Learning to Run the Power Network using a Reinforcement Learning Actor Agent

Team : Learning_RL

Team Members:
Amar Ramapuram Matavalam – Iowa State University
Kishan Guddanti – Arizona State University (Advised by Dr. Yang Weng)
Soumya Indela – University of Maryland

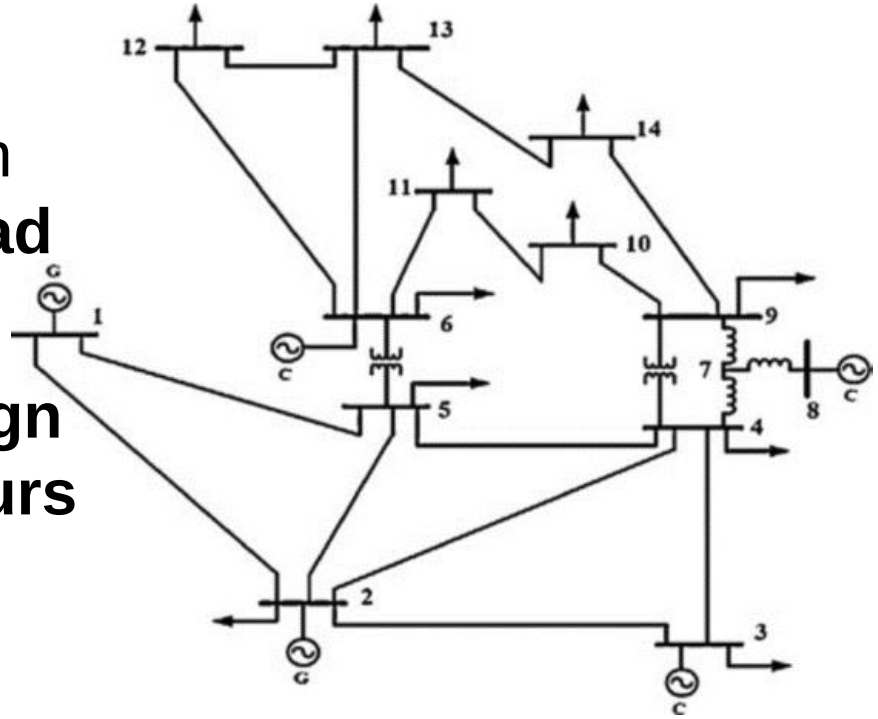
Thanks to the L2RPN team for releasing pypownet and the datasets !!

Outline

- Introduction to L2RPN
- Challenge from an RL Perspective
- Problem Analysis Using Domain Knowledge
- Reasons to Use A3C
- Training Methodology
- Conclusion & Lessons To Take Forward

Introduction to L2RPN

- A power grid consists of loads & generators that are **sparsely interconnected** via transmission lines.
- Overloading of the lines causes them to break –can cause **cascading failure and loss of power to load**
- The overall goal of the L2RPN challenge is to **design an agent that ensures no failure of the grid occurs by switching elements in each substation**



IEEE 14 bus test system

Challenge from an RL Perspective

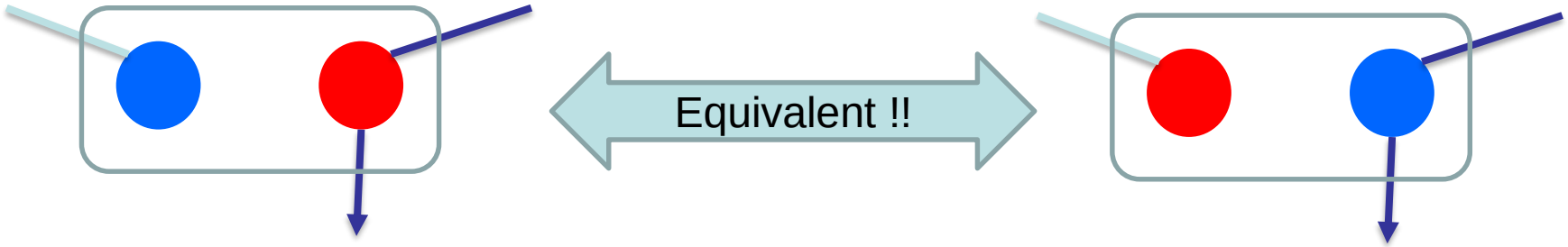
- The state space has **both continuous variables and discrete variables**
  - Continuous Variables - line flows, MW generation, MW load, etc.
  - Discrete Variables - line status, line switchable flag, element configuration in each substation, etc.

