Sequential Probability Ratio Tests with a Moving Average using the Generalized Likelihood Ratio (SMAGLR) for Change Point Detection on Graphs

Christopher Robbiano — Due: December, 2017 — ECE651 Final

1 Introduction

The purpose of this project is to investigate different methods of change point detection for identifying the change in either the variance, mean or both from observations of independent and identically distributed (IID) signal samples buried in Gaussian noise from the wave equation on a graph, Γ. An example of a graph of interest is seen in Figure (1). The graph contains no cycles and each outer vertex obeys a set of boundary conditions while each edge is defined as a wave equation. A known forcing term is placed on only one of the outer vertices and observations are taken at the remaining outer vertices. Assuming that sensing the forcing term at each vertex introduces some known Gaussian noise structure into the observations, the goal is to identify when a change in the distribution of the noise structure occurs.

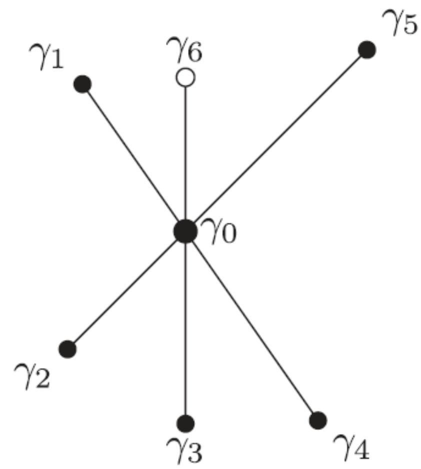


Figure 1: Tree Graph

The network described above has simple real world analogs, such as an airplane wing or a bridge[1]. The case of detecting a change in the distribution of a signal that is transmitted from one end of a bridge to another is akin to performing non-destructive evaluation (NDE) for addressing safety concerns. One end of the bridge would have a transducer affixed that provides a known forcing term, while at the other end of the bridge a transceiver would record the signal that has propagated through the bridge. The model used in this project assumes that both the forcing term and variance of the measurement device are known. These assumptions lead to a completely specified distribution of the observed samples and any minor deviation in the parameters of the distribution should indicate that the structural integrity of the bridge has been changed. Abnormalities in the structure can be responsible for an increase in the variance, a change in the mean, or a combination of both but still have a extremely small overall effect on the observation. Due to this, any technique for detection must be sensitive to changes in both the signal-to-noise ratio (SNR) of the signal as well as deviations in the mean.

In this project, several well known change point detectors are implemented and evaluated for performance. A Sequential Probability Ratio Test (SPRT) with a Moving Average (MA) using the Generalized Likelihood Ratio (GLR), coined the SMAGLR, is also developed and the performance is compared against the performance of well known detectors. The results show that for this type of problem, the detector developed in this project outperforms all other evaluated detectors in average detection delay (ADD) and probability of detection (P_D). The developed detector outperforms all other detectors expect one in probability of false alarm (P_{FA}).

2 Previous Work

It is shown by Avdonin *et al* in [2] and [3] that for the case when a graph, Γ, is a single interval that the model defining the graph is given as follows

$$u_{tt} - u_{xx} + q(x)u = f(t)g(x) \qquad 0 < x < l,\ 0 < t < T \tag{1}$$
$$u(0,t) = u(l,t) = 0 \qquad 0 < t < T \tag{2}$$
$$u(x,0) = u_t(x,0) = 0 \qquad 0 < x < l \tag{3}$$

where $q \in L^\infty(0,l)$ and $f \in H^1(0,T)$, $f(0) \neq 0$ and $g \in L^2(0,l)$ are known functions. The subscripts on u denote the first and second derivatives with respect to that variable. It has been shown that

$$u_x(0,t) = \int_0^t g(x)w(x,t)dx \qquad \forall\ 0 \le t \le l \tag{4}$$

where $w(x,t) = f(t-x) + \int_x^t k(x,s)f(t-s)ds$ for $0 < x \le t$ and 0 otherwise. Robbiano showed in [9] that

$$u_x(0,t) = \int_0^t f(t-x)g(x)dx = (f * g)(x) \tag{5}$$

when the trivial potential, q, is the zero function.

Change point detection is a well studied topic in statistical signal processing. Most notably, Wald developed the SPRT[10] and developed a method for calculating the ADD. There have been advancements in the area of sequential detection theory including the development of the Shiryaev-Roberts procedure[8], and Shewart 3-sigma and CUSUM control charts. The evaluation of these detectors has been performed for on-line environmental radiation monitoring under similar assumptions that are made for this project[11].

