J. Sia, L. Shalalfeh, and P. Bogdan, ``PMU Change Point Detection of imminent voltage collapse and stealthy attacks," *57th IEEE Conference on Decision and Control (CDC)*, Miami Beach, FL, December 17-19, 2018.

PMU Change Point Detection of Imminent Voltage Collapse and Stealthy Attacks

J. Sia[1], E. Jonckheere[2,*], L. Shalalfeh[3], and P. Bogdan[4,†]

***Abstract*— Phasor Measurement Units (PMUs) provide significant value towards ensuring autonomous cognition in the power grid by enabling the abnormal events to be fault-detected and so as to trigger proactive measures to avoid large catastrophic system states. For instance, a change in the baseline distribution of PMU signals can indicate imminent voltage collapse, false data injection, and other security threats. Fractal geometry inspired analysis of PMU signals (via the Hurst exponent) reveals that an imminent voltage collapse is preceded by a significant increase in the Hurst exponent. We propose a novel change point detection strategy that optimally anticipates the fractal geometry change point from the PMU signals subject to a pre-specified false alarm rate.**

I. Introduction and Novel Contributions

Fundamental for endowing the smart grid with autonomous cognitive intelligence is the capability to monitor and analyze in real-time the mathematical characteristics of the PMU signals and identify the change points through rigorous statistical techniques. Towards this end, a pioneering effort [13] demonstrated that PMU time series exhibits long-range dependence (memory) and fractal characteristics. From a mathematical perspective, this long-range memory and fractal behavior is investigated through Hurst exponent quantification. A Hurst exponent of 0.5 indicates a short-range memory behavior (implying independence between consecutive events). In contrast, the Hurst exponents observed in the PMU analysis were significantly greater than 0.5 and demonstrated a long-range memory behavior [13].

Moreover, an extension of this work has shown that the increasing trend in the Hurst exponent of the frequency time series is a good indicator of the proximity of the power system to blackout and thus can be used as an early warning signal [14]. These prior research efforts [13][14] not only offer novel ways for power systems operators to detect a blackout given the PMU frequency time series data, but also enable new mathematical and algorithmic strategies to monitor and predict the chance of an abnormal event with catastrophic implications in real time.

We make the following contributions: First, we propose a novel and robust change point detection strategy capable of detecting the chance of an imminent blackout ahead of time such that proactive measures can be taken. Second, we provide a rigorous mathematical framework to quantify the confidence on the change point detection algorithm by analyzing the false alarm rate. Third, we evaluate the proposed mathematical framework and algorithmic strategy on real PMU datasets which demonstrate that we can predict a power grid blackout at least 12 minutes in advance. Of note, the computational complexity of our algorithmic strategy is O(NlogN), where is N is the dataset length, and thus can be computed in real time.

1,2,4: Dept. of Elec. Eng., Univ. of Southern California, Los Angeles, CA 90089 {jsia, jonckhee, pbogdan}@usc.edu
3: German Jordanian University, Amman, Jordan, laith.shalalfeh@gju.edu.jo
*: corresp. author, supported by NSF CCF-1423624; †: supported by DARPA

II. Change point detection

A. Fundamentals

Consider an i.i.d. sequence $\{X_k\}_{k=1}^n$ with "normal" regime probability density p_0 from $k=1$ up to and including $k=\lambda-1$, and with "abnormal" probability density p_1 as of $k=\lambda$ up to and including n. Change Point Detection (CPD) endeavors to find the change point λ in the fastest possible way subject to some acceptable false alarm rate. There is a vast literature on the subject (see [3] and references cited therein), which can be partitioned into, on the one hand, the Shiryaev (Bayesian) procedure [15] and, on the other hand, Page's CUmulative SUM (CUSUM) (minimax) procedure [9]. In the Shiryaev procedure, λ is assumed to have an a priori distribution and the goal is to minimize the expected detection delay subject to a false alarm rate. In the CUSUM procedure, λ is deterministic, but unknown, and the goal is to minimize the worst case detection delay subject to an acceptable false alarm rate. Here we follow the CUSUM, consistently with the early work on application of CPD to security problems [4][5][12].

Given a change point λ, let P_λ denote the probability measure defined as p_0 on $\{X_k\}_{k=1}^{\lambda-1}$ and p_1 on $\{X_k\}_{k=\lambda}^n$. Let $\mathbb{E}_\lambda$ be the corresponding mathematical expectation. Let $\mathbb{E}_{p_{0,1}}$ be the mathematical expectation relative to the probability density $p_{0,1}$ on $\{X_k\}_{k=1}^n$. Note that $\mathbb{E}_\infty = \mathbb{E}_{p_0}$.