- A **large action space** – line connection/disconnection & changing element configuration in each substation.

- **Critical Infrastructure -** agent should not cause grid failure **over several scenarios** and a **large number of time steps**

Problem Analysis Using Domain Knowledge -1

- Reduction of the **action space and state space** using domain knowledge
- To simplify, we tried to see if only one kind of action works- based on our tests, only line switching did not always work. Thus, we looked at **only controlling the substation configuration**.
- There are a total of 312 substation node configuration switching actions – these are **reduced by half due to the symmetry** in the problem



- Thus, a **total of 157 actions** are considered (including no action) – determined offline

Problem Analysis Using Domain Knowledge -2

- From the power system behavior, the **planned productions and consumptions are highly correlated to the present productions and consumptions** and are not considered – same for other states

- The following observations are used as states – a total of 164 states
  - Nodes to which the various elements are connected
  - Line status, line thermal limits and line flows
  - P & V of generators, P & Q of loads
  - Time – Hour and minute
  - Time steps before nodes reactionable

Reasons to use A3C RL model

- Due to the large number of system configurations, states and actions, a **deep-RL framework** is necessary

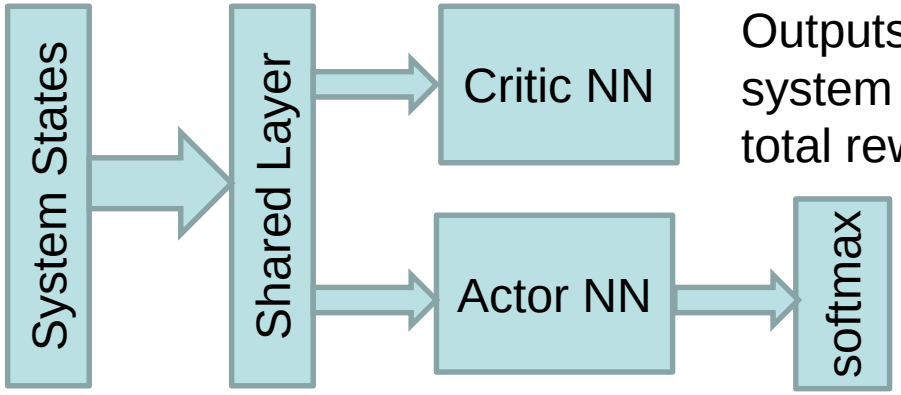
- At present, **the advantage actor-critic framework executing multiple agents in parallel** is state-of-the-art [a] – simple implementation and has good learning stability

- This is called as the **Asynchronous Advantage Actor-Critic (A3C) method**

[a] Mnih, Badia, Mirza, Graves, Lillicrap, Harley, Silver, Kavukcuoglu (2016) - Asynchronous methods for deep reinforcement learning: A3C -- parallel online actor-critic

