



Ultra Low Power Ambient Artificial Intelligence

Antoine Bonneau, Frédéric Le Mouël, Fabien Mieyeville

► To cite this version:

Antoine Bonneau, Frédéric Le Mouël, Fabien Mieyeville. Ultra Low Power Ambient Artificial Intelligence. Journées sur la recherche en apprentissage frugal (JRAF), Grenoble INP; Inria; EcoInfo, Nov 2022, Grenoble, France. hal-03986104

HAL Id: hal-03986104

<https://hal.science/hal-03986104>

Submitted on 13 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Ultra Low Power Ambient Artificial Intelligence

Antoine Bonneau^{1,2} Frédéric Le Mouël¹ Fabien Mieyeville²

Contact : `antoine.bonneau@insa-lyon.fr`

¹Univ Lyon, INSA Lyon, Inria, CITI, EA3720, 69621 Villeurbanne, France

²Univ Lyon, Université Claude Bernard Lyon 1, INSA Lyon, Ecole Centrale de Lyon, CNRS, Ampère, UMR5005, 69622 Villeurbanne, France

JRAF - Grenoble - Nov. 2022

1 Use Cases & Context

- Ampere's active control platform
- CITI's Cortex Lab
- Ambient / Frugal AI
- Thesis Subject

2 Research key points / State of the art (SoA)

- On-Device Learning
- Intermittent Learning
- Federated Learning

3 Work plan

Ampère's active control platform

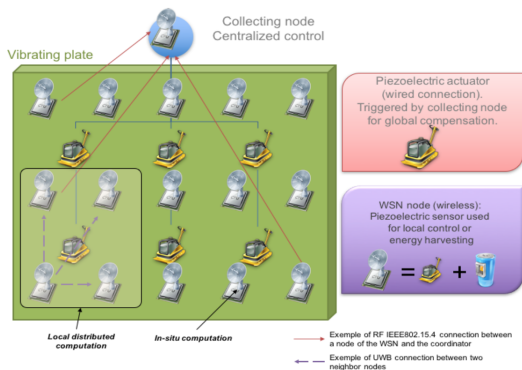


Figure 1: Ampère's vibrating bench

► Issues

- Vibration Reduction
- Energy Harvesting

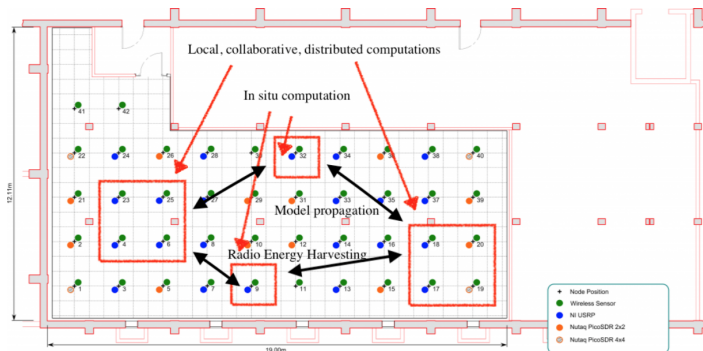


Figure 2: Nodes and activation example in Cortex Lab platform

► Issues

- reallocate / power off antennas
- max covering / bandwidth / homogeneity
- min interferences

- ▶ Ambient Intelligence (AmI)¹
 - ▶ embedded
 - ▶ context-aware
 - ▶ adaptive
 - ▶ anticipatory
- ▶ Frugal Artificial Intelligence (Frugal AI)²
 - ▶ input : less data
 - ▶ learning process : less computational resources
 - ▶ model : less memory
 - ★ reduced energy consumption

¹E. Aarts [Jan. 2004]. “Ambient intelligence: a multimedia perspective”. In: *IEEE Multimedia*. DOI: [10.1109/MMUL.2004.1261101](https://doi.org/10.1109/MMUL.2004.1261101).

²Mikhail Evchenko et al. [Nov. 2021]. *Frugal Machine Learning*. DOI: [10.48550/arXiv.2111.03731](https://doi.org/10.48550/arXiv.2111.03731).

