When the Internet Sleeps: Correlating Diurnal Networks With External Factors(extended)

USC/ISI Technical Report ISI-TR-2014-691, May 2014 *

Lin Quan[1,2] John Heidemann[1] Yuri Pradkin[1]
1: USC/Information Sciences Institute 2: Bank of China
{linquan, johnh, yuri}@isi.edu

ABSTRACT

As the Internet matures, policy questions loom larger in its operation. When should an ISP, city, or government invest in infrastructure? How do their policies affect use? In this work, we develop a new approach to evaluate how policies, economic conditions and technology correlates with Internet use around the world. First, we develop an adaptive and accurate approach to estimate *block availability*, the fraction of active IP addresses in each /24 block over short timescales (every 11 minutes). Our estimator provides a new lens to interpret data taken from existing long-term outage measurements, this requiring no no additional traffic. (If new collection was required, it would be lightweight, since on average, outage detection requires less than 20 probes per hour per /24 block; less than 1% of background radiation.) Second, we show that spectral analysis of this measure can identify *diurnal usage*: blocks where addresses are regularly used during part of the day and idle in other times. Finally, we analyze data for the entire responsive Internet (3.7M /24 blocks) over 35 days. These global observations show *when* and *where* the Internet sleeps—networks are mostly always-on in the US and Western Europe, and diurnal in much of Asia, South America, and Eastern Europe. ANOVA testing shows that diurnal networks correlate negatively with country GDP and electrical consumption, quantifying that national policies and economics relate to networks.

1. INTRODUCTION

The Internet is a large and complex combination of technology, policy, and human behavior. Separating these effects is difficult, yet businesses, governments, and individuals must make decisions every day, even with incomplete and contradictory information.

One metric of Internet maturity is network availability and the shift from on-demand on *always-on* network availability. Increasing use of broadband networks brings *always-on* networks. In fact, always-on is part of the U.S. Federal Communication Commission's *definition* of broadband [8]. Always-on networks have practical differences on how networks are used—the ease of not having to turn on a computer and connect is one enabler of the shift of home telephony and entertainment to IP. A similar shift is seen in the deployment of high-speed mobile phones, where always-on 3G and 4G networks are fast enough to take on data-intensive tasks like social media and streaming video. For many, an always-on phones is their primary interface to social media

Government policy makers and ISPs evaluate how the choices they make affect network usage. Until recently, tools to study the results of policy choices have been limited, with telephone-based surveys reaching a few thousand of randomly chosen people (for example, [36]), and in-home routers such as SamKnows [32] and BISmark [31] providing visibility to a few hundred, statically-chosen volunteers. While important, these studies consider a very sparse sample of individuals. A broader perspective would better quantify how policies and technologies affect Internet usage, and it is necessary to compare global differences and understand natural experiments that arise when different countries adopt different policies. For example, we know that some countries use network blocks on a diurnal basis because of economic constraints (saving power bills), or because of different access technologies. To move beyond these anecdotal we would like to systematically explore what factors contribute different network usage.

To reach this goal, the first contribution of our work is to develop a new method to estimate *block availability*, the fraction of addresses in each /24 block that are in use and will reply to an ICMP probe. We provide a new analytic tool to interpret existing long-term measurement studies of network outages [29]. We take this approach to both to leverage existing data, and because

*This research is partially sponsored by the Department of Homeland Security (DHS) Science and Technology Directorate, HSARPA, Cyber Security Division, BAA 11-01-RIKA and Air Force Research Laboratory, Information Directorate under agreement number FA8750-12-2-0344, and via SPAWAR Systems Center Pacific under Contract No. N66001-13-C-3001. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views contained herein are those of the authors and do not necessarily represent those of DHS or the U.S. Government.

these outage measurement are already shown to impose minimal cost (in probing traffic) on the subject networks. On average, outage detection requires less than 20 probes per hour, per /24, a value that is less than 1% of the amount of Internet background radiation observed in 2010. Even if our analysis required new data collection, its impact would be minimal.

Second, we show how this measurement can be interpreted to detect *diurnal* blocks. We use spectral analysis to identify blocks whose usage shows large daily changes. We also show that analysis of the *phase* of these changes relates to the longitude of blocks. We develop these two methods (§ 2) and validate them with controlled experiments (§ 3).

Our final contribution is to use these methods to study both diurnal behavior of 3.7M responsive blocks (each a /24 IPv4 prefix that replies to active probes) over more than one month (§ 4). This analysis provides the first *long-term*, *wide-scale* evaluation these phenomena over the global Internet. Using ANOVA testing, we correlate these direct observations with economic and technical factors to *examine how policy and national constraints* relate to Internet use, showing correlations by country and link technologies, as well as by economic factors (GDP and electricity consumption) (§ 5).

All data from this paper is currently available at no cost to other researchers.

2. METHODOLOGY

Our analysis begins with low-rate estimation of block-level availability (§ 2.1), allowing detection of diurnal blocks (§ 2.2). We then relate these measures to geolocation, organization, and link technologies (§ 2.3).

2.1 Dynamic Tracking of Block Availability

Our first new method estimates block *availability* [9], the probability that addresses in the block will respond to ICMP probes. While prior work defined availability A over the whole block, here we estimate three variants $\hat{A}_E$: short- and long-term estimates of the availability of the ever-active IP addresses ($\hat{A}_s$ and $\hat{A}_l$), and an *operational* ($\hat{A}_o$) value that is consistently designed to avoid overestimating A.

2.1.1 Problem Statement

Our goal is to derive an estimate $\hat{A}$ from analysis of probe data already taken as part of outage detection in Trinocular. Trinocular already minimizes probes, with each observer sending on average to less than 1% of background traffic as of 2010. We wish to estimate A with *no* additional probes beyond what outage detection already requires.

We define the ground truth A as the fraction of addresses that respond when all our probed. While easy to define, a full measurement ground truth is too expensive for operation.

There are several challenges that make A estimation difficult to add to outage detection. First, sampling in outage detection is *biased*. The goal of outage detection is to make the minimum number of probes to reduce stress on the target, so a few or even one positive response is usually sufficient to terminate probing. Probes are thus biased in favor of positive responses.

Second, outage detection itself depends on $\hat{A}$; we require an operational $\hat{A}_o$ estimate that is be *too high*. When $\hat{A}_o > A$, a few negative probes are a stronger indication that the block is down and Trinocular will produce false outages.

Two additional challenges are that observations are quantized, and that our initial A estimate may be quite inaccurate. Because we take from 1 to 15 observations per round, precision of a new A value is at most ± 0.07 and could be as coarse as 0.5 or 1. Short-term estimates therefore show significant jitter. Our initial estimates for A are based on historical data over several years. They may be off significantly if block usage has changed.

2.1.2 Approach

Each of our estimates use exponentially weighted moving average (EWMA) to smooth raw observations.

For each block, in each round of adaptive probing, we observe p positive responses and t total responses.

We first model short-term EWMA of positive and negative responses to quickly adapt to real values.

$$\hat{p}_s = \alpha_s p + (1 - \alpha_s)\hat{p}$$

$$\hat{t}_s = \alpha_s t + (1 - \alpha_s)\hat{t}$$

We use a gain of $\alpha_s = 0.1$ to quickly adapt to changes, and calculate the short-term availability $\hat{A}_s$:

$$\hat{A}_s = \hat{p}_s / \hat{t}_s$$

It is important that we estimate p and t separately. Because A is their ratio, applying EWMA directly to observations of A skews the data, for the same reason that one must use geometric mean to estimate normalized results [16]. (Our experiments in dataset A_{12w} use a variation on this approach, where short-term availability $\hat{A}'_s$ is tracked directly with EWMA, rather than tracking $\hat{p}_s$ and $\hat{t}_s$ separately. This method consistently over-estimates $\hat{A}$. While we accommodate for this variation in $\hat{A}_o$ and verify that it does not affect periodicity in $\hat{A}_s$, our current approach provides tighter results.)

An example of $\hat{A}_s$ to is shown as a cyan line in Figures 1 through 3. Compared to the true A taken from complete survey data (the black line), we see that $\hat{A}_s$ is quite noisy (see Figure 1). However, this rapidly changing estimate is useful to understand diurnal block changes (§ 2.2).

We model long-term estimates of availability ($\hat{A}_l$) with the same process using a gain $\alpha_l = 0.01$.

Our goal in a conservative $\hat{A}_o$ is that it should not exceed the true value. To reflect this goal, we evaluate the absolute deviation of each sample from the estimate and then intentionally underestimate the operational value by a half of this amount. We model a deviation term d_l, and use it to compute a more conservative $\hat{A}_o$ (note that we limit $\hat{A}_o$ to a minimum of 0.1):

$$\hat{d}_l = \alpha_l |\hat{A}_l - p/t| + (1 - \alpha_l)\hat{d}$$

$$\hat{A}_o = \max(\hat{A}_l - \frac{1}{2}\hat{d}_l, 0.1)$$

2.2 Detecting Diurnal Blocks

Block availability relates to the number of *active IPs* in the block. Blocks that change diurnally will show regular daily changes to A as computers come and go.

To look for periodic changes we estimate A as a timeseries of $\hat{A}_s$ observations taken each 11 minutes. We then take the Fast-Fourier Transform of this timeseries to determine periodicity and phase.