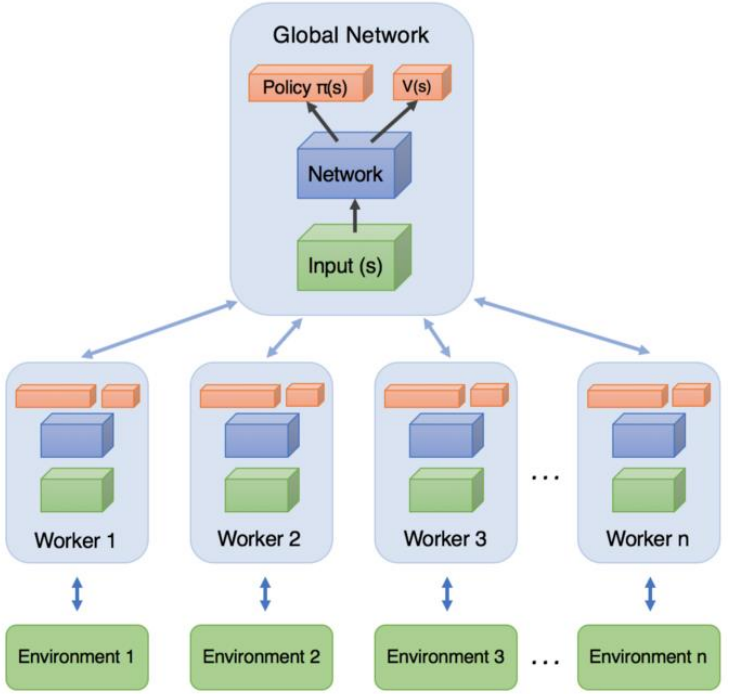
Training Methodology - 1

- The actor and critic are represented by neural networks whose weights need to be learnt – a shared layer improves the performance



Outputs the value $V(s)$ of the system state which is expected total reward from state s

Outputs probabilities of each action – the policy $\pi(s)$ samples from this distribution



- Modified an existing A3C Keras code to use 48 workers that loaded the various chronics in parallel.
- ISU HPC cluster is used for the training over 3 weeks

Training Methodology -2

- To stabilize learning, the initial settings of the training are:
  - The game mode is ‘easy’ with a large penalty on the overloading of lines
  - The simulate method is used in the actor to evaluate a few (4) high probability actions along with no-action.
  - The data is subsampled (by a factor of 7) so that each time step is larger

- These settings train the actor to minimize line overloads and the entire dataset can be seen by the actor/critic to enforce robustness indirectly.

Training Methodology -3

- Several print statements are inserted to track the progress of the learning as the training is proceeding – **each thread is a different CPU**.
- After a few days of training on the ‘easiest’ settings:

```
Continue Thread: 36 / train episode:  41 / score :  378 / with recent time: 100 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [60840.]
Continue Thread: 0 / train episode:  41 / score :  1716 / with recent time: 400 / with recent action 0 / number of non-zero actions 36 / day_hour_min: [81340.]
----STOPPED Thread: 4 / train episode:  41 / score :  -108 / with final time: 77 / with final action 31 / number of non-zero actions 7 / day_hour_min: [40840.]
Continue Thread: 30 / train episode:  42 / score :  1295 / with recent time: 100 / with recent action 0 / number of non-zero actions 19 / day_hour_min: [40920.]
Continue Thread: 22 / train episode:  42 / score :  794 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [51020.]
Continue Thread: 37 / train episode:  42 / score :  -87 / with recent time: 100 / with recent action 0 / number of non-zero actions 29 / day_hour_min: [81340.]
----STOPPED Thread: 0 / train episode:  42 / score :  904 / with final time: 407 / with final action 0 / number of non-zero actions 38 / day_hour_min: [81540.]
----STOPPED Thread: 5 / train episode:  43 / score :  1146 / with final time: 255 / with final action 0 / number of non-zero actions 30 / day_hour_min: [71300.]
----STOPPED Thread: 41 / train episode:  44 / score :  1740 / with final time: 269 / with final action 31 / number of non-zero actions 34 / day_hour_min: [101740.]
----STOPPED Thread: 8 / train episode:  45 / score :  -1284 / with final time: 131 / with final action 0 / number of non-zero actions 24 / day_hour_min: [101740.]
----STOPPED Thread: 37 / train episode:  46 / score :  -928 / with final time: 106 / with final action 0 / number of non-zero actions 29 / day_hour_min: [81520.]
----STOPPED Thread: 25 / train episode:  47 / score :  607 / with final time: 127 / with final action 31 / number of non-zero actions 33 / day_hour_min: [81400.]
Continue Thread: 21 / train episode:  48 / score :  1879 / with recent time: 200 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [61420.]
----STOPPED Thread: 10 / train episode:                                                                  number of non-zero actions 35 / day_hour_min: [100800.]
saved NN model at episode 49
Continue Thread: 34 / train episode:  49 / score :  79 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [101740.]
Continue Thread: 1 / train episode:  49 / score :  863 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [61620.]
----STOPPED Thread: 19 / train episode:  49 / score :  2374 / with final time: 273 / with final action 31 / number of non-zero actions 37 / day_hour_min: [81900.]
Continue Thread: 40 / train episode:  50 / score :  392 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [91840.]
----STOPPED Thread: 7 / train episode:  50 / score :  -646 / with final time: 35 / with final action 0 / number of non-zero actions 4 / day_hour_min: [81940.]
----STOPPED Thread: 14 / train episode:  51 / score :  1104 / with final time: 267 / with final action 31 / number of non-zero actions 34 / day_hour_min: [81700.]
----STOPPED Thread: 11 / train episode:  52 / score :  1609 / with final time: 272 / with final action 0 / number of non-zero actions 40 / day_hour_min: [81840.]
----STOPPED Thread: 42 / train episode:  53 / score :  -695 / with final time: 73 / with final action 0 / number of non-zero actions 9 / day_hour_min: [102000.]
Continue Thread: 20 / train episode:  54 / score :  1286 / with recent time: 100 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [61700.]
Continue Thread: 47 / train episode:  54 / score :  894 / with recent time: 100 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [41700.]
Continue Thread: 9 / train episode:  54 / score :  282 / with recent time: 100 / with recent action 0 / number of non-zero actions 13 / day_hour_min: [41700.]
```