3 Theory

The SMAGLR detector assumes a model where IID observations of a known signal in noise from a completely specified Gaussian distribution are collected before the change point. The variance and/or mean of the distribution after the change point is unknown in both magnitude and direction, but it assumed that the variance always increases if it does change while the mean is free to change arbitrarily. It is also assumed that there is only one change point, and the detection problem ends once the decision to choose $\mathcal{H}_1$ has been made. The observation model is seen below for each hypothesis

$$\mathcal{H}_0 :\ z(n) = u_x(0,n) + v(n) \qquad \mathcal{H}_1 :\ z(n) = u_x(0,n) + \alpha(n) + v(n) \tag{6}$$

where $u_x(0,n)$ is known and deterministic and is the convolution between the driving source and other known function as described in (5), $v(n) \sim \mathcal{N}(0, \sigma_0^2)$ with σ_0^2 known and $\alpha(n) \sim \mathcal{N}(\mu(n), \sigma_\alpha^2(n))$ with both $\sigma_\alpha^2(n)$ and $\mu(n)$ unknown. In the case where $\sigma_\alpha^2(n) \to 0$, $\alpha(n) \to$

$\mu(n)\delta(n)$ where $\delta(\cdot)$ is the Dirac delta measure. For ease of notation, $\mu(n) = \mu$, $\sigma_\alpha^2(n) = \sigma_\alpha^2$ and $\sigma_1^2 = \sigma_\alpha^2 + \sigma_0^2$. This leads to the following distributions for $z(n)$

$$\mathcal{H}_0 : \; z(n) \sim \mathcal{N}(u_x(0,n), \sigma_0^2) \qquad\qquad \mathcal{H}_1 : \; z(n) \sim \mathcal{N}(u_x(0,n) + \mu, \sigma_1^2) \tag{7}$$

We develop the GLR test based on a window containing the k most recent observations, $z(n-k), z(n-(k+1)), \ldots, z(n)$[7]. Since all observations in the window are IID we get the following form for the joint distribution of the observations within the window for $j = 0, 1$

$$\mathbf{P}\big(z(n-k), \ldots, z(n) | \mathcal{H}_j\big) = \prod_{i=n-k}^{n} \mathbf{P}\big(z(i) | \mathcal{H}_j\big) \tag{8}$$

Given that the mean function, $u_x(0,n)$, is known *a priori* we choose to work with a zero mean random variable for the following derivations to simplify notation. The likelihood ratio based on the Neyman-Pearson lemma becomes [6]

$$\begin{aligned}
L\big(z(n-k), \ldots, z(n)\big) &= \prod_{i=n-k}^{n} \frac{\mathbf{P}\big(z(i)|\mathcal{H}_1\big)}{\mathbf{P}\big(z(i)|\mathcal{H}_0\big)} && (9)\\
&= \prod_{i=n-k}^{n} \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\big(-\frac{1}{2\sigma_1^2}(z(i)-\mu)^2\big)}{\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\big(-\frac{1}{2\sigma_0^2} z^2(i)\big)\big)} && (10)\\
&= \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\big(-\frac{1}{2\sigma_1^2}\sum_{i=n-k}^{n}(z(i)-\mu)^2\big)}{\frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\big(-\frac{1}{2\sigma_0^2}\sum_{i=n-k}^{n} z^2(i)\big)} && (11)\\
&= \sqrt{\frac{\sigma_0^2}{\sigma_1^2}} \exp\Big(\frac{1}{\sigma_0^2}\sum_{i=n-k}^{n} z^2(i) - \frac{1}{\sigma_1^2}\sum_{i=n-k}^{n}(z(i)-\mu)^2\Big) && (12)
\end{aligned}$$

Taking the natural log of both sides provides the log-likelihood ratio

$$\begin{aligned}
\Lambda\big(z(n-k), \ldots, z(n)\big) &= \ln\Big(L\big(z(n-k), \ldots, z(n)\big)\Big) && (13)\\
&= \frac{1}{2}\ln\frac{\sigma_0^2}{\sigma_1^2} + \Big(\frac{1}{\sigma_0^2}\sum_{i=n-k}^{n} z^2(i) - \frac{1}{\sigma_1^2}\sum_{i=n-k}^{n}(z(i)-\mu)^2\Big) && (14)
\end{aligned}$$

