Power Systems Stability Control – Using A Reinforcement Framework

Picture from the blackout of NYC
High costs.

How did happen – series of small failures that combined.
4 harmless events (which no one knew about):

1. unexplained voltage swings brought down coal fired generator
2. power line outages – one caused by trees
3. ???
4. Alarm rang, nobody noticed.

triggered cascading shutdown.

Power Systems Stability Control – Using A Reinforcement Framework

Electronic Power System – three subsystems
Power reacant storage/distribution
Fuel cell power places to produce power
Electrical power distribution and control

Power System Stability = remain in state of balance under normal operating conditions.
On disturbance move ASAP to stable state.

Not enough power plants being built to meet growing power consumption needs, in part due to environmental concerns. Originally system had some **margin for error**. If have problem, consumption increases, overworking, less room for error. Stability margin decreasing.
Also people can independently built power system, not regulated by govt.

Problems:
Large unpredictable fluctuation.
Freq differences between weakly tied power systems.
Interaction bet.

Leads to uncertainty.
One way to avoid phenomena such as power system breakdown is to increase margin for error.

3 elements:

- Agents (made intelligent via reinforcement learning framework)
- Devices
- Power System

System needs be able to: access power system vulnerability, monitor hidder failures of protection devices, and Exception handling.

Now, good communication over broadband, better data collection methods, and higher computational ability – as justification for reinforcement framework.

<IMAGE showing interactions between various components of the system>

Agents trained via reward and punishment.
No need to specify how to perform the task.
☺ how many blackouts will need to train?? A: done via simulation.
Discussant??: mentions paper says would not want to put untrained agent in field.

O: A set of observations
D: a set of devices

$CS^{n=}$ { (o, a, d). … } observations, actions, devices.
An n agent control scheme where o and d are subsets of O and D, where a denotes an agent that process info in o_i to control devices of d.

Note: while set of combos is infinite, can restrict by looking at the problem, and separating into sets specific to say, a certain power grid.

Use a simulation. The agents observe the system state, take actions, and observe effects. We want to choose the best possible action for a particular observation. They learn better control law on the basis of more experience.

RL: learn both from interaction with system and simulator (hybrid. First through simulation, then on the job), goal oriented aspect, and ability to deal with stochastic nature.
Skip over specific implementation of dynamic programming,. State space discretization to reduce the roblem space into regions – taking a group of states as a region.

Learning the q function. Either estimates transition prob and rewards. Model based learning method in this paper – because more efficient use of data, find better policies.

Q: how does simulation model account for fact that trees in Ohio would cause problem?
A: That is something that needs to be learned online.
D: Model doesn't know about trees. Reflected in model by voltage drop.

Practical Problems:

Curse of dimensionality problem. Same issue that we want to reduce state space, but afraid of leaving out something important. So use engineering expect, and have algorithm to implement that task.

Partially observable state. Usually noisy. But based on past observation, can guess better what observation was.

Non-stationary. EPS does not totally deend on time. Changes slowly with respect to speed of learning, and weights far old observations less.

<IMAGE SHOWING DESCRIPTION OF TEST CONTROL SYSTEM – 4 machine control system>
Off Line Design (Simulation) of Dynamic Brake Controller (A device that controls the power system
<Image showing scattered strategy after 100 scenarious, but after 1000 clusters and hads learnt.>

ON-Line learning, with diff device called TCSC

State and Reward function, etc. see slide are all covered in the paper though not discussed.
IMAGE:figures 7,8, Case study showing fluctuations after 10 minutes, 1 hour, 10 hours gets better.

Related work: in past was difficult to implement, now easier. Did papers on fuzzy logic and non-model based system.

Conclusion – looking at uncertain area and trying to predict better.
Paper talked about closed vs open loop. Closed=learn from simulation, from data already know. Open loop, on real system, learning on the fly. They say in general open loop performs ~ closed loop. Here, closed loop performed better.
Say also use together with thraditional control methods.

D: at end of paper, testing phase, showing nice graphs. Showing improving. But read carefully, in predictable system where repeat the data. And says but does not demonstrate it in a stochastic system. And that was point of whole paper, how great it is in unpredictable situations.
A: many papers wrote, and in Chapter 10, they deal with this particular area.

Q: glitch: alarm rang, no one informed. how will adding complexity help.
D: will talk to each other
Q: repeat question, now have agent complexity beside that of system, won't this introduce more problems.
A: System updating policy. And will work with better policy. But does add complexity.
D: key is the communication. Not human listening to bell, but an agent, and so less of failures saw, get quicker reaction time. Don't say, well we haven't gotten a call from Cleveland so what do we do?

Q: seen often situation, a problem and people built system to handle problem. But no way engineers will hand over power to machine. Can it not instead figure this out and present information. In end, if agent fails, will be legal ramifications, so how will engineers and businesses hand over these. Part of learning should be what info is relevant and to whom to provide the info to.
S: People have. Spanish power company.
A: whole world going to automation
Q: too much control is ludicrous though. There are some limit.

D: If phrase it as a cost issue, might get it through. With automation, can fire people. Less damage, less lawsuits, and they might not care about safety.
S: Will be figuring out optimal policy, and often humans perform suboptimally. May be case that these things never get fielded though.

S: Impression of paper: at beginning, this is great. whenever write a paper, is a sales job. Can oversell or undersell. Must steer a course. Stake out claims what trying to do. Good paper in terms of reporting results and showing the limits of these results. There is more that we could do. Not bad in terms of how presented it.

Q: blackout rare and cause very different. Very sparse.
A: true, this is the problem.
Q: so should gather data about the blackout
S: A benefit of simulation is if a blackout situation is fairly rare, can figure out what factors went into it and feed that same scenario often.

S: Problem: should compare results with something else and compare mean squared error. Quantifiably better than something before.
Q: gave references
S: but should show other system, and compare results.

Q: note in article, they claim nobody tried to tackle whole problem using RL before.
S: but perhaps have other control methods
Q2: after all, have systems which are working now