Al for 5G Sustainability: Green Savior or Great Pretender?

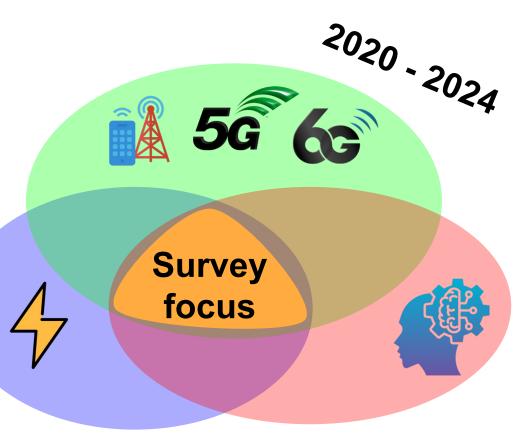
Kishan Sthankiya*, Nagham Saeed[§], Greg McSorley[^], Mona Jaber*, Richard G. Clegg*

> * Queen Mary University of London § University of West London ^ BT Research

36th Multi-Service Networks workshop (MSN 2024) 4th July 2024

Introduction

- Survey to assess:
- Energy impact on 5G by AI algorithms.
- Operational cost of AI.
- Does the Al cost more than it saves?



Motivation

ML Inference Energy



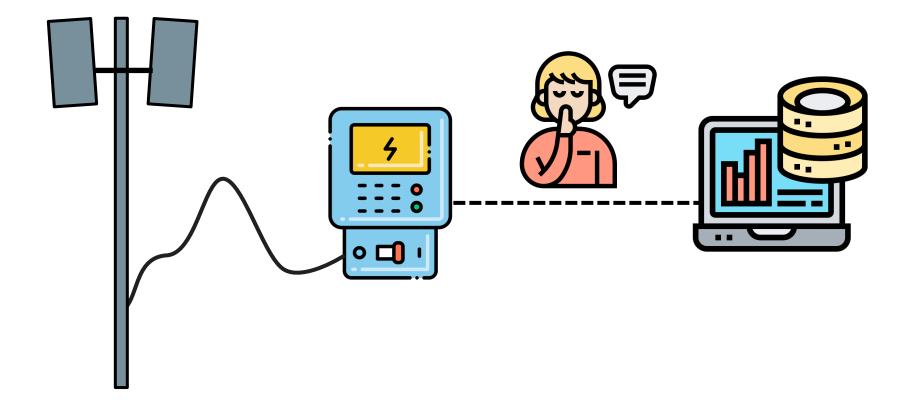




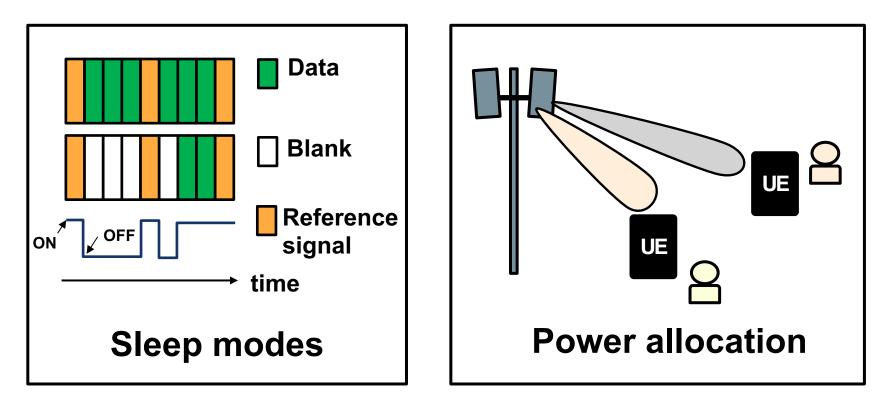
~ 14.2 MWh UK household (annual)

~ 205,860 MWh ~ 1,387,000 MWh GPT-3 ML pipeline (annual) (annual)

Power models



Energy Saving Techniques



Question #1

How much energy can be saved in the RAN using ML techniques?

Observation #1

Consumption savings: -9 - 65%

Efficiency improvement: -3 – 15%

Problem #1

Most studies don't report the energy used by their ML model.

Suggestion #1

Green Algorithms How green are your computations?							
Check out the new Green Algorithms website: <u>www.green-algorithms.org</u>							
Details about your algorithm To understand how each parameter impacts your carbon footprint, check out the formula below and the <u>methods article</u>			CO 2	253.64 g CO2e Carbon footprint		4	2.28 kWh Energy needed
Runtime (HH:MM)	12	0					
Type of cores	CPU				æ		ç
Number of cores		12	0.28 tree-	0.28 tree-months		km	0.51%
Model	Xeon E5-	2683 v4 🔻	Carbon sequestration		in a passenger car		of a flight Paris-London

Question #2

How we can compare the energy impact of Algorithm A vs. Algorithm B?

Observation #2

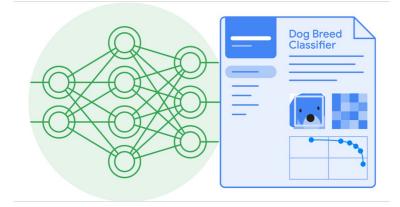
No standard algorithm performance

metrics that incorporate energy.

ML Energy Factors

- Computational load imposed by the model.
- Number of iterations.
- CPU frequency / time to completion.

Suggestion #2



Model Report Cards

Metadata that may be useful:

- FLOPS/Watt
- System configuration report
- Location
- Cooling system
- Total energy consumption for training **and inference**
- Mean inference duration (ms)
- Training duration

Key Points

- Accurate energy profiling of 5G RAN
 remains an open question.
- Estimating energy performance of algorithms is non-trivial.
- 'Report cards' could help compare model credentials.



TIME, JULY 28, 1952

Questions?