Save the Neural Network at regular intervals

Training Methodology -3

- Several print statements are inserted to track the progress of the learning as the training is proceeding – **each thread is a different CPU**.
- After a few days of training on the ‘easiest’ settings:

```
Continue Thread: 36 / train episode:  41 / score :  378 / with recent time: 100 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [60840.]
Continue Thread: 0 / train episode:  41 / score :  1716 / with recent time: 400 / with recent action 0 / number of non-zero actions 36 / day_hour_min: [81340.]
----STOPPED Thread: 4 / train episode:  41 / score :  -108 / with final time: 77 / with final action 31 / number of non-zero actions 7 / day_hour_min: [40840.]
Continue Thread: 30 / train                                  with recent time: 100 / with recent action 0 / number of non-zero actions 19 / day_hour_min: [40920.]
Continue Thread: 22                                          th recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [51020.]
Continue Thread: 37 / train                                  th recent time: 100 / with recent action 0 / number of non-zero actions 29 / day_hour_min: [81340.]
----STOPPED Thread: 0 / train episode:  42 / score :  904 / with final time: 407 / with final action 0 / number of non-zero actions 38 / day_hour_min: [81540.]
----STOPPED Thread: 5 / train episode:  43 / score :  1146 / with final time: 255 / with final action 0 / number of non-zero actions 30 / day_hour_min: [71300.]
----STOPPED Thread: 41 / train episode:  44 / score :  1740 / with final time: 269 / with final action 31 / number of non-zero actions 34 / day_hour_min: [101740.]
----STOPPED Thread: 8 / train episode:  45 / score :  -1284 / with final time: 131 / with final action 0 / number of non-zero actions 24 / day_hour_min: [101740.]
----STOPPED Thread: 37 / train episode:  46 / score :  -928 / with final time: 106 / with final action 0 / number of non-zero actions 29 / day_hour_min: [81520.]
----STOPPED Thread: 25 / train episode:  47 / score :  607 / with final time: 127 / with final action 31 / number of non-zero actions 33 / day_hour_min: [81400.]
Continue Thread: 21 / train episode:  48 / score :  1879 / with recent time: 200 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [61420.]
----STOPPED Thread: 10 / train episode:  48 / score :  -463 / with final time: 142 / with final action 0 / number of non-zero actions 35 / day_hour_min: [100800.]
saved NN model at episode 49

Continue Thread: 34 / train episode:  49 / score :  79 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [101740.]
Continue Thread: 1 / train episode:  49 / score :  863 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [61620.]
----STOPPED Thread: 19 / train episode:  49 / score :  2374 / with final time: 273 / with final action 31 / number of non-zero actions 37 / day_hour_min: [81900.]
Continue Thread: 40 / train episode:  50 / score :  392 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [91840.]
----STOPPED Thread: 7 / train episode:  50 / score :  -646 / with final time: 35 / with final action 0 / number of non-zero actions 4 / day_hour_min: [81940.]
----STOPPED Thread: 14 / train episode:  51 / score :  1104 / with final time: 267 / with final action 31 / number of non-zero actions 34 / day_hour_min: [81700.]
----STOPPED Thread: 11 / train episode:  52 / score :  1609 / with final time: 272 / with final action 0 / number of non-zero actions 40 / day_hour_min: [81840.]
----STOPPED Thread: 42 / train episode:  53 / score :  -695 / with final time: 73 / with final action 0 / number of non-zero actions 9 / day_hour_min: [102000.]
Continue Thread: 20 / train episode:  54 / score :  1286 / with recent time: 100 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [61700.]
Continue Thread: 47 / train episode:  54 / score :  894 / with recent time: 100 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [41700.]
Continue Thread: 9 / train episode:  54 / score :  282 / with recent time: 100 / with recent action 0 / number of non-zero actions 13 / day_hour_min: [41700.]
```

