A Novel Machine Learning Approach for Solving Optimal Transmission Switching

Asunción Jiménez-Cordero asuncionjc@uma.es

JOINT WORK WITH: Juan Miguel Morales González Salvador Pineda Morente









oasys.uma.es

3rd EUROYoung Workshop

June 6th, 2023

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 755705)

Outline

- Motivation
- 2 Methodology
- 3 Computational Experience
- 4 Conclusions and Further Research

Outline

- Motivation
- 2 Methodology
- 3 Computational Experience
- 4 Conclusions and Further Research

Mathematical Optimization

Machine Learning

Combine knowledge from both worlds

Recent reviews: Bengio et al. [2021]; Gambella et al. [2021]

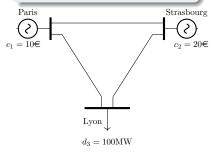
Mathematical Optimization

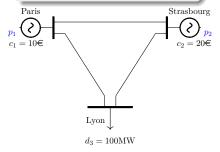
Machine Learning

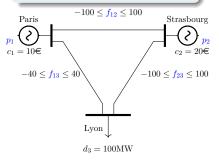
Combine knowledge from both worlds

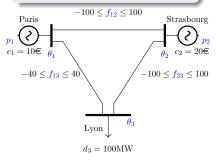
Recent reviews: Bengio et al. [2021]; Gambella et al. [2021]

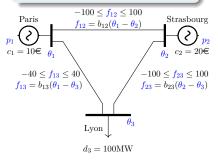
Power Systems Operations

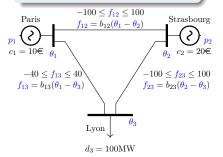






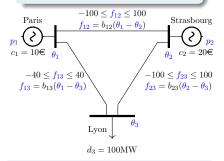






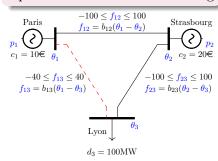
- Balance constraints.
- Assume $b_{nm} = 1, \forall (n, m)$.
- Minimizing costs. $Cost^* = 1800 \in$.
- $p_1^* = 20, p_2^* = 80.$
- $\theta_1^* = 40, \, \theta_2^* = 60, \, \theta_3^* = 0.$
- $f_{12}^* = -20, f_{13}^* = 40, f_{23}^* = 60.$

Optimal Power Flow (OPF)

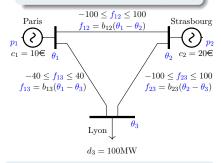


- Balance constraints.
- Assume $b_{nm} = 1, \forall (n, m)$.
- Minimizing costs. $Cost^* = 1800 \in$.
- $p_1^* = 20, p_2^* = 80.$
- $\theta_1^* = 40, \, \theta_2^* = 60, \, \theta_3^* = 0.$
- $f_{12}^* = -20, f_{13}^* = 40, f_{23}^* = 60.$

Optimal Transmission Switching

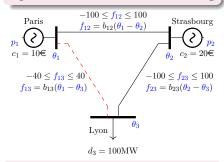


Optimal Power Flow (OPF)



- Balance constraints.
- Assume $b_{nm} = 1, \forall (n, m)$.
- Minimizing costs. $Cost^* = 1800 \in$.
- $p_1^* = 20, p_2^* = 80.$
- $\theta_1^* = 40, \, \theta_2^* = 60, \, \theta_3^* = 0.$
- $f_{12}^* = -20, f_{13}^* = 40, f_{23}^* = 60.$

Optimal Transmission Switching



- Switchable lines, \mathcal{L}_S .
- Binary variables $x_{nm}, \forall (n,m) \in \mathcal{L}_S$.
- Minimizing costs. $Cost^* = 1000 \in$.
- $p_1^* = 100, p_2^* = 0.$
- $\theta_1^* = 200, \, \theta_2^* = 100, \, \theta_3^* = 0.$
- $f_{12}^* = 100, \, f_{13}^* = 0, \, f_{23}^* = 100.$

