Infonet Group
University of Namur
www.infonet.fundp.ac.be

# 3D-LD: a graphical Wavelet-based method for Analyzing Scaling Processes


Steve UHLIG

suh@infonet.fundp.ac.be

Infonet Group



15th ITC Specialist Seminar on Internet Traffic Engineering and Traffic Management

Wurzburg, Germany

22-24 July 2002

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Schedule

- Scaling basics
- Wavelet basics
- Scaling detection
- Conclusion

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Scaling basics

**Scaling** $\sim$ no particularly important timescale

## Two examples:

1. Self-similarity (SS):

   - formally: $\{X(t), t \in \mathbb{R}\} \stackrel{d}{=} \{c^H X(t/c), t \in \mathbb{R}\}, \forall c > 0$
     $\Leftrightarrow$ affinity between original and zoom (all timescales behave in the same way)
   - property: $E|X(t)|^q = E|X(1)|^q \ |t|^{qH}$
     $\Rightarrow$ dependence of moments on time $\equiv$ **non-stationarity**

Note: H is “self-similarity” parameter

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Scaling basics (contd)

2. Long-range dependence (LRD):

   - definition: $r(k) \sim c_r |k|^{2H-2} \ \ as \ \ |k| \to \infty$
     $\Leftrightarrow$ power-law in auto-correlation $(r(k))$
   - persistence in correlations but **requires 2nd-order stationarity**

Relationships between self-similarity and LRD:

- $H > 0.5$ : self-similarity $\Rightarrow$ LRD
- ~~**LRD $\Rightarrow$ self-similarity**~~

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Scaling basics (contd bis)

Refining SS: strict SS $\neq$ statistical SS:

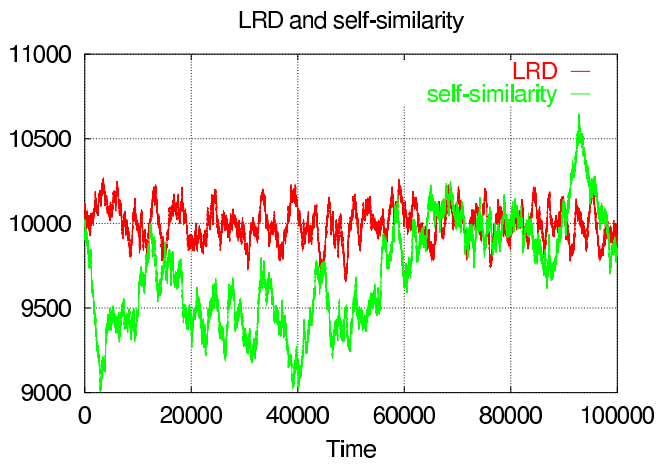
- statistical SS (discrete-time processes) often encountered in the literature (see Park and Willinger (2000)) :
  1. build aggregated sequence
     $X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \;\; k = 1, 2, ...$
  2. then $X \stackrel{d}{=} m^{1-H} \; X^{(m)}$ as $m \; \rightarrow \infty$
- statistical SS $\equiv$ LRD for $H > 0.5$

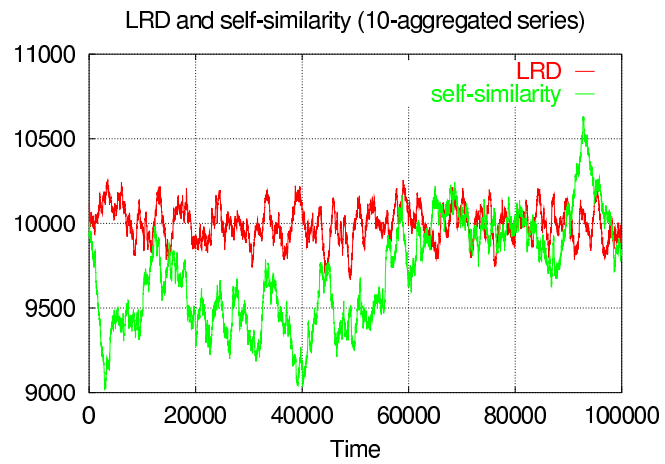
statistical SS requires 2nd-order stationarity !

