Trend Analysis of Regime Change and Social Unrest with Laplace Test Statistic

Christian Díaz-Cáez and Charles Kim

Department of Electrical Engineering and Computer Science, Howard University, Washington, DC 20059, USA christian.diaz@bison.howard.edu; ckim@howard.edu

Abstract—This research focuses on the integration of the Laplace trend statistic into a web-based collaborative platform to analyze regime changes and social unrest. The Laplace statistic, known for discerning trend, is applied to unveil any trend of the changes and unrest in data sets. The platform accommodates real-time analysis and encourages inter-agency collaboration. The paper details the development of the system, including the mathematical derivation of the Laplace statistic, coding techniques, and application examples. By combining statistical methods with a versatile collaboration tool, the work aims to contribute to the evolving understanding of complex socio-political dynamics, offering a more responsive and robust approach to trend analysis and prediction of such dynamics

Index Terms—Collaborative platform, fragment, Laplace statistics, trend analysis,

I. INTRODUCTION

Analyzing the dynamics of regime change and social unrest requires a deeply interdisciplinary approach, blending tools from political science, sociology, and statistics [1]. Such intricacy underscores the necessity for collaboration and consensus-building across multiple organizations. Yet, these collaborations often face obstacles due to the diverse nature of data sources and varying viewpoints on the same issue. Recognizing this, we have taken our previously created collaborative web-based platform a step further [?]. In this work, we integrate trend analysis, paving the way for interactive participation and richer contributions from various stakeholders [2].

Beyond the establishment of a web-based platform, a key aspect of this research is the innovative integration of the Laplace statistic for robust trend analysis [3]. This comprehensive approach seeks to take advantage of the Laplace statistic's unique capability to effectively distinguish between a constant and increasing rate of occurrence of complex, multifaceted events. The goal is to leverage this capability for detecting early, significant signs of regime change and social unrest, thereby enhancing the overall predictive power and accuracy of the system. The choice of the Laplace statistic is informed by its simplicity, efficiency, adaptability, and reliability, making it an excellent tool for analyzing trends in a constantly changing and unpredictable socio-political environment. As we further unravel the complexities associated with regime shifts and social unrest, it is increasingly evident that single-source data and traditional methods of analysis may not yield satisfactory results. By adopting a more integrated and interactive approach, we can more effectively utilize the vast amounts of data generated by diverse agencies, organizations, and stakeholders. Therefore, the development and implementation of a web-based platform for collaborative data collection is not merely a choice; it is a necessity. This platform must not only facilitate the gathering of information but also be capable of real-time analysis, providing a more dynamic and insightful understanding of regime changes and social unrest.

The volatile nature of political regimes and social movements necessitates an agile and responsive analysis system capable of quickly recognizing emerging patterns, trends, and anomalies. This paper's application of the Laplace statistic in trend analysis adds a refined layer of sophistication to traditional approaches, offering the potential to detect nuances, subtleties, and complexities within large data sets. This statistical method, coupled with the web-based collaborative platform, represents a significant stride toward creating a predictive tool that's both responsive to the multidimensional factors influencing societal change and robust enough to confidently accommodate variations in data quality and availability.

Furthermore, the emphasis on collaboration and interagency communication reflects a broader paradigm shift towards collective intelligence and concerted action. In the age of information overload, the capacity to filter, interpret, and act upon relevant data is paramount. The system described in this paper offers not just a technical solution but also a conceptual framework that seeks to harmonize different perspectives, methodologies, and objectives [4]. By creating a platform that encourages shared understanding and common purpose, this work contributes to the evolving discourse on technology's role in understanding and shaping our complex socio-political landscape. Moreover, fostering a culture of inclusivity and open dialogue within the platform can foster a sense of ownership and commitment among stakeholders, leading to more effective and sustainable solutions for addressing societal challenges.

This paper is structured into several key sections: Chapter II delves into the development of an enhanced decision support system using trend analysis; Chapter III offers an in-depth examination of the mathematical derivation, interpretation, and coding techniques for trend analysis; Chapter IV investigates

Financial support provided by U. S. Army through MSI STEM Research & Development Consortium with the project entitles "Irregular Warfare Decision Assist System for Determining Threat," (Award Number: D01 _W911SR-14-2-0001-0015)

the application and testing of the Laplace statistic, complemented by examples and validation; and finally, Chapter V provides the conclusion to the paper.