Given that both the mean and variance are unknown after a change in the distribution, we must provide the Maximum Likelihood Estimates (MLE) of both μ and σ_1^2 to determine the GLR test. It has been shown in [5] that

$$\hat{\mu} = \frac{1}{k}\sum_{i=n-k}^{n} z(i) \qquad\qquad \hat{\sigma}_1^2 = \frac{1}{k}\sum_{i=n-k}^{n}\big(z(i)-\hat{\mu}\big)^2 \tag{15}$$

Plugging in the values from (15) into (14) gives the following test statistic

$$\hat{\Lambda}(z(n-k), \ldots, z(n)) = t = \frac{1}{2}\ln\frac{\sigma_0^2}{\hat{\sigma}_1^2} + \Big(\frac{1}{\sigma_0^2}\sum_{i=n-k}^{n} z^2(i) - \frac{1}{\hat{\sigma}_1^2}\sum_{i=n-k}^{n}(z(i)-\hat{\mu})^2\Big) \gtrless \gamma_1' \tag{16}$$

The moving average portion of the detector is related to the MLE estimate of μ, which is a moving average over k observations. The value of k that is chosen is ultimately up for debate and can be implemented as a tuning parameter.

The alarm time a_t, the time at which the detector decides $\mathcal{H}_1$ over $\mathcal{H}_0$, can be quantified as $a_t = \inf\{n|\mathcal{H}_1 \text{ is chosen}\}$ and the detection delay time d_t can be quantified as $d_t = \upsilon - a_t$, where υ is the true change time.

4 Testing Procedure

In order to evaluate the performance of the SMAGLR detector versus the others mentioned previously, a series of tests was run and an evaluation of the ADD, P_D and P_{FA} was done. Each test consisted of 1000 time steps, with change points at $n = 200$ for variance and $n = 500$ for mean. Each test had either no change, a change at $n = 200$ or $n = 500$ or both.

There was no available data set to examine, and thus an artificial data set needed to be generated according to the problem as described in Section 2. To generate the data that is observed under $\mathcal{H}_0$, $u_x(0, n) = (g * f)(n)$ was produced per (5) as to satisfy the wave