No failure in the grid

Training Methodology -3

- Several print statements are inserted to track the progress of the learning as the training is proceeding – **each thread is a different CPU**.
- After a few days of training on the ‘easiest’ settings:

```
Continue Thread: 36 / train episode:  41 / score :  378 / with recent time: 100 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [60840.]
Continue Thread: 0 / train episode:  41 / score :  1716 / with recent time: 400 / with recent action 0 / number of non-zero actions 36 / day_hour_min: [81340.]
----STOPPED Thread: 4 / train episode:  41 / score :  -108 / with final time: 77 / with final action 31 / number of non-zero actions 7 / day_hour_min: [40840.]
Continue Thread: 30 / train episode:  42 / score :  1295 / with recent time: 100 / with recent action 0 / number of non-zero actions 19 / day_hour_min: [40920.]
Continue Thread: 22 / train episode:  42 / score :  794 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [51020.]
Continue Thread: 37 / train episode:  42 / score :  -87 / with recent time: 100 / with recent action 0 / number of non-zero actions 29 / day_hour_min: [81340.]
----STOPPED Thread: 0 / train episode:  42 / score :  904 / with final time: 407 / with final action 0 / number of non-zero actions 38 / day_hour_min: [81540.]
----STOPPED Thread: 5 / train episode:  43 / score :  1146 / with final time: 255 / with final action 0 / number of non-zero actions 30 / day_hour_min: [71300.]
----STOPPED Thread: 41 / tra                                                    number of non-zero actions 34 / day_hour_min: [101740.]
----STOPPED Thread:                                                             number of non-zero actions 24 / day_hour_min: [101740.]
----STOPPED Thread:                                                             number of non-zero actions 29 / day_hour_min: [81520.]
----STOPPED Thread: 25 / train episode:  47 / score :  607 / with final time: 127 / with final action 31 / number of non-zero actions 33 / day_hour_min: [81400.]
Continue Thread: 21 / train episode:  48 / score :  1879 / with recent time: 200 / with recent action 0 / number of non-zero actions 22 / day_hour_min: [61420.]
----STOPPED Thread: 10 / train episode:  48 / score :  -463 / with final time: 142 / with final action 0 / number of non-zero actions 35 / day_hour_min: [100800.]
saved NN model at episode 49

Continue Thread: 34 / train episode:  49 / score :  79 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [101740.]
Continue Thread: 1 / train episode:  49 / score :  863 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [61620.]
----STOPPED Thread: 19 / train episode:  49 / score :  2374 / with final time: 273 / with final action 31 / number of non-zero actions 37 / day_hour_min: [81900.]
Continue Thread: 40 / train episode:  50 / score :  392 / with recent time: 100 / with recent action 0 / number of non-zero actions 15 / day_hour_min: [91840.]
----STOPPED Thread: 7 / train episode:  50 / score :  -646 / with final time: 35 / with final action 0 / number of non-zero actions 4 / day_hour_min: [81940.]
----STOPPED Thread: 14 / train episode:  51 / score :  1104 / with final time: 267 / with final action 31 / number of non-zero actions 34 / day_hour_min: [81700.]
----STOPPED Thread: 11 / train episode:  52 / score :  1609 / with final time: 272 / with final action 0 / number of non-zero actions 40 / day_hour_min: [81840.]
----STOPPED Thread: 42 / train episode:  53 / score :  -695 / with final time: 73 / with final action 0 / number of non-zero actions 9 / day_hour_min: [102000.]
Continue Thread: 20 / train episode:  54 / score :  1286 / with recent time: 100 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [61700.]
Continue Thread: 47 / train episode:  54 / score :  894 / with recent time: 100 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [41700.]
Continue Thread: 9 / train episode:  54 / score :  282 / with recent time: 100 / with recent action 0 / number of non-zero actions 13 / day_hour_min: [41700.]
```