II. ENHANCED DECISION SUPPORT SYSTEM VIA TREND ANALYSIS

The innovative construct of a web-based collaborative platform is a significant design principle that focuses on enabling cooperative interactions between stakeholders from various entities across geographical locations. This dynamic concept is a versatile tool employed in numerous applications, ranging from production, development, and marketing to service sectors, and is utilized to facilitate comprehensive collaboration and joint endeavors [5].

The imperative nature of a reliable web-based collaborative platform can hardly be overstated. It empowers resource consolidation and encourages synergistic collaboration, which allows for the identification of common patterns in data sets, thus fostering a shared understanding. This heightened demand for tailored systems motivates organizations to allocate substantial resources toward their design and development. Guided by this context, we have embarked on the development of our own platform. This platform's aim is to accommodate data collection from diverse sources, spanning multiple agencies and fields. Additionally, it strives to formulate global decision rules, thereby providing a comprehensive insight into a population and societal statuses **[?]**.

In the preceding stage of our research [?], we successfully pinpointed dominant variables or attributes for threat prediction using the entropy minimum principle. Following this principle, a decision rule was formulated based on these dominant variables. This rule dictated that a specific value exceeding or falling below a determined threshold would signify a particular outcome, such as a threat. In situations where the dominant variable was a time series, monitoring its progression over time allowed real-time tracking of the rule's output, especially in terms of threat prediction [?].

For trend analysis, both from a qualitative and a statistical standpoint, it is noteworthy that when a variable is observed in a chronological sequence over a set duration, its random appearances denote its average time of occurrences midperiod. However, if the variable shows up more frequently towards the period's end, it indicates a distribution with its mean located at or near the observation period's end [6]. One can employ the Laplace test statistic for quantitative assessment of distribution shapes. With the Laplace trend test, distinguishing between a constant event occurrence rate and an escalating one is both straightforward and potent [7].

Figure 1 illustrates the proposed web-based user interface for our platform, designed to enable users to upload their data and gain access to the resulting global decision rules.

III. TREND ANALYSIS

A. Mathematical Formulations—Laplace Statistic

The Laplace statistic, also referred to as the Laplace estimator or Laplacian smoothing, is a powerful statistical tool



Fig. 1. Proposed Webpage Interface.

for estimating probabilities, particularly when available data is sparse or limited. Its applications are widespread, ranging from natural language processing to machine learning. The core principle of the Laplace statistic is the adjustment of probability estimates by the addition of a small constant to both the numerator and denominator. This adjustment helps prevent zero probabilities, thereby accommodating unseen or infrequent events. This smoothing technique ensures non-zero probabilities for all possible events, enhancing the robustness and generalization capability of statistical models.

In trend analysis, the Laplace statistic emerges as an indispensable tool, shedding light on the intricacies of temporal datasets. Specifically, it quantifies the degree of smoothness or regularity within a time series, acting as a bridge between past observations and the likelihood of upcoming data points. When one embarks on the task of analyzing the Laplace statistic values plotted across a time series, it becomes possible to discern significant shifts, fluctuations, or underlying trends that might otherwise go unnoticed. This statistic is not just a mere number; it plays a pivotal role in trend analysis, offering a comprehensive lens through which we can comprehend the dynamics and recurring patterns of a given variable as it evolves over time. By harnessing the insights derived from this statistical measure, we can pinpoint both the subtle, gradual shifts and the more pronounced, abrupt changes within a dataset. This, in turn, equips decision-makers with a robust analytical foundation, ensuring decisions are data-driven, precise, and well-informed.

