loi de probabilité de l'écart maximum," *Ann. Sot. Pal. Math*. (Krakow), vol. 6, pp. 93-116, 1927
[3] R. A. Fisher & L. H. C. Tippett, "Limiting forms of the frequency distributions of the largest or smallest member of a sample". *Proceedings of the Cambridge Philosophical Society* **24**, pp. 180-190 (1928), doi: 10.1017/S0305004100015681
[4] E. J. Gumbel, "Les valeurs extrêmes des distributions statistiques", *Annales de l'Institut Henri Poincaré*, 5 (2): 115–158, 1935
[5] W. Weibull, "A Statistical Theory of the Strength of Materials." (Generalstabens Litografiska Anstalts Förlag, Stockholm) 1939
[6] T. Fine and M. Levin, "Extreme-value theory applied to false-alarm probabilities," *IRE Transactions on Information Theory*, vol. 8, no. 3, pp. 259-260, April 1962. doi: 10.1109/TIT.1962.1057712
[7] E. C. Posner, "The application of extreme-value theory to error- free communication," *Technometrics*, vol. 7, Nov. 1965
[8] J. C. Ashlock and E. C. Posner, "Error Probability Estimation Using Bivariate Extreme-Value Theory," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-2, no. 4, pp. 463-468, July 1966. doi: 10.1109/TAES.1966.4501877
[9] M. Schwartz, S. H. Richman, "Extremal statistics in computer simulation of digital communication systems", *Proc. Spring Joint Computer Conf*, 1968. doi: 10.1145/1468075.1468146
[10] L. Milstein, D. Schilling and J. Wolf, "Robust detection using extreme-value theory," *IEEE Transactions on Information Theory*, vol. 15, no. 3, pp. 370-375, May 1969. doi: 10.1109/TIT.1969.1054300
[11] S. Weinstein, "Theory and application of some classical and generalized asymptotic distributions of extreme values," *IEEE Transactions on Information Theory*, vol. 19, no. 2, pp. 148-154, March 1973. doi: 10.1109/TIT.1973.1054987
[12] M. Jeruchim, "On the estimation of error probability using generalized extreme-value theory," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 108-110, January 1976. doi: 10.1109/TIT.1976.1055494
[13] L. Milstein, "Performance Monitoring of Digital Communication Systems Using Extreme-Value Theory," *IEEE Transactions on Communications*, vol. 24, no. 9, pp. 1032-1036, September 1976. doi: 10.1109/TCOM.1976.1093414
[14] J. B. Broadwater and R. Chellappa, "Adaptive Threshold Estimation via Extreme Value Theory," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 490-500, Feb. 2010. doi: 10.1109/TSP.2009.2031285
[15] A. Cevrero et al., "Fast and accurate BER estimation methodology for I/O links based on extreme value theory," *Design, Automation & Test in Europe Conference & Exhibition* (DATE), Grenoble, France (IEEE, 2013), pp. 503-508. doi: 10.7873/DATE.2013.114
[16] D. H. Barnes and J. J. O'Connor, "Total Network Data System: Small Office Network Data System," *Bell System Technical Journal*, vol. 62, no. 7, pp. 2397-2431, Sept. 1983. doi: 10.1002/j.1538-7305.1983.tb03544.x
[17] H. C. Ratz, "Extreme value engineering for local network traffic," *IEEE Transactions on Communications*, vol. 36, no. 12, pp. 1302-1308, Dec. 1988. doi: 10.1109/26.8943
[18] I. Berberana, "Application of Extreme Value Theory to the Analysis of a Network Simulation," *Proceedings of 23rd Annual Simulation Symposium*, Nashville, TN, USA, 1990, pp. 105-121. doi: 10.1109/SIMSYM.1990.717283
[19] F. Bernabei, M. Listanti, D. Matrullo and G. Zingrillo, "A combined Monte Carlo/GEVT extrapolating method for the estimate of buffer length distribution tails," *Proceedings of SUPERCOMM/ICC '92*, Chicago, IL, USA, 1992, pp. 1206-1211 vol.3. doi: 10.1109/ICC.1992.268049
[20] Jinwoo Choe and N. B. Shroff, "New bounds and approximations using extreme value theory for the queue length distribution in high-speed networks," *Proceedings of IEEE INFOCOM '98*, San Francisco, CA, USA, 1998, pp. 364-371 vol.1. doi: 10.1109/INFCOM.1998.659674
[21] M. Uchida, "Traffic data analysis based on extreme value theory and its applications," *Proceedings of IEEE Global Telecommunications Conference*, 2004. GLOBECOM '04., Dallas, TX, 2004, pp. 1418-1424 Vol.3. doi: 10.1109/GLOCOM.2004.1378217
[22] R. D. Maurer, "Strength of fiber optical waveguides," *Appl. Phys. Lett.* 27, 220 (1975); doi: 10.1063/1.88435
[23] R. Olshansky and R. D. Maurer, "Tensile strength and fatigue of optical fibers", *J. Applied Physics* 47, 4497 (1976); doi: 10.1063/1.322419
[24] D. Kalish and B. K. Tariyal, "Probability of static fatigue failure in optical fibers", *Appl. Phys. Lett*. 28, 721 (1976); doi: 10.1063/1.88628
[25] C. R. Kurkjian, J. T. Krause and M. J. Matthewson, "Strength and fatigue of silica optical fibers," *J. Lightw. Technol.*, vol. 7, no. 9, pp. 1360-1370, Sept. 1989. doi: 10.1109/50.50715
[26] S.J. Savory, F.P. Payne and A. Hadjifotiou, "Estimating Outages Due to PMD Using Extreme Value Statistics" *Proceedings European Conference on Optical Communications*, Stockholm, 2004, paper We1.4.3
[27] Y. Kopsinis, J. Thompson and B. Mulgrew, "Performance evaluation of optical communication systems using extreme value theory," *IEE Seminar on Optical Fibre Communications and Electronic Signal Processing*, London, 2005, pp. 0_31-10/10. doi: 10.1049/ic:20050548
[28] S. Savory, F. Payne and A. Hadjifotiou, "Estimating Outages Due to Polarization Mode Dispersion Using Extreme Value Statistics," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 3907-3913, Nov. 2006. doi: 10.1109/JLT.2006.883605
[29] K. Shibahara and K. Yonenaga, "Experimental demonstration of PDL penalty reduction by wavelength-interleaving transmission," *Opt. Express* 20, B479-B484 (2012) doi: 10.1364/OE.20.00B479
[30] I. Neokosmidis, S. Marakis, and D. Varoutas, "Semianalytical estimation of the four-wave mixing noise based on extreme value theory," *Opt. Express* 21, 23075-23086 (2013), doi: 10.1364/OE.21.023075
[31] S. J. Savory, "Congestion Aware Routing in Nonlinear Elastic Optical Networks," *IEEE Photonics Technology Letters*, vol. 26, no. 10, pp. 1057-1060, May, 2014. doi: 10.1109/LPT.2014.2314438
[32] J. C. Cartledge, F. Matos, C. Laperle, A. Borowiec, M. O'Sullivan and K. Roberts, "Use of Extreme Value Statistics to Assess the Performance Implications of Cascaded ROADMs," *J. Lightw. Technol.*, vol. 35, no. 23, pp. 5208-5214, 2017. doi: 10.1109/JLT.2017.2768962