Internet Delay Statistics: Measuring Internet Feel Using a Dichotomous Hurst Parameter

Mark DeVirgilio, W. David Pan and Laurie L. Joiner
Department of Electrical and Computer Engineering
University of Alabama in Huntsville
Huntsville, AL, USA

Dongsheng Wu
Department of Mathematical Sciences
University of Alabama in Huntsville
Huntsville, AL, USA

Abstract. As Internet delay times affect the perceptions of users, the ability to differentiate the feel of Internet sites using statistical measures beyond average packet delays may be of commercial value. One such statistical measure is the Hurst parameter of a long-range dependent (LRD) process. A dichotomous Hurst parameter derived from segmented day and night delay data suggests that variations in Internet delay statistics are real and may be caused by diurnal human activity.

Keywords: Internet delays, Hurst parameter, network feel, fingerprinting

I. INTRODUCTION

Internet activities such as Facebook, YouTube, and World of Warcraft (WoW) are enhancing social connectedness and are also asserting claims for more bandwidth and stability from Internet servers and the ramifying backbone of autonomous systems (AS). The perceived early evening slowdown of the Internet is now part of Americana and suggests that a significant portion of the Internet's capability is dedicated to social interactions. This perceived Internet feel may be associated with the statistical properties of Internet delays. The research question pursued in this paper is as follows: What are the deeper time series statistics behind Internet packet delays when examined from a socially relevant time scale measured in minutes and hours?

Previous researchers have left this time scale relatively unexplored perhaps because the famous Bellcore network delay data from the early 1990s used sampling time scales in the milliseconds. Leland et al. used this delay data to demonstrate in a visual sense the self-similarity (SS) of network delays [1]. In addition, the sampled time series data were short enough to limit suspected diurnal variations in their first and second moments, which would have unsettled the mathematical assumptions behind SS and related long-range dependent (LRD) processes. A time series is said to have LRD properties if the shape of its autocorrelation function exhibits a slow or hyperbolic decay. A Hurst parameter can be used to describe this decay, and Hurst parameter can theoretically vary between 0.5 and 1.0. With actual time series data, the Hurst parameter may be out of this range because of the real world interactions that affect the statistical processes being measured. For example, Internet router configurations change dynamically and mean delay times may change due to periodic maintenance events. However, elevated Hurst parameters above 0.75 over many hours may suggest busier or bursty connections to Internet sites and servers.

Two motivational elements for this study are the sparse research on suspected diurnal variations in packet delay statistics and the dramatic changes in the Internet since Leland et al. conducted their study almost two decades ago. We searched the literature and found that in 2005, Iranian researchers identified diurnal and weekly delay time patterns while analyzing Internet data taken among interconnected computers within their country [2]. However, we examined Internet delay data taken in 2009 and 2010 and did not find significant diurnal variations in delay times when communicating with local and international sites and servers. The lack of recently observed diurnal variations suggested that we search deeper into the statistical properties of Internet delay data with a goal of finding a statistical parameter that may quantify the feel of the Internet.

II. LONG RANGE DEPENDENCE FROM THEORY TO PRACTICAL MEASUREMENTS

Many physical phenomena exhibit LRD, and the work of Hurst on describing the historical flooding patterns found in the Nile Basin is a classic. He realized that floods were not adequately characterized by Markovian or memoryless processes and subsequently calculated that the levels of Nile Basin floods had a Hurst parameter well above 0.5, which strongly suggested long-range dependence [3]. The eponymous Hurst parameter, which in the context of our study can vary around 0.5 to 1.0, will be used to describe LRD processes behind Internet delays. We will treat SS and LRD processes as being equivalent when applied to Internet delay statistics, although there are LRD processes which are not SS. We took this liberty as real-world time series data have enough quirks to render the mathematical distinctions between LRD and SS moot.

978-1-4673-1375-9/12/$31.00 ©2012 IEEE

A decade after Hurst's revelation, Mandelbrot found SS patterns when he examined clusters of errors in electronic communication traffic. Mandelbrot's contribution was the introduction of a property called conditional stationarity, which implies that future probabilities are dependent on events or conditions that have already occurred [5]. In 1983, Mills wrote a comprehensive Request for Comments (RFC) 889 on Internet delay measurements. He used a tool called Packet InterNet Groper (PING) that could send Internet Control Message Protocol (ICMP) packets to probe Internet Transmission Control Protocol (TCP) transmission delays. He found that "The incidence of long-delay bursts, or glitches, varied widely during the experiments" [6]. Mills' work was taken as early evidence of LRD phenomena in Internet packet delays, but it was not until the early 1990s that researchers mathematically associated LRD and SS phenomena with Internet delay statistics.

