Focus Group: Big Data Analytics for Smart Manufacturing Systems

Report by
Sudarsan Rachuri
sudarsan@nist.gov
Improving Manufacturing Efficiency through Predictive Analytics

- 5% decrease in batch cycle time
- 10% improvement in machine reliability
- 10% reduction in water consumption
- 5% reduction in energy costs

Source: www.ge-ip.com
The new Program
Smart Manufacturing Systems Design and Analysis

- **Objective:** The objective is to deliver measurement science, standards and protocols, and tools needed to predict, assess, optimize, and control the performance of smart manufacturing systems.

**Major Projects:**

1. Reference architecture and open solution stack to enable and assess the composable SMS
2. Modeling methodology and associated tools to predict, assess, and optimize the operational performance
3. Data analytics and associated methods and tools to enable adaptive system
4. Methods and tools for system performance assurance.
We need to understand the Predictive Analytics Workflow

- Standardize the predictive models
  - Model definition
  - Model Composition
  - Model chaining

Standards and protocols for this information flow:

1. Data extraction/Data stream
2. Input validation
3. Data Pre-processing
4. Predictive Model
5. Data Post-processing
6. Decision Storage/Decision Processing
7. Data visualization

Outliers, missing values, invalid values
Normalize, Discretize, Filter etc.
Scaling, Decision, Scores etc.

Sender
Receiver
Sender
Receiver

Interface Standard
Protocol Standard
ensures compatibility
define both the transmitter and receiver function at the same time.

Standard
Promise of Big Data Analytics Solution!

<Concept Architecture>

**Application Layer**

- **CAPP**, MES, FDC, YMS, ...

**Integration Layer**

- **Model Life Cycle**
  - Creation → Deployment → In-Use → Disposal
- **Life Cycle Control**
  - Duration Control, Uncertainty Resolution

**Analytics Modeling Layer**

- **Statistics Approach**
  - R, ...
- **Machine Learning Approach**
  - Neural Network, SVM, Decision Tree, ...

**Big Data Infrastructure Layer**

- **R Hive, Hadoop, HDFS, MapReduce, ...**

**Data Layer**

- **Static Data**
  - Process Plan (STEP-NC), Production Plan, Master, ...
- **Dynamic Data**
  - Monitoring (MTConnect), Metrology, Defect, ...

**Shop Floor Layer**

- Manufacturing Process
Focus Group Discussion Points

• Data acquisition issues
  – In Health care, Manufacturing (discrete, continuous)
  – Cost of collecting data
  – Availability of data (real world data and data simulator)
  – What are the real implications of volume, velocity, variety, and veracity?
  – Methods of collecting data (manual, automated)
  – Open Data Repository
Focus Group Discussion Points

• Standards for data acquisition
  – Data attributes (meta data), unstructured (extracted from text,…), structured (standards), data sampling
  – Data access and query
  – Data modeling and data science
  – Safety data, privacy of data (data masking??)
  – Open Data Initiative ??
  – Measurement and metrics for V&V
Focus Group Discussion Points

Analytics Modeling

- Problem classification: No need for DA, Good to have DA, must have DA
- Data driven Models, Architecture for Data Analytics (common issues for Manufacturing and Health Care)
- Model Composition, chaining, reuse
- Correlation to Cause-Effect Analysis
- Analytics Workflow – Standards and Protocols
- Moving analytics to the data
- More research needed in understanding feature vector (minimal or optimal)
- Computing and IT infrastructure for DA
Data Analytics – Past, Present and Future

Reduce the information overload.
Can we get same level of insights with less data?

Data Volume

Past                      Present                      Future

Current DA

Big Data

Relevant and Useful Data

Reduce the information overload.
Can we get same level of insights with less data?