 $\begin{cases} \min_{p_n, f_{nm}, \theta_n, x_{nm}} & \sum_n c_n p_n \\ \text{s.t.} & \underline{p}_n \le p_n \le \overline{p}_n, \quad \forall n \in \mathcal{N} \end{cases}$ $\sum_{(n,m)\in\mathcal{L}_n^-} f_{nm} - \sum_{(n,m)\in\mathcal{L}_n^+} f_{nm} = p_n - d_n, \quad \forall n \in \mathcal{N}$ $f_{nm} = x_{nm}b_{nm}(\theta_n - \theta_m), \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$ $-x_{nm}\overline{f}_{nm} \leq f_{nm} \leq x_{nm}\overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$ $f_{nm} = b_{nm}(\theta_n - \theta_m), \quad \forall (n, m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}}$ $-\overline{f}_{nm} \le f_{nm} \le \overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}}$ $x_{nm} \in \{0,1\}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$ $\theta_1 = 0$

$$\begin{cases} \min_{p_n, f_{nm}, \theta_n, x_{nm}} & \sum_n c_n p_n \\ \text{s.t.} & \underbrace{p_n \leq p_n \leq \overline{p}_n, \quad \forall n \in \mathcal{N}} \\ & \sum_{(n,m) \in \mathcal{L}_n^-} f_{nm} - \sum_{(n,m) \in \mathcal{L}_n^+} f_{nm} = p_n - d_n, \quad \forall n \in \mathcal{N} \\ & \underbrace{f_{nm} = x_{nm} b_{nm} (\theta_n - \theta_m), \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}} \\ & -x_{nm} \overline{f}_{nm} \leq f_{nm} \leq x_{nm} \overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}} \\ & f_{nm} = b_{nm} (\theta_n - \theta_m), \quad \forall (n,m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}} \\ & -\overline{f}_{nm} \leq f_{nm} \leq \overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}} \\ & x_{nm} \in \{0,1\}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}} \\ & \theta_1 = 0 \end{cases}$$

$$\begin{cases} \min \limits_{p_n,f_{nm},\theta_n,x_{nm}} & \sum_n c_n \, p_n \\ \text{s.t.} & \underbrace{p_n \leq \overline{p}_n, \quad \forall n \in \mathcal{N}} \\ & \sum_{(n,m) \in \mathcal{L}_n^-} f_{nm} - \sum_{(n,m) \in \mathcal{L}_n^+} f_{nm} = p_n - d_n, \quad \forall n \in \mathcal{N} \\ & \underbrace{f_{nm} = x_{nm} b_{nm} (\theta_n - \theta_m), \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}} \\ & -x_{nm} \overline{f}_{nm} \leq f_{nm} \leq x_{nm} \overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}} \\ & f_{nm} = b_{nm} (\theta_n - \theta_m), \quad \forall (n,m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}} \\ & -\overline{f}_{nm} \leq f_{nm} \leq \overline{f}_{nm}, \quad \forall (n,m) \in \mathcal{L} \setminus \mathcal{L}_{\mathcal{S}} \\ & x_{nm} \in \{0,1\}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}} \\ & \theta_1 = 0 \end{cases}$$

- MINLP because $x_{nm}(\theta_n \theta_m)$.
- NP-hard problem.

Original constraint

$$f_{nm} = x_{nm}b_{nm}(\theta_n - \theta_m)$$

Reformulation using big-Ms

$$b_{nm}(\theta_n - \theta_m) - \overline{M}_{nm}(1 - x_{nm}) \le f_{nm} \le b_{nm}(\theta_n - \theta_m) - \underline{M}_{nm}(1 - x_{nm})$$

How to compute bigM values?

$$\underline{M}_{nm} \le \underline{M}_{nm}^{\text{OPT}} := b_{nm} \times \min_{x_{nm}=0} (\theta_n - \theta_m)$$

$$\overline{M}_{nm} \ge \overline{M}_{nm}^{OPT} := b_{nm} \times \max_{x_{nm}=0} (\theta_n - \theta_m)$$

How to compute bigM values?

$$\underline{M}_{nm} \le \underline{M}_{nm}^{\text{OPT}} := b_{nm} \times \min_{x_{nm}=0} (\theta_n - \theta_m)$$

$$\overline{M}_{nm} \ge \overline{M}_{nm}^{OPT} := b_{nm} \times \max_{x_{nm}=0} (\theta_n - \theta_m)$$

Are we done?

How to compute bigM values?

$$\underline{M}_{nm} \le \underline{M}_{nm}^{\text{OPT}} := b_{nm} \times \min_{x_{nm}=0} (\theta_n - \theta_m)$$

$$\overline{M}_{nm} \ge \overline{M}_{nm}^{OPT} := b_{nm} \times \max_{x_{nm}=0} (\theta_n - \theta_m)$$

Are we done?

• Computing optimal bigM values is as hard as solving the original problem, [Fattahi et al., 2019].

