Collaborative Filtering for Coordinated Monitoring in Sensor Networks

Janne Toivola & Jaakko Hollmén

Department of Information and Computer Science, Aalto University School of Science, Finland
An idea...

- Collaborative filtering useful in a wireless sensor network?
Outline

- **Background**
  - Structural Health Monitoring (SHM)
  - Wireless Sensor Networks (WSN)

- **Coordinated monitoring problem**
  - Collaborative Filtering (CF)
  - Random assignment and Majority voting

- **Experiments and results**
  - Wooden model bridge
  - Aluminum bookshelf

- **Summary**
Background

- Multidisciplinary research project
  - Intelligent Structural Health Monitoring System (ISMO)
  - Applied mechanics, data mining, sensor networks...

- Structural Health Monitoring
  - Assessing the condition of man-made structures

- Data mining
  - Extracting relevant information from data

- Wireless Sensor Networks
  - Distributed computing environment
Structural health monitoring

- No sensors that directly indicate damage
  - Extracting features from vibration measurements
  - Sensitive to damages, not to environmental variability
  - Transmissibility magnitude: how well vibrations of certain frequency propagate between two sensors

- Novelty detection problem
  - Not allowed to damage the (unique) structure to collect training samples
  - Supervised feature selection not possible
  - Maybe some additional information available?
Wireless sensor networks

- Eliminate costly and error-prone wires
- Restricted bandwidth and computation
  - Simple features can be computed locally on a node
  - Number of transmitted features should be low
- Each sensor node should concentrate on measuring relevant features!
- Relevance depends on what the other nodes are measuring => coordination problem!
Coordinated monitoring

- Monitor at most D local features / node
  - An upper limit on computation & communication
  - Sparsity helps also the detection algorithm

- Each local feature can be given a rating
  - Encodes *additional application-specific information*
  - The above restriction holds: only D ratings at a time

- Computation of combined features
  - Requires measuring certain sets of local features
  - Local ratings combined to select monitored features
Collaborative filtering

- Traditionally in recommendation systems
  - *Users* give ratings to *items*
  - Recommend new items to a user based on ratings from other similar users

- Assumes structure in the rating data
  - Similar users like prefer same items

- Needs to deal with the sparsity of ratings

- WSN: recommend features to nodes!
  - some similarities & differences to conventional CF...
Baseline: random & voting

- What if it doesn’t matter which features to monitor / all are equally good?
  - Let’s try also random selection, at most D/node

- What if the sensors don’t need to specialize / just one global choice enough?
  - Let’s try majority voting based on the same ratings
  - Number of combined features may still be high due to the number of combinations...
The proposed architecture

- **Accelerometers:**
  - time series $x_s[n]$

- **Local features:**
  - power spectrum $X_s[k]$

- **Combined features:**
  - transmissibility
  $$T[s_1,s_2,k] = \frac{X_{s_1}[k]}{X_{s_2}[k]}$$

- **Novelty detector:**
  - k-NN or Gaussian
Frequency domain ratings

- Mean power spectrum $X_s[k]$
  - run FFT etc.
  - over $t$ time windows

- Sort by power

- For top-D bins
  - $r[s,k] += 1$
  - Default vote 0

- Repeat $M_{CF}$ times
Combined ratings and CF

- After collecting local ratings $r[s,k]$
- Compute ratings for the combined features
  - $w[s_1,s_2,k] = corr_j(r[s_1,j], r[s_2,j]) \cdot r[s_1,k] \cdot r[s_2,k]$
  - similarity of sensors $s_1$ and $s_2$ measured by correlation
    - symmetric, since transmissibility is "symmetric"
- Sort and select top-D combined features for each sensor node
  - some sensors may be left with less than D features, if other nodes are fully utilized
  - no eliminated items like in recommendation systems
Data: Wooden bridge [19]

- 15 accelerometers and a shaker
- Added weights to simulate damages
- 1998 + 265 time series measurements
  - 32 seconds • 256 Hz = 8192 samples each
Example: Random, D=4
Example: Majority vote, D=4
Example: CF, D=4
Example: CF, D=20
Novelty detection accuracy

- Iterated 10 times
- CF, vote, random
- D=2...20
- k-NN detector
- min, med, max AUC
Novelty detection accuracy

- Iterated 10 times
- CF, vote, random
- \( D = 2 \ldots 20 \)
- Gaussian detector
- \( \min, \med, \max \) AUC

AUC values with Gaussian

Iterated 10 times
- CF, vote, random
- \( D = 2 \ldots 20 \)
- Gaussian detector
- \( \min, \med, \max \) AUC
Data: LANL bookshelf [20]

- 24 accelerometers
- Loosened and removed bolts
- Shaker at the bottom
- 150 + 120 time series measurements
- 5.12 sec • 1600 Hz = 8192 samples each
Novelty detection accuracy

- Iterated 10 times
- CF, vote, random
- \( D = 2 \ldots 20 \)
- k-NN detector
- min, med, max AUC

AUC values with k-NN detector: Iterated 10 times, CF, vote, random, \( D = 2 \ldots 20 \), k-NN detector, min, med, max AUC.
Novelty detection accuracy

- Iterated 10 times
- CF, vote, random
- \(D=2...20\)
- Gaussian detector
- \(\text{min, med, max } AUC\)
Summary

- Problem of coordinated monitoring
- Collaborative filtering as a solution
- Applied as part of SHM system
- Demonstrated with two data sets
  - 8192•S acceleration samples => one detection result
  - CF performed well with wooden bridge data
  - Differences smaller with the LANL bookshelf data
- Some CF problems remain: coverage