Thesis Subject : Ultra Low Power Ambient Artificial Intelligence

- ▶ Started last month : $\sim$ 20 papers selected
- ▶ Goals
 - ▶ Deployment of a WSN as a monitoring and adaptive action solution
 - ▶ Minimal footprint energy aiming for autonomy
 - ▶ Exploiting networking AI to improve and maintain itself
- ★ Energy efficiency as a cross-cutting topic
- ▶ Focus on Time Series data (reduced feature set and samples)

Outline : State of the Art

1 Use Cases & Context

- Ampere's active control platform
- CITI's Cortex Lab
- Ambient / Frugal AI
- Thesis Subject

2 Research key points / State of the art (SoA)

- On-Device Learning
- Intermittent Learning
- Federated Learning

3 Work plan

SoA On-Device Learning : Resources constrained learning

- ▶ Target : WSN Nodes $\rightarrow$ IoT low-end devices [Ojo et al. 2018]

MCU Arch	Clock speed	On-Board Storage	RAM	Current
16 or 32 bit	$\leq 50\text{ MHz}$	$\leq 512\text{ KB}$	$\leq 64\text{ KB}$	$\leq 10\text{ mA}$

- ▶ Memory constraints $\rightarrow$ Compressed Models [Profentzas et al. 2022]
 - ▶ Pruning
 - ▶ Quantization
 - ▶ Encoding
- ▶ Computation constraints $\rightarrow$ Preprocessed Models [Lin et al. 2022]
 - ▶ Transfer Learning
 - ▶ Fine Tuning

SoA On-Device Learning : Resources constrained learning

- ▶ Target : WSN Nodes $\rightarrow$ IoT low-end devices [Ojo et al. 2018]

MCU Arch	Clock speed	On-Board Storage	RAM	Current
16 or 32 bit	$\leq 50\text{ MHz}$	$\leq 512\text{ KB}$	$\leq 64\text{ KB}$	$\leq 10\text{ mA}$

- ▶ Memory constraints $\rightarrow$ Compressed Models [Profentzas et al. 2022]

- ▶ Pruning
- ▶ Quantization
- ▶ Encoding

Use 2-3 $\times$ more RAM

- ▶ Computation constraints $\rightarrow$ Preprocessed Models [Lin et al. 2022]

- ▶ Transfer Learning
- ▶ Fine Tuning

Tested on IoT middle-end device
No energy consumption evaluation

SoA Intermittent Learning : Duty cycling

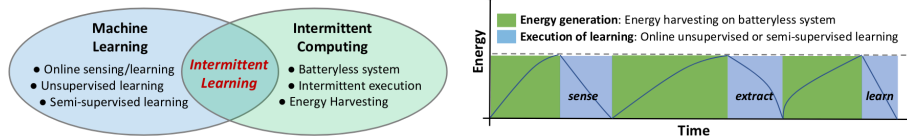


Figure 3: Intermittent Learner duty [Lee et al. 2019]

- ▶ Energy Harvesting → Extreme power scarcity
 - ▶ checkpointing approaches
 - ▶ task-based approaches [Lee et al. 2019]
- ▶ Changes in energy availability → Energy-aware reconfiguration [Bakar et al. 2021]
 - ▶ peripheral selection / sensor sampling
 - ▶ communication pattern

SoA Intermittent Learning : Duty cycling

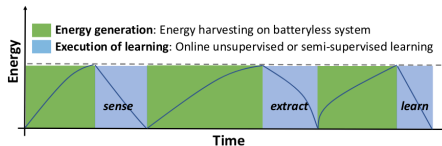
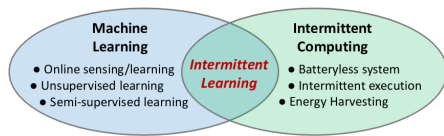


Figure 3: Intermittent Learner duty [Lee et al. 2019]

