



Fourier Analysis and Change Detection

Prof. L. Richard Carley

carley@ece.cmu.edu



Carnegie Mellon

Center for Computational Analysis of
Social and Organizational Systems
<http://www.casos.cs.cmu.edu/>



Dynamic Network Analysis

- Key focus – Networks change over time
- Summary statistics – typically average all data
 - Useless for seeing changes over time
- Longitudinal Networks and Change
 - Getting longitudinal networks from communications logs
 - Stability, Evolution, Shock, Mutation
- Statistical Models of Networks to Detect Change
 - Link Probability Model (LPM) for Stability
 - Actor-Oriented Models for Evolution
 - Multi-Agent Simulation for Evolution, Shock, and Mutation
- Network Change Detection Algorithms
- Fourier Analysis to remove periodic variations



Carnegie Mellon
IST Institute for Software Research

Types of Changes in Network Data

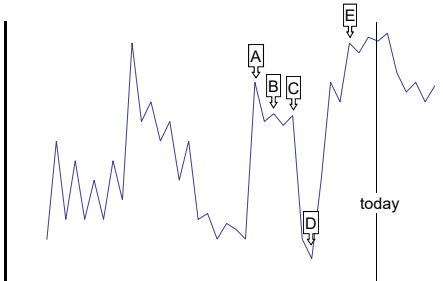
- Stability: Relationships remain statistically the same over time
 - If you are a signal processing person, the Network is “Ergodic”
- Evolution: Interaction among agents cause the relationships to change over time.
 - All link weights / costs are evolving over time during observations
- Shock: Change is exogenous to the social group.
 - E.g., like an earthquake hits Southern California
- Mutation: A shock stimulates evolutionary behavior.
 - E.g., after earthquake, people form many new links trying to survive

CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 3

Carnegie Mellon
IST Institute for Software Research

Changes in Network Data

- Various measures of a network are taken for a window at each time point.
- Change detection: quickly determine *that* a change occurs.
- Change point identification: *when* did the change occur.



CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 4

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Change Detection Hands-On

- Based on Roger Federer 2010 data

ORA 3.0.9.71
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations System Help

Meta-Network Manager
Dynamic-Meta-Network: Roger Federer 2010

Name: Roger Federer 2010
Filename: /Users/rfc/Dropbox/DATA_LRC_SI-2018/Roger Federer 2010-Dates Fixed.xml

Keyframes and deltas time stamped by: Date Period

Statistics:
Keyframe count: 12
Delta count: 0
Earliest date: 2010-01-01 00:00:00
Latest date: 2010-12-01 00:00:00

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 5

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Change Detection Hands-On

- Analysis uses over-time changes in “measures” based on the network data

ORA 3.0.9.71
Simulations Visualizations System Help

Measure Charts...
View Measures Over Time...
View Networks Over Time...
View Networks Over Time in 3D...
View Network Distance over Time...
Vector Maps
Network Drill Down...
Node Cloud...
Color Grid...
Network Block...
GeoSpatial Networks...
Region Viewer...
View Trails...
View Trails in GIS...
View Networks

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 6



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Select The Metanetwork

Start: 2010-01-01 00:00:00 Skip:

Measures Over Time

Select Meta-Networks

Select the meta-networks in which to compute measures.

- 10 - Roger Federer 2010

Run Close < Back Next >

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 7

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Custom Measure Selection

Computation Parameters

Choose which measures to compute, and whether to combine and transform datasets using the controls below.
The meta-networks will be identified by date.

Measures Aggregate Select Transform

- All measures
- Only fast measures
- Centrality measures
- Custom [Click to select...](#)

CASOS

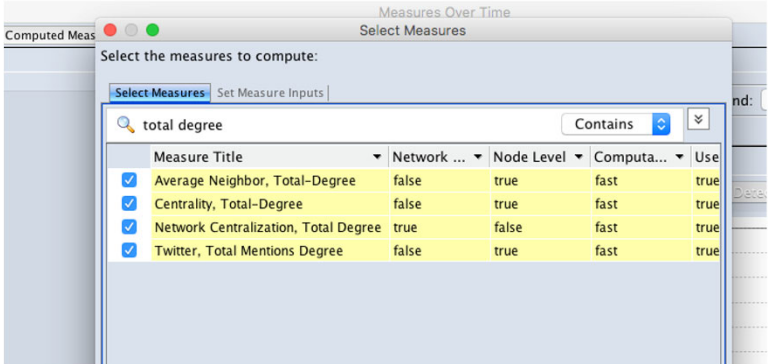
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 8