In 1994, Leland et al. [2] analyzed Bellcore network data and used a Hurst parameter to characterize the self-similarity of the observed patterns. Their primary mathematical assumption was to treat their time series delay data, X(k), as wide-sense stationary (WSS). This allowed them to approximate the autocorrelation function (ACF) as the following: $\rho(k) \propto k^{(2H-2)}L(t)$; where k = 0, 1, 2, ...; H is the Hurst parameter and limited such that $0.5 < H < 1$, and L(t) is a slowly varying time function. Next, these researchers created a new aggregated process $X_k^{(m)}$ from the first time series by averaging elements of X(k) in blocks of size m and by ensuring that these blocks were non-overlapping and sequential. They claimed that the new sequence was "second order self-similar" because of the WSS conditions placed on the original time series. This allowed them to specify $\mathrm{var}(X^{(m)}) = \sigma^2 m^{(2H-2)}$ and $\rho_k^{(m)}(k) = \rho(k)$, where σ^2 is the variance of original time series; k = 0, 1, 2, ... ; and m = 1, 2, 3, ... In addition, Leland et al. suggested that if the WSS condition was not rigorously met, then the sequence could still be "asymptotically self-similar" giving $\rho^{(m)}(k) \longrightarrow \rho(k)$, as $m \longrightarrow \infty$ and $k \longrightarrow \infty$. This equation suggests the need to collect as much data as practical to satisfy the limit concerns. The notion of using very long block sizes, which is the frequency domain equivalent of favoring lower frequency spectral components, produced two options for measuring Hurst parameters.

The relation between the autocorrelation function and the spectral power density according to the Wiener-Khinchin-Einstein theorem suggested that a frequency domain technique could yield statistically valid Hurst parameters. Beran, in his classic book *Statistics for Long-Memory Processes*, discussed measuring the Hurst parameter from a periodogram [7]. In one approach, as the number of samples increased, the expected value of the periodogram approached the power spectral density: $lim_{n\to\infty} E[I(\lambda)] = f(\lambda)$, where λ represents the individual frequency steps. Exploiting the shape of the spectral density at the origin as $|\lambda| \to 0$ reveals the value of the Hurst parameter H according to $E[I(\lambda)] \simeq f(\lambda) \sim c_f |\lambda|^{1-2H}$. Taking the logarithm of both sides and accounting for an error term yields $\log I(\lambda_k) \approx \log c_f + (1-2H)\log \lambda_k + \log \xi_k$; where $I(\lambda_k)$ are the Fourier frequencies; k = 1, 2, 3, ...; c_f is a scaling factor; and ξ_k are the error terms. Putting this into the form of a linear regression equation produces some mathematical questions about averaging exponents, but yields $y_k = \beta_0 + \beta_1 x_k + e_k$, where $\beta_0 = \log c_f$, $\beta_1 = (1-2H)$, and $e_k = \log \xi_k$. An estimate of the Hurst parameter is related to the regression coefficient for the slope by $\widehat{H} = \left(\frac{1-\widehat{\beta}_1}{2}\right)$ with Beran's stipulation that this number is valid "in a small neighborhood of zero only." Averaging exponents produces a small systematic error that under-reports the Hurst parameter, and this did not adversely affect our goal of distinguishing Hurst parameters.

We chose a frequency domain approach to measuring Hurst parameters and used periodograms or power spectral density plots. These plots can be readily created using a discrete Fourier transform of time series data. Beran warned that the estimated H parameter could be biased toward higher values when lower frequency components are preferentially used. However, he did not provide advice on the optimal number of low frequency components to reduce bias. Yet, the simplicity and rapidity of the log-log periodogram technique outweighed its shortcomings. We just had to empirically determine the number of lower frequency components to use.