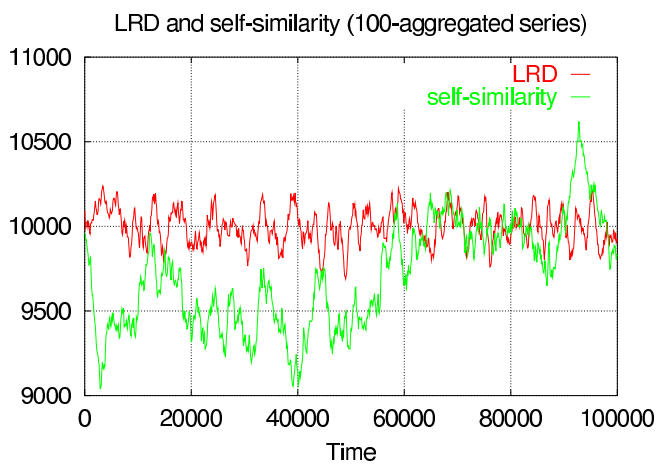
Infonet Group
University of Namur
www.infonet.fundp.ac.be

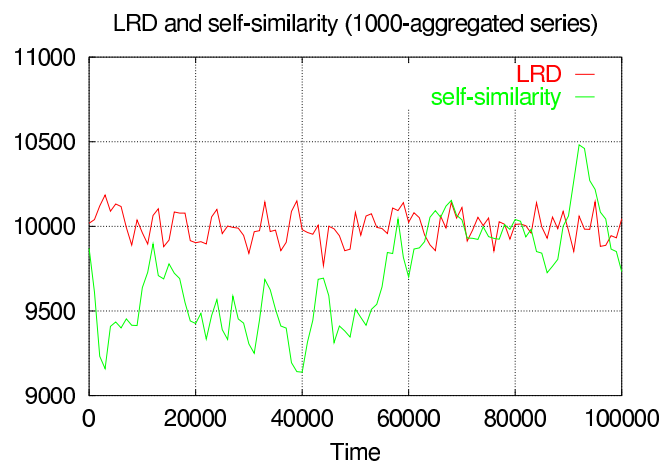
# Scaling examples: LRD vs. SS

LRD through large and persistent correlations vs. SS through fBm (gaussian noise increments)



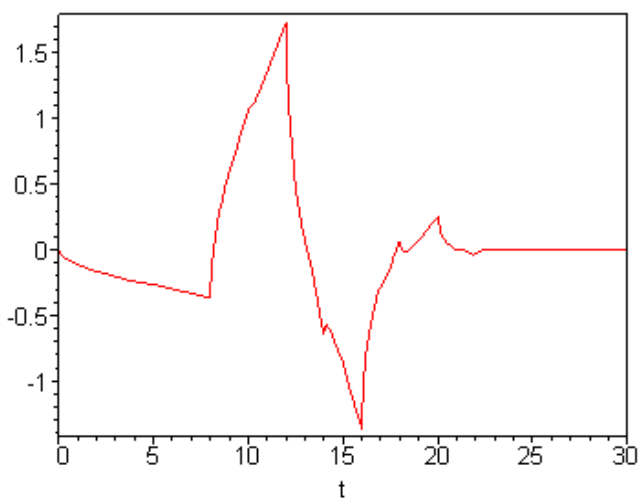



Infonet Group
University of Namur
www.infonet.fundp.ac.be

# A Wavelet Primer

- Wavelet $\psi_{j,k}$: bandpass oscillating function where $j$ is timescale and $k$ is time
- $\psi_{j,k}(t)$ form an orthonormal basis of $L_2$
- internal product $<X(t), \psi_{j,k}(t)>$ matches irregularities in $X(t)$ at time $t$ and timescale $j$
- A look at the beast:



Infonet Group
University of Namur
www.infonet.fundp.ac.be

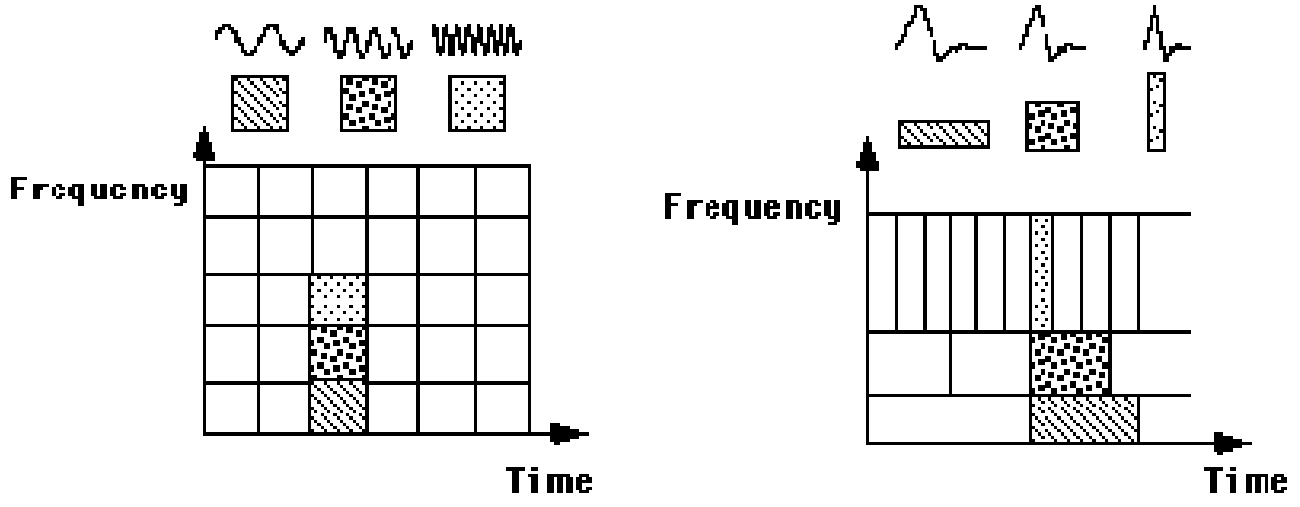
# A Wavelet Primer (contd)

- dyadic grid signal decomposition:

$$X(t) = \underbrace{\sum_k c_X(j_0,k)\phi_{j_0,k}}_{} + \underbrace{\sum_{j\le j_0}\sum_k d_X(j,k)\psi_{j,k}(t)}_{}$$

signal = approximation + $\sum$ details at each timescale

- time-frequency plane tiling: Fourier vs. wavelets




- study $d_X(j,k)$ instead of process increments $X_\delta(k)$

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# A Wavelet Primer (contd bis)

"Design" constraints on the wavelet function $\psi_{j,k}$:

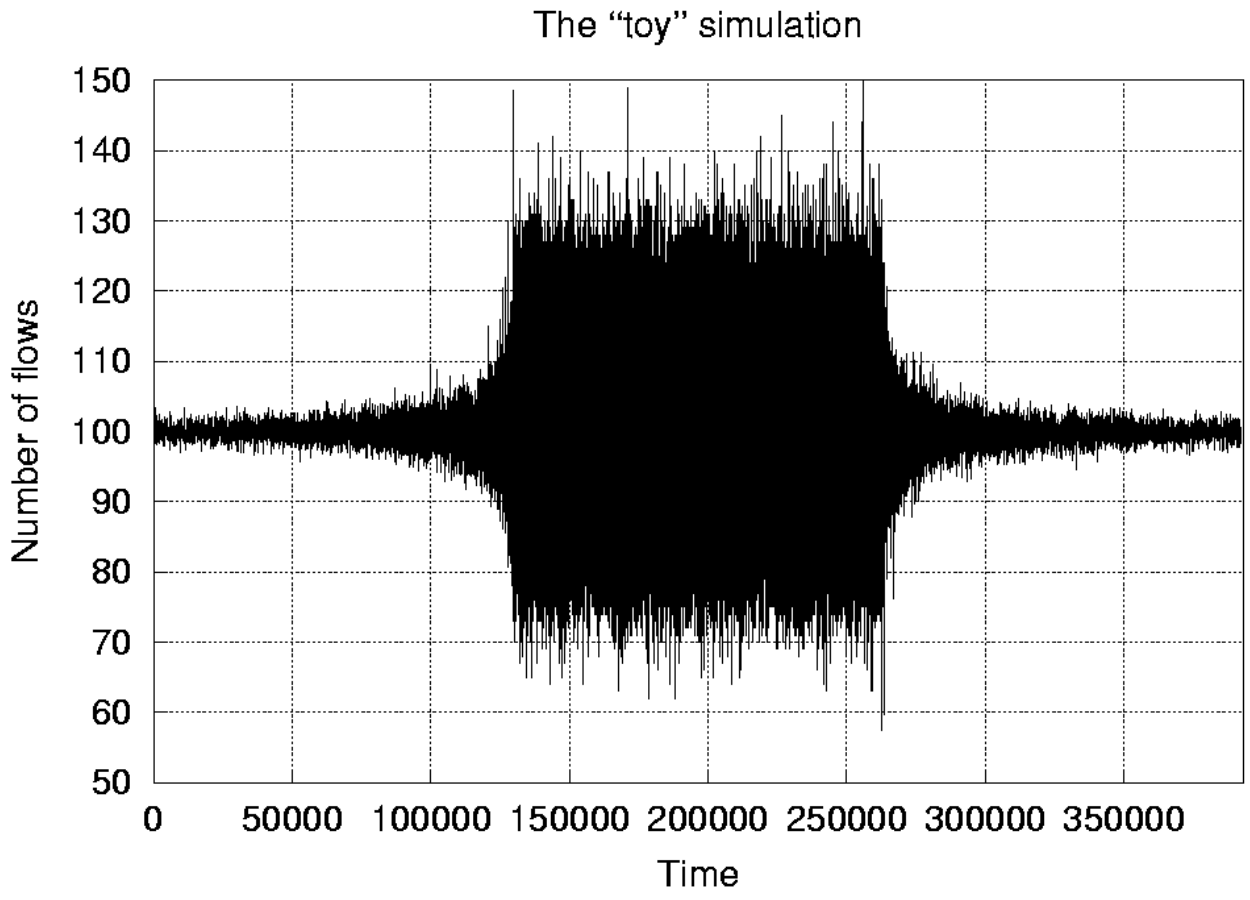
- **built-in scaling**: $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$
  $\Rightarrow$ scaling in process automatically captured
- **vanishing moments**: $\int t^k \psi_0(t)dt \equiv 0,\ k = 0, ..., N-1$
  $\Rightarrow$ polynomial non-stationarity automatically removed

Another (very) nice property:

- **almost decorrelation**: under $N \geq H + 1/2$, LRD in increments become SRD in coefficients
  $\Rightarrow$ "almost" independence among timescales $j$

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Toy example




3 different “periods” in this synthetic process, but what’s in there ?

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Scaling detection

Logscale diagram:

1. Compute $\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^2$,
2. then plot $y_j = log_2(\mu_j)$ against $j$ (+ confidence intervals)

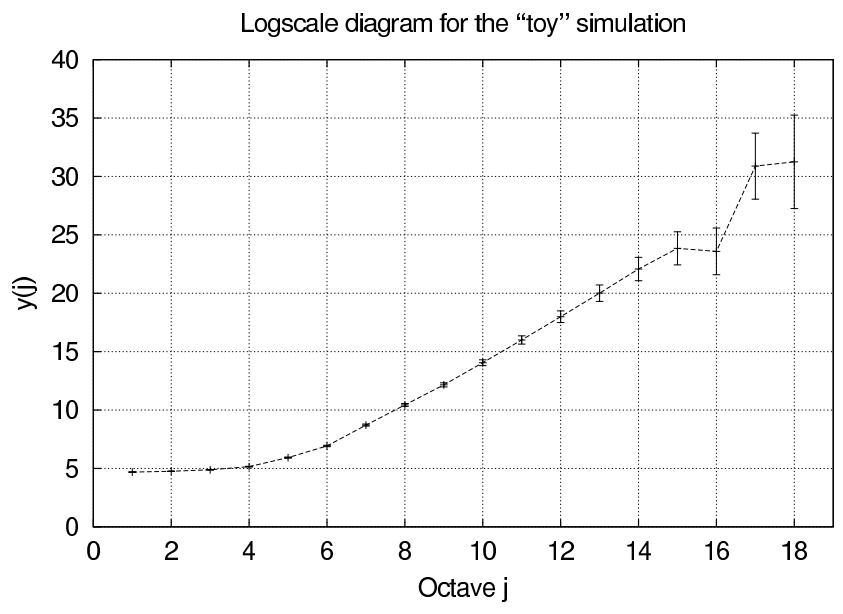
Scaling detection:

- straight line in slope of the LD for some range of $j$
- alignment in confidence intervals

⚠ LD is a second-order statistic (variance of wavelet coefficients), hence we know nothing about local scaling properties.

