

Visualize and Analyze Time-series IoT Data

Dr. Hui Zhang, Ranjini Subramanian

RET18



Data Science

Data Science

Extraction of knowledge from large volumes of data that are structured or unstructured

1

Data Collection

Raw Data

2

Processing

*Clean bad format,
missing data, useless
variables, wrong
data, transform
variables*

3

Data Exploration

*Understanding data,
modeling,
visualization*

4

Analysis and
Prediction

*Extract knowledge,
models to describe or
predict data*

Project

Introduction - Visualize and Analyze Time-series IoT Data

- The internet of things has numerous applications in healthcare
 - remote monitoring ,smart sensors and medical device integration
- Healthcare IoT has the potential to
 - keep patients safe and healthy
 - improve how physicians deliver care as well
 - boost patient engagement and satisfaction by allowing patients to spend more time interacting with their doctors.

What is a Time Series?

- Set of evenly spaced numerical data
 - Obtained by observing response variable at regular time periods
- Forecast based only on past values
 - Assumes that factors influencing past, present, & future will continue
- Example

Year:	2001	2002	2003	2004	2005
Sales:	78.7	63.5	89.7	93.2	92.1

Objective

- To find of the temporal relationship of “events” identified in the data collected from different sources.
- **The underlying hypothesis** is that the sequential patterns will spring from the large amount of multi-dimensional data sources if visualized and analyzed in the right and smart way

Weekly Plan

Week 1

- Get familiar with R by taking [Datacamp Free R course](#)
- Install [R](#) and [RStudio](#) (see links above) and try R programming
 - [Try sample R code to create data frames](#)
 - [Try sample R code to read and write CSV files](#)
 - [How to create an R notebook on RStudio](#)
- Paper reading: [Medical Internet of Things and Big Data in HealthCare](#)

Week 2

- Complete [Datacamp Free R course](#)
- Research and hands-on experience: from raw accelerometer CSV data to R data frame
- Paper reading: [Medical Internet of Things and Big Data in HealthCare](#)

The Background

1. Wearables and mobile apps today support symptom tracking, disease management, and care coordination

2. IoT allows different devices to send and receive data enabling better connectivity, data processing, and analytics

3. Data challenges in IoT:

- Large-scale machine generated time series
- resolution, scale, complexity
- visualization and analysis
 - Feature engineering + Data Science