Irrespective of the technique used to measure the Hurst parameter, we had to examine reports of grossly non-stationary data. If such data were common, then the routine measurement of a Hurst parameter would not be fruitful. Mukherjee's 1992 work is frequently cited as being one of the first to recognize significant diurnal variations in Internet delay times [8]. Diurnal variations would imply that Internet delay times are not stationary in their first moment, which violates a condition for LRD. Therefore, we inspected our Internet delay data sets for significant diurnal variations.

III. Recursive Development and Calibration of the Measurement Tool

The first task was generation of time series data that could reflect Hurst parameters between 0.5 and 1.0. In 1980, Granger and Joyeux introduced the use of fractional differencing of an autoregressive integrated moving average (ARIMA) model to simulate long-memory time series [9]. Likewise, we used a Fractional Autoregressive Integrated Moving Average (FARIMA) model with parameters p, d, and q. Parameters p and q represent the powers of the autoregressive and moving average functions and were set to zero. Beran implied that the FARIMA generator would create a time series determined solely by the fractional differencing variable d and defined by $(X_t - X_{t-1})^{-d} = \epsilon_t$, where ϵ_t is a sampled Gaussian random process and d is the fractional differencing exponent [7]. This exponent is related directly to the Hurst parameter by $d = H - \frac{1}{2}$.

There are several ways to use a supplied d to create fractional differencing, and our chosen method used 100 integrations of a gamma function to create the effect. The time series data generator was implemented in software, and a noise-like periodogram of the generator output for $H = 0.51$ or $d = 0.01$ suggested fractional Brownian motion. The periodogram for $H = 1.0$ or $d = 0.5$ was that of a strong LRD process, and the expected hyperbolic decay of the spectral envelope was apparent.

The development of MATLAB code for the Hurst parameter measurement tool was a straightforward task, because the periodogram and linear regression routines are built-in functions and MATLAB efficiently handles large data arrays. When the time series data is partitioned into weekly day and night segments, 30,240 points are contained in each segment, when a sampling time of 10 seconds between each delay measurement is used. Lessons learned from recursively using the tool against generated and real data sets showed that two practical modifications were needed.

The first modification involved differencing the time sequence data from the median delay instead of the mean delay in order to center the distribution on zero delay. A rationale for this decision was that our actual time series data were more symmetrical around their medians as compared to their means. This tool modification had minimal effect when processing FARIMA generated sequences because the mean and median of the symmetrical underlying normal distribution are theoretically the same. Yet, differencing asymmetrical real-world data from the median ensured that most of our Hurst parameter estimates fell between 0.5 and 1.0.

The second modification of the measurement tool involved selecting the lowest frequency components of the periodogram, and these components were then used for the regression estimate of the slope. This was done in accordance with Beran's limitations previously discussed. Using the lower 1/32 of the available frequency components of a FARIMA (0, $H = 0.95$, 0) calibration sequence yielded a Hurst parameter of 0.8365. This 12% inaccuracy at a high Hurst parameter was due to an idiosyncrasy of our FARIMA generation method. However, numerous runs of the FARIMA generator at settings of $H = 0.6$, 0.7, 0.8 and 0.9 proved that our Hurst parameter measurement tool could still differentiate among the resulting calibration sequences.

IV. Collecting Data and Using the Measurement Tool

The selection of Internet sites from which to collect delay data included sites that appeared to have an international distribution. We attempted data collection from over twenty sites, and each attempt was done in isolation to prevent overload or queuing bias of our engineering department's Blackhawk Linux server. Our server's IP address was 146.229.162.184, and it was located at the University of Alabama in Huntsville. Collection periods averaged two weeks apiece, and eleven sites were successfully sampled. A success was indicated when at least a week's worth of uninterrupted data was collected. Table 1 lists the successes and contains the site names, IP addresses, and collection dates. The Google site required the use of its URL, as it rotated through several IP addresses. Thus, the last octet is marked with an asterisk. The Japanese traceroute server had stability problems, so the penultimate server in its path was used.