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# LD of toy example




| LD Slope | Scaling property |
|---|---|
| $\alpha \sim 0$ | uncorrelated |
| $0 < \alpha \leq 1$ | correlation or LRD |
| $\alpha > 1$ | self-similarity |

Guessing scaling of the signal with LD: uncorrelated for [1,4], self-similar for [6,15] ?

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Back to scaling detection

Principle for 3D-LD:

1. break the process into constant-size time intervals
2. compute the LD over each interval
3. plot in 3D the time-evolution of the LD

Scaling detection:

- same principle as for LD
- but additional time dimension...
- now checking for stationarity in scaling.

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Back to scaling detection (contd)

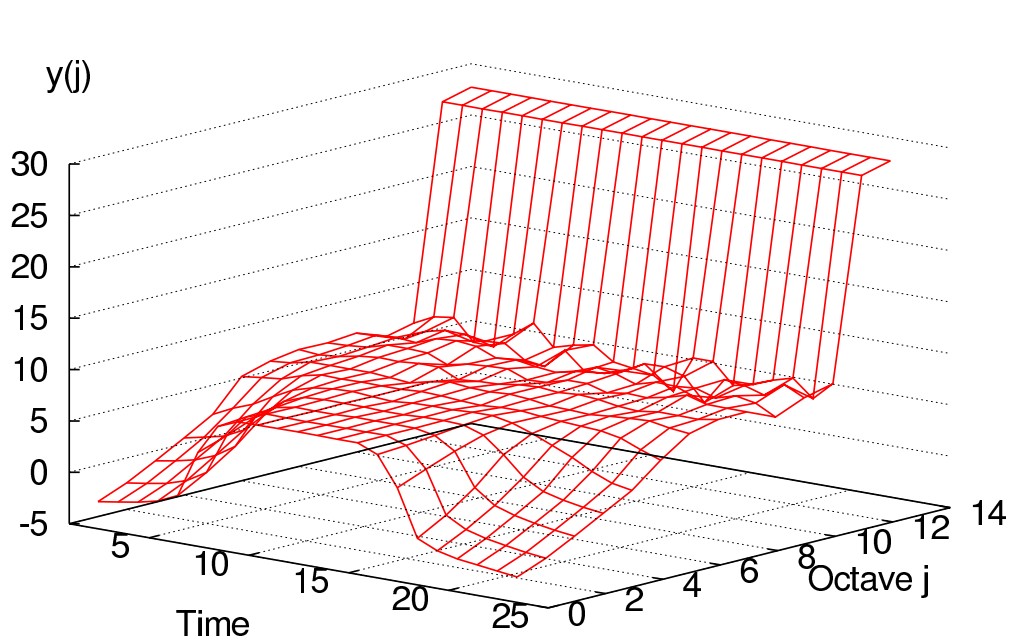
Statistical nature of the process? (3D-LD version)

| Estimator | LD time-dependency | Type of stationarity |
|---|---|---|
| 3D-LD | none | strict stationarity |
| | change in LD level | quantitative non-stationarity |
| | no change in LD slope | qualitative stationarity |
| | slope of LD changes with time | qualitative non-stationarity |

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# 3D-LD of toy example

3D-LD for the “toy” simulation




A better guess for scaling ?

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Conclusions

- LD considers each timescale as a homogeneous (stationary) process
- 3D-LD allows for **identifying changes in scaling with time**
- 3D-LD is **more than just a binary stationarity check**: it also tells what type of (non-)stationarity

Infonet Group
University of Namur
www.infonet.fundp.ac.be

# Case study: TCP flow arrivals

See paper for details...also longer tech-report available at
http://www.info.fundp.ac.be/∼suh/3D-LD/tech-report.ps.gz

A scaling model for flow arrivals:

1. “time of the day” and “day of the week” seasonalities for timescales longer than hours,
2. self-similar process for timescales between minutes and hours,
3. non-stationary correlations for timescales between seconds and minutes,
4. non-stationary Poisson process following the first two components.