To exemplify the Laplace statistic's utility, consider a scenario where numerous precursor events precede a final event, such as a permanent fault. The frequency of these events increases as we approach the fault. However, if similar events occur but are not precursors to the fault, their distribution over the observation period would be random. Let's assume an observation period, denoted as T or [0, T], with n events uniformly distributed on [0, T] occurring at t_1, t_2, \ldots, t_n . In the case of a uniform distribution over [0, T], the probability density function (PDF) is given by:

$$f(x) = \frac{1}{T-0}, \qquad 0 \le x \le T.$$
 (1)

By employing the definition of expectation, we can obtain the mean value (μ) or expectation $\mathbb{E}(X)$ for the aforementioned

PDF:

$$\mu = \int_0^T x * f(x) \, dx = \int_0^T \frac{x}{T} \, dx = \frac{1}{2T} \left[x^2 \right]_0^T = \frac{T}{2}.$$
 (2)

Now, the variance $\mathbb{V}(X)$ or s with the definition of $\mathbb{V}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ can be obtained as

$$s = \mathbb{E} \left(X^2 \right) - \left[\mathbb{E} \left(X \right) \right]^2$$
$$= \int_0^T x^2 f(x) \, \mathrm{d}x - \left[\frac{T}{2} \right]^2$$
$$= \frac{1}{T} \int_0^T x^2 \, \mathrm{d}x - \left[\frac{T}{2} \right]^2$$
$$= \frac{T^2}{12}. \tag{3}$$

If we sum up all the occurrence times, the resulting sum is approximately equal to n times the mean of PDF, that is

$$\sum_{i=1}^{n} t_i = n * \mu = \frac{nT}{2}.$$
 (4)

Likewise, the variance of the sum closely approximates n times the variance of the PDF, which is

$$\mathbb{V}\left(\sum_{i=1}^{n} t_i\right) = n * s = \frac{nT^2}{12}.$$
(5)

Hence, the total sum of occurrence times in a uniform distribution can be approximated by a normal distribution with the mean and the variance in Equation (3). That can be expressed as

$$\sum_{i=1}^{n} t_i \approx N\left(\frac{nT}{2}, \frac{nT^2}{12}\right). \tag{6}$$

If we subtract the mean occurace time ${}^{nT/2}$ from the sum of occurrances times $\sum_{i=1}^{n} t_i$, the resulting difference would be zero:

$$\sum_{i=1}^{n} t_i - \frac{nT}{2} \approx 0. \tag{7}$$

Moreover, when we divide the aforementioned difference by the standard deviation (or the square root of the variance) of the occurrence times, the resulting equation transforms into a standard normal distribution such as

$$\frac{\sum_{i=1}^{n} t_i - \frac{nT}{2}}{\sqrt{\frac{nT^2}{12}}} = \frac{\sum_{i=1}^{n} \left(t_i - \frac{T}{2}\right)}{T\sqrt{\frac{n}{12}}} \approx N(0, 1).$$
(8)

The equation mentioned above represents the Laplace test statistic, which has numerous variations. In our work, we will be utilizing a specific variation of this statistic as expressed below

$$U_{L} = \frac{\frac{1}{n} \sum_{i=0}^{n} \left(t_{i} - \frac{T}{2} \right)}{T \sqrt{\frac{1}{12n}}}$$

$$=\frac{\mu - \frac{T}{2}}{T\sqrt{\frac{1}{12n}}}.$$
(9)

where μ is the mean of the event occurrence times.

Equation (9) provides valuable insights into the distribution of occurrence times, enabling the identification of patterns and trends in discrete events. Nevertheless, it is crucial to acknowledge the constraints of U_L when working with smaller sample sizes. Detecting trends may pose challenges, especially when the changes in U_L occur gradually. We delve deeper into this topic in the upcoming section.

B. Interpretation for Trend Analysis

In this section, we delve deeply into the interpretation and profound significance of the statistic U_L as a valuable tool for trend analysis in discrete events. The underlying interpretation of U_L is relatively straightforward and intuitive. By meticulously examining the distribution of occurrence times, we can accurately gauge the trend exhibited by the data. If a majority of occurrence times are found to be after the crucial midpoint T/2, the mean of these occurrence times (which constitutes the first term of the numerator in the U_L equation) will inevitably be larger than T/2. Consequently, the difference between the first and second terms of the numerator, and therefore the value of U_L , tends to be positive and often of significant magnitude. On the other hand, if a significant number of occurrence times precede the midpoint T/2, the mean of the occurrence times will be smaller than T/2. As a result, the difference between the first and second terms of the numerator, and subsequently U_L , tends to be negative and relatively small.