Rejecting the Null Hypothesis H_0 that there has been no changes from λ up to and including n could be based on positive value of the log-likelihood ratio statistic. However, in this security context, the statistic is computed for the worst position of the change point and for the same conservativeness reason it is reset to 0 in case it takes negative values[1]:

$$U^n = \left\{ \max_{0\le\lambda\le n} Z_\lambda^n \right\}_+, \quad Z_\lambda^n = \sum_{k=\lambda}^n \log \frac{p_1(x_k)}{p_0(x_k)}, \quad (1)$$

[1]Here, the exposition departs from the traditional one to avoid the heuristic argument justifying the reset of the recursive algorithm (2) to 0 in case the U^{n+1} statistic takes negative values [3, Chap. 2, Sec. 2.2]. Another reason for this departure is to prepare the ground for simultaneous detection and identification of p_1.

where $\{z\}_+ = \max\{0, z\}$. To put U^n in recursive form, observe the following:

$$U^{n+1} = \left\{ \max_{0 \le \lambda \le n+1} \left(Z_\lambda^n + \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})} \right) \right\}_+ \\ = \left\{ \max \left\{ U^n + \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})}, \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})} \right\} \right\}_+ .$$

Since U^n is forced to be nonnegative, the first term in the argument of $\max\{\cdot, \cdot\}$ is greater than or equal to the second. Hence, the recursion takes the form

$$U^{n+1} = \max \left\{ 0, U^n + \log \frac{p_1(x_{n+1})}{p_0(x_{n+1})} \right\}. \tag{2}$$

The decision that there has been a change is taken at the first time τ the CUSUM statistic U^n goes above a threshold h:

$$\tau(h) = \min\{n : U^n \geq h\}. \tag{3}$$

With this threshold, assuming that the algorithm is restarted at a false alarm and that the distribution remains p_0, one would expect the next false alarm at time $T_{\text{FA}} = \mathbb{E}_{p_0}(\tau(h))$, so that the False Alarm Rate, confronted with its upper admissible bound, is

$$\text{FAR} = 1/\mathbb{E}_{p_0}(\tau(h)) \leq \overline{\text{FAR}}. \tag{4}$$

The threshold $\bar{h}$ is selected by solving the above inequality.

B. Unknown "abnormal" density p_1

While it is fair to assume that the "normal" regime density p_0 is known, the same cannot be said for the "abnormal" regime density p_1. One approach is to choose a family $\{f_\theta\}$ of distributions parameterized by θ and adjust θ consistently with some empirical knowledge on p_1. But this already raises the first question on how to objectively choose the family and, given a family, how to estimate θ.

1) Weibull distribution: The Weibull distribution with known shape parameter β,

$$f_\theta(x) = \tfrac{\beta}{\eta} \left(\tfrac{x}{\eta} \right)^{\beta - 1} \mathrm{e}^{-\left(\frac{x}{\eta}\right)^\beta}, \tag{5}$$

is a Koopman-Darmois distribution with sufficient statistic $T(x) = x^\beta$ and natural parameter $\theta = -1/\eta^\beta$, where η is the scale parameter. It is known to be the probability density that takes the least amount of data to be correctly identified [7]. So its utilization is wholly justified in this problem where it is imperative to correctly identify p_1 in the shortest amount of time.

2) Simultaneous detection and estimation: Here, instead of (1), we proceed from the *double* maximization [3, Chap. 2, Sec. 4.3.1],

$$U^n = \left\{ \max_{0 \le \lambda \le n} \max_\theta Z_\lambda^n \right\}_+, \tag{6}$$

where, in the definition of Z_λ^n, p_1 is replaced by f_θ. Again, in this security context, especially in case of stealthy attack [11][16], another justification for $\max_\theta$ is to assume that the density f_θ is the worst possible given the data. Whatever the motivation, the detection rule remains $U^n \geq h$, but with U^n now defined by Eq. (6) instead of Eq. (1). The problem is that this approach does not easily lend itself to a recursive formulation. A remedy is to smooth over $\arg\max_\theta \log \frac{f_\theta(x_{k+1})}{p_0(x_{k+1})}$ by combining the last one at time $k = n$ with the previous estimate of θ:

$$\widetilde{U}^{n+1} = \max \left\{ 0, \widetilde{U}^n + \log \frac{f_{\widetilde{\theta}^{n+1}}(x_{n+1})}{p_0(x_{n+1})} \right\}, \ \widetilde{U}^0 = 0, \\ \widetilde{\theta}^{n+1} = (1 - \kappa)\widetilde{\theta}^n + \kappa \arg\max_\theta \log \frac{f_\theta(x_{k+1})}{p_0(x_{k+1})}, \tag{7}$$

with $0 < \kappa < 1$ is some gain.