Carnegie Mellon
 ist Institute for SOFTWARE RESEARCH

Use Search to Find Measure

Hint: Click Select Box at bottom to deselect all measure,
 Then use search to find the ones you want



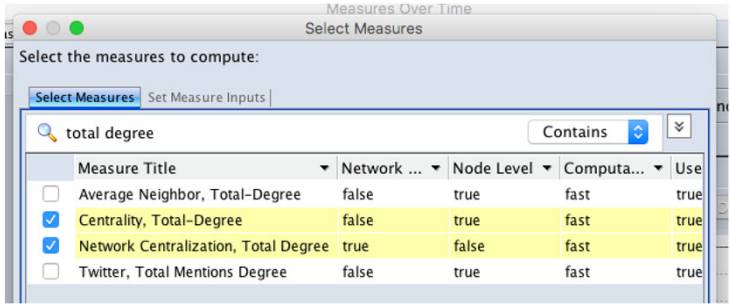
The screenshot shows a window titled 'Measures Over Time' with a sub-window 'Select Measures'. A search bar contains 'total degree'. Below the search bar is a table of measures with columns for Measure Title, Network, Node Level, Computation, and Use. All measures have their 'Use' checkboxes checked.

Measure Title	Network ...	Node Level	Computation	Use
Average Neighbor, Total-Degree	false	true	fast	<input checked="" type="checkbox"/>
Centrality, Total-Degree	false	true	fast	<input checked="" type="checkbox"/>
Network Centralization, Total Degree	true	false	fast	<input checked="" type="checkbox"/>
Twitter, Total Mentions Degree	false	true	fast	<input checked="" type="checkbox"/>

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 9

Carnegie Mellon
 ist Institute for SOFTWARE RESEARCH

Two measures selected to run



The screenshot shows the same 'Measures Over Time' window. The search bar still contains 'total degree'. In the table, the 'Use' checkboxes for 'Centrality, Total-Degree' and 'Network Centralization, Total Degree' are checked, while the others are unchecked.

Measure Title	Network ...	Node Level	Computation	Use
Average Neighbor, Total-Degree	false	true	fast	<input type="checkbox"/>
Centrality, Total-Degree	false	true	fast	<input checked="" type="checkbox"/>
Network Centralization, Total Degree	true	false	fast	<input checked="" type="checkbox"/>
Twitter, Total Mentions Degree	false	true	fast	<input type="checkbox"/>

Add Measure – Agent Based Measure – select “Centrality, Total Degree”

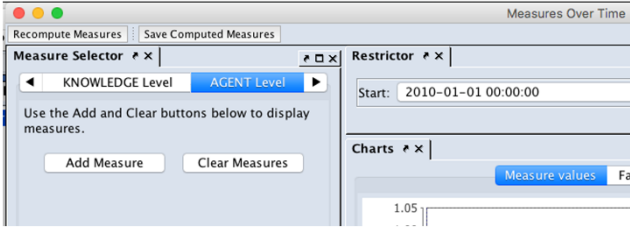
CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 10



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Now Select Display

First step is to select type of variables to display
 – AGENT Level in this case



Then click on “Add Measure” to add a new plot line

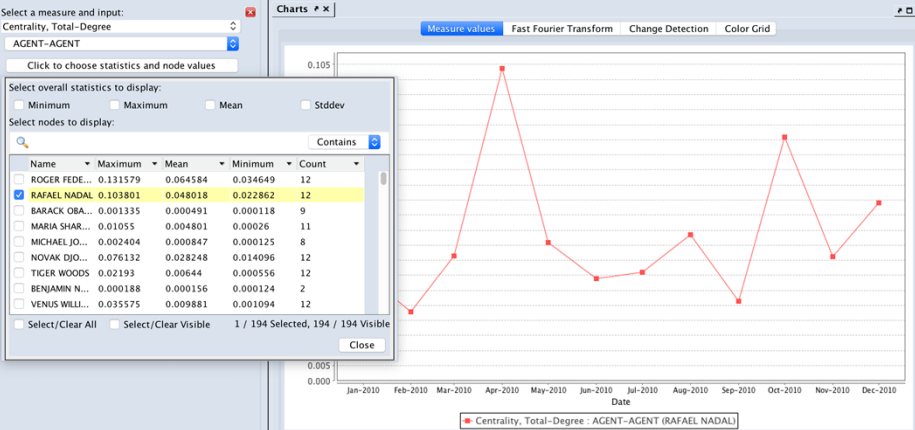
CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 11