Spectral analysis of timeseries data wants regular input aligned with the sampling interval (in our case 11 minutes). However, because our probing software is not perfectly aligned with 11 minute rounds, sometimes we see missing or duplicate observations in a round. Like previous work [28], we correct this minor drift in analysis, extrapolating single missing estimates, and trusting most recent observation when duplicates occur.

We use Discrete Fast-Fourier transform to analyze availability. Given a timeseries $A = a_m$, $m \in [0, n-1]$ of n input samples, we compute its $k \in [0, n-1]$ frequency-domain components as:

$$\alpha_k = \sum_{m=0}^{n-1} a_m e^{-2\pi i m k/n}$$

The magnitude of α_k shows the strength kth frequency component (sub-signal) of a_m. With our 11-minute ($R = 660$ s) sampling period, α_k shows the frequency at $k/(Rn)$ Hz. To identify diurnal patterns for an experiment lasting N_d days we look in the $k = N_d$ component of α_k. In our work we discard the DC component (α_0) of the frequency.

Detecting diurnal blocks: Diurnal activity in a block appears as periodicity in α_k corresponding to 24 hour intervals (and multiples of this frequency). For a *strictly diurnal* block, the highest amplitude frequency occurs at 1 cycle per day (when $k = N_d$), *and* its strength is at least twice the next strongest frequency.

We also consider *relaxed diurnal* blocks, where the highest amplitude frequency is at 1 cycle per day, without comparison to the second strongest frequency.

Block phase: In addition to the the amplitude of the FFT (the "strength" of a given periodicity), spectral analysis also provides the phase (when the period occurs relative to the start of measurements). Phase is shown in the angle of the complex-valued FFT coefficients.

We examine phase in the 1-cycle-per-day component, and consider it valid only for strong or relaxed diurnal blocks. (For non-diurnal blocks, phase of this component is effectively random.)

Phases range from $[-\pi, \pi]$, with the exact value depending on when a measurement begins. We consider relative phases and compare them to physical location (§ 5.2); future work may tie phase to time-of-day.

Data appropriateness: FFT over data too short or non-stationary can distort analysis of periodic behavior. Our datasets are sufficiently long (two or more weeks) to capture many diurnal periods.

We verified our data is roughly stationary for our measurement periods by doing a linear fit of A over the observation and confirming slopes are near-zero. We confirmed with all 29,001 blocks in Survey S_{51w} that about 80.3% of these blocks are stationary, with a slope equivalent to less than 1 address changes per day (Appendix A).

2.3 Other Network Factors: Geolocation, Organizations, and Link Types

We analyze network behavior in /24 blocks, but to put these in context, we relate diurnal behavior to other factors: physical location, logical location (the organization operating the address), and last-mile link technology.

2.3.1 Geolocation

Different countries and regions have different policies and economics for networking. To understand these effects we use IP geolocation to get city-level physical locations. While there have been many approaches to IP geolocation (see for example [12, 26]), we use MaxMind's city-level database [23] as a free and widely used source. Although this source is not the most accurate (claimed accuracy is 40 km), it is sufficient to demonstrate at least country-level correlation.

We map each /24 block to location. Although sometimes blocks span multiple locations, such instances are fairly rare. Our analysis of existing per-IP data [10] shows about 93% of /24 blocks have homogeneous locations, and our MaxMind dataset specifies location with at most /24 precision.

2.3.2 Organizations

To understand how the policies of different organizations affect how they use IP addresses, we map IP addresses to Autonomous System (AS) numbers, and ASes to organizations.

We use Team Cymru's IP to AS number mapping [33]. We map each /24 to an AS based on its .0 address. Dividing ASes across /24s is rare, with only 0.005% of

blocks in in Team Cymru's data showing differing .0 and .128 ASes. Their data provides AS numbers and names for 99.41% of /24 blocks.

We map ASes to organizations using prior work that uses WHOIS and string-based clustering [4]. For a given organization or ISP P (for example, Time Warner Cable), we first use keyword matching (ex. "Time Warner") to find relevant clusters, then find all ASes within same cluster(s). Finally, for all ASes within P, we join with IP/AS mapping and find all relevant IP blocks for P.

The above method assumes good accuracy of IP/AS and AS/organization mappings. Future work may consider how accuracy of these mappings affects our correlations, or to compare the behavior of different ASes within the same organization.

2.3.3 Link Technology

We define the *link type* as the technology connecting the final hop to the destination. Identification of link types for blocks is difficult because such information is not readily available, and because in principle different parts of a /24 block could be connected differently (for example, a router at .1 and the rest of the addresses with dial-up). Previous work has shown one can often infer link types from reverse domain names [4,15,25,30]. We build on the work of Thunderping [30], but cannot reuse their data directly because their classification was done manually and only for the United States; we wish to study millions of international blocks. Other data is also sparse or old, so we identify block-level link types as follows.

First, we look up the reverse domain name of each address in each analyzable block. We then use string matching against pre-defined keywords to non-exclusively identify features (corresponding to link types, see § 5.5) for each address. For example, a reverse name of `dhcp-dialup-001.example.com` is marked as both DHCP and dial-up. Each block is therefore defined by a vector of 256 addresses, each with zero or more features. We suppress minor features in each block by filtering out features that are less than $\frac{1}{15}$th of the most frequent feature. Finally, we label the block with all remaining features that have non-zero counts.

We consider 16 keywords (sta, dyn, srv, rtr*, gw*, dhcp, ppp, dsl, dial, cable, ded*, res, client*, sql*, wireless*, wifi*). Of these, we discard the seven marked with asterisks because they are dominant in less than 1000 blocks. These keywords are all English based; future work should explore non-English networking keywords.

For our 3.7M blocks in dataset A_{12w}, and find that 46.3% of these blocks have some feature (4× more than Thunderping), and 11.4% have multiple features.

2.4 Factorial Analysis with ANOVA

To quantify the correlation of policies, places, and technologies on Internet usage, we need to weigh the factors that relate to network outcomes such as diurnal address usage and network reliability. While anecdotal and personal evidence suggests that individuals in some countries turn off computers at night to save power, we wish to measure these effects systematically.

To discover correlations between a range of possible factors and network outcomes we use analysis of variance or ANOVA, a form of statistical hypothesis testing. It tests the probability (p-value) of a factor or factors (separately or in combination) being correlated with an observation. A high p-value indicates that the factor has little correlation with outcome, while a low p-value indicates that correlation of the factor and outcome are unlikely to have been by chance alone. In practice, p-values less than a threshold (0.05 or 0.01) are typically considered statistically significant, because a low p-value (with high probability) rejects the null-hypothesis: that the input variable(s) and output variable are unrelated.

A full discussion of ANOVA is found in many textbooks (for example, [24]). We use the open source ANOVA package in R for our analysis.

2.5 Datasets

Our work draws on two active measurements: IPv4 surveys and Trinocular outage measurements.

To validate our results with complete information we re-evaluate *Internet IPv4 surveys* [9]. These datasets have probes to every address in about 2% of IPv4 /24 blocks, taken every 11 minutes for 2 weeks. With complete data for all IPs in the block, surveys provide ground truth. We primarily use Survey S_{51w}, starting 2012-11-16. To detect trends over the last three years (Figure 11), we analyze historical surveys taken since 2010 (Survey S_{31w}).

Our new estimators to derive short-timescale block availability and diurnalness are based on data from Internet outage observation. We use outage data collected from Trinocular [29]. Our analysis here uses the A_{12w} dataset, taken from Los Angeles, California, USA [35] and covering 3.7 M Internet edge network blocks over 35 days, starting at 2013-04-24 17:18 UTC. We observe similar results from two other sites in Fort Collins, Colorado, USA, and Keio, Japan.

These datasets are publicly available [34]. We add new public datasets for link technology (§ 2.3.3) and our new availability and diurnal analysis.

3. VALIDATION

To validate our new measurement approaches we re-evaluate Internet surveys (summarized above § 2.5). We add controlled simulations of diurnal blocks to systematically investigate our sensitivity to variation in diurnal behavior. We also evaluate diurnal detection against

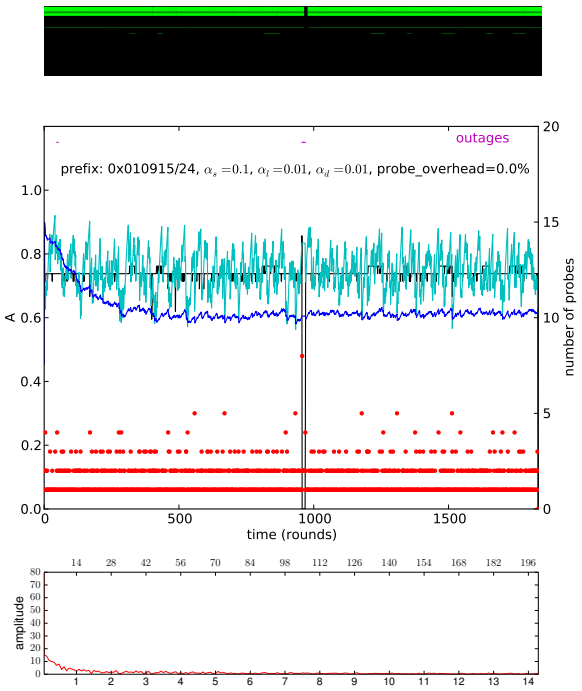


Figure 1: Block 1.9.21/24 from S_{51w}, a sparse but high availability block (42 every responding addresses, $A = 0.735$). Top: raw survey data: green is response, black is no response. Middle: compares true A (black), short-term estimate $\hat{A}_s$ (cyan), operational estimate $\hat{A}_o$ (dark blue), and report the number probes per round (red dots and right axis). An outage occurs at round 957 (the magenta line). Bottom: amplitude of FFT ($N_d = 14$), with no strong diurnal signals.