III. Distribution of the Frequency Scaling Exponent under Normal Conditions

In this section, we study the statistical characteristics of the frequency (f) time series collected in EPFL campus grid in 2016 [1]. The data was measured at normal conditions using PMUs installed to monitor the EPFL campus grid. The sampling rate of the PMUs is 50 samples/second.

Detrended Fluctuation Analysis (DFA) is one of the most reliable and robust methods to calculate the scaling exponent. We calculate the scaling exponent using the function 'dfa' from package 'nonlinearTseries' in R software. We applied the DFA method on 467 frequency time series (100,000 samples each) chosen from several months in 2016. The histogram of the frequency scaling exponents (Fig. 1(a)) is centered approximately around 1.48 with scaling exponents between 1.36 and 1.65.

Our main goal is identifying the best distribution to fit the histogram of the frequency scaling exponents.

A. Weibull Distribution

Using the maximum likelihood estimation, we fit the histogram of the scaling exponents to Weibull distribution. The maximum likelihood estimation is implemented in the function 'fitdistr' from package 'MASS'. The Weibull distribution with the best fit has shape (β) of 26.44 and scale (η) of 1.52. The probability density function (PDF) and the cumulative distribution function (CDF) are shown in green color in Fig. 1(a) and Fig. 1(b), respectively. It is clear from Fig. 1(a) that the Weibull distribution is not a good fit for the histogram of scaling exponents. To test the goodness of fit, we use the Kolmogorov-Smirnov test (KS test) with the null hypothesis (H_0) that the data follows a Weibull distribution. The KS test is implemented in the function 'ks.test' from package 'stats'. Applying the KS test on the histogram of scaling exponent shows that we can reject the null hypothesis with p-value of $3.54 \times 10^{-3} < 0.05$.

B. Normal Distribution

Similarly, the function 'fitdistr' uses the maximum likelihood to calculate the parameters of the Normal distribution. The parameters of the best Normal distribution fitting for the EPFL PMU data are

$$\mu_0 = 1.488, \quad \sigma_0 = 0.055 \quad (\text{EPFL}).$$

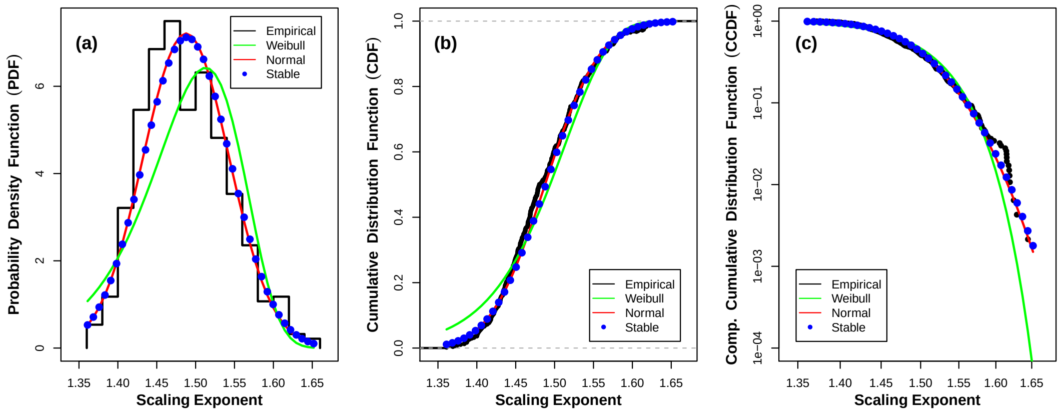

Fig. 1: (a) Probability density function (PDF) of the frequency scaling exponent (Empirical (black), Normal (red), Stable (blue), and Weibull (green)) (b) Cumulative distribution function (CDF) of the frequency scaling exponent (Empirical (black), Normal (red), Stable(blue), and Weibull (green)) (c) Log-log plot of the complementary cumulative distribution function (CCDF) of the frequency scaling exponent ((Empirical (black), Normal (red), Stable(blue), and Weibull (green))

The PDF and CDF of the Normal distribution are shown in red color in Fig. 1(a) and Fig. 1(b), respectively. We use the KS test to identify the goodness of the Normal distribution fitting. The KS test compares the empirical CDF of the data with CDF generated from the best distribution fit to check if the data samples come from the Normal distribution. Applying the KS test on the frequency scaling exponent, the result shows that we cannot reject the null hypothesis with p-value equal to $0.1469 > 0.05$. So, we accept that the data follows a Normal distribution with μ = 1.488 and σ = 0.055.