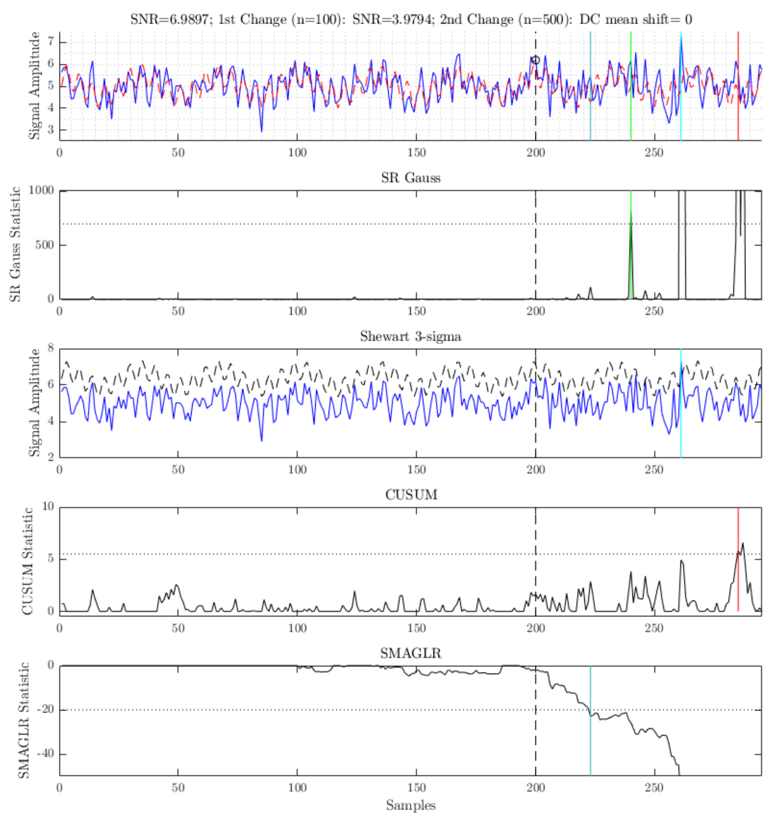


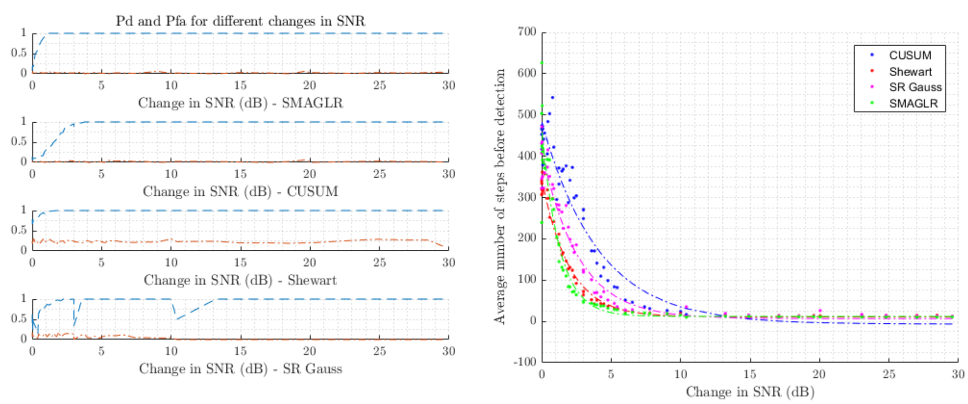
Figure 2: Control charts

equation. Noise is then added with a known variance such that $z(n) = u_x(0,n) + v(n)$. The forcing term is $g(n) = \sin(2\pi f_0 n + \phi)$ and the other known function is $f(n) = \cos(n)$. To generate the data that is observed under $\mathcal{H}_1$, $z(n) = u_x(0,n) + \alpha(n) + v(n)$ for $n >= 200$. The parameters dictating the distribution of $z(n)$ were swept such that $\sigma_0^2 \in [0.001, 1]$, $\sigma_\alpha^2 \in [0, 1]$ and $\mu \in [-3, 3]$. One hundred realizations of $z(n)$, $n = 1, \ldots, 1000$, for each set of $(\sigma_0^2, \sigma_\alpha^2, \mu)$ was then evaluated and the performance over the one hundred realizations for each set of $(\sigma_0^2, \sigma_\alpha^2, \mu)$ was averaged. Four detectors were compared for performance: Shiryaev-Roberts, Shewart 3-sigma, CUSUM and SMAGLR. In total, 180,000 tests were run on each detector.

5 Results and Discussion

The data collected from the tests were separated into three different categories. The first category was no change occurring throughout the entire test. This category of tests was an explicit test for false alarm rate. The second and third categories are zero mean change and zero variance change, respectively. Some auxiliary tests were run with gradual ramps in σ_1^2, μ and a combination of the two, but no performance insights were gleaned and thus only tests that exhibited an abrupt change were evaluated. An example of the control charts produced by the testing procedure is shown in Figure (2).

Due to the real world application of sensing small abnormalities in relatively large signals, performance of the detector is compared against the SNR in the case of a change in variance and against the magnitude of change in mean for a change in mean[4].



(a) P_D and P_{FA} versus SNR (b) ADD performance

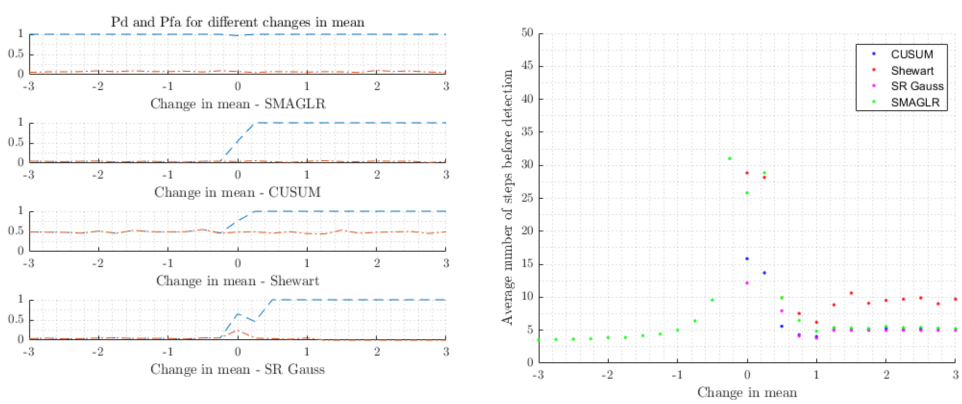
Figure 3: Detection performance for change in variance

5.1 Change in Variance

Figure (3a) shows the P_D in blue and P_{FA} in orange plotted against the change in SNR for each detector. All changes in SNR degrade the overall signal, as it is assumed that no abnormality will improve the noise variance. As one would expect, the detection performance increases with larger changes of SNR coinciding with a larger J divergence. The probability of false alarm is essentially constant for all of the detectors. It can be seen that the detection performance of the Shewart detector is slightly better than that of the SMAGLR for small (< 1 dB) change in SNR. This performance discrepancy can also be seen in the ADD in Figure (3b) by looking at the data points and the fit lines for both the Shewart and SMAGLR detectors. For changes in SNR less than 1 dB, the ADD of the SMAGLR detector is greater than that of the Shewart. The SMAGLR detector is superior for all other changes in SNR.