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Select Agent for Measure

Click on the “Click to choose” button
 and select second agent for analysis (Federer is always primary)



Name	Maximum	Mean	Minimum	Count
ROGER FEDE...	0.131579	0.064584	0.034649	12
RAFAEL NADAL	0.103801	0.048018	0.022862	12
BARACK OBA...	0.001335	0.000491	0.000118	9
MARIA SHAR...	0.01055	0.004801	0.00026	11
MICHAEL JO...	0.002404	0.000847	0.000125	8
NOVAK DJO...	0.076132	0.028248	0.014096	12
TIGER WOODS	0.02193	0.00644	0.000556	12
BENJAMIN N...	0.000188	0.000156	0.000124	2
VENUS WILL...	0.035575	0.009881	0.001094	12

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 12



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Add Another Agent Measure Plot

Click on Add a Measure button again to add another line

Select a measure and input:
 Centrality, Total-Degree
 AGENT-AGENT

Click to choose statistics and node values

Select overall statistics to display:
 Minimum Maximum Mean Stddev

Select nodes to display:

Name	Maximum	Mean	Minimum	Count
ROGER FEDE...	0.131579	0.064584	0.034649	12
RAFAEL NADAL	0.103801	0.048018	0.022862	12
BARACK OBA...	0.001335	0.000491	0.000118	9
MARIA SHAR...	0.01055	0.004801	0.00026	11
MICHAEL JO...	0.002404	0.000847	0.000125	8
NOVAK DJO...	0.076132	0.028248	0.014096	12
TIGER WOODS	0.02193	0.00644	0.000556	12
BENJAMIN N...	0.000188	0.000156	0.000124	2
VENUS WILLI...	0.035575	0.009881	0.001094	12

Select/Clear All Select/Clear Visible 1 / 194 Selected, 194 / 194 Visible

Save Chart As... Date markers...

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 13

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Change Detection Hands-On The Shewhart X-Bar Chart

Charts x x Measure values Fast Fourier Transform Change Detection Color Grid

Shewhart X-Bar
 Use filtered Fourier val...
 Networks in control: 5
 Risk: 0.05

of networks used to fit normal distribution
 False positive probability

No Change detected

Nadal
 Venus

Value
 Date

Centrality, Total-Degree : AGENT-AGENT (RAFAEL NADAL)
 Centrality, Total-Degree : AGENT-AGENT (VENUS WILLIAMS)

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 14



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Change Detection Hands-On The Shewhart X-Bar Chart

Measure values | Fast Fourier Transform | **Change Detection** | Color Grid

Shewhart X-Bar
 Use filtered Fourier val...
 Networks in control:
 Risk:

of networks used to fit normal distribution
 False positive probability

Change detected

Nadal
 Venus

Value
 Date

Centrality, Total-Degree : AGENT-AGENT (RAFAEL NADAL)
 Centrality, Total-Degree : AGENT-AGENT (VENUS WILLIAMS)

CASOS
 June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 15

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Change Detection Hands-On CUMSUM Method

Measure values | Fast Fourier Transform | **Change Detection** | Color Grid

CUSUM
 Use filtered Fourier val...
 Networks in control:
 Standardized Change:
 Set sensitivity to false alarm:
 Decision interval:
 Risk: 0.05
 Observations:

The δ parameter

Change detected

Value
 Date

Centrality, Total-Degree : AGENT-AGENT (RAFAEL NADAL) decrease
 Centrality, Total-Degree : AGENT-AGENT (RAFAEL NADAL) increase
 Centrality, Total-Degree : AGENT-AGENT (VENUS WILLIAMS) decrease
 Centrality, Total-Degree : AGENT-AGENT (VENUS WILLIAMS) increase

CASOS
 June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 16



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Fast Fourier Transform (FFT)