C. Stable Distribution

The family of stable distributions are defined by four parameters: stability (α), skewness (β), scale (γ), and location (δ) parameters. The stability parameter (α) determines the existence of mean and variance of the stable distribution. The mean is undefined for $\alpha \leq 1$ and the variance is undefined for $\alpha < 2$. There is no closed formula for the PDF of stable distribution except for Gaussian distribution ($\alpha = 2$), Cauchy distribution ($\alpha = 1$ *and* $\beta = 0$), and Lévy distribution ($\alpha = 0.5$ and $\beta = 1$). The stable distributions are typically defined by their characteristic function as shown in Eq. (8),

$$\phi(t) = \exp\left(i\delta t - |\gamma t|^{\alpha}[1 + i\beta\,\mathrm{sgn(t)}\,\omega(t,\alpha)]\right). \quad (8)$$

The function $\omega(t,\alpha)$ is equal to $\tan(\frac{\pi\alpha}{2})$ for $\alpha \neq 1$ and equal to $\frac{2}{\pi}\log|t|$ for $\alpha = 1$. We fit the histogram of scaling exponents to a stable distribution using Koutrouvelis regression method [6]. This method is implemented using the function 'Estim' from package 'StableEstim' in R software. The parameters of the best stable distribution are $\alpha = 2$, $\beta = 1$, $\gamma = 0.039$, and $\delta = 1.488$. Given these parameters, $\omega(t,\alpha)$ becomes equals to zero and Eq. (8) reduces to $\phi(t) = \mathrm{e}^{i1.488t - 1.52\times10^{-3}t^2}$ which is the same as the characteristic function of Normal distribution with mean $\mu = 1.488$ and standard deviation $\sigma = (2 \times 1.52 \times 10^{-3})^{0.5} = 0.055$. That means the best fit for the histogram of frequency scaling exponents is also a Normal distribution.

IV. Change Point Detection of Frequency PMU for Blackout Detection

In this section, we will utilize the derivations from Sec. II and empirical results from Sec. III to implement the change point detection on the PMU frequency data obtained before the Indian blackout which occurred on the 30th and 31st July 2012 [2]. The frequency time series data collected before the 2012 Indian blackout has a length of 167,600 samples (sampled at 50 samples/second) and spans approximately 56 minutes as shown in Fig. 2(a).

The Hurst exponent of the frequency data has been calculated using the DFA procedure with a moving window of length 110,000 samples (37 minutes) and shift of 900 samples (18 seconds). Fig. 2(b) shows the Hurst exponent time series from the PMU frequency data.

A. Empirical estimate for pre- and post-distributions

Although a separate distribution parameter fitting could be done particularly for the normal operating regime of the Indian blackout data, this dataset contains only around 50 mins of information prior to the blackout. In comparison, the EPFL campus has put in place a medium-voltage grid installed with PMUs running for extended periods of time under normal operations with reliable and accurate data collection protocols [10]. The parameters obtained in Sec. III were considered to be the "normal" regime (p_0) parameters $\mu_0 = 1.488$ and $\sigma_0 = 0.055$, and the "abnormal" regime (p_1) was empirically defined as a normal distribution but with a shift in mean keeping same variance: $\mu_1 = 1.7$ and $\sigma_1 = 0.055$.

Fig. 2(c) and (d) show the change point algorithm Eq. (2) simulation results with the log-likelihood ratio values and the change point CUSUM statistic $U(k)$, respectively, at each time step based on the estimated distributions p_0 and p_1.

B. Threshold for False Alarm Rate

We first consider the simplified case where $b = 0$ in Model (12) of Appendix. Relative to this model, the average

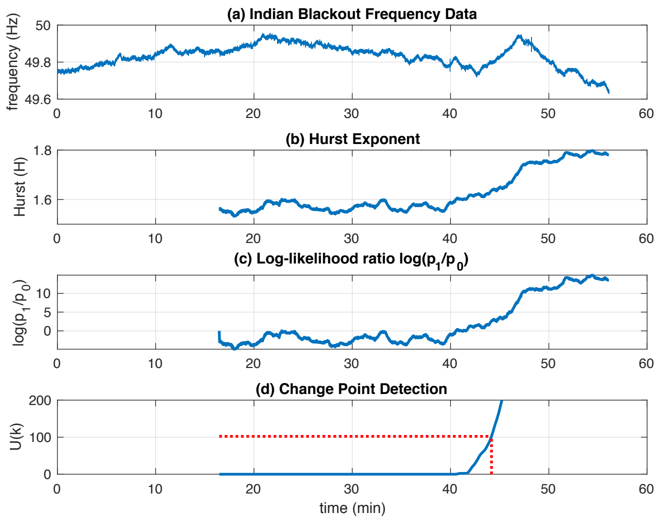

Fig. 2: Change point detection on the 2012 Indian blackout

time $T(x)$ for the 1-D random walk starting at x and reflecting at $-\epsilon$ to cross the absorption barrier at h is easily given by the solution of the differential equation of Theorem 1:

$$T(x) = \frac{1}{\sigma^2}(h^2 - x^2) + \frac{2\epsilon}{\sigma^2}(h - x). \tag{9}$$