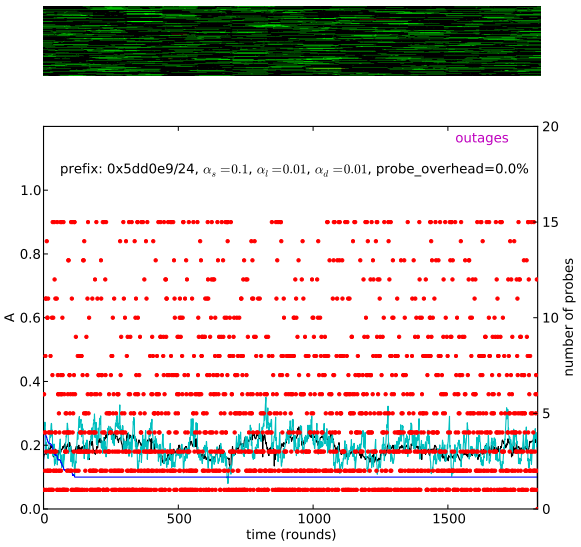


Figure 2: Block 93.208.233/24 from from S_{51w},$|E(b)| = 245$, $A = 0.191$. Dense but low availability block. (Block is non-diurnal, so FFT is omitted.)

blocks at USC that are or perhaps should be diurnal.

3.1 Tracking Block Availability

We first validate that we accurately estimate availability (A), since detection of diurnal blocks uses this estimate. We show our measure with examples and evaluation of all 29k blocks in S_{51w}.

3.1.1 Sample Blocks

To illustrate estimation of $\hat{A}$, we first compare $\hat{A}$ to true A for several representative sample blocks drawn from Survey S_{51w} (Figures 1 to 3). The survey provides complete data for all 256 addresses in the block for two weeks. We use the Trinocular algorithm to downsample this data, considering from 1 to 15 samples per round (shown as red dots on the graphs). We apply Trinocular's adaptive sampling to consider from 1 to 15 samples per 11 minute round. We then compare short-term and operational estimates ($\hat{A}_s$ and $\hat{A}_o$, the cyan and dark blue lines) and compare them to the true value (A, the black line) measured from complete data. Full data is shown in the image above each graph, with light green dots showing responses and black showing non-responses. We have examined hundreds of blocks; the examples we show here represent different conditions.

Figure 1 shows a block that is sparse, with only a few responding addresses (42 active over time), but relatively high availability ($A = 0.735$). The middle graph shows that the short-term estimate is noisy around the true value (the jagged cyan $\hat{A}_s$ line, compared to the steady black A line). We also see the operational value never overestimates once it leaves its inaccurate initial value ($\hat{A}_o < A$, the dark blue line is lower than the black line). Finally, we detect an outage at round 957.

Figure 2 shows a dense block where nearly all addresses in use ($|E(b)| = 245$), but availability is low ($A = 0.191$). Again, the short-term estimate tracks truth reasonably well, but with some noise, and the operational does not exceed the true value. The low availability of this block means several probes are often used per round (mean 5.08 probes/round). We conclude that our approach works on low-availability blocks as well.

Finally, Figure 3 shows a diurnal block. We count 14 "bumps" of A corresponding with the 14-days survey. The short-term estimate lags the truth slightly (compare cyan $\hat{A}_s$ to black A) , and the operational value ($\hat{A}_o$) is generally below the worst-case true A.

These samples show that our algorithms track A accurately, as we observe in many other blocks.

3.1.2 Comparison in Full Survey Blocks

To generalize from these examples we next systematically compare *all* 29k /24 blocks in Survey S_{51w}. We apply our availability tracking algorithm with sparse sampling and estimation to each block and compare to full data as ground truth.

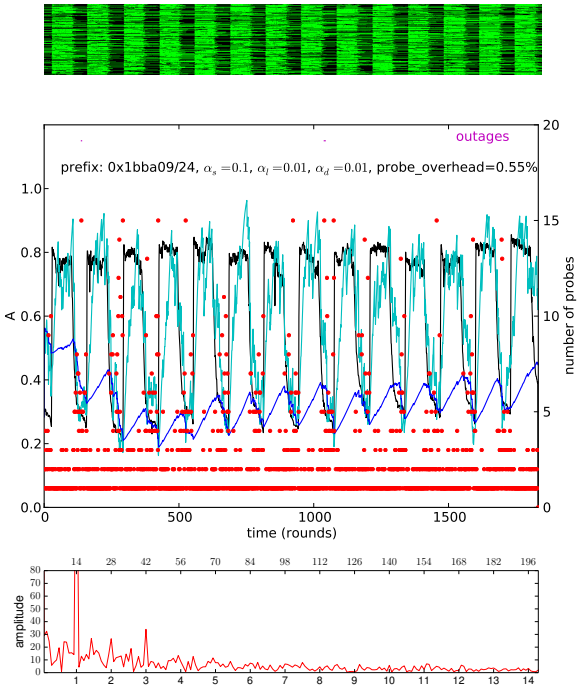


Figure 3: Block 27.186.9/24 from S_{51w}, $|E(b)| = 256$, $A = 0.598$. Diurnal block. Bottom: amplitude of FFT ($N_d = 14$) showing strong diurnal amplitude.

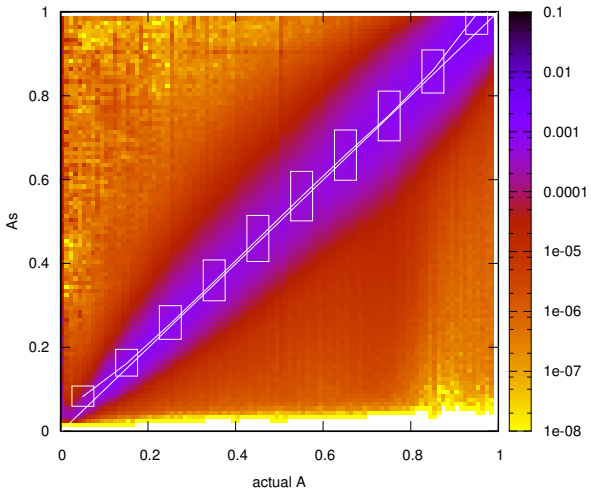


Figure 4: Correlation graph showing actual availability A and estimated availability $\hat{A}_s$. Density is normalized by the product of number of blocks and rounds. Quartiles show A_s, binned by $0.1A$.

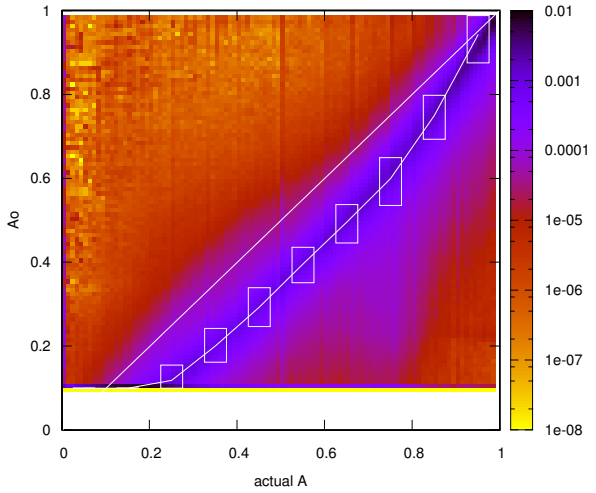


Figure 5: Correlation graph showing actual availability A and operational availability $\hat{A}_o$. Density is normalized by the product of number of blocks and rounds. Quartiles show A_o, binned by $0.1A$.

Figure 4 compares our short-term estimates ($\hat{A}_s$) against ground truth (A) as a density plot. We normalize the density with the product of number of blocks and number of rounds. The dense cluster near the "perfect correlation" of $x = y$ line shows this correlation visually. Over the density plot, we plot the median (white line) and quartiles for all values in each group of 0.1 true A (white boxes), confirming our estimator is unbiased.

Figure 5 shows the same visualization for operational $\hat{A}_o$. Our goal is that the operational value not overestimate. This data shows our operational A_o values are almost always (94% of the time) under true A. Since we do not probe very-sparse blocks, we omit cases where $\hat{A}_o < 0.1$.

Finally, we quantify the accuracy of our estimates by computing the correlation coefficient over true block availability A and estimated values. We find good correlation between true A value and estimated $\hat{A}_s$ with an overall correlation coefficient of 0.95685.

3.2 Diurnal Blocks

We next consider sample diurnal blocks, then systematically explore how different factors detect of diurnalness.