- ▶ Energy Harvesting → Extreme power scarcity
 - ▶ checkpointing approaches
 - ▶ task-based approaches [Lee et al. 2019]
- Do not consider powering off by itself
- ▶ Changes in energy availability → Energy-aware reconfiguration [Bakar et al. 2021]
 - ▶ peripheral selection / sensor sampling
 - ▶ communication pattern
- Use tiny machine learning model instead of heuristics

SoA Federated Learning : Bypass computational weaknesses

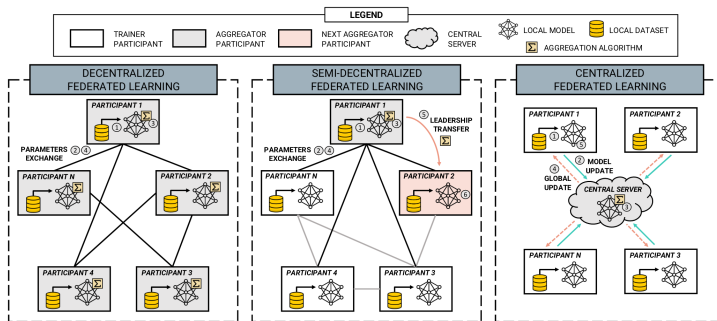


Figure 4: DFL architectures [Beltrán et al. 2022]

- ▶ Learning in a spatially correlated group → Adaptive Neighbor Matching [Z. Li et al. 2022]
- ▶ Costly communication → reduce shared weights [C. Li et al. 2021]

SoA Federated Learning : Bypass computational weaknesses

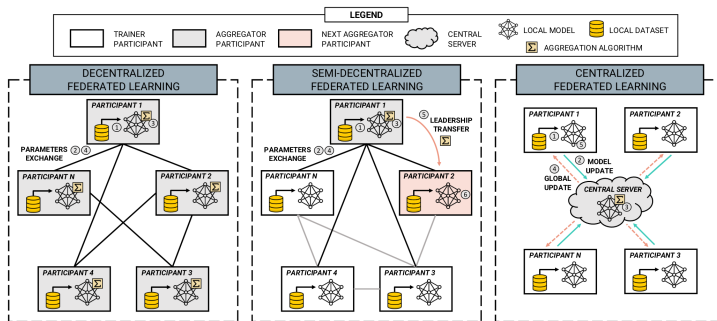


Figure 4: DFL architectures [Beltrán et al. 2022]

Do not evolve with changes & resource-consuming

- ▶ Learning in a spatially correlated group → Adaptive Neighbor Matching [Z. Li et al. 2022]
 - ▶ Costly communication → reduce shared weights [C. Li et al. 2021]
- Use a centralized FL Algorithm

1 Use Cases & Context

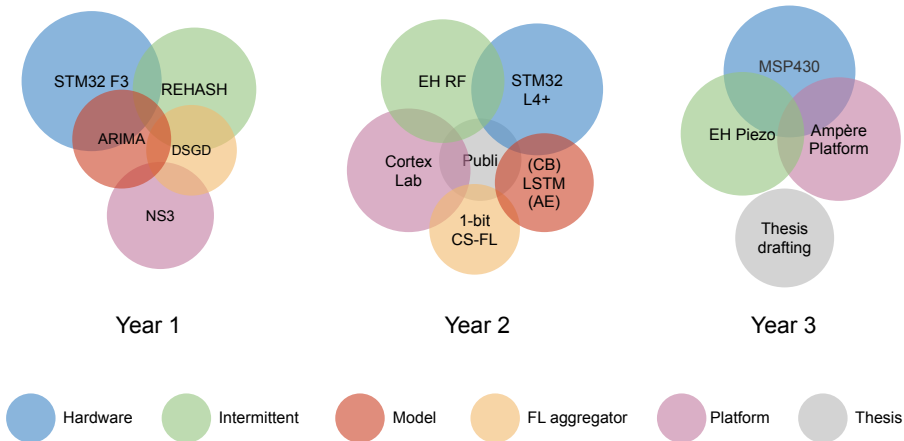
- Ampere's active control platform
- CITI's Cortex Lab
- Ambient / Frugal AI
- Thesis Subject