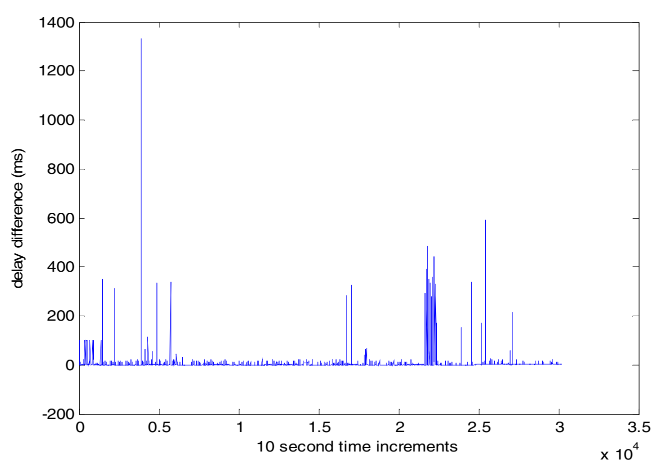
Table 1. Successful data collection from eleven sites with IP addresses current during 2009-2010.

Site	IP Address	Collection date	Comments
Apple	96.6.77.15	17 Aug 10	commercial site
Baidu	220.181.6.175	25 Apr 10	Chinese search company
BBC	212.58.244.142	01 Mar 10	British news
Google	74.125.157.*	13 Feb 10	US search company
Japan server	211.79.42.140	20 Jan 10	penultimate IP was used
Pravda	209.50.249.218	19 Jul 10	Russian news
South Africa	174.143.53.58	07 Jul 10	tourism site during World Cup
Tata Motors	66.132.222.54	28 May 10	Indian car company
Thailand	122.155.17.64	03 Aug 10	tourism site
US Congress	140.147.249.9	10 May 10	Library of Congress access
US time server	132.163.4.22	15 Apr 10	NIST

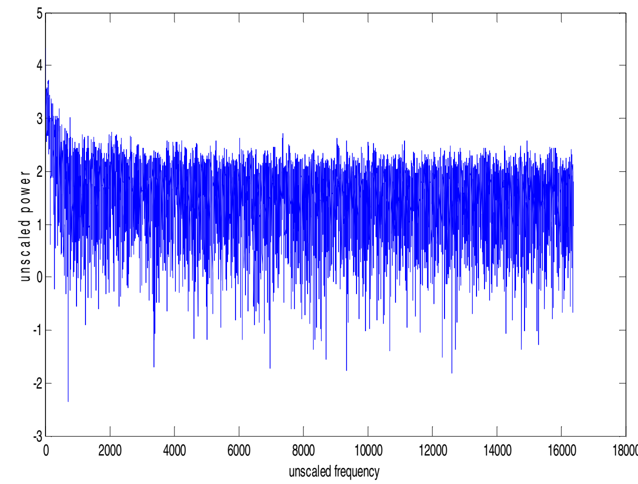
Data collection was straightforward using a Linux shell script. As a TCP connection was desired, the traceroute command was forced to use this option. The data collection sampling period was once every ten seconds. We recorded delay data and time of day information to a hard drive for subsequent retrieval and analysis. Several traceroute option parameters were used to limit the number of packets sent, the number of repeats at the same hop count, and the amount of delay data from each intermediate node. These steps ensured that the command would finish in well less than ten seconds.

Processing the raw data measurements, which averaged close to 100,000 samples, involved truncating the sequence to a local starting time of 2100 hours, US Central Time. We collected raw data for least a week and often up to two weeks. The extra data was not normally used by the measurement tool, but came in handy for experimentation on changing the day and night starting times. Data for the eleven Internet sites successfully sampled were processed to remove extraneous characters and imported into MATLAB. For each site, the Hurst parameter measurement tool was designed to partition a week's worth of data into a 0900-2100 hrs Central Time day sequence and a 2100-0900 hrs Central Time night sequence. Daylight savings time was not compensated for, as we felt that urban human activity would follow the shift. Data falling at the time boundaries were added to a particular sequence as to ensure day and night sequences each contained 30,240 elements. Thus, seven days of 12-hour samples measured every 10 seconds formed a sequence of 30,240 elements. The two sequences per site aptly were named night and day.

