Invigorating the Mass Spectrometer: Software Solution for Tandem MS


Pacific Northwest National Laboratory, Richland, WA, USA

Overview

- We have developed a first generation version of an analysis pipeline to enhance the capabilities of the mass spectrometer using an entirely software-based solution.
- Using an in-house developed implementation of enterprise service bus (ESB), MeDICI\(^1\), we have coupled a number of existing software modules to operate on incoming spectra directly from the mass spectrometer.
- Our pipeline is highly flexible and extensible and allows the connection of disparate software modules to enhance the efficiency of the mass spectrometer.
- Additional benefits include the connection of computing resources such as field programmable gate-arrays, and other high performance computing hardware.

Introduction

- Most studies to date, treat new experiments as an individual entity and no prior experimental knowledge is retained.
- This results in a data explosion with huge amounts of unwanted redundancies and unresolved ambiguities.
- In tandem MS-MS experiments, it is a common phenomenon to regularly fragment the higher abundance signals while the low abundance ones are missed.

Methods

- Based on the principles of enterprise service bus (ESB) architectures and distributed computing, we have developed a pipeline of software modules (Figure 1) that operate on incoming spectra directly from the mass spectrometer.
- A feature discovery module is run to discover expected patterns for isotopic distributions present within the spectra. The list of isotopic profiles is distributed simultaneously to a central visualization component as well as to another component for online alignment in the mass and time dimensions.
- A decision-making module determines whether the isotopic distribution currently under consideration is one that has been frequently observed or needs further investigation.
- Feedback to the mass spectrometer is achieved via a list of mass-to-charge ratios to be selected for fragmentation at the next fragment scan interval.

Conclusions

- First generation pipeline will facilitate the experimental setup in two modes: complete coverage and maximum overlap
- MeDICI allows us to connect disparate software modules for different experimental setups as well as connect computing resources of varying processing capacity
- Extensible, flexible and purely software-based solution to enhance the performance of existing mass spectrometers for tandem MS experiments.

Acknowledgements

This work was funded by The Data-Intensive Computing Initiative at Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is operated for the DOE by Battelle under contract DE-AC05-76RLO-1830.

References


CONTACT: Anuj R. Shah, Ph. D., Scientific Data Management, K7-92, 9020 South 30th Street, P.O. Box 999, Richland, WA 99352
E-mail: anuj.shah@pnl.gov