2 Research key points / State of the art (SoA)

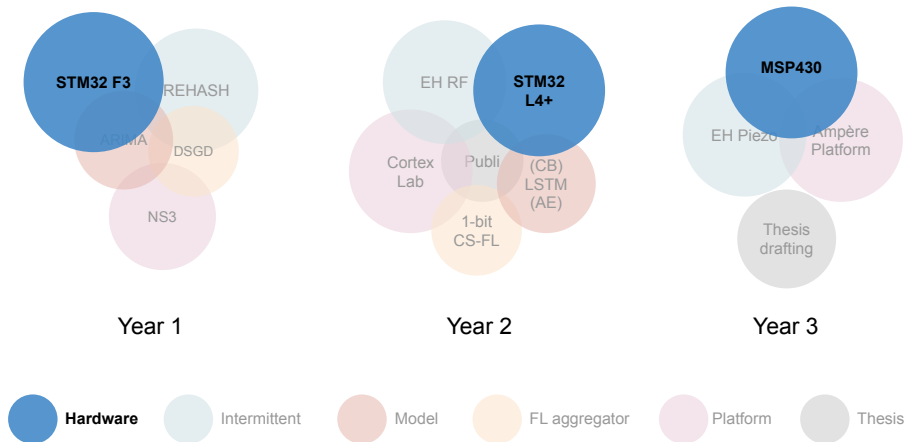
- On-Device Learning
- Intermittent Learning
- Federated Learning

3 Work plan

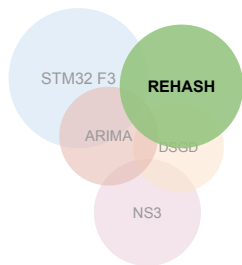
Work plan ... will most likely age poorly



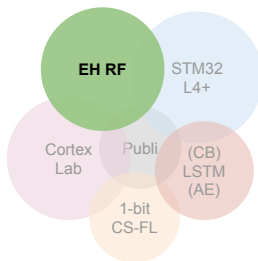
Work plan ... will most likely age poorly



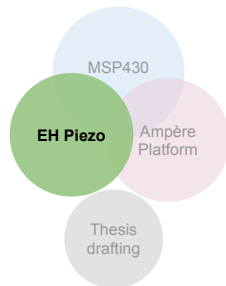
Work plan ... will most likely age poorly



Year 1



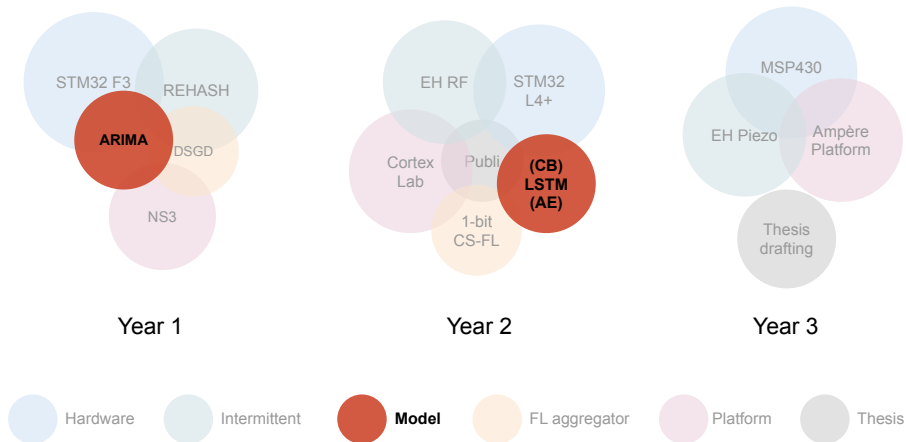
Year 2



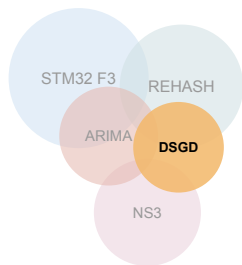
Year 3