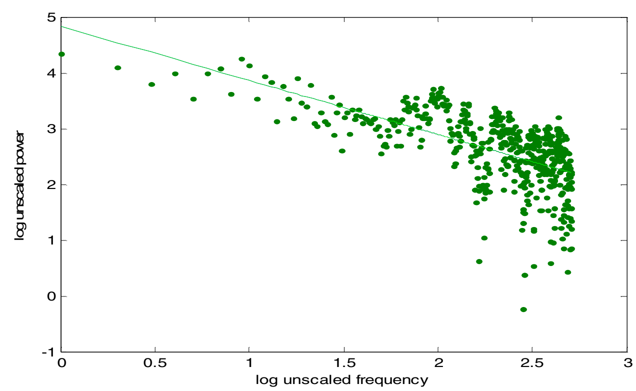
The three graphic outputs of the Hurst parameter measurement tool are plots of the delay time differences from the median value, a log-log periodogram of these differences, and a least squares regression plot of the lower 1/32 of the log-log periodogram. Fig. 1(a) shows a week's worth of Baidu data sampled from 2100-0900 hours US Central Time and differenced from its median. No evidence of cyclic diurnal variations was apparent. In addition, the ten other sites used in our study did not support the notion of significant diurnal variations in delay times. Fig. 1(b) shows a log-log periodogram for the Baidu night data as referenced to Central Time. As only the decay shape of the spectrum near the origin is used by the measurement tool for the Hurst parameter calculation, the information in the higher frequency components is discarded. Perhaps the most important graphical output of the tool is shown in Fig. 1(c). This log-log periodogram is from the lower 1/32 of the frequency points, and the least squares regression line runs down and outwards from near zero frequency. The slope of the regression line is an estimate of the Hurst parameter. Approximately 500 low frequency points are used in the regression. The magnitude and frequency parameters were left unscaled because there was no use for the estimate of the constant term $\hat{\beta}_0$. $\hat{\beta}_1$ was -0.9685 and $\hat{H}$ was 0.9842. Thus, the Baidu delay data collected at US Central time night exhibited strong long-range dependence.



(a)



(b)



(c)

Figure 1. Hurst parameter measurement tool standard output graphs: (a) Baidu delay difference from the median measured at night Central Time, (b) Baidu delay difference log-log periodogram, and (c) Baidu log-log regression line using the lower 1/32 of the available spectral components.

Data from eleven sites were processed by the tool, and the results are shown in Table 2. Some of the median delay time results for distant sites were as expected; Baidu, the Japanese traceroute server, and the Thai tourism sites showed median delay times above 200 milliseconds. The geographically closer BBC site had a median delay time around 100 milliseconds. All the sites located in the United States had median delay times below 50 milliseconds. There were three apparent anomalies in delay times for sites that were supposed to be located overseas. The Pravda, South African tourism, and Tata Motors sites had delay times between 20 and 35 milliseconds. A subsequent "whois" investigation of their IP addresses, which were resolved by a domain name server using their URLs, showed that each of these addresses belong to servers located in the United States.

Table 2. Hurst parameter measurements.

Site Name	Collection at Central Time	Median delay (ms)	Hurst parameter estimate	95% CI	Significant difference
Apple	day	28.9020	1.0770	1.0214 1.1326	yes
	night	28.3550	0.5777	0.5213 0.6342	
Baidu	day	227.537	0.7845	0.7262 0.8429	yes
	night	227.272	0.9842	0.9304 1.0381	
BBC	day	97.354	0.7437	0.6891 0.7983	yes
	night	97.135	0.4884	0.4453 0.5314	
Google	day	7.163	0.5809	0.5438 0.6181	yes
	night	7.028	0.7766	0.7156 0.8376	
Japan server	day	203.531	0.6956	0.6607 0.7305	yes
	night	203.878	1.0776	1.0237 1.1315	
Pravda	day	23.663	0.6260	0.5616 0.6905	no
	night	23.459	0.6473	0.6234 0.6712	
South Africa	day	34.004	0.9797	0.9217 1.0377	yes
	night	33.739	0.6932	0.6524 0.7340	
Tata Motors	day	20.306	0.6022	0.5448 0.6597	no
	night	20.238	0.5754	0.5180 0.6328	
Thailand	day	284.706	0.6901	0.6293 0.7510	yes
	night	284.767	0.9642	0.9064 1.0220	
US Congress	day	23.098	0.6403	0.5821 0.6984	yes
	night	22.952	1.0754	1.0200 1.1308	
US time server	day	48.650	0.5582	0.4968 0.6197	no
	night	48.399	0.5770	0.5284 0.6255	

We noted two interesting delay time features. One site with surprising delay times below 10 milliseconds was the Google site. As mentioned earlier, the Google site traceroute command was run against the "www.google.com" URL because specific IP addresses were not always active during the collection period. Another interesting observation was that all sites located in the United States showed slightly longer median delay times during the Central Time day sampling period when compared to their night sampling period. This suggested that the Internet was slightly slower in the day, and the suspected cause could be changing human activity on the Internet. However, we did not pursue this claim, as there was no obvious diurnal pattern in the raw delay time plots.