How to compute bigM values?

$$\underline{M}_{nm} \leq \underline{M}_{nm}^{\text{OPT}} := b_{nm} \times \min_{x_{nm}=0} (\theta_n - \theta_m)$$

$$\overline{M}_{nm} \ge \overline{M}_{nm}^{OPT} := b_{nm} \times \max_{x_{nm}=0} (\theta_n - \theta_m)$$

Are we done?

- Computing optimal bigM values is as hard as solving the original problem, [Fattahi et al., 2019].
- We have to be happy with bounds.

How to compute bigM values?

$$\underline{M}_{nm} \le \underline{M}_{nm}^{\text{OPT}} := b_{nm} \times \min_{x_{nm}=0} (\theta_n - \theta_m)$$

$$\overline{M}_{nm} \ge \overline{M}_{nm}^{OPT} := b_{nm} \times \max_{x_{nm}=0} (\theta_n - \theta_m)$$

Are we done?

- Computing optimal bigM values is as hard as solving the original problem, [Fattahi et al., 2019].
- We have to be happy with bounds.
- Objective: Find good bounds for bigMs.

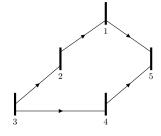
Outline

- 1 Motivation
- 2 Methodology
- 3 Computational Experience
- 4 Conclusions and Further Research

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].

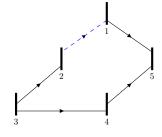
$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in \mathrm{SP}_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].



$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

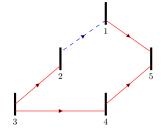
- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].



$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

$$\overline{M}_{21} \ge b_{21}(\theta_2 - \theta_1)$$
 if $x_{21} = 0$

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].

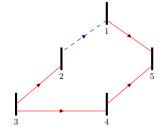


$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

$$\overline{M}_{21} \ge b_{21}(\theta_2 - \theta_1) \text{ if } x_{21} = 0$$

 $(\theta_2 - \theta_1) = (\theta_2 - \theta_3) + (\theta_3 - \theta_4) + (\theta_4 - \theta_5) + (\theta_5 - \theta_1)$

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].



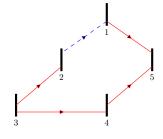
$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

$$\overline{M}_{21} \ge b_{21}(\theta_2 - \theta_1) \text{ if } x_{21} = 0$$

$$(\theta_2 - \theta_1) = (\theta_2 - \theta_3) + (\theta_3 - \theta_4) + (\theta_4 - \theta_5) + (\theta_5 - \theta_1)$$

$$f_{23} \le \overline{f}_{23} \implies b_{23}(\theta_2 - \theta_3) \le \overline{f}_{23} \implies (\theta_2 - \theta_3) \le \overline{f}_{23}/b_{23}$$

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].



$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

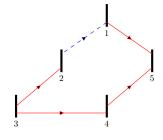
$$\overline{M}_{21} \ge b_{21}(\theta_2 - \theta_1) \text{ if } x_{21} = 0$$

$$(\theta_2 - \theta_1) = (\theta_2 - \theta_3) + (\theta_3 - \theta_4) + (\theta_4 - \theta_5) + (\theta_5 - \theta_1)$$

$$f_{23} \le \overline{f}_{23} \implies b_{23}(\theta_2 - \theta_3) \le \overline{f}_{23} \implies (\theta_2 - \theta_3) \le \overline{f}_{23}/b_{23}$$

$$(\theta_2 - \theta_1) \le \overline{f}_{23}/b_{23} + \overline{f}_{34}/b_{34} + \overline{f}_{45}/b_{45} + \overline{f}_{51}/b_{51}$$

- Exact approach.
- Benchmark.
- Shortest path problem (Dijkstra).
- [Fattahi et al., 2019].



$$-\underline{M}_{nm} = \overline{M}_{nm} = b_{nm} \sum_{(k,l) \in SP_{nm}} \frac{\overline{f}_{kl}}{b_{kl}}, \quad \forall (n,m) \in \mathcal{L}_{\mathcal{S}}$$

$$\overline{M}_{21} \ge b_{21}(\theta_2 - \theta_1) \text{ if } x_{21} = 0$$

$$(\theta_2 - \theta_1) = (\theta_2 - \theta_3) + (\theta_3 - \theta_4) + (\theta_4 - \theta_5) + (\theta_5 - \theta_1)$$

$$f_{23} \le \overline{f}_{23} \implies b_{23}(\theta_2 - \theta_3) \le \overline{f}_{23} \implies (\theta_2 - \theta_3) \le \overline{f}_{23}/b_{23}$$

$$(\theta_2 - \theta_1) \le \overline{f}_{23}/b_{23} + \overline{f}_{34}/b_{34} + \overline{f}_{45}/b_{45} + \overline{f}_{51}/b_{51}$$

$$\overline{M}_{21} \ge b_{21}(\overline{f}_{23}/b_{23} + \overline{f}_{34}/b_{34} + \overline{f}_{45}/b_{45} + \overline{f}_{51}/b_{51})$$

Strategy 2, Knn-D

- Data-driven approach (Knn).
- Naive approach. Learning this problem is a challenge.