3.2.1 Sample Blocks

We first consider what spectral analysis (§ 2.2) reports about our sample blocks. In Figure 1, a few addresses are always used, and in Figure 2 all addresses are used intermittently. Both blocks show a flat spec-

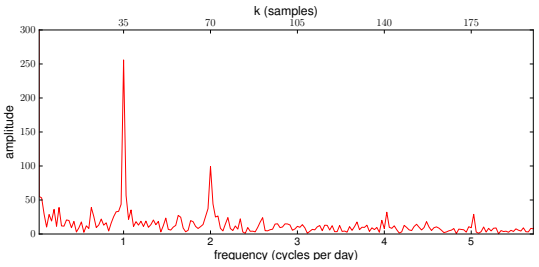


Figure 6: Amplitude of FFT for block 27.186.9/24, in 35-day A_{12w} ($N_d = 35$); shows strong diurnal pattern.

trum, as illustrated by the bottom graph in Figure 1 (the second example is similar but omitted for space).

By contrast, the daily pattern in Figure 3 is visible in the raw data (top) and observed $\hat{A}_s$ (middle). It produces a strong peak corresponding to daily activity (bottom graph, here at $k = 17$). This sample block in our longer dataset A_{12w} also shows a strong daily peak in Figure 6 (at $k = 35$ due to the 35-day observation).

We have looked at many other example blocks; these cases are representative of diurnal and non-diurnal spectra. We report on diurnalness in A_{12w} in § 4. Of the 3.7M blocks we study in A_{12w} we classify 412k (11%) as strictly diurnal and 926k blocks (25%) as relaxed diurnal.

3.2.2 *Controlled Diurnal Blocks with Simulation*

We next turn to simulation to systematically study how easily we can detect diurnalness in the face of different kinds of variation and noise.

We simulate one /24 block (256 addresses) to provide exact ground truth against which to test our algorithm. Our simulation evaluates responses every 11-minutes for 4 weeks. In that block, 50 addresses are stable and always responding, and $n_d = 100$ addresses are diurnal, and the remaining addresses are not active. Diurnal addresses are responsive for 8 hours and down for 16 hours each day. Each diurnal address i turns on at a certain time during the day, the phase ϕ_i.

We evaluate the effects of three kinds of variation in this model. First, we vary phase, selecting ϕ_i for each address (once, at simulation start) uniformly from the range $[0, \Phi]$. We also add normally-distributed noise, either adding variance σ_s to the start time of the up period, or changing its duration with variance σ_d. In both cases, noise is drawn each day for each address. In each simulation below, we run 10 batches of experiments, each batch with 100 experiments to report accuracy: percentage of correct detection of diurnal blocks within the 100 experiments. The error bars show median and quartiles in the 10 batches.

For each simulation we observe $\hat{A}_s$ (§ 2.1) and diurnal detection (§ 2.2). We detect 100% of these simulated diurnal blocks in the simplest case when there is no noise ($\Phi = \sigma_s = \sigma_d = 0$).

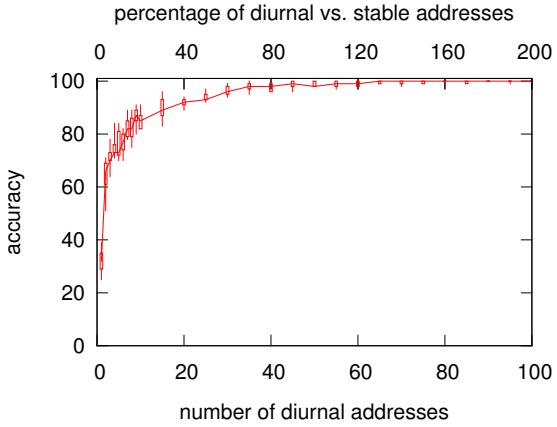


Figure 7: Accuracy of diurnal detection when the number of diurnal behavior addresses (n_d) per block varies from 2% to 67% of responsive addresses ($\Phi = \sigma_s = \sigma_d = 0$). Data: simulation.

Number of diurnal addresses: We first explore effects of *how many* addresses are diurnal. In Figure 7, we vary n_d, the number of diurnal addresses, from 1 to 100 (or 2% to 67% of responsive addresses). Our detection accuracy improves quickly. After 10 addresses (or 17% of responsive addresses), the accuracy is more than 85%. Examination of several cases shows the misses with $n_d < 10$ are because n_d is only a small fraction of the 50 always-on addresses, and our probing will usually pick an address within the stable ones and stop.

Varying phase: To vary phase we return to $n_d = 100$ but now select ϕ_i randomly, picking a value linearly distributed between 0 and Φ. to decide when it starts during the time of each day. Figure 8 shows a sharp drop when maximum Φ at 14 hours. This sharp threshold is due to the twice next strongest amplitude requirement of strict diurnalness; although variations in phase decay gradually, the signals from different addresses to blur together around this period, as confirmed in simulations where we vary maximum Φ from 0 to 24 hours (see Appendix B for details).

Varying duration: We vary duration by changing σ_d in Figure 9. We vary σ_d from 0 to 24 hours to evaluate our sensitivity to variations in up-time. We find that this variance only slightly affects accuracy for large σ_d (>10 hours). This is because we synchronize the up periods daily, similar to real world settings where people synchronize clocks and go to work. With a normal distribution, variations in up-time for the sub-signals will cancel each other out over time. Considering that ordinary people's schedules have variations within only a few hours, this shows that our algorithm works well for wide range of up-times.

After studying sensitivity of n_d, ϕ, σ_d, we are confi-

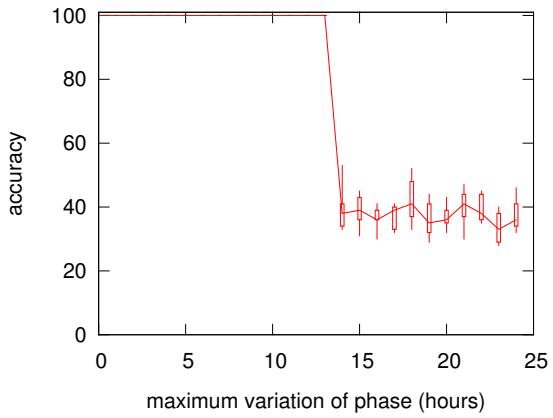


Figure 8: Accuracy of diurnal block detection when maximum Φ of diurnal addresses varies from 0 to 24 hours ($n_d = 100, \sigma_s = \sigma_d = 0$). Data: simulation.

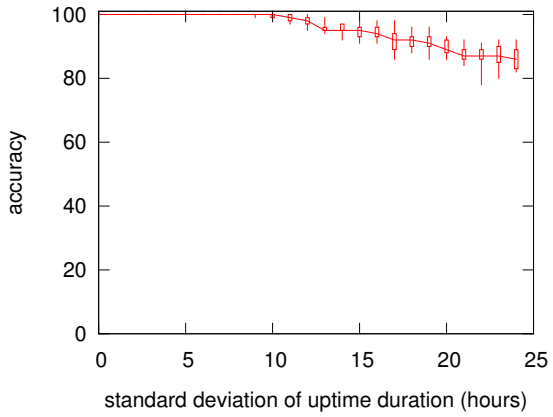


Figure 9: Accuracy of diurnal block detection when standard deviation of uptime duration (σ_d) varies from 0 to 24 hours ($n_d = 100, \Phi = \sigma_s = 0$). Data: simulation.

dent with our availability tracking and diurnal detection algorithms. Because we can track blocks with small n_d (less than 20% of stable addresses), large Φ (typical human phase is less than 4 hours), and large σ_d (typical up time is 6 to 10 hours).

3.2.3 Comparison in Full Survey Blocks

We next evaluate diurnal behavior with and without estimated A. Following § 3.1.2, we define ground truth as diurnal detected with true availability (A) taken from full data in Survey S_{51w}; we cannot contact thousands of network operators to get actual ground truth. We compare that to diurnal computed from our lightweight estimate of availability ($\hat{A}_s$).

The comparison in Table 1 shows that our method has good precision (82.48%): we rarely falsely predict a wrong diurnal block. We also provide good overall accuracy (90.99%), which represents the ability to correctly track both positives and negatives. Our measurement is conservative in detecting diurnalness, with a fairly high false negative rate. For the comparisons we make in § 5, this bias seems preferable to false positives.

outcome		**dataset**	
with A	**with $\hat{A}_o$**	S_{51w}	
(correct) d	$\hat{d}$	2,890	9.97%
n	$\hat{n}$	23,497	81.02%
(error) d	$\hat{n}$	1,999	6.89%
n	$\hat{d}$	614	2.12%
precision: 82.48%; accuracy: 90.99%			

Table 1: Validation of diurnal blocks in Survey S_{51w} (29k blocks), using true availability (A) to compute ground truth, and Trinocular estimated availability ($\hat{A}_s$) to predict.

3.2.4 Validation with USC Ground Truth

Although we cannot contact thousands of network operators, we are able to discuss our results with the operators at USC. We next verify that the diurnal blocks we discover there are correct (no false positives) and evaluate completeness (false negatives).

Our experimental results are from A_{12w}, where we find four strictly diurnal blocks and 44 relaxed diurnal blocks. We evaluate block usage using both reverse DNS names and discussions with network operators. USC is large and network configuration is distributed and done partially on physical location, so DNS names provide the most precise information; we then confirm our results with operators.

Of the four strictly diurnal blocks, two are wireless blocks with dynamic address assignment. The other two have mixed use, mostly assigned to general use (assigned to buildings or departments). One block includes a range of 16 dynamically-assigned addresses. Examining the additional 40 relaxed diurnal blocks, keyword assignment identifies 20 as dynamic, 4 as wireless, and the other 16 are general use. In addition to hostnames, for 39 blocks (2 strong and 37 relaxed diurnal) we have full raw data from Survey S_{51w}.