Failure in the grid – load loss or power flow diverged

Training Methodology -4

- A successful agent goes for ~1000 steps without causing grid failure. When multiple threads are successful for several episodes, the hardness in increased by:
  - Increasing the game mode
  - Reducing number of probable actions evaluated by simulate
  - Removing including no-action explicitly in the evaluation by simulate
  - Reducing subsampling rate

Day 7 of training – Soft Mode

```
Continue Thread: 28 / train episode:  218 / score :  4760 / with recent time: 700 / with recent action 0 / number of non-zero actions 17 / day_hour_min: [11520.]
Continue Thread: 18 / train episode:  218 / score :  10389 / with recent time: 800 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [262140.]
Continue Thread: 21 / train episode:  218 / score :  8771 / with recent time: 700 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [211120.]
Continue Thread: 3 / train episode:  218 / score :  6181 / with recent time: 500 / with recent action 0 / number of non-zero actions 12 / day_hour_min: [221040.]
Continue Thread: 8 / train episode:  218 / score :  7187 / with recent time: 500 / with recent action 0 / number of non-zero actions 9 / day_hour_min: [241140.]
Continue Thread: 42 / train episode:  218 / score :  9817 / with recent time: 700 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [281520.]
Continue Thread: 46 / train episode:  218 / score :  6272 / with recent time: 500 / with recent action 0 / number of non-zero actions 13 / day_hour_min: [221300.]
Continue Thread: 15 / train episode:  218 / score :  7596 / with recent time: 600 / with recent action 0 / number of non-zero actions 18 / day_hour_min: [260220.]
Continue Thread: 34 / train episode:  218 / score :  10085 / with recent time: 700 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [271420.]
Continue Thread: 27 / train episode:  218 / score :  5579 / with recent time: 500 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [60100.]
Continue Thread: 10 / train episode:  218 / score :  8227 / with recent time: 600 / with recent action 0 / number of non-zero actions 13 / day_hour_min: [281240.]
saved NN model at episode 218

Continue Thread: 14 / train episode:  218 / score :  8732 / with recent time: 600 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [281420.]
Continue Thread: 33 / train episode:  218 / score :  2190 / with recent time: 200 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [261020.]
Continue Thread: 2 / train episode:  218 / score :  5727 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [250040.]
Continue Thread: 26 / train episode:  218 / score :  7264 / with recent time: 500 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [261340.]
Continue Thread: 35 / train episode:  218 / score :  5806 / with recent time: 400 / with recent action 0 / number of non-zero actions 14 / day_hour_min: [240920.]
Continue Thread: 9 / train episode:  218 / score :  1942 / with recent time: 400 / with recent action 0 / number of non-zero actions 33 / day_hour_min: [231920.]
Continue Thread: 11 / train episode:  218 / score :  7176 / with recent time: 500 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [250120.]
Continue Thread: 17 / train episode:  218 / score :  10085 / with recent time: 800 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [222140.]
saved NN model at episode 218
```