Continuity relative to ϵ is trivial. Therefore, we can set $\epsilon = 0$ (initial condition and reflection barrier coincide at $x = 0$).

In the genuine model (2) of the process, let $p_0 = \mathcal{N}(\mu_0, \sigma_0)$ and $p_1 = \mathcal{N}(\mu_1, \sigma_1)$, consistently with the scaling exponent identification. It then follows that, up to a good approximation, Eq. (2) becomes

$$U^{n+1} - U^n \approx \frac{\mu_1 - \mu_0}{\sigma_\alpha^2}\left(X_{n+1} - \frac{\mu_0 + \mu_1}{2}\right), \tag{10}$$

where the approximation stems from equating σ_0 and σ_1 to σ_α. Practically, we could take $\sigma_\alpha^2 = (\sigma_0^2 + \sigma_1^2)/2$. The above equation has to be confronted with the continuous-time model (12) where $\mathbb{E}(B_{t+\Delta t} - B_t)^2 = \Delta t$, where $1/\Delta t$ here is the PMU sampling rate. Identifying the discrete-time and continuous-time processes hence yields

$$\sigma^2 = \frac{2(\mu_1 - \mu_0)^2}{\sigma_\alpha^2 \Delta t}.$$

Finally, setting $\epsilon = 0$ in Eq. (9), recalling that $\mathrm{FAR} = 1/T(x = 0) = \sigma^2/h^2$, one gets the estimate

$$h = \frac{\sqrt{2}(\mu_1 - \mu_0)}{\sigma_\alpha \sqrt{\Delta t}} \frac{1}{\sqrt{\mathrm{FAR}}}. \tag{11}$$

Given the empirical results for the parameters μ_0, μ_1 and $\sigma_\alpha = \sigma_0 = \sigma_1$, the sampling rate $\Delta t = 0.033$, and a 10% false alarm rate, $\mathrm{FAR} = 0.1$, we obtained the threshold value from Eq. (11) as $h = 101.9$. The horizontal dotted line in Fig. 2(d) shows the threshold value h and it crosses the change point statistic $U(k)$ at $t = 44.17$ min (11.83 mins before the blackout).

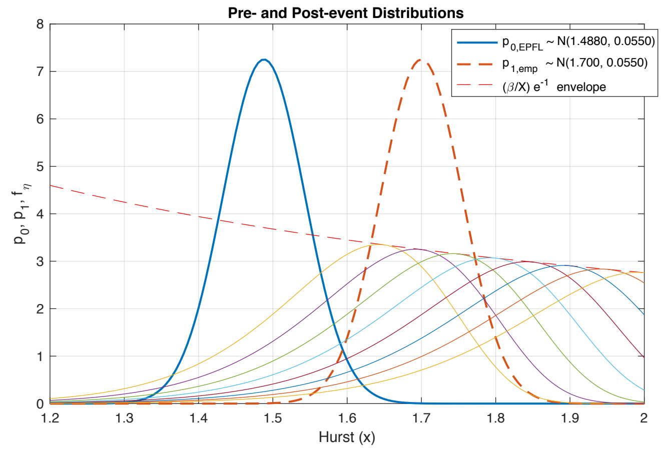

Fig. 3: Normal distributions $p_0(x)$ and $p_{1,\mathrm{emp}}(x)$, family of Weibull distributions $f_\eta(x)$ with $\beta = 15$, and envelope of Weibull distributions

C. Unknown post-distribution

For this case, the post distribution p_1 is assumed to be unknown and is assigned the Weibull distribution with natural parameter $\theta = -1/\eta^\beta$ as suggested in Eq. (5). For a fixed shape parameter β, the distribution is only parameterized by the scaling parameter η as $f_\eta(x)$, where the scaling parameter is related to the mean as $\mathbb{E}(x) = \eta\Gamma(1 + 1/\beta)$.