Strategy 2, Knn-D

- Data-driven approach (Knn).
- Naive approach. Learning this problem is a challenge.

Algorithm

- 1) Training set $\mathcal{T} = \{(\mathbf{d}^t, \mathbf{x}^t)\}$ for $\forall t$.
- 2) For a given test demand $\mathbf{d}^{\hat{t}}$, compute K closest neighbors, \mathcal{T}_K .
- 3) Binary $\mathbf{x}^{\hat{t}}$ as the rounded mean of the binary decision values from \mathcal{T}_K to the closest integer.
- 4) Solve an LP from the OTS by fixing variables.

Strategy 3, Knn-BM

- Data-driven approach (Knn).
- One of our proposals.

Strategy 3, Knn-BM

- Data-driven approach (Knn).
- One of our proposals.

Algorithm

- 1) Training set $\mathcal{T} = \{(\mathbf{d}^t, \mathbf{x}^t)\}$ for $\forall t$.
- 2) For a given test demand $\mathbf{d}^{\hat{t}}$, compute K closest neighbors, \mathcal{T}_K .
- 3) Fixing the binary variables if there unanimity in the value among the instances in \mathcal{T}_K .
- **4)** The bigM values of the remaining variables are updated using the shortest path.
- **5)** Solve the resulting MILP.

Strategy 4, Knn-B \widehat{M}

- Data-driven approach (Knn).
- One of our proposals.

Strategy 4, Knn-B \widehat{M}

- Data-driven approach (Knn).
- One of our proposals.

Algorithm

- 1) Training set $\mathcal{T} = \{(\mathbf{d}^t, \mathbf{x}^t)\}$ for $\forall t$.
- 2) For a given test demand $\mathbf{d}^{\hat{t}}$, compute K closest neighbors, \mathcal{T}_K .
- 3) Fixing the binary variables if there is unanimity in the value among the instances in \mathcal{T}_K .
- 4) The bigM values of the remaining variables are updated using the angles information.

$$\overline{M}_{nm} = b_{nm} \times \max_{t \in \mathcal{T}: x_{nm}^t = 0} (\theta_n^t - \theta_m^t)$$

$$\underline{M}_{nm} = b_{nm} \times \min_{t \in \mathcal{T}: x_{nm}^t = 0} (\theta_n^t - \theta_m^t)$$

5) Solve the resulting MILP.

Outline

- 1 Motivation
- 2 Methodology
- 3 Computational Experience
- 4 Conclusions and Further Research

Experimental Setup

- Realistic network, [Blumsack, 2006].
- 118 buses 186 lines.
- $|\mathcal{L}_S| = 69$.
- 500 instances.
- Demand follows uniform distribution in $[0.9d_n, 1.1d_n]$.
- Gurobi 9.1.2.
- Gap = 0.01%.
- Time limit: 1 hour.
- Github.

	# opt	# sub	gap-max	time (s)
BEN	500	0	-	145.44
Knn-D	0	500	14.13	0.0
Knn-BM	500	0	-	12.33
K nn-B $\widehat{\mathbf{M}}$	495	5	0.39	0.7

Table: All approaches (K = 50)

	# opt	# sub	gap-max	time (s)
BEN	500	0	-	145.44
Knn-D	0	500	14.13	0.0
Knn-BM	500	0	-	12.33
K nn-B $\widehat{\mathbf{M}}$	495	5	0.39	0.7

Table: All approaches (K = 50)

BEN vs Knn-D

- Knn-D has no optimal instances.
- \bullet Knn-D is faster.
- Max gap: 14%

	# opt	# sub	gap-max	time (s)
BEN	500	0	-	145.44
Knn-D	0	500	14.13	0.0
Knn-BM	500	0	-	12.33
K nn-B $\widehat{\mathrm{M}}$	495	5	0.39	0.7

Table: All approaches (K = 50)

BEN vs Knn-BM

- All optimal instances
- \bullet Knn-BM is faster.
- No gap.