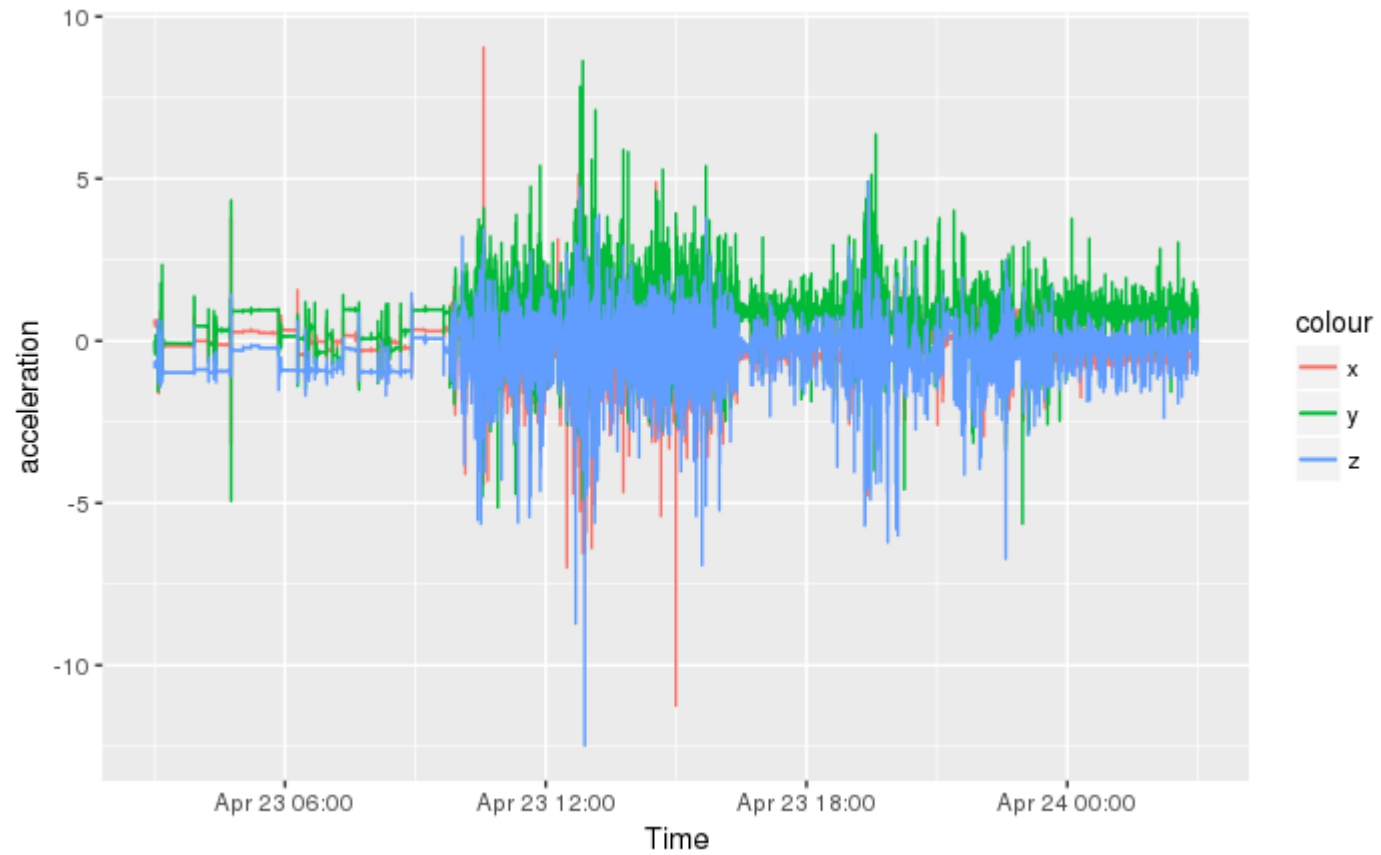
Name	Date modified	Type	Size
4B1AA5_04232017_accelerometer.protobin	1/26/2018 2:56 PM	Microsoft Excel C...	275,033 KB
4B1AA5_04232017_ambient_temp.protobin	1/26/2018 3:08 PM	Microsoft Excel C...	43 KB
4B1AA5_04232017_eda.protobin	1/26/2018 3:07 PM	Microsoft Excel C...	43,000 KB
4B1AA5_04232017_gyroscope.protobin	1/26/2018 3:06 PM	Microsoft Excel C...	346,524 KB
4B1AA5_04232017_humidity.protobin	1/26/2018 3:05 PM	Microsoft Excel C...	44 KB
4B1AA5_04232017_ppg.protobin	1/26/2018 3:03 PM	Microsoft Excel C...	88,371 KB
4B1AA5_04232017_pressure.protobin	1/26/2018 3:00 PM	Microsoft Excel C...	46 KB
4B1AA5_04232017_wearing.protobin	1/29/2018 4:22 PM	Microsoft Excel C...	1 KB
4B1AA5_04242017_accelerometer.protobin	1/26/2018 3:45 PM	Microsoft Excel C...	274,996 KB
4B1AA5_04242017_ambient_temp.protobin	1/26/2018 3:44 PM	Microsoft Excel C...	43 KB
4B1AA5_04242017_eda.protobin	1/26/2018 3:44 PM	Microsoft Excel C...	42,999 KB
4B1AA5_04242017_gyroscope.protobin	1/26/2018 3:44 PM	Microsoft Excel C...	346,811 KB
4B1AA5_04242017_humidity.protobin	1/26/2018 3:44 PM	Microsoft Excel C...	44 KB
4B1AA5_04242017_ppg.protobin	1/26/2018 3:43 PM	Microsoft Excel C...	88,414 KB
4B1AA5_04242017_pressure.protobin	1/26/2018 3:43 PM	Microsoft Excel C...	46 KB