- Goal: detect periodicity in over-time data
- Examples
 - Weekly periodicity in email data
 - Time of the day effects
- Fourier's theorem
 - Any time signal can be represented by a sum of sinusoidal functions with different frequencies, amplitudes and phase shifts
- Fourier transform finds sinusoids that decompose a signal
 - Analogy: given a dish, find the ingredients
 - Sinusoids have the advantage that they are orthogonal

CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 17

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Sinusoidal Function

- A sinusoidal function $y = A \sin(2\pi ft + \phi)$ has
 - A amplitude
 - f frequency ($T = \frac{1}{f}$ is the period)
 - ϕ phase

$y = 3 \sin(2\pi(0.25)(t + 0.5))$

CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 18

Carnegie Mellon
 Institute for SOFTWARE RESEARCH

Frequency Domain: The FFT

- Fast Fourier Transform (FFT)
- Decompose time waveform into sum of sinewaves

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 19

Carnegie Mellon
 Institute for SOFTWARE RESEARCH

FFT Example: Sum of Sinusoids

- Fast Fourier transform of sinusoidal function is a spike at the sinusoidal frequency
- Example $y = \sin(2 \pi 0.25 t)$

Over-time signal

FFT

CASOS

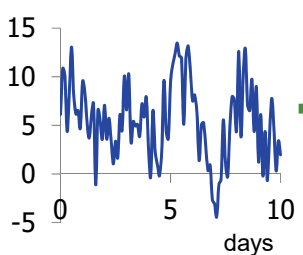
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 20



Carnegie Mellon
 Institute for SOFTWARE RESEARCH


FFT Example 2

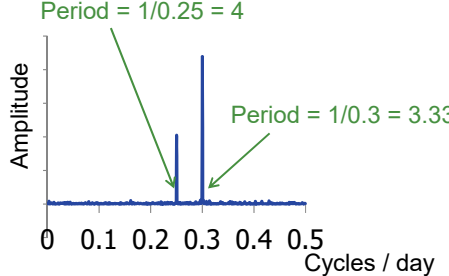
- FFT finds periodicities that may be unclear in over-time signal



Over-time signal

FFT





FFT

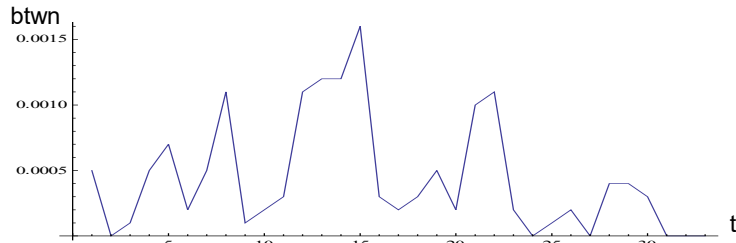
Hidden "recipe": over – time signal computed as
 $y(t) = 2 \sin(2 \pi 0.25 t) + 3 \sin(2 \pi 0.3 t + 0.2) + \text{noise}$