Day 9 of training – No simulate & Hard mode

```
Continue Thread: 29 / train episode:  6 / score :  7010 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [70925.]
Continue Thread: 21 / train episode:  6 / score :  2475 / with recent time: 200 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40605.]
Continue Thread: 24 / train episode:  6 / score :  4622 / with recent time: 300 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [50735.]
Continue Thread: 25 / train episode:  6 / score :  4923 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [60925.]
Continue Thread: 11 / train episode:  6 / score :  5532 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [60925.]
Continue Thread: 7 / train episode:  6 / score :  5640 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [60925.]
Continue Thread: 43 / train episode:  6 / score :  5915 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [60925.]
Continue Thread: 26 / train episode:  6 / score :  3877 / with recent time: 400 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [60925.]
Continue Thread: 46 / train episode:  6 / score :  6416 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [60925.]
Continue Thread: 37 / train episode:  6 / score :  3609 / with recent time: 400 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [60925.]
Continue Thread: 27 / train episode:  6 / score :  4767 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [60925.]
Continue Thread: 2 / train episode:  6 / score :  5854 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [60925.]
Continue Thread: 15 / train episode:  6 / score :  3383 / with recent time: 200 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [81205.]
Continue Thread: 0 / train episode:  6 / score :  4118 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 5 / train episode:  6 / score :  5686 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [50925.]
Continue Thread: 35 / train episode:  6 / score :  5286 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [50925.]
Continue Thread: 36 / train episode:  6 / score :  5780 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 20 / train episode:  6 / score :  5863 / with recent time: 400 / with recent action 0 / number of non-zero actions 5 / day_hour_min: [40925.]
Continue Thread: 33 / train episode:  6 / score :  3888 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 45 / train episode:  6 / score :  3249 / with recent time: 400 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [40925.]
Continue Thread: 6 / train episode:  6 / score :  5838 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 3 / train episode:  6 / score :  6496 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 13 / train episode:  6 / score :  3192 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40925.]
Continue Thread: 1 / train episode:  6 / score :  3467 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [40940.]
Continue Thread: 32 / train episode:  6 / score :  3413 / with recent time: 400 / with recent action 0 / number of non-zero actions 8 / day_hour_min: [40925.]
Continue Thread: 38 / train episode:  6 / score :  6179 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [30925.]
Continue Thread: 39 / train episode:  6 / score :  4723 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [30925.]
Continue Thread: 30 / train episode:  6 / score :  5389 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 8 / train episode:  6 / score :  6409 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [80925.]
Continue Thread: 4 / train episode:  6 / score :  5459 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 44 / train episode:  6 / score :  5854 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [30925.]
Continue Thread: 22 / train episode:  6 / score :  4240 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [30925.]
Continue Thread: 9 / train episode:  6 / score :  5427 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 17 / train episode:  6 / score :  4513 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 47 / train episode:  6 / score :  4152 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 23 / train episode:  6 / score :  6185 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [20925.]
Continue Thread: 18 / train episode:  6 / score :  5781 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [60925.]
Continue Thread: 31 / train episode:  6 / score :  4852 / with recent time: 400 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [30925.]
Continue Thread: 19 / train episode:  6 / score :  5207 / with recent time: 400 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [60925.]
Continue Thread: 34 / train episode:  6 / score :  665 / with recent time: 100 / with recent action 0 / number of non-zero actions 16 / day_hour_min: [81520.]
Continue Thread: 16 / train episode:  6 / score :  4972 / with recent time: 400 / with recent action 0 / number of non-zero actions 6 / day_hour_min: [30925.]
Continue Thread: 14 / train episode:  6 / score :  6544 / with recent time: 500 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [61745.]
Continue Thread: 10 / train episode:  6 / score :  9131 / with recent time: 500 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [71745.]
Continue Thread: 12 / train episode:  6 / score :  5684 / with recent time: 400 / with recent action 0 / number of non-zero actions 10 / day_hour_min: [40925.]
Continue Thread: 29 / train episode:  6 / score :  9007 / with recent time: 500 / with recent action 0 / number of non-zero actions 3 / day_hour_min: [71745.]
Continue Thread: 40 / train episode:  6 / score :  9032 / with recent time: 500 / with recent action 0 / number of non-zero actions 4 / day_hour_min: [71745.]
```

Day 9 of training – No simulate & Hard mode

No failure in the grid over various time series data

Conclusion & Lessons To Take Forward

- The A3C based actor trained using multiple chronics in parallel is able to learn the right actions for the substation configurations
- Starting the learning problem from a ‘low’ level and gradually rising the ‘toughness’ is key to enable learning – Doman knowledge
- A HPC/server computing infrastructure is key for state-of-the art RL methods
- The code for training the models will be up on github at the link: https://github.com/amar-iastate/L2RPN-using-A3C

Thank You!

Contact – Amar (amar@iastate.edu)
https://github.com/amar-iastate