4B1AA5_04242017_wearing.protobin	2/12/2018 11:28 AM	Microsoft Excel C...	2 KB
4B1AA5_04252017_accelerometer.protobin	1/26/2018 3:11 PM	Microsoft Excel C...	273,924 KB
4B1AA5_04252017_ambient_temp.protobin	1/26/2018 3:12 PM	Microsoft Excel C...	43 KB
4B1AA5_04252017_eda.protobin	1/26/2018 3:13 PM	Microsoft Excel C...	43,000 KB
4B1AA5_04252017_gyroscope.protobin	1/26/2018 3:13 PM	Microsoft Excel C...	346,513 KB
4B1AA5_04252017_humidity.protobin	1/26/2018 3:14 PM	Microsoft Excel C...	44 KB
4B1AA5_04252017_in_bed.protobin	1/29/2018 4:38 PM	Microsoft Excel C...	1 KB
4B1AA5_04252017_in_bed_heart_rate.prot...	1/26/2018 3:15 PM	Microsoft Excel C...	331 KB
4B1AA5_04252017_ppg.protobin	1/26/2018 3:15 PM	Microsoft Excel C...	88,624 KB
4B1AA5_04252017_pressure.protobin	1/26/2018 3:16 PM	Microsoft Excel C...	46 KB
4B1AA5_04252017_wearing.protobin	2/12/2018 11:51 AM	Microsoft Excel C...	2 KB
4B1AA5_04262017_accelerometer.protobin	1/26/2018 3:21 PM	Microsoft Excel C...	155,777 KB
4B1AA5_04262017_ambient_temp.protobin	1/26/2018 3:22 PM	Microsoft Excel C...	24 KB
4B1AA5_04262017_eda.protobin	1/26/2018 3:22 PM	Microsoft Excel C...	24,297 KB
4B1AA5_04262017_gyroscope.protobin	1/26/2018 3:23 PM	Microsoft Excel C...	196,572 KB
4B1AA5_04262017_humidity.protobin	1/26/2018 3:24 PM	Microsoft Excel C...	25 KB