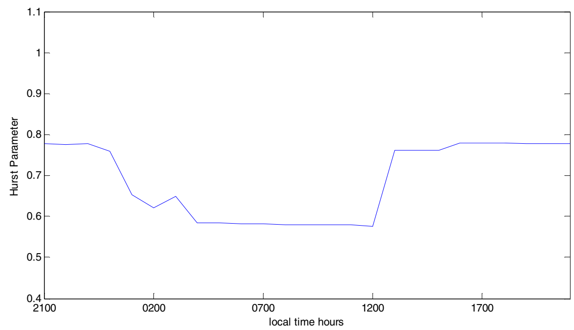
The main feature of the Table 2 is the confirmation of dichotomous Hurst parameters for eight of the eleven sites tested. We defined a dichotomous Hurst parameter as one that exhibits a significant difference between its day and night values. Some sites exhibited estimated Hurst parameters slightly above 1.0 or below 0.5, which mathematically makes their time series not SS. However, the results were within 10% of the targeted range and they were useful for looking at differences in Hurst parameters. Three sites did not exhibit dichotomous Hurst parameters during their one-week observation periods, and they will be discussed first.

Based on the overlap of their 95% confidence intervals, the Pravda, Tata Motors, and US timeserver sites did not have significantly different day and night Hurst parameters. All three did have Hurst parameters that tended toward the low side of the scale and were between 0.56 and 0.65. This range suggested weak long-range dependence, and their Hurst parameters did not change significantly between day and night. One possible explanation is that these sites have low popularities, which may have masked social use fluctuations. More research is needed to confirm the idea that less popular sites have lower Hurst parameters.

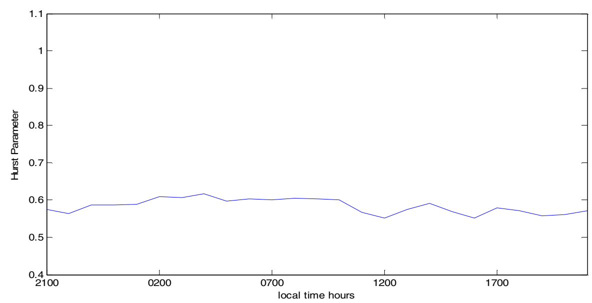
All four overseas sites exhibited dichotomous Hurst parameters. Two of these sites, a Japanese internal server and a Thai tourism site, were presumed to be low on the social interest order. The Baidu and BBC sites presumably had much greater use due to their popularities. Thus, social interest may not be the only cause of the observed dichotomous Hurst parameters. In these four overseas cases, changes caused by the autonomous systems connecting these sites to the United States may also be important. One conclusion from the overseas data is that physical distance and long delay times appear interrelated and may contribute to dichotomous Hurst parameters when viewed from the United States. More research is needed to examine the delay times of these sites from within their local time zones.

Four popular sites within the United States exhibited dichotomous Hurst parameters. The Apple, Google, and United States congressional sites are known for their heavy social usage. The South African tourism site was hosted on a server in the United States. This was done presumably for faster access to travel information and match results for the World Cup soccer games being played in South Africa. Thus, our data suggested that there is a relationship between popular sites and sites exhibiting dichotomous Hurst parameters. Extending this logic, a time varying Hurst parameter may indicate a diurnal change in social activity for that site.

In order to examine some details behind dichotomous Hurst parameters and to allay concerns that an arbitrary Central Time day and night demarcation time choice created artifacts, two runs on a modified measurement tool were employed. The measurement tool was modified to use 24 different starting times, one hour apart, on eight consecutive days worth of data; the results are shown in Fig. 2.