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 21

Carnegie Mellon
 Institute for SOFTWARE RESEARCH

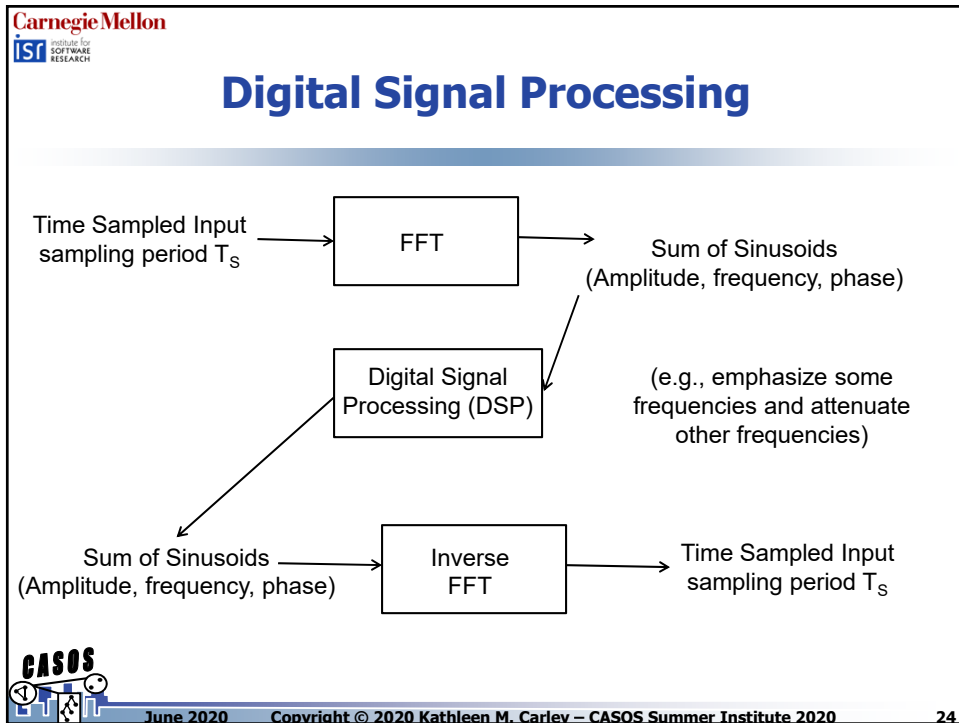
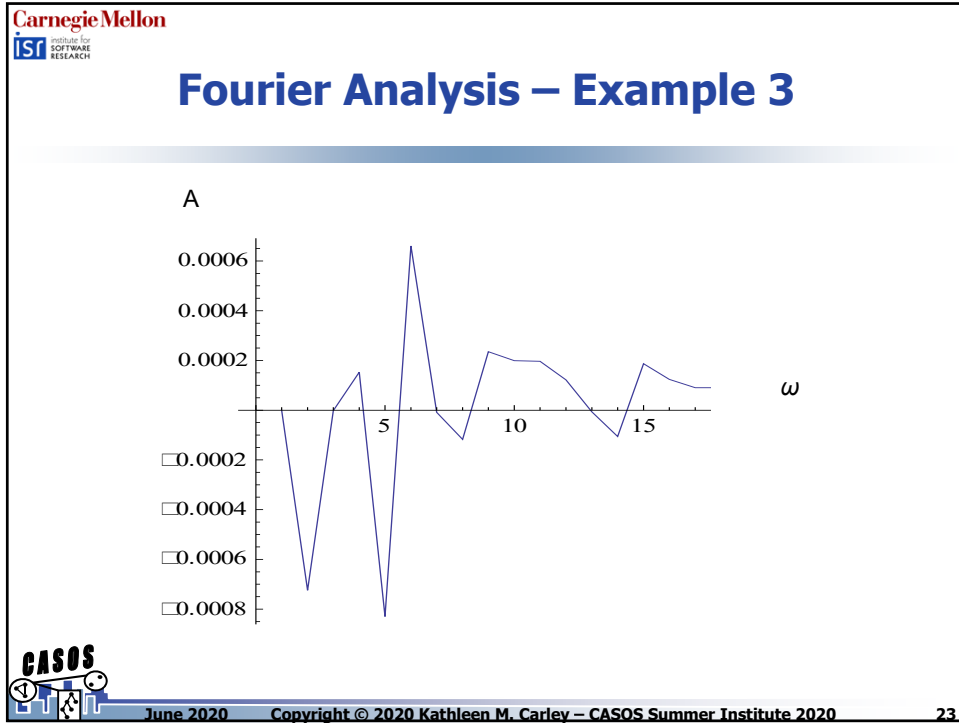
Fourier Analysis Example 3

- 24 cadets in a regimental chain of command agreed to have their email monitored to form a social network data set known as IkeNet3.
- The betweenness was calculated based on the e-mail communications observations over the first month in their duty positions.



CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 22





Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

One Key Issue: Aliasing Sampling Limits

Sampling freq. $F_s = 100\text{kHz}$

If we want DSP to work unambiguously, we choose to limit maximum input frequency to $\leq 1 / (2 T_s) = F_s / 2 = \text{"Nyquist Frequency"}$

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 25

Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

Digital Signal Processing

How do we use this?