Therefore, U_L emerges as an effective indicator of trends in discrete events. Its sign conveys valuable information about the direction of the trend. A positive value indicates an upward or increasing trend, implying that the occurrence of events is predominantly happening more frequently in the latter half of the time interval T. Conversely, a negative value suggests a downward or decreasing trend, signifying that the events are occurring more frequently in the first half of the time interval.

For a more quantitative approach to trend analysis, we can determine the statistical significance of the observed U_L value by comparing it to the z-value associated with the standard normal distribution, which is determined by the desired confidence level. For instance, at a standard 95% confidence level, the critical z-value is approximately 1.96. If the calculated value of U_L surpasses this threshold, it provides 95% confidence that there exists a significant increasing trend in event occurrence. Conversely, if U_L falls below -1.96, there is 95% confidence in the presence of a substantial decreasing trend. For other confidence levels, such as 90% or 99%, the corresponding threshold values for U_L are approximately 1.645 (-1.645) and 2.576 (-2.576), respectively.

It is worth noting that for situations where the sample size is small, the test for trends using U_L may encounter challenges, especially when the change in U_L is relatively slow. To address this concern, a suggested approach is to employ an adjusted Laplace test statistic that closely approximates a standard

Table I. Pseudocode Output From Sample Data.

T_i	$Sum_i(T_i)$	m_i	$\mu_i \left(Sum_i / m_i \right)$	L_i	s_i	La_i
1	1	1	1	1.73		
5	6	2	3	0.48	2.82	0.51
7	13	3	4.3	0.71	3.05	1.01
14	27	4	6.75	-0.12	5.43	-0.15
15	42	5	8.4	0.46	5.98	0.65
16	58	6	9.66	0.88	6.18	1.38
17	75	7	10.71	1.19	6.29	2.03
18	93	8	11.625	1.42	6.36	2.60
19	112	9	12.44	1.61	6.44	3.11
20	132	10	13.2	1.75	6.52	3.54
22	154	11	14	1.56	6.73	3.25
23	177	12	14.75	1.69	6.92	3.60
24	201	13	15.46	1.80	7.11	3.91
25	226	14	16.14	1.88	7.29	4.17
26	252	15	16.8	1.96	7.47	4.40

Gaussian distribution. This adjusted statistic can be obtained by multiplying the mean (μ) by the original test statistic U_L and dividing the result by the standard deviation (s) of the timed event. The resulting adjusted Laplace test statistic, denoted as U_{AL} , can be expressed as follows:

$$U_{AL} = \frac{U_L \cdot \mu}{s}.$$
 (10)

Both the original U_L and the adjusted U_{AL} share the same critical threshold determined by the corresponding *z*-value, making them equally reliable tools for trend analysis, regardless of the sample size.

C. Coding for Laplace Statistic

In this section, we explore the implementation of the Laplace statistic, denoted as U_L and U_{AL} . Their calculation follows from Equation (9), where T represents the most recent event time and n accounts for the total number of such events until time T or t_n . The derivation of U_{AL} also involves the standard deviation (s). These computations, while simple, need to be performed in real-time. In essence, each new timed event leads to an update in the statistics, ensuring that the calculated Laplace statistic incorporates all past timed events, providing a comprehensive snapshot of the data.

To illustrate, we present a pseudo-code where m keeps a running total of events and Time records the corresponding timestamps. Sample data demonstrating the behavior of the proposed method is showcased in Table I. A corresponding plot, Figure 2, traces the evolution of U_L and U_{AL} over time. Here, U_L reaches 1.96 at the 25th event, confirming an upward trend. Conversely, U_{AL} exhibits a faster response, crossing the 1.96 mark at the 17th event, indicating a quicker identification of the upward trend.

Building on the entropy principle discussed earlier, we draw two critical pieces of information: (1) identification of dominant variables (or attributes) instrumental in predicting a threat, and (2) a decision rule associated with these dominant variables. The rule is framed around a specific threshold, either above or below, which indicates an outcome of a threat or no threat. When dealing with a time-series dominant variable, the Laplace test statistic tracks its occurrences over time, thereby serving as a real-time threat state monitor.