Week 3

- Research and hands-on experience: convert R data frame to plotting
- Paper reading: [Health Monitoring and Management Using IoT Sensing with Cloud-based Processing: Opportunities and Challenges](#)

The RET Goal

Research and evaluate visualization and analysis methods for time-series data

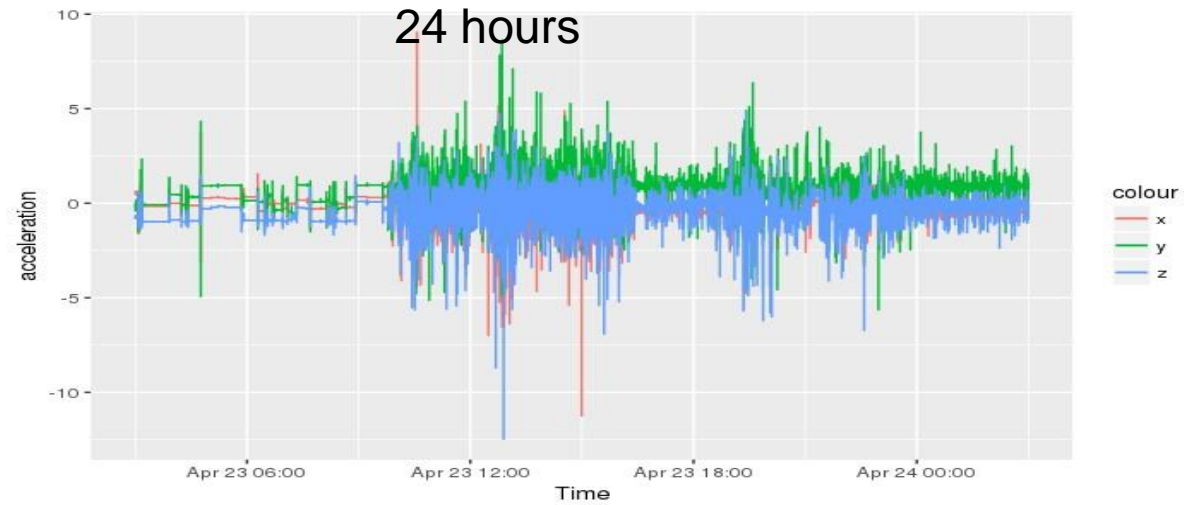


Week 4

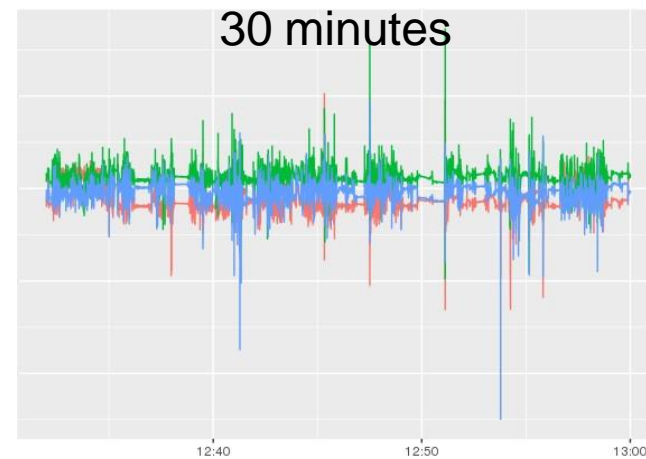
- Research and hands-on experience: R interactive plotting
- Paper reading: [Health Monitoring and Management Using IoT Sensing with Cloud-based Processing: Opportunities and Challenges](#)