	# opt	# sub	gap-max	time (s)
BEN	500	0	-	145.44
Knn-D	0	500	14.13	0.0
Knn-BM	500	0	-	12.33
K nn-B $\widehat{\mathbf{M}}$	495	5	0.39	0.7

Table: All approaches (K = 50)

BEN vs Knn-B \widehat{M}

- Knn-B \widehat{M} has almost all optimal instances
- Knn-B \widehat{M} is very fast.
- Small gap.

	# opt	# sub	gap-max	time (s)
BEN	500	0	-	145.44
Knn-D	0	500	14.13	0.0
Knn-BM	500	0	-	12.33
K nn-B $\widehat{\mathbf{M}}$	495	5	0.39	0.7

Table: All approaches (K = 50)

Knn-BM vs Knn-B \widehat{M}

- Knn-BM is more conservative than Knn-B \widehat{M} but slower.
- Trade-off decision.
- Competitive against existing approaches.

More details

Learning-Assisted Optimization for Transmission Switching

Salvador Pineda^{1,4*}, Juan Miguel Morales^{2,4} and Asunción Jiménez-Cordero^{3,4}

 $\label{lem:corresponding} $$ $$ \operatorname{corresponding author(s). E-mail(s): spineda@uma.es;} $$ $$ \operatorname{contributing authors: juan.morales@uma.es; asuncionjc@uma.es;} $$$

Available at:

S. Pineda, J. M. Morales and A. Jiménez-Cordero, Learning-Assisted Optimization for Transmission Switching, Submitted. Link: https://www.researchgate.net/publication/ 370058669_Learning-Assisted_Optimization_for_ Transmission_Switching.



^{1*}Dept. of Electrical Engineering, University of Málaga, Spain.
²Dept. of Applied Mathematics, University of Málaga, Spain.
³Dept. of Statistics and Operations Research, University of Málaga, Spain.

 $^{^4{\}rm OASYS}$ Research Group, University of Málaga, Spain.

Outline

- Motivation
- 2 Methodology
- 3 Computational Experience
- 4 Conclusions and Further Research

Conclusions

- OTS is a challenging problem.
- Useful to reduce costs.
- Find good bigM values for the reformulation.
- Learning strategies.
- Tested on a real-world network.

Conclusions

- OTS is a challenging problem.
- Useful to reduce costs.
- Find good bigM values for the reformulation.
- Learning strategies.
- Tested on a real-world network.

Further research

- All switchable lines.
- Other ML approaches.

References

- Bengio, Y., Lodi, A., and Prouvost, A. (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. European Journal of Operational Research, 290(2):405–421, doi:10.1016/j.ejor.2020.07.063.
- Blumsack, S. (2006). Network topologies and transmission investment under electric-industry restructuring. Carnegie Mellon University, Pittsburgh, Pennsylvania.
- Fattahi, S., Lavaei, J., and Atamtürk, A. (2019). A bound strengthening method for optimal transmission switching in power systems. *IEEE Transactions on Power Systems*, 34(1):280-291, doi:10.1109/TPWRS.2018.2867999.
- Gambella, C., Ghaddar, B., and Naoum-Sawaya, J. (2021). Optimization problems for machine learning: A survey. European Journal of Operational Research, 290(3):807–828, doi:10.1016/j.ejor.2020.08.045.
- Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction, volume 2. Springer, New Yourk.
- Lodi, A. and Zarpellon, G. (2017). On learning and branching: A survey. TOP, 25(2):207–236, doi:10.1007/s11750-017-0451-6.
- Ruiz, P. A., Goldis, E., Rudkevich, A. M., Caramanis, M. C., Philbrick, C. R., and Foster, J. M. (2016). Security-constrained transmission topology control milp formulation using sensitivity factors. *IEEE Transactions on Power Systems*, 32(2):1597–1605.
- Yang, Z. and Oren, S. (2019). Line selection and algorithm selection for transmission switching by machine learning methods. In 2019 IEEE Milan PowerTech, pages 1–6. IEEE.

A Novel Machine Learning Approach for Solving Optimal Transmission Switching



Asunción Jiménez-Cordero asuncionjc@uma.es

JOINT WORK WITH: Juan Miguel Morales González Salvador Pineda Morente

Thank you very much for your attention!









oasys.uma.es

3rd EUROYoung Workshop

June 6th, 2023

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 755705)