In the context of binary classification, if a rule suggests that a value of 1 for a top variable leads to a Positive Transition,



Fig. 2. Laplace Statistic vs. Adjusted Laplace Statistic Over Time Axis.

we then track this top variable over time. When its value is 1 at any given time, we compute all statistical values for U_L and U_{AL} . If either statistic reaches 1.96, we infer an emerging Positive Transition trend. Alternatively, if the rule indicates a Negative Transition, we track the trend accordingly. Our future plans involve applying this methodology to the top variable and rule derived from the Polity IV data [8] pertaining to regime transitions.

IV. TESTING OF LAPLACE STATISTIC

A. Polity IV Data Set

An evaluation of decision assistance is carried out utilizing the Polity IV dataset, accessible online via the Center for Systematic Peace (CSP) [8]. This institution is committed to the innovation and enhancement of research into political violence methodologies, contextualized within the everchanging framework of the global system. The CSP also provides aid for empirical research and quantitative scrutiny into human interactions and processes of socio-systemic evo-

Algorithm 1: Real-time Laplace statistic calculation					
Input: A series of timed events Time					
Output: Prints the calculated statistics for each event					
1 Initialize tempsum to 0					
2 for each event <i>i</i> in Time do					
3	Add the event's time to <i>tempsum</i>				
4	Calculate the cumulative sum and average time up				
	to this event				
5	Calculate the Laplace statistic $L[i]$ for this event				
6	if this is not the first event then				
7	Calculate the standard deviation $s[i]$ of the time				
	up to this event				
8	if standard deviation is not zero then				
9	Adjust the Laplace statistic $La[i]$ to				
	account for the standard deviation				
10	end				
1	end				
2	Print the statistics for this event				
3 end					



Fig. 3. Testing results for trend analysis, showing (a) the Laplace statistic results for Venezuela, (b) the Laplace statistic results for Bahrain, (c) console output for Venezuela, and (d) console output for Bahrain.

lution. Further, it is responsible for tracking and trend analysis of social-system performance at global, regional, and state levels, including governance aspects. The center persistently observes political behavior across all major nations, focusing on those with a population exceeding 500,000 (167 as of 2014), and reports on matters related to political violence and state failure. The CSP's data is utilized in our research due to their commitment to open and unrestricted dissemination of their research findings [8].

Prepared by researchers at the Integrated Network for Societal Conflict Research (INSCR), the Polity IV dataset, using open-source information, is a valuable service to the broader research community. The data resources of the INSCR are rigorously cross-verified with multiple other resources, ensuring they are accurate, reliable, and complete [8]. The Polity IV project, being a treasured resource in research, has become the go-to tool for comparative and quantitative analysis, particularly useful in monitoring changes in regimes and authority and examining their effects.

The term "polity", used as a unit of analysis, refers to an organized governmental, political, or societal entity or institution, which can simply be understood as distinct patterns within an authority class. The Polity IV project amalgamates crucial information to align with the approach of examining state failures. Researchers have put significant effort into differentiating the attributes of a regime and the effectiveness of state authority, separate from the employment of organized and anti-regime armed forces that may challenge and potentially overthrow said authority.

B. Validation

In the section under discussion, we illuminate the utilization of Polity IV data for unearthing top variables and formulating rules, confronting the inherent challenges posed by the dataset's structure, and probing strategies to deal with them. The Polity IV dataset encompasses the yearly political events across 194 countries, thus allowing multiple entries from different nations within the same year. Although this dataset organization facilitates the derivation of variables and rules, it simultaneously complicates trend analysis.

Our devised solution to this complication involves the formation of individual subsets for each country to ease trend analysis. A major obstacle in this approach is the presence of missing values and specific variable entries within these subsets. Two subsets that nearly fulfill our criteria are those related to Venezuela (with a country code of 101) and Bahrain (with a country code of 692).

We harness a decision-rule algorithm to identify the top variable, known as "Fragment". This variable assesses the existence of autonomous polities within the recognized state borders, capturing the degree to which the state has no effective authority over these entities [8]. The variable spans a continuous scale from 0 (no fragmentation) to 3 (serious fragmentation). It undergoes binary categorization at a threshold value of 0.0559, where values above this threshold are marked as 1, and those below it are marked as 0, signifying a Negative Transition. The starting point, or t_0 , is denoted by the Polity_Begin_Year column, reflecting the inception year of the state.