(a)



(b)

Figure 2. Hurst parameter measurements at different starting times for 12 hour segments: (a) Google Hurst parameters, and (b) Tata Motors Hurst parameters.

The modified tool was applied to the Google data, and the 0900 hours day demarcation point was within the 0200 to 1200 hours window where the Hurst parameters were lower. The 2100 hours demarcation point was on a broad plateau where the Hurst parameters were higher. This is shown in Fig. 2(a). This observation suggests that different demarcation times and data segments less than 12 hours may improve the sensitivity of the Hurst parameter measurement tool. Next, we chose the data from the Tata motors site, where a dichotomous Hurst parameter was absent. The results are shown in Fig. 2(b). The measured Hurst parameters stayed in a range between 0.55 and 0.62. These findings suggested that diurnal site activity can be described by a dichotomous Hurst parameter.

V. CONCLUSION AND POSSIBLE APPLICATION

None of the eleven Internet sites exhibited rhythmic diurnal changes in packet delay times as reported by a few researchers. Eight of the eleven Internet sites tested exhibited significantly different day and night Hurst parameters at the 95% confidence level. All three overseas sites and five popular sites in the United States had dichotomous Hurst parameters that changed according to day and night collection periods based off Central Time. These findings suggested that there is merit in segmenting delay time data into periods relevant to diurnal human activity. In addition, we found that the arbitrary 0900 and 2100 hours Central Time demarcation points for day and night periods could be varied to get time varying Hurst parameter estimates. Demarcation points based on local time at the site may better show the social patterns of users.

Another purpose for examining dichotomous Hurst parameters was to find statistical measures beyond average delay times and associated variances that could be used to characterize or fingerprint activity at Internet sites. Median delay times often cluster together depending on geographic regions and cannot be used alone to differentiate some sites. For example, the United States congressional site and the Pravda site had Central time night delays of 23.5 and 23.0 milliseconds, respectively; however, their respective Hurst parameter estimates were significantly different at the 95% confidence level. A database of Hurst parameters kept by each user and updated daily could be used to distinguish and profile his or her favorite sites.

A software application that we call *Internet radar* could enhance a user's perception of the Internet by first displaying delay times for favorite sites and then by displaying a time varying Hurst parameter to differentiate sites and to indicate site activity. Although consistently low delay times are paramount for "twitch" games played on the Internet, the feel of an Internet site may also be determined by the burstiness of its TCP connection. The Hurst parameter can measure this burstiness. Thus, an Internet radar application could learn about a user's site preferences and delay tolerances and then could create an itinerary of optimal times to access each site.

REFERENCES

[1] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. on Networking*, vol. 2, no. 1, pp. 1-15, Feb. 1994.

[2] E. Kamrani and M. Mehraban, "Modeling internet delay dynamics using system identification," IEEE Int. Conf. on Industrial Technology, Mumbai, India, December 15-17, 2006, pp. 430-438.

[3] H. E. Hurst, *The Nile*. London: Whitefriars Press, 1952.

[4] T. Karagiannis, M. Molle, and M. Faloutsos, "Long range dependence: Ten years of Internet traffic modeling," *IEEE Internet Computing*, vol 8, no. 5, pp. 2-9, 2004.

[5] B. Mandelbrot, "Self-similar error clusters in communication systems and the concept of conditional stationarity," *IEEE Trans. Commun. Technol.*, vol. 13, pp. 71-90, 1965.

[6] D. L. Mills, (1983). *Internet delay experiments (RFC 889)*. Available: <http://www.ietf.org/rfc/rfc889.txt>.

[7] J. Beran, *Statistics for long-memory process*. New York, NY: Chapman and Hall, 1998.

[8] A. Mukherjee, "On the dynamics and significance of low frequency components of Internet load," Dept. Comp. and Info. Sci., University of Pennsylvania, Philadelphia, PA, Rep. MS-CIS-92-83/DSL-12, Dec. 1992.

[9] C. W. J. Granger and R. Joyeux, "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis,* vol. 1, pp. 15-29, 1980.

[10] J-M. Bardet, G. Lang, G. Oppenheim, A. Philippe, and M. S. Taqqu, "Generators of long-range dependent processes: A survey," In *Theory and Applications of Long-range Dependence*. Boston, MA: Birkhäuser, 2003, pp. 579-623.