The RET Goal

Research and evaluate visualization and analysis methods for time-series data



exploration

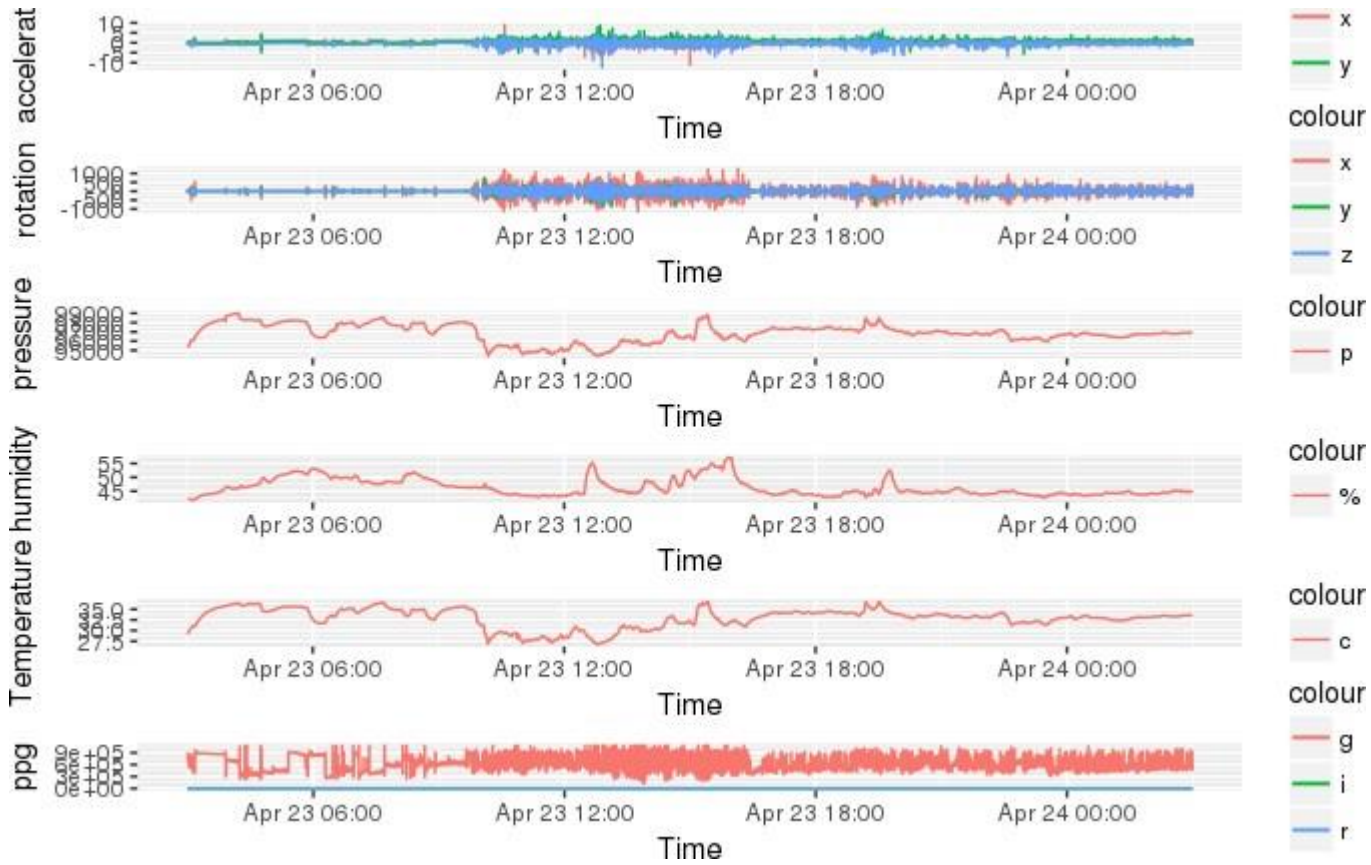


Week 5

- Research and hands-on experience: plotting IoT data from multiple sources
- Paper reading: [Revolutionizing the HealthCare Industry with Big Data, Analytics, and Visualization](#)

The RET Goal

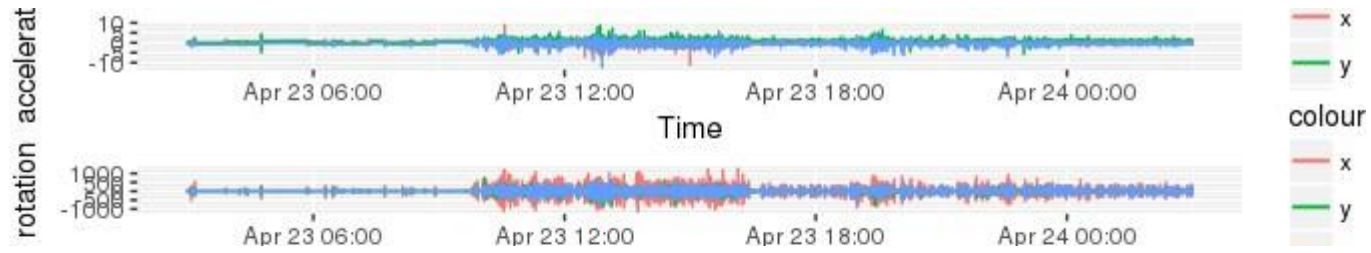
Research and evaluate visualization and analysis methods for time-series data



synchronization
Re-grid

The RET Goal

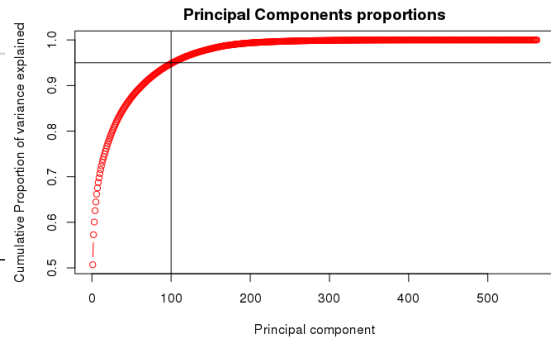
Research and evaluate visualization and analysis methods for time-series data