We initially apply the Laplace statistic to the Venezuelan subset. The years in Venezuela with a "Fragment" value of 0 include 2000, 2001, 2006, 2007, 2008, 2009, and 2013. The trend, as depicted in Figure 3a, reveals the first appearance of the "Fragment" value 0 at the 120^{th} year, succeeded by multiple instances of the same value. The U_L for Negative Transition transcends the 1.96 threshold and exhibits a steep upward trajectory until the final actual positive transition at the 133^{rd} year, which culminates in the rising trend. This interval of growth in U_L corresponds to the real transition phase from the 121^{st} year to the 129^{th} year. The outcome produced by the Laplace statistic code is illustrated in Figure 3c.

In the following scenario, we administer the Laplace statistic to the Bahrain subset. The years in Bahrain that feature a "Fragment" value of 0 are 2001, 2002, 2010, 2011, 2012, 2013, and 2014. The trend, highlighted in Figure 3b, discloses that at the 29^{th} year from the Polity_Begin_Year, the "Fragment" value 0 initially manifests, followed by several more such instances. The U_L for Negative Transition climbs above the 1.96 line, and subsequent to the 39^{th} year, it sharply ascends. This incrementing phase of U_L aligns with the actual negative transition period from the 40^{th} year to the 42^{nd} year. The output generated by the Laplace statistic code, based on this data, is exhibited in Figure 3d.

V. CONCLUSION

For expanded utilization of the Laplace trend test, we designed and tested a web-based collaboration platform to gather data from collaborators and generate decision rules on a global scale. Additionally, we thoroughly discussed the Laplace trend test's essential components, including variable binarization, optimal variable selection using the entropy minimum principle, and the calculation of trend certainty using the maximum entropy principle. In light of evolving global concerns, the web server's testing involved creating a user interface for seamless data upload, database updates, and rule adjustments using Polity IV data. By integrating the Laplace trend test into the decision-assist system with multiple data sources, our web-based platform facilitates collaborative data collection and updated decision rules, enabling global trend analysis and prediction of social unrest.

ACKNOWLEDGMENT

The authors would like to thank the financial support provided by U. S. Army through MSI STEM Research & Development Consortium with the project entitles "Irregular Warfare Decision Assist System for Determining Threat," (Award Number: D01 _W911SR-14-2-0001-0015).

REFERENCES

- Chris Ansell and Alison Gash. Collaborative governance in theory and practice. *Journal of public administration research and theory*, 18(4):543– 571, 2008.
- [2] Shubha Kadambe, Leandro G Barajas, Youngkwan Cho, and Pulak Bandyopadhyay. System and method for signal prediction, March 1 2011. US Patent 7,899,761.
- [3] Yan Luo, Torsten Bergander, and A Ben Hamza. Software reliability growth modelling using aweighted laplace test statistic. In 31st Annual international Computer Software and Applications Conference (COMP-SAC 2007), volume 2, pages 305–312. IEEE, 2007.
- [4] Ryan Steed, Terrance Liu, Zhiwei Steven Wu, and Alessandro Acquisti. Policy impacts of statistical uncertainty and privacy. *Science*, 377(6609):928–931, 2022.
- [5] Weiming Shen, Qi Hao, and Weidong Li. Computer supported collaborative design: Retrospective and perspective. *Computers in Industry*, 59(9):855–862, 2008.
- [6] Charles J Kim. Identification of symptom parameters for failure anticipation by timed-event trend analysis [power distribution systems]. *IEEE power engineering review*, 20(9):48–49, 2000.
- [7] Harold E Ascher and Christian K Hansen. Spurious exponentiality observed when incorrectly fitting a distribution to nonstationary data. *IEEE transactions on reliability*, 47(4):451–459, 1998.
- [8] IV Polity. Polity iv: Regime authority characteristics and transitions datasets. Statistical database. Vienna: Center for Systemic Peace. Available at: www.systemic peace. org/inscrdata. html (accessed July 3, 2017). Development and poverty in SSA, 105, 2017.