Breakout Session Discussion

Breakout #3: Show me the Data

Objective: *Identify the critical data requirements for a competition. Discuss potential pathways for the creation and/or acquisition of a large-scale, realistic, challenging dataset.*

November 24, 2014

Breakout #3: Show me the Data

‣ Our current hypothesis is that a competition would require two separate datasets:

  – "Power System Description" – A description of the physical world to be optimized.

  – "Scenarios" – Specific, individual cases to be optimized (loads, equipment availability, wind/solar generation, etc.)

  – Is this a reasonable description of the dataset challenge? Is this terminology clear/correct?

Breakout #3: Show me the Data

‣ What MUST be included in the "power system description" dataset?

- Network topology (including thermal limits)?
- Contingency Lists? How detailed?
- Including substation equipment, FACTs, etc.?
- Transmission only or Transmission + Distribution?
- Controllable LTCs, PSTs, etc.?
- Energy storage?
- What constraints are particularly important to include?
- Controllable loads?
- Dynamics?
- Protection Systems?
- Remedial Action Schemes?
- Cost curves? (How complex?)
- Voltage/Transient stability limited?
- Line switching allowed? How do you set limits?
- Wind/Solar Generation? Forecasts?
- Any other items on the "wish list"?

‣ What is the optimal size (i.e. # of buses, # of generators, etc.) for the power system?

Breakout #3: Show me the Data

‣ What "power system description" datasets could be leveraged for a competition?

- To what extent is it important to make the competition datasets publically available?

- To what extent might we be able to use utility/ISO provided datasets? How? (I.e. Real data!)

- In the absence of real data, how could a synthetic dataset be created for this purpose?

- What types of organizations might have the expertise and prominence needed to generate a dataset that would be useful and trusted?

- What might this cost?

Breakout #3: Show me the Data

‣ What MUST be included in the "scenarios" dataset?

- MW/MVAr demand at each bus?

- Wind/Solar generation?

- Transmission/Generator Outages/Availability?

- Weather? (What details?)

– To what degree is it reasonable and desirable to include infeasible scenarios in the dataset?

– How many "scenarios" might be needed for a competition?

Breakout #3: Show me the Data

‣ To what extent could the data specified in the previous question be located and integrated into a competition data set?

- – What "scenario" datasets could be leveraged for a competition?
- – To what extent is it important to make the competition datasets publically available?
- – To what extent might we be able to use utility/ISO provided datasets? How? (I.e. Real data!)
- – In the absence of real data, how could a synthetic dataset be created for this purpose?
- – What types of organizations might have the expertise and prominence needed to generate a "scenarios" dataset that would be useful and trusted?
- – What might this cost?

Breakout #3: Show me the Data

‣ To what extent might dataset creation be an ARPA-E hard problem itself? Why?

  – What technical miracles must be overcome to enable the creation of comprehensive, realistic datasets?

Breakout #3 Readout Summary