Mining Recurrent Activities: Fourier Analysis of Change Events

Abram Hindle, Michael W. Godfrey, Richard C. Holt

Software Architecture Group
David R. Cheriton School of Computer Science
University of Waterloo
Canada
http://swag.uwaterloo.ca/

\{ahindle,migod,holt\}@cs.uwaterloo.ca
Fourier Transform of MySQL 5.0

Frequency

Magnitude

Time
Developers
create
software
together
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Devs change software incrementally. They commit revisions to the project.
Aggregation
Aggregation
revision count

time
Timeseries

Revision

Time
Discrete Events ➔ Signal
Does software follow earth cycles, or its own?
Maybe instead of assuming we should go ahead..
we should go ahead and look
Recurrent Behaviour
Let us rotate this...
An iteration
Local recurrence?
Fourier transform:
* recurrent behaviour
* periodicities
Color / Magnitude

Time/X axis

Magnitude / Y axis
Fourier Transform of MaxDB 7.500
Fourier Transform of MySQL 3.23

Frequency Domain

Frequency

Magnitude

Time
The smears seen on the MySQL plot have a period of about 7-8 days (1 week).
Similar Smears, Different Projects

* Evolution
* Mozilla
* Xerces
* MaxDB 7.6
Future Work
What caused this?
A signal composed of slices

A signal can be composed of slices
A behaviour explained by slices

Different slices can be responsible for subsignals
Contributions:

* applied the Fourier transform to software repositories
Contributions:
* showed recurrent behaviour in FLOSS projects.
(mailing lists)