With this data, we use the recursive form of the change point algorithm for the simultaneous detection and estimation as defined in (7). For the Weibull distribution, the $\arg\max_\eta$ term in (7) results in

$$\arg\max_\eta \log \frac{f_\eta(X)}{p_0(X)} = \arg\max_\eta f_\eta(X).$$

Taking the derivative with respect to η and equating to zero results in a simple expression $\eta_{\max} = x$, where x is the *numerical value* recorded. As such the η is re-adjusted every single step, which creates some oscillation in the statistic, which can be somewhat smoothed over by the second equation of (7).

Fig. 3 shows the plot of the probability distributions for the empirical estimates for p_0 and $p_{1,\mathrm{emp}}$, which is used in Sec. IV-A, and the unknown distribution taking the form f_η. Note that by setting $\eta_{\max} = X$, the *variable* rather than the numerical value, and plugging this in $f_\eta(X)$ yields $(\beta/X)\,\mathrm{e}^{-1}$, the envelope of the Weibull distributions shown in Fig. 3.

One can choose the shape parameter β so that the distributions have higher peak values. However, setting β too high can lead to higher false alarm rates. Additionally, the range of values for the scaling parameter η, related to the mean, is limited to a minimum value of 1.65. By setting the minimum value for η, this ensures that f_η would not completely overlap with the known pre-event distribution p_0.

Fig. 4 shows the simulation results for simultaneous detection and estimation algorithm with $\beta = 15$ for the shaping parameter. For the same threshold value $h = 101.9$, the algorithm resulted in 3 false alarms before it correctly raised

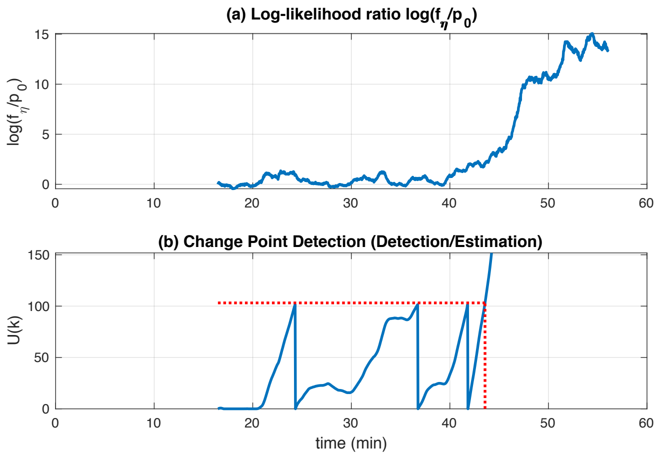

Fig. 4: Change point detection for simultaneous detection and estimation with chosen parameter $\mu_0 = 1.488$ and $\sigma_0 = 0.055$, and $\beta = 15$

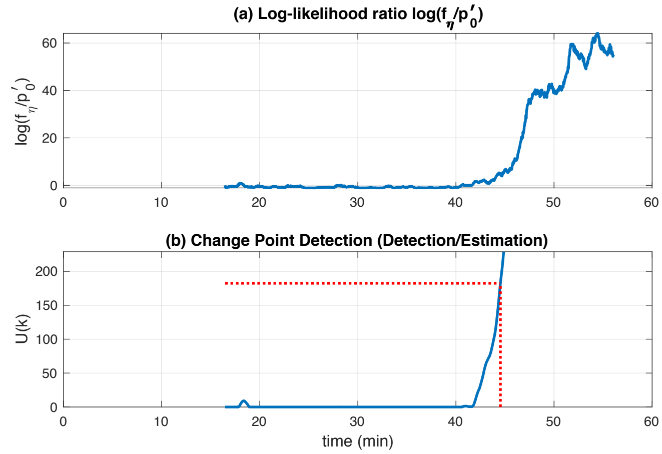

Fig. 5: Change point detection for simultaneous detection and estimation with chosen parameter $\mu_0' = 1.5722$ and $\sigma_0' = 0.0198$, and $\beta = 20$

the alarm at $t = 43.56$ min (12.44 mins before the blackout). Similarly, simulations were run for higher values of β and these resulted for an even higher numbers of false alarms.

Finally, simulations were run with an estimated p_0 by using the first 40 minutes of the Indian blackout data which are considered operating under normal conditions, and we have obtained the following normal distribution p_0' parameters

$$\mu_0' = 1.5722, \quad \sigma_0' = 0.0198 \quad \text{(Indian pre-blackout)}.$$

For this case, the threshold value at these p_0' parameters is $h = 178.76$, the algorithm resulted in no false alarms and it raised the alarm at $t = 44.53$ min (11.47 mins before the blackout) as shown in Fig. 5.

It is worth noting that for this implementation of the CPD, the algorithm relies in a more accurate model for the "normal" regime distribution p_0.