Feature computation

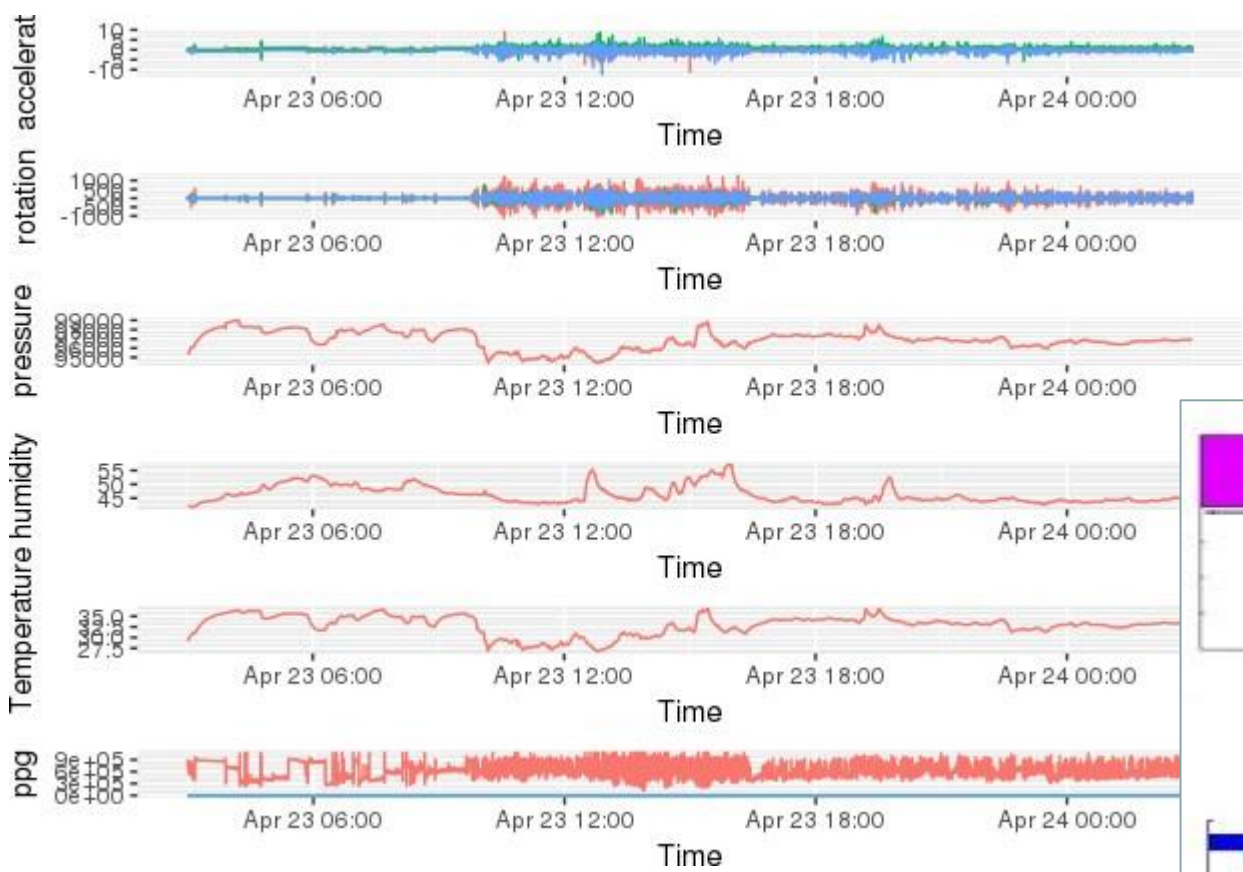
	tBodyAcc.mean...X <dbl>	tBodyAcc.mean...Y <dbl>	tBodyAcc.mean...Z <dbl>	tBodyAcc.std...X <dbl>	tBodyAcc.std...Y <dbl>
1	0.2885845	-0.02029417	-0.1329051	-0.9952786	-0.9831106
2	0.2784188	-0.01641057	-0.1235202	-0.9982453	-0.9753002
3	0.2796531	-0.01946716	-0.1134617	-0.9953796	-0.9671870

3 rows | 1-6 of 563 columns

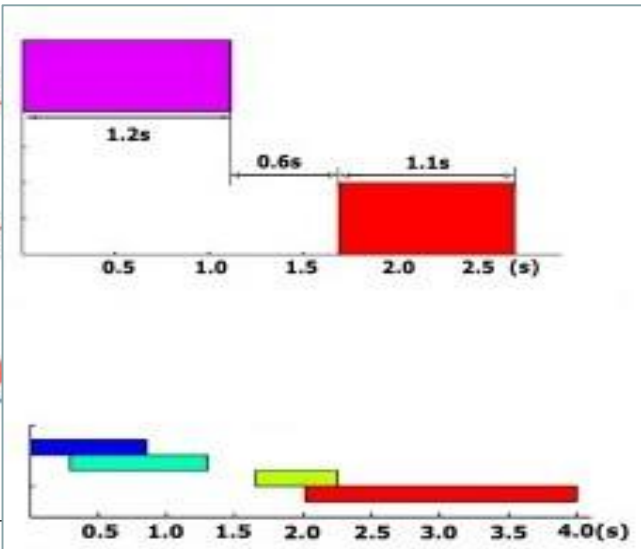


The RET Goal

Framework of visualization and analysis methods for time-series data



Correlation analysis



Week 6

- Research and hands-on experience: experiment with classifying time series
- Paper reading: [Revolutionizing the HealthCare Industry with Big Data, Analytics, and Visualization](#)

Technique, Languages, Platform

- Prototype of basic methods in R and its packages (ts, ggplot)
- Towards a Shiny/R app (Shiny/R)
- Resources - <http://cecsresearch.org/vcl/RET18/>