False Positives are blocks detected as diurnal when survey shows consistent availability. Examination of all 39 blocks with full data shows clear diurnal trends in 38 blocks: A changes by at least 10% over multiple days in the two week survey. DNS data for the one block without diurnal trends shows a group of 15 dynamic addresses for mobile wired devices. We believe these addresses were activated in the four months between ground truth (S_{51w}) and detection (A_{12w}). We conclude that there are *at most 3% false positives* for USC.

False Negatives are diurnal blocks that we miss.

DNS names at USC identify 32 blocks as dynamic and 142 as wireless, blocks that should show strong diurnal trends. We identify only 20 dynamic and 6 wireless blocks as diurnal, suggesting a high false negative rate.

Examination of these blocks shows that USC's wireless is heavily overprovisioned, with 142 blocks (roughly one wireless address for every student). Although many addresses in these blocks are active over months of observation, *on average* each block has around ten live addresses, a load confirmed by operators. To avoid excessive per-address probing, by default Trinocular removes blocks with fewer than 15 active addresses, and it is difficult to detect diurnal trends in very sparse blocks (for example, see Figure 7). Thus, of the 142 wireless blocks, 119 were excluded from probing. Of the 23 blocks that are probed, we identify 6 as diurnal. We conclude that diurnal detection requires either moderately dense blocks (more than 15 active addresses) or tolerance for higher probing rates than 19 probes per hour per /24. These self-imposed *policy constraints* result in false positives for USC's sparely operated wireless network.

Sparse blocks cause only false negatives, not false positives. We therefore claim that our Internet-wide measurements are a lower-bound on the fraction of diurnal blocks.

Surprises: We found two surprising results examining USC. First, while DNS names show a few large, continuous areas that are diurnal (centralized wireless), decentralized address management also results in pockets of dynamic addresses in many /24 blocks. These dynamic addresses result in diurnal trends in otherwise general-use blocks. Second, while we expected diurnal usage of wireless and dynamically assigned addresses, we were surprised by the strong diurnal use in the 16 general-use blocks. We speculate that diurnal use reflect either broad use of laptop computers that are away from USC outside business hours, or a policy of turning off desktop computers at night.

4. DIRECTLY OBSERVED RESULTS

In this section, we show direct observations of diurnal blocks in A_{12w}. We compare these observations with other factors to identify correlations in the next section § 5). We find 412k blocks (11%) that are strictly diurnal in A_{12w}, and 926k blocks (25%) that are relaxed diurnal of the 3.7M blocks we measure. These blocks are distributed all over the address space.

Daily or other periodicity? Our test for diurnalness looks strictly for regularity at 24 hours, requiring that to be the strongest periodicity in the block. We focus on daily patterns because they are likely to be related to human use of the Internet. There may be other sources of periodicity, for example related to DHCP lease cycles. If dynamic addresses are allocated for some period p, and given out sequentially across a region that spans multiple /24 blocks, then those blocks will see usage that changes with period p.

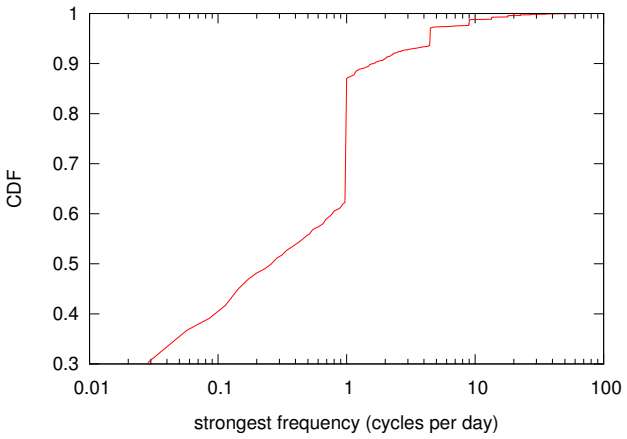


Figure 10: Cumulative distribution of the highest frequency in 35-day A_{12w}.

To understand the how prominent different periodicities are, Figure 10 shows the distribution of the strongest frequency for each of the 3.7 M blocks in A_{12w}. As expected, we observe a strong peak at 1 cycle per day (24 hours), about 25% of all the blocks. We only declare 11% to be strongly diurnal, where the 24-hour amplitude is twice the next strongest frequency; the remaining are relaxed diurnal. The second group (3%) is at about 4.3 cycles per day. This periodicity is a probing artifact, because we restart our probing software every 5.5 hours (4.3 times per day) to recover from possible prober failure. Our measurements starting in 2014-04 (A_{16all}) use restart times around one week to reduce this effect.

Long-term status: We next consider how block diurnalness changes over time. In Figure 11, we apply our spectral diurnal detection algorithm over more than 3 years of Internet surveys (each a sample of about 2% of IPv4 blocks). We observe that the percentage of diurnal blocks is relatively stable, but that it shows a marked decline since 2012. While we have observed in increasing trend in dynamic address allocation this downward trend may indicate that these dynamic addresses are slowly shifting to always-on activity (even if they move around). These results are generally consistent the 11% diurnal blocks in A_{12w}.

5. INDIRECTLY OBSERVED RESULTS

We next compare diurnal blocks to other factors: location, time-of-day, block allocation dates, economic factors, and link technologies.

5.1 Where Does the Internet Sleep? (Locating diurnal blocks)

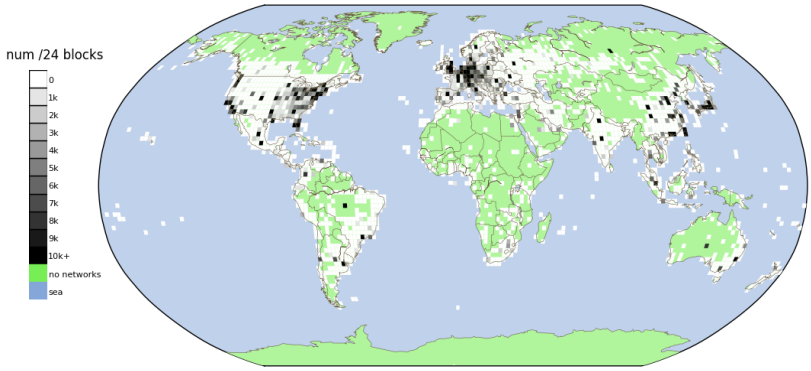


Figure 12: Number observable blocks (linear grayscale). Dataset: A_{12w}

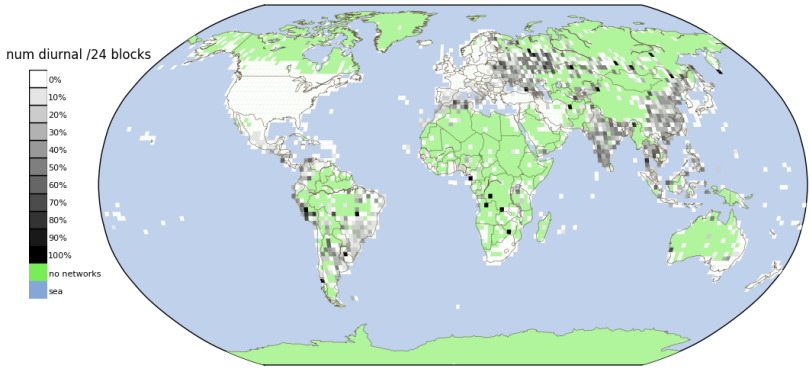


Figure 13: Percent of observable blocks that are diurnal. Dataset: A_{12w}.

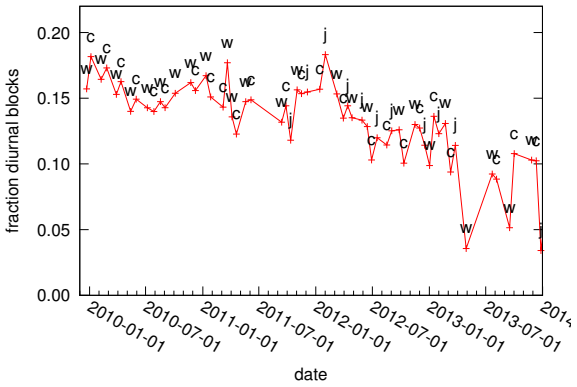


Figure 11: Fraction of diurnal blocks from long-term surveys. Uses 63 datasets, starting with S_{30w} in Dec. 2009, letters indicate location (w: Los Angeles, c: Colorado, j: Japan).

Our first question is to identify *where diurnal Internet use is common*. We study this question using A_{12w} with MaxMind geolocation (succeeding for 3.45M of blocks, 93%). We first put our results in context by showing where the geolocatable blocks are located (Figure 12). In this world map, we count /24s blocks geolocated into a 2×2 degree grid. We see large numbers of addresses in North America and Europe, as well as concentrations in Japan, China, and several other countries. We also see some geolocation anomalies: when MaxMind knows only country, it locates the blocks in the geographic center of the country, falsely placing many networks away from population in Brazil, Russia, and Australia.

Of the 412k strictly diurnal blocks in this dataset, Figure 13 shows what percentage of blocks in each grid cell are diurnal. We see significant variation by country, with very few diurnal blocks in the U.S., western Europe, and Japan, but *many* diurnal blocks in Asia, Eastern Europe, and South America. Table 2 lists the countries with the most diurnal blocks.