V. Conclusion

In this paper, a CUSUM change point detection approach to the prediction of imminent voltage collapse and possibly stealthy security attacks to the power grid was explored. Given empirical data, distribution fitting of the frequency scaling (Hurst) exponent under normal (pre-event) operating conditions was shown to follow a normal distribution. Two recursive formulations were derived are based on an empirical p_1, the other as simultaneous detection and identification of p_1. Simulation results for the 2012 Indian blackout data has shown that the change point detection algorithm is capable of detecting imminent voltage collapse as early as 12 minutes before the blackout event.

For future work, the threshold value h based on the complete convective-diffusion PDE model will be solved analytically to obtain an optimal threshold value.

Acknowledgements

The authors gratefully acknowledge the support by the U.S. Army Defense Advanced Research Projects Agency (DARPA) under grant no. W911NF-17-1-0076, DARPA Young Faculty Award under grant no. N66001-17-1- 4044, and the National Science Foundation under CAREER Award CPS-1453860 support. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the Department of Defense or the National Science Foundation.

Appendix

Here, we solve an essentially stochastic problem via a partial differential equation (PDE) solution, instead of solving the partial differential equation via Monte Carlo procedure, as is commonly done with the Dirichlet problem.

The idea is to model U^n as an Itô diffusion process over some domain $D \subset \mathbb{R}$:

$$dU^t = b(U^t)dt + \sigma(U^t)dB_t, \quad U^0 = x \in D, \qquad (12)$$

where B_t is a Brownian motion and $b(U) < 0$ and $\sigma(U)$ are assumed to satisfy Conditions 5.14 and 5.15 of [8, Th. 5.5]. In this specific escape problem, the absorbing barrier is at $h > 0$, and both the initial condition U^0 and the reflecting barrier are at 0.

In the general escape problem, the reflecting barrier is on the boundary of the domain (hence not in the domain) and the initial condition is in the domain. To circumvent this difficulty, we temporarily consider the domain $D = (-\epsilon, h)$, where the reflecting barrier is at $-\epsilon < 0$, and then show continuity of the escape time as $\epsilon \downarrow 0$.

Even though in this specific problem $U^0 = 0$, we need to take $U^0 = x \in D$, not for the sake of generality but because the solution involves a PDE in the initial condition x.

Theorem 1: Consider the process (12) running over the domain $(-\epsilon, h)$, with $h > 0$ the absorbing barrier and $-\epsilon$ the reflecting barrier. Define $\tau_{-\epsilon,x}(h)$ such that $U^{\tau_{-\epsilon,x}(h)} = h$. Then $\mathbb{E}(\tau_{-\epsilon,x}(h)) = T(x)$, where $T(x)$ is given by the mixed

Neumann-Dirichlet boundary value problem

$$\begin{aligned}\left(\frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}-\frac{\partial}{\partial x}b(x)\right)T(x) &= -1,\\ T(h)=0,\qquad \left.\frac{\partial T(x)}{\partial x}\right|_{x=-\epsilon} &= 0.\end{aligned}$$

Proof: Let $p(x,y,t)$ be the probability density of $y=U^t$ given the initial condition $U^0=x$. As is well known, p is solution to the Kolmogorov Fokker-Planck (KFP) forward equation:

$$\begin{aligned}\left(\frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2}-\frac{\partial}{\partial y}b(y)\right)p(x,y,t) &= \frac{\partial p}{\partial t},\\ \left.\frac{\partial p(x,y,t)}{\partial y}\right|_{y=-\epsilon}=0,\quad p(x,h,t) &= 0, \forall t\geq 0,\\ p(x,y,0) &= \delta(x-y).\end{aligned}$$

Next, we show that $p(x,y,t)$ decays exponentially fast as $t\to\infty$. Use the method of separation of variables: $p(x,y,t)=f(y)g(t)$. Plug the latter in the KFP equation and we get

$$\frac{1}{2}\sigma^2 f''(y)-b(y)f'(y)=cf(y),\quad g'(t)=cg(t),$$

where c is a constant, an eigenvalue of the KFP PDE. Multiply both sides of the equation for f by $f(y)dy$, integrate by parts, set $b(y)=\beta y$ with $\beta<0$ to simplify, and use the boundary conditions to get

$$-\frac{1}{2}\sigma^2\|f'\|_D^2+\frac{\beta}{2}\|f\|_D^2=c\|f\|_D^2,$$

where $\|\cdot\|_D$ denotes the L^2-norm over D. It follows that $c<0$. The equation for g hence yields the exponential decay.