One possible approach – big peaks are periodic “background”

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 26



Carnegie Mellon Institute for SOFTWARE RESEARCH

Nonlinear Filtering

e.g., identify 3 main (high magnitude) components
keep them and remove FFT components
at all other frequencies

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 27

Carnegie Mellon Institute for SOFTWARE RESEARCH

Inverse Fast Fourier Transform

This is the inverse Fourier transform of just the 3
selected components, which are then reconverted to time waveform

There is a weekly, two week and three week cycle

CASOS June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 28



Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

Anomaly Detection

The filtered pattern has been subtracted from the original
 The red is what is left – the anomalies

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 29

Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

FFT Example Hands-On

- IkeNet data (IkeNet3-dynamic.xml)
 - Email exchange data among mid-career officers in a one-year graduate program at Columbia University
 - Granularity: day; Duration: month

File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations System Help

Meta-Network Manager

Dynamic-Meta-Network: Roger Federer 2010

Name Roger Federer 2010

Filename /Users/lrc/Dropbox/DATA_LRC_SI-2018/Roger Federer 2010-Dates Fixed

Generate Reports... Visualize Measure Charts... View T

Open

DATA_LRC_SI-2018

Load everything
 Load some things:
 Load sources
 Load networks

Name	Date Mo.
Flightpaths.good.xml	Friday...
IkeNet3-dynamic.xml	Wedn...
Matrix Data	Mond...
Raiders of the Lost Ark - Dynamic.xml	Tuesd...

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 30



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Over Time Measures

Again View Measures Over Time

ORA 3.0.9.9.71

File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations System Help

Meta-Network Manager Dynamic-Meta-Network: IkeNet3

Name: IkeNet3
Filename: /Users/lrc/Dropbox/DA...
Generate Reports...
Keyframes and deltas time...
Statistics:
Keyframe count: 30
Delta count: 0
Earliest date: 2008-09-01 00:00:00
Latest date: 2008-09-30 00:00:00

Visualizations menu:
Measure Charts...
View Measures Over Time...
View Networks Over Time...
View Networks Over Time in 3D...
View Network Distance over Time...
Vector Maps
Network Drill Down...
Node Cloud...
Color Grid...
Network Block...
GeoSpatial Networks...
Region Viewer...
View Trails...
View Trails in GIS...
View Networks

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 31

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Selecting MetaNetwork

Measures Over Time

Select Meta-Networks

Select the meta-networks in which to compute measures.

IkeNet3

Run Close < Back Next >

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 32



Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

Selecting Measures

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 33

Carnegie Mellon
 ISI Institute for SOFTWARE RESEARCH

Selecting Measures to Plot

Density, Weighted & eAgent2eAgent

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 34



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Run Shewhart to Detect Change

Run Change Detection Shewhart
 10 points for basis
 0.05 for false positives

Lots of apparent events
 That could be Change

All upward changes

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 35

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

FFT Can Help with Periodic Patterns

Network Level / Density, Weighted / eAgent2eAgent

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 36



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

FFT Power Spectral Density

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 37

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

FFT Power vs. Period

Adjust Power Threshold to just select “background” periodic variations to be removed

NOTE – viewing the Dominant Periods is often more intuitive than PSD

The long period values are artifact of finite length of data

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 38



Carnegie Mellon
IST Institute for Software Research

Original vs. FFT subtracted data

Remaining FFT components are converted back to time values (inverse FFT) and subtracted

This removes some of the periodic patterns covering up the fundamental changes to be detected

Charts x

Measure values | Fast Fourier Transform | Change Detection | Color Grid

Power Spectral Density
Maximum: 0.111737
Average: 0.050062
Std.dev: 0.029496

Power threshold: 0.068397

Dominant Frequencies
 Dominant Periods
 Dominant Values
 Filtered Values
 Original Values
 Original and Filtered Values

Measure Value

Date

Density, Weighted: eAgent2eAgent original
Density, Weighted: eAgent2eAgent filtered

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 39

Carnegie Mellon
IST Institute for Software Research

Shewhart on FFT subtracted data

Now use the Filtered Data

Run Change Detection Shewhart
10 points for basis
0.05 for false positives

Only 1 + event and 2 - events

Open Question – is this better than original or not?