5.2 Change in Mean

Figure (4a) shows the P_D in blue and P_{FA} in orange plotted against the change in mean for each detector. Again as one would expect, the detection performance increases with larger changes of the mean coinciding with a larger J divergence. The probability of false alarm is essentially constant for all of the detectors, and is particularly bad for the CUSUM detector. The CUSUM detector was developed specifically for detecting small deviations in the mean, and could likely see improved results by tweaking some of the tuning parameters. It can be seen that the detection performance of the SMAGLR detector is better in every regard when compared to the other three detectors. This is especially evident when comparing a negative change in the mean, as the SMAGLR detector is the only one that is able to operate and provide results. All of the other detectors have a 100% missed detection rate when the change in mean is negative. There are ways to adopt the other detectors to allow for a detection of a negative mean change, such as a two sided CUSUM, but those changes were not made for this comparison as the SMAGLR detector needed no changes to attain its current performance. The ability to detect a negative mean change can also be see in the ADD in Figure (4b) by looking at the left hand side of the graph.



(a) P_D and P_{FA} versus change in mean (b) ADD performance

Figure 4: Detection performance for change in mean

6 Future Work and Conclusion

A new detector, SMAGLR, was developed specifically for the problem of identifying small changes in the parameters of a known Gaussian distribution. It was shown that under most operating conditions, the SMAGLR detector outperformed some well known counter parts. In particular, it was the only detector that was evaluated that could detect a negative change in mean without a fundamental change to its theory. Although there are some impressive results in the case of a single segment graph, there is still much work that can be done to improve the detector as well as extend the theory to arbitrary tree graphs. They theory for generating an artificial data set for an arbitrary tree graph exists in a number of papers by Avdonin, but the known mean and variance at each detector would need to be calculated separately. Also, the detection problem for an arbitrary tree graph would seem to lead its self towards a combination of distributed and sequential sensing. The test statistic used can be simplified and the distribution of the test statistic can be found under both hypothesis. An attempt to do this, in order to explicitly solve for the expected value of the detection delay $E(K^*)$, was stopped short when calculations led to negative expected detection delay times. Also, examining the P_D and P_{FA} versus SNR and change in mean plots, it appears that the detectors are all constant false alarm rate (CFAR) detectors. Future work will aim to address whether the SMAGLR detector is truly a CFAR detector as well as any other areas of study that were not covered in this project.

References

[1] Sergei Avdonin, Bill Moran, and Ali Pezeshki. Sensing and estimation of differential equation networks. Project proposal with practical applications, 2015.

[2] Sergei Avdonin and Serge Nicaise. Source identification for the wave equation on graphs. *Comptes Rendus Mathematique*, 352(11):907–912, 2014.

[3] Sergei Avdonin and Serge Nicaise. Source identification problems for the wave equation on graphs. *Inverse Problems*, 31(9):095007, 2015.

[4] Michèle Basseville, Igor V Nikiforov, et al. *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs, 1993.

[5] Steven M Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. PTR Prentice-Hall, Englewood Cliffs, 1993.

[6] Steven M Kay. *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. PTR Prentice-Hall, Englewood Cliffs, 1998.

[7] Robert Nowak. The sequential probability ratio test. Lecture 9, 2015.

[8] Aleksey S Polunchenko and Alexander G Tartakovsky. On optimality of the shiryaev-roberts procedure for detecting a change in distribution. *The Annals of Statistics*, pages 3445–3457, 2010.

[9] Christopher Robbiano. Source identification for the wave equation based on noisy measurements on tree graphs. Final Presentation and Paper, 2015.

[10] Abraham Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, 1945.

[11] Mara Mae Watson. *Evaluation of Shiryaev-Roberts Procedure for On-line Environmental Radiation Monitoring*. PhD thesis, Clemson University, 2015.