Using the exponential convergence of p to 0 as $t\to\infty$, we define

$$G(x,y)=\frac{1}{2}\int_0^\infty p(x,y,t)dt.$$

With this definition, is not hard to show that G is the Green function of the KFP equation, that is,

$$\left(\frac{1}{2}\sigma^2\frac{\partial^2}{\partial y^2}-b(y)\frac{\partial}{\partial y}\right)G(x,y)=\delta(x-y),$$

subject to the boundary conditions.

Finally, the probability that the motion hasn't reached the threshold h as of time t is

$$\pi(x,t)=\int_{-\epsilon}^h p(x,y,t)dy.$$

Hence the expected lifetime of the motion before it reaches h is

$$\begin{aligned}\mathbb{E}(\tau_{-\epsilon,x}(h)) &= \int_0^\infty t(\pi(x,t)-\pi(x,t+dt))dt\\ &= \int_{-\epsilon}^h G(x,y)dy=T(x).\end{aligned}$$

■

Corollary 1: $\mathbb{E}(\tau_{-\epsilon,x}(h))$ is continuous as $\epsilon\downarrow 0$, $\forall x\geq 0$.

Proof: The solution to the differential equation for $T(x)$ is given by

$$\begin{aligned}T(x) &= -\frac{1}{B}+\frac{e^{-z^2}}{2B}\left(2Bc_2+\sqrt{\pi B}\sigma\mathrm{Erfi}(z)\right),\\ T'(x) &= c_1-\frac{e^{-z^2}}{\sigma^2}2Bxc_2-2\sqrt{B}xc_1D(z),\end{aligned}$$

where $B=-\beta>0$, $z=\sqrt{B}x/\sigma$, and $D(\cdot)$ denotes the Dawson integral. Setting the boubndary conditions $T(h)=0$, $T'(-\epsilon)=0$ yields a system of linear equations for c_1,c_2 with a solution continuous for ϵ.

■

References

[1] Pmu data from epfl campus. `http://nanotera-stg2.epfl.ch/`. [Online; accessed 3-March-2018].

[2] Report on the Grid Disturbance on 30th July 2012 and Grid Disturbance on 31st July 2012. Technical report, Central Agency Regulatory Commision, 2012.

[3] Michèle Basseville and Igor V. Nikiforov. *Detection of Abrupt Changes - Theory and Application*. Englewood Cliffs, NJ, 2010.

[4] R. Blazek, H. Kim, B. Tozovskii, and A. Tartakovsky. A novel approach to detection of 'denial-of-service' attacks via adaptive sequential and batch sequential change-point detection. In *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, pages 220–226, U.S. Military Academy, West Point, NY, June 5-6 2001.

[5] H. Kim, B. L. Rozovskii, and A. G. Tartakovsky. A nonparametric multichart CUSUM test for rapid detection of DOS attacks in computer networks. *International Journal of Computing & Information Sciences*, 2(3):149–158, December 2004.

[6] Ioannis A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association*, 75(372):918–928, 1980.

[7] R. Martin del Campo Vera and E. Jonckheere. Neuronal rhythmicity and its rate variability. Technical report, University of Southern California, Ming Hsieh Dept. of Electrical Engineering, 2018. Available at `http://eudoxus2.usc.edu`.

[8] B. Oksendal. *Stochastic Differential Equations*. Springer, Berlin, 1995.

[9] E. S. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.

[10] Marco Pignati et al. Real-time state estimation of the EPFL-campus medium-voltage grid by using PMUs. *2015 IEEE Power and Energy Society Innovative Smart Grid Technologies (ISGT) Conference*, 2015.

[11] H. Sedghi and E. Jonckheere. Statistical structure learning to ensure data integrity in smart grid. *IEEE Transaction on Smart Grid*, 6(4):1924–1933, 2015. Available at http://eudoxus2.usc.edu.

[12] Khushboo Shah, Edmond Jonckheere, and Stephan Bohacek. Dynamic modeling of internet traffic for intrusion detection. *EURASIP Journal on Advances in Signal Processing*, 2007:Article ID 90312, 14 pages, 2007. doi:10.1155/2007/90312.

[13] L. Shalalfeh, P. Bogdan, and E. Jonckheere. Evidence of long-range dependence in power grid. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5, July 2016.

[14] Laith Shalalfeh, Paul Bogdan, and Edmond Jonckheere. Kendall's tau of frequency Hurst exponent as blackout proximity Margin. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 466–471. IEEE, nov 2016.

[15] A. N. Shiryaev. On optimum methods in quickest detection problems. *Theory Probab. Appli.*, pages 22–46, 1963.

[16] A. Teixeira, G. Dan, H. Sandberg, and K H. Johansson. A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator. In *IFAC World Congress*, September 2011.