Charts x

Measure values | Fast Fourier Transform | Change Detection | Color Grid

Shewhart X-Bar

Use filtered Fourier values

Networks in control: 10

Risk: 0.06

Value

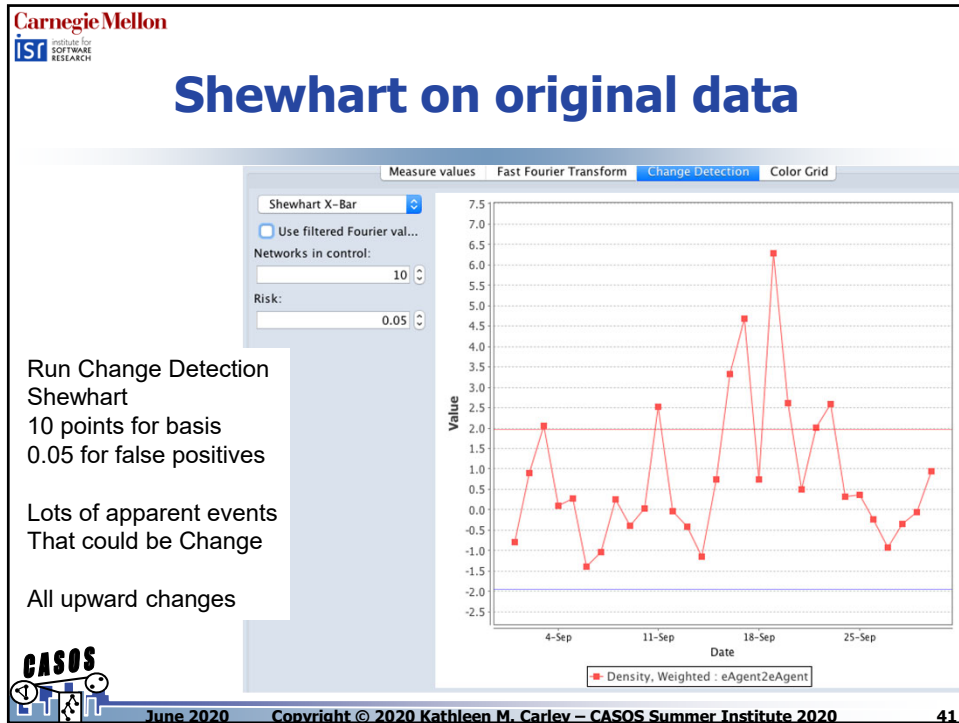
Date

Density, Weighted: eAgent2eAgent

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 40





Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Fourier Analysis to Handle Periodicity

- Fourier analysis can effectively identify periodic trends in longitudinal network data.
- Identification of periodic trends can allow the analyst to aggregate relational data over the period to remove over-time dependence.
- The inverse Fourier transform of the significant period can be used to subtract off periodicity from longitudinal network data measures over time.
- Further exploration of wavelets may produce greater insights in to network dynamics.

CASOS

June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 42



Carnegie Mellon
IST Institute for Software Research

Scalability

- The change detection algorithm is linear, thus the time consuming part is calculating network measures.
- Networks with less than 20 nodes tend to have a higher variance in over time measures. When a link is added or removed, it affects $(n-1)(n-2)$ triads.
- Requires at least 3 time periods: >2 to determine typical behavior and 1 to compare at each time point. In practice, 10+ network time points are preferred.
- No difference in number of required networks for each technique: CUSUM, EWMA, Scan Statistic, x-bar, etc.
- Wavelet/Fourier based approach needs many more time periods and complexity grows roughly as $\#T(\log(\#T))$

CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 43

Carnegie Mellon
IST Institute for Software Research

Limitations

- View findings on data with caution
- Slicing and dicing can distort conclusions
- Examine errors associated with technique through extensive simulations.
- Investigate more real world data sets.
- Investigate the degree to which network measures are correlated to understand the effects of compounding error.
- Investigate multi-dimensional network properties such as the cosine similarity between the triad census at different time periods.


CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 44



Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Summary of Change Detection


- Rapid change detection may allow an analyst to get inside a decision cycle and shape network evolution.
- Simulation is important for modeling longitudinal network behavior.
- Isolating when networks change enables more focused study on the causes of evolution, shock, and mutation, which may lead to future predictive analysis.
- Statistical process control is a useful tool for understanding social behavior.

 CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 45

Carnegie Mellon
IST Institute for SOFTWARE RESEARCH

Conclusions

- Change detection
 - Detect occurrence of shocks i.e. change due to reasons exogenous to the network
- Fourier analysis
 - Detect periodicity in over-time data

 CASOS
June 2020 Copyright © 2020 Kathleen M. Carley – CASOS Summer Institute 2020 46