These strong regional differences show the ability of our approach to identify global trends in Internet usage. In coming sections we explore how these differences correlate with of a combinations of cultural, policy, and economic causes.

5.2 When Does the Internet Sleep? (Phase of Diurnal Blocks)

Amplitude in spectral analysis of A shows *which* blocks are diurnal, and geolocation showed *where* these blocks are. We next show that *phase* in the FFT coefficients shows *when* these blocks are active.

We next look at the phase of geolocated diurnal blocks

country code	region	blocks (/24s)	frac. diurnal	GDP (US$)
AM	W. Asia	1075	0.630	5900
GE	W. Asia	1395	0.546	6000
BY	E. Europe	1748	0.512	15900
CN	E. Asia	394244	0.498	9300
PE	S. America	4600	0.401	10900
KZ	Cent. Asia	3832	0.400	14100
RS	S. Europe	4429	0.393	10600
AR	S. America	20382	0.339	18400
TH	S.E. Asia	10986	0.336	10300
SV	Cent. America	1145	0.311	7600
UA	E. Europe	16575	0.289	7500
CO	S. America	9379	0.261	11000
MY	S.E. Asia	9747	0.247	17200
PH	S.E. Asia	5721	0.239	4500
IN	S. Asia	36470	0.225	3900
MA	N. Africa	2115	0.185	5400
BR	S. America	79095	0.185	12100
VN	S.E. Asia	8197	0.183	3600
ID	S.E. Asia	7617	0.166	5100
RU	E. Europe	53048	0.159	18000
...	...	...	...	
US	N. America	672104	0.002	50700

Table 2: Fraction of diurnal blocks, top 20 countries (with at least 1000 blocks in our study), and United States. Diurnal analysis data is from A_{12w}; Geolocation data is from MaxMind [23]; Per-capita GDP data is from CIA world factbook [5].

region	blocks (/24s)	frac. diurnal
Northern America	721716	0.002
Southern Africa	11255	0.0108
W. Europe	275224	0.0109
Northern Europe	133911	0.0131
Caribbean	2174	0.016
Oceania	27206	0.0349
W. Asia	25570	0.0765
Northern Africa	9984	0.0992
Southern Europe	134933	0.124
Central America	44644	0.133
Eastern Europe	146552	0.135
Southern Asia	44524	0.200
South America	133493	0.208
South-Eastern Asia	48885	0.219
Eastern Asia	757352	0.279
Central Asia	3832	0.401

Table 3: Fraction of diurnal blocks grouped by regions. Diurnal analysis data is from A_{12w}; Geolocation data is from MaxMind [23].

to see that it correlates strongly with longitude and therefore timezone. We measure how many blocks occur at each combination of longitude and phase. Since both phase and longitude are continuous, we count similar blocks by quantizing longitude and phase each into 100 bins ($\pm\pi$ and $\pm 180°$). Since both wrap around the circle, we "unroll" them by plotting phase in the range $[-\pi + L, \pi + L]$, where L is the longitude in radians.

Figure 14a shows a density plot of the 287k geolocatable, strictly diurnal blocks in A_{12w} as described above. We first observe that many blocks at longitude zero have phase near zero, and longitude and phase are strongly correlated—larger phases correspond with more eastward longitudes.

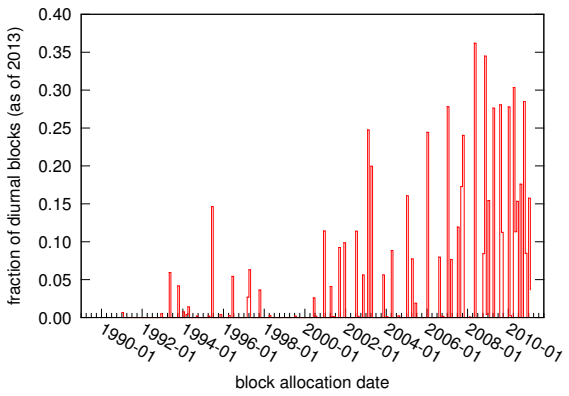


Figure 15: Percentage of diurnal blocks in each month, based on block allocation date. Dataset: 2013 diurnalness: A_{12w}; block allocation: IANA).

The correlation coefficient of unrolled phase against longitude is 0.835, showing they are strongly related. We did not attempt to align phase zero with longitude zero; their proximity is accidental and calibrating phase with local time of day is future work.

In Figure 14b we repeat this analysis with 590k blocks that are relaxed diurnal and geolocatable. The fit here is not quite as good (correlation coefficient is 0.763), but it is very close. This strong similarity suggests that our strict test for diurnal may be overly conservative.

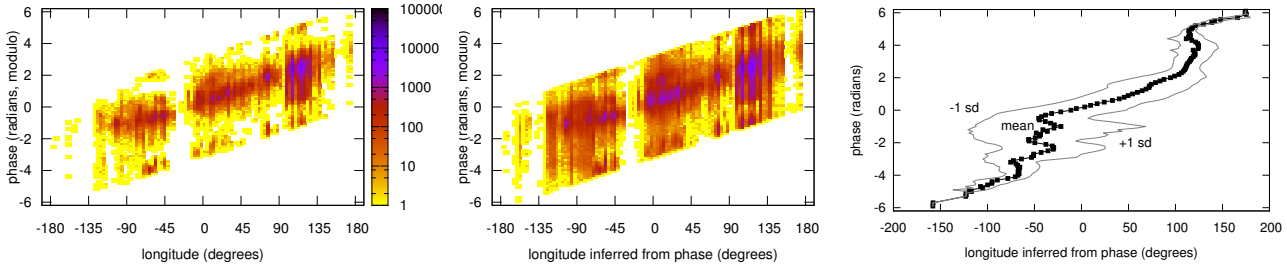
In both Figure 14 we observe that the correlation between phase and longitude is fairly strong, except for longitudes between 100°E to 140°E. This range corresponds with Mongolia, China and southeast Asia. We believe this variation is primarily because our IP geolocation is not particularly accurate in China, and because it is a geographically wide country that uses a single timezone.

Finally, the relationship between phase and longitude suggests that phase may help geolocate diurnal blocks. Figure 14c shows mean and standard deviations when this mapping is derived from relaxed diurnal data. We see that the quality of the predictor varies significantly: a phase from -2 to 0 indicates only the hemisphere, while most other phases predict longitude within ± 20 degrees.

5.3 Correlating Block Diurnalness and Allocation Date

Policies for IP address allocation have evolved since IPv4 was introduced, with stricter requirements for density and reuse that encourage dynamic allocation through full allocation in May 2012 [14]. Prior work has documented increased dynamic addressing over time [3].

We can see how evolution in policies results in in-



(a) Strictly diurnal blocks. (b) Relaxed diurnal blocks. (c) Predicting longitude from phase (relaxed diurnal shown, strict is similar).

Figure 14: Comparing phase and longitude. Dataset: A_{12w}.

creased diurnal Internet blocks by correlate the fraction of diurnal blocks with the date when each /8 block was allocated to a regional registry by ICANN [13]. We show this relationship in Figure 15 by plotting a histogram of how many blocks are currently used diurnally, based on the date of allocation. The trend is for increased diurnal use as allocations proceed, corresponding with more careful policies in address use. (Linear regression shows a positive slope of 0.08% per month, with 0.609 correlation coefficient.) Even though the overall fraction of diurnal blocks shows a slight downward trend (Figure 11), this data shows that newer blocks are more often diurnal.

5.4 Effects of Economic Conditions

We hypothesize that economic factors affect Internet use, since dynamic address assignment and turning devices off at night is more efficient than static or always-on dynamic assignment and always-on devices, and economic factors affect investment in telecommunications infrastructure. We evaluate these questions next by correlating diurnalness and outages with economic factors.

5.4.1 Correlating Diurnal Blocks with Economics

Why do some countries have a greater fraction of diurnal blocks (§ 5.1)? Cultural differences across countries prompt different rates of savings, and different economic situations leave citizens different levels of disposable income. These difference and anecdotal evidence suggest that people of some countries, such as as China and India, to turn off their computers at night to save electricity bills.

Although we cannot directly model cultural differences, we can observe correlations between disposable income (measured in per-capita GDP, taken from the CIA World Factbook [5]) and diurnal network usage. Table 2 previously gave specific values. Figure 16 visualizes this correlation, including the weak linear fit (confidence coefficient: -0.526). The top 20 countries with diurnal fractions above 0.15 generally have a per-capita GDP less than USD $15,000, less than one-third

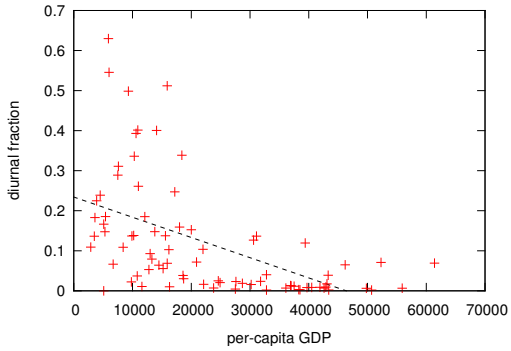


Figure 16: Scatter plot of diurnalness and per-capita GDP for all countries, for Table 2, with weak linear fit (dashed line). Dataset: A_{12w}, CIA.

of the United States. We conclude that economic factors correlate with differences in Internet use.

Other Possible Factors: GDP is one of many economic indicators. To understand correlations with other factors we carry out ANOVA factorial analysis (§ 2.4) comparing five factors: per-capita GDP, Internet users per host, electricity consumption per capita. and age of first (and mean) block allocation, and fraction of diurnal blocks.

Table 4 shows this analysis for all one- and two-factor combinations. A p-value less than 0.05 suggests a strong correlation. We find there are three combinations meet this requirement: per-capita GDP, with p-value of 6.61×10^{-8}. The second is the combination of per-capita electricity consumption and mean age of allocation, with p-value of 0.001476. The third is mean age of allocation, with a p-value of 0.031354.

The result from ANOVA analysis supports our previous hypothesis that per-capita GDP is the main factor for diurnalness, and suggests that electrical usage and mean allocation age are also correlated. Electricity consumption is correlated with economic activity. Mean allocation age suggests that newer countries have more diurnal use, as shown in § 5.3.

	per-capita gdp	p.cpt. elec. cons.	I-net users/ host	age/ first alloc.	mean age/ alloc.
gdp	$\mathbf{6.61\times10^{-8}}$	0.306	0.822	0.789	0.995
elec.		0.703	0.609	—	**0.00148**
users/host			0.0366	0.959	0.849
first alloc.				0.830	—
mean alloc.					**0.0314**

Table 4: ANOVA analysis of correlations between diurnal and individual factors (the diagonal) and pairwise combinations of factors (off the diagonal). Bold combinations are statistically significant. Dataset: A_{12w}.

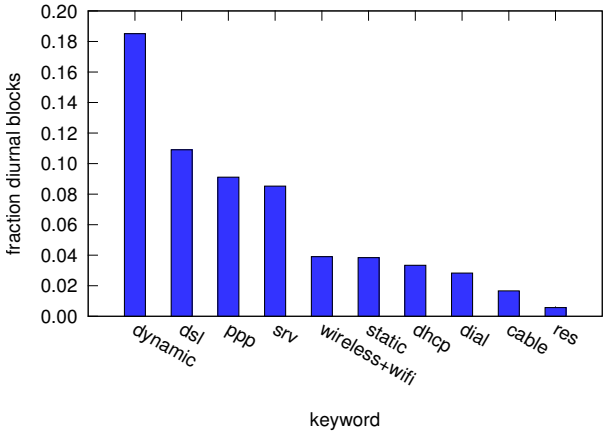


Figure 17: Fraction of diurnal blocks for 9 access keywords. Dataset: A_{12w}.

Although ANOVA analysis is powerful, we must take care that factors not necessarily independent—low GDP may imply less developed infrastructure and less available power. Additional work is needed to show a causal relationship between these (or other) factors and network use. However, our work suggests the potential for such relationships and is the first to suggest this area as a direction for study.

5.5 Effects of Access-Link Technology

We next consider correlations between access-link technology and diurnalness. National policies are often set to encourage new link technologies and promote always-on networks. In addition, while such policies and economic conditions are difficult for individuals to influence, users and ISPs often can choose different "last-mile" methods to connect to the Internet.

We infer link type (§ 2.3.3) for all blocks in A_{12w}, classifying 22.4% of those blocks. For these blocks, we correlate link type with diurnalness. Each of these link types have more than 30k blocks; we omit the wireless keyword because we see only 691 wireless/wifi blocks, not enough to provide statistically strong results.

Figure 17 shows the fraction of diurnal networks for each different link type. As expected, dynamic addresses are strongly correlated with diurnal behavior, with *dynamic* at 19%. Somewhat surprisingly, *dialup* is not strongly diurnal ($< 3\%$), while *dsl* is more diurnal (11%). These results suggest the importance of measuring network behavior rather than assuming.

6. RELATED WORK

ISP comparison: Several third-party systems exist to compare performance of ISPs, including Keynote [18], top ten reviews [18], and the popular speed test tool offered by OOKLA [27]. They share a common drawback: lack of scalability. Perhaps closest to our work is Netdiff [22], which measures performance of hundreds of paths for 18 backbone ISPs. They find that ISP performance depends on geographic properties of traffic and popularity of destinations. Our work is more focused on nodes instead of paths, and thus complements Netdiff.

Adaptive sampling: Several prior studies have used adaptive probing techniques to find network faults (for example, [2]), and use simulations for validation. Similar to these, we use adaptive sampling inside network blocks to achieve accurate availability information ($\hat{A}_s$), with low probing cost.

Node-based studies: Barford et al. use node-based methods to detect network performance anomalies [1], with clever weighting algorithms to find which links to probe next; they simulate 15 types of topologies based on ns-2. Several other previous studies also use similar strategies, including Hubble [17] and iPlane [21]. Such node-based studies are precise in understanding performance on link paths between the node mesh. However, they usually "only" cover hundreds or thousands of nodes. We complement such systems with a much larger coverage (3.7M /24 blocks), and with measurements on a regular basis every 11 minutes.

Tomography studies: Network tomography is interested in finding the exact location of problems [6, 7, 11, 19, 20]. We correct for correlated error by removing outages near our probers. We are interested in tracking block-level network reachability in the whole Internet edge, but without exact localization of outages. Our work and tomography studies can benefit from each other. Our future work could be more precise by localizing problems. Tomography studies can use our system as a pre-step for finding symptoms.

7. CONCLUSIONS

In this paper, we develop an accurate adaptive approach to estimate network block availability over time. Based on this, we perform the first long-term, wide-scale analysis of diurnal network behavior and outages. We also begin to evaluate how policies, economic conditions and technologies affect Internet use in the world. With statistical ANOVA analysis, we correlate diurnal

usage and outages to economic factors such as GDP and power consumption.

Acknowledgments

This work was classified by USC's IRB as non-human subjects research (IIR00001648).

We thank Brian Yamaguchi (USC/ITS) for guidance about USC networks, and Christos Papadopoulos (CSU) and Rod Van Meter, Yohei Kuga, and Midori Kato (WIDE) for supporting data collection from their sites. We thank John Wroclawski for comments on the paper.

8. REFERENCES

[1] P. Barford, N. Duffield, A. Ron, and J. Sommers. Network performance anomaly detection and localization. In *INFOCOM 2009, IEEE*, pages 1377–1385, 2009.

[2] M. Brodie, I. Rish, and S. Ma. Intelligent probing: A cost-effective approach to fault diagnosis in computer networks. *IBM Systems Journal*, 41(3):372 –385, 2002.

[3] Xue Cai and John Heidemann. Understanding Block-level Address Usage in the Visible Internet. In *Proc. of SIGCOMM*, 2010.

[4] Xue Cai, John Heidemann, Balachander Krishnamurthy, and Walter Willinger. Towards an AS-to-Organization map. In *Proceedings of the ACM Internet Measurement Conference*, pages 199–205, Melbourne, Australia, November 2010. ACM.

[5] CIA World Factbook. List of countries by GDP (PPP) per capita. `https://www.cia.gov/library/publications/the-world-factbook`.

[6] Ítalo Cunha, Renata Teixeira, Nick Feamster, and Christophe Diot. Measurement methods for fast and accurate blackhole identification with binary tomography. In *Proc. of 9thACM IMC*, 2009.

[7] Amogh Dhamdhere, Renata Teixeira, Constantine Dovrolis, and Christophe Diot. NetDiagnoser: troubleshooting network unreachabilities using end-to-end probes and routing data. In *Proc. of ACM Conference on Emerging Networking Experiments and Technologies*, pages 18:1–18:12. ACM, 2007.

[8] Federal Communication Commission. Collecting America: the national broadband plan. Technical report, March 2010.

[9] John Heidemann, Yuri Pradkin, Ramesh Govindan, Christos Papadopoulos, Genevieve Bartlett, and Joseph Bannister. Census and survey of the visible Internet. In *Proc. of ACM IMC*, pages 169–182, Vouliagmeni, Greece, October 2008. ACM.

[10] Zi Hu and John Heidemann. Towards geolocation of millions of IP addresses. In *Proc. of the ACM IMC*, pages 123–130, Boston, MA, USA, 2012. ACM.

[11] Yiyi Huang, Nick Feamster, and Renata Teixeira. Practical issues with using network tomography for fault diagnosis. *ACM CCR*, 38:53–58, September 2008.

[12] Bradley Huffaker, Marina Fomenkov, and k. c. claffy. Geocompare: a comparison of public and commercial geolocation databases. Technical report, CAIDA, May 2011.

[13] Internet Assigned Numbers Authority. IANA IPv4 address space registry. web page `http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.txt`, May 2013.

[14] Internet Corporation for Assigned Names and Numbers. Global policy for post exhaustion IPv4 allocation mechanisms by the IANA, May 2012.

[15] Internet Software Consortium. Internet Domain Survey. web page `http://www.isc.org/solutions/survey`, January 2007.

[16] Philip j. Fleming and John J. Wallace. How not to lie with statistics: the correct way to summarize benchmark results. *Communications of the ACM*, 29(3):218–221, March 1986.

[17] Ethan Katz-Bassett, Harsha V. Madhyastha, John P. John, Arvind Krishnamurthy, David Wetherall, and Thomas Anderson. Studying black holes in the internet with Hubble. In *NSDI*, 2008.

[18] Keynote. Internet Health Report. `http://www.internetpulse.net`.

[19] Ramana Rao Kompella, Jennifer Yates, Albert Greenberg, and Alex C. Snoeren. IP fault localization via risk modeling. In *NSDI*, 2005.

[20] Ramana Rao Kompella, Jennifer Yates, Albert Greenberg, and Alex C. Snoeren. Detection and Localization of Network Black Holes. In *Proc. of IEEE Infocom*, 2007.

[21] Harsha V. Madhyastha, Tomas Isdal, Michael Piatek, Colin Dixon, Thomas Anderson, Arvind Krishnamurthy, and Arun Venkataramani. iPlane: An information plane for distributed services. In *Proc. of 7th OSDI*, pages 367–380. USENIX, November 2006.

[22] Ratul Mahajan, Ming Zhang, Lindsey Poole, and Vivek Pai. Uncovering performance differences among backbone isps with netdiff. In *Proc. 5th USENIX NSDI*, pages 205–218, 2008.

[23] MaxMind. GeoIP Geolocation Products. `http://www.maxmind.com/en/city`.

[24] Rupert G. Miller. *Beyond ANOVA: Basics of Applied Statistics (Texts in Statistical Science Series)*. Chapman & Hall/CRC, January 1997.

[25] P. Mockapetris. Domain names—concepts and

facilities. RFC 1034, Internet Request For Comments, November 1987.
[26] James A. Muir and Paul C. Van Oorschot. Internet geolocation: Evasion and counterevasion. *ACM Computing Surveys*, 42(1), December 2009.
[27] OOKLA. Speedtest. `http://www.speedtest.net`.
[28] Lin Quan, John Heidemann, and Yuri Pradkin. Detecting internet outages with precise active probing (extended). Technical Report ISI-TR-2012-678, USC/ISI, February 2012.
[29] Lin Quan, John Heidemann, and Yuri Pradkin. Trinocular: Understanding internet reliability through adaptive probing. In *Proc. ACM SIGCOMM*, pages 255–266, August 2013.
[30] Aaron Schulman and Neil Spring. Pingin' in the rain. In *Proc. of ACM IMC*, pages 19–25, Berlin, Germany, November 2011.
[31] Srikanth Sundaresan, Sam Burnett, Nick Feamster, and Walter de Donato. BISmark: A testbed for deploying measurements and applications in broadband access networks. In *Proc. of.* USENIX, June 2014.
[32] Srikanth Sundaresan, Walter de Donato, Nick Feamster, Renata Teixeira, Sam Crawford, and Antonio Pescapè. Broadband Internet performance: A view from the gateway. In *Proc. of SIGCOMM*, pages 134–145, Toronto, Ontario, Canada, August 2011. ACM.
[33] Team Cymru. IP to ASN Mapping. `http://www.team-cymru.org/Services/ip-to-asn.html`.
[34] USC/LANDER Project. LANDER: Los Angeles network data exchange and repository. web page `http://www.isi.edu/ant/lander/`, November 2007.
[35] USC/LANDER project. Internet address survey dataset, predict id `usc-lander/internet_address_survey_adaptive_reprobing_a12w`. web page `http://www.isi.edu/ant/lander`, April 2013.
[36] Kathryn Zickuhr and Aaron Smith. Home broadband 2013. Technical report, Pew Research Center, August 2013.

APPENDIX

A. CHECKING STATIONARITY OF AVAILABILITY TRACKING

To validate the stationarity of blocks we study, we check all 29,001 blocks in Survey S_{51w}. We do linear regression on the true availability A, getting a slope per block. We plot the cumulative distribution of slopes in Figure 18. We observe that 80.3% of these blocks are stationary, with a slope equivalent to less than 1 address changes per day. We randomly selected 10 extreme cases with large slopes, and find that they are mostly due to a period of no data, or a sudden change of availability which is most likely a repurpose of the block. Because blocks are randomly selected, this experiment means that most Internet blocks are behaving in a stationary manner, thus our FFT analysis is not distorted by non-stationarity.

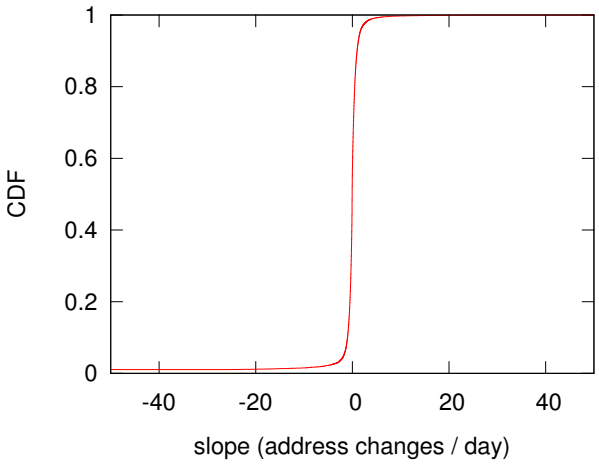


Figure 18: Checking stationarity of all blocks in Survey S_{51w}, showing the slope of linear regression on true availability A.

B. EFFECT OF MAXIMUM PHASE

To verify the effect of phase on our ability to detect diurnal blocks, we vary the maximum phase Φ in our simulations. We believe with a large Φ, many sub-signals (each with varying phases) form a smoothed combined signal, and is hard to detect. We plot and confirm this effect varying maximum Φ from 0 to 24 hours. In Table 5, we observe that at $\Phi = 14$ hours, the combined signal starts to smooth out, without strong diurnal patterns. Note that human phase behavior is usually within a few hours, in which cases we are able to detect diurnal blocks correctly.

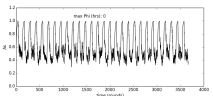
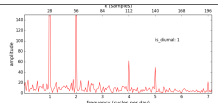
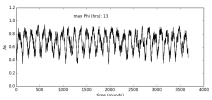
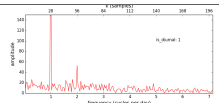
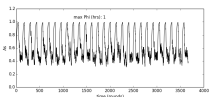
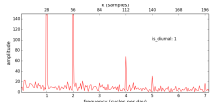
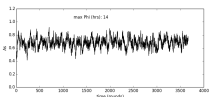
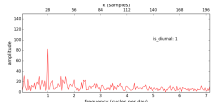
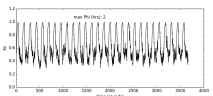
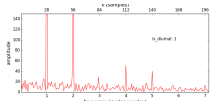
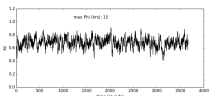
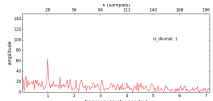
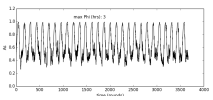
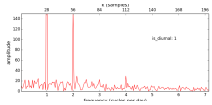
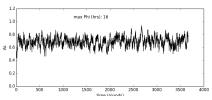
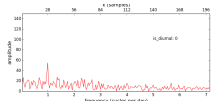
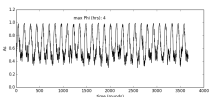
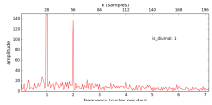
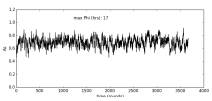
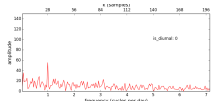
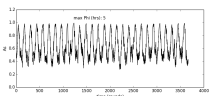
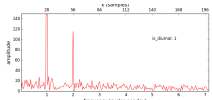
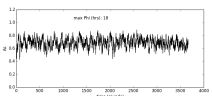
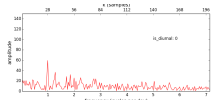
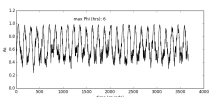
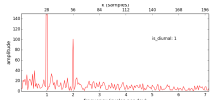
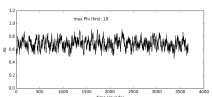
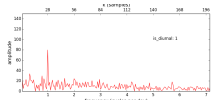
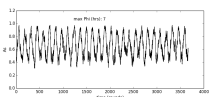
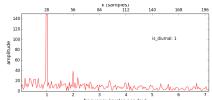
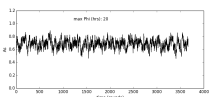
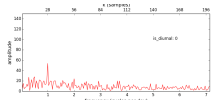
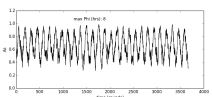
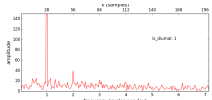
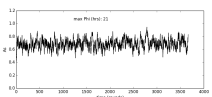
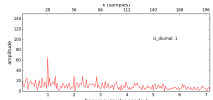
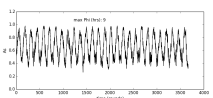
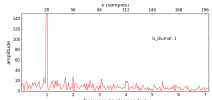
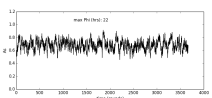
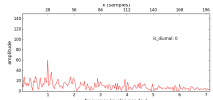
Φ (hrs)	raw A_s	FFT	Φ (hrs)	raw A_s	FFT
0			13		
1			14		
2			15		
3			16		
4			17		
5			18		
6			19		
7			20		
8			21		
9			22		
10			23		
11			24		
12					

Table 5: Vary maximum phase Φ, showing the effect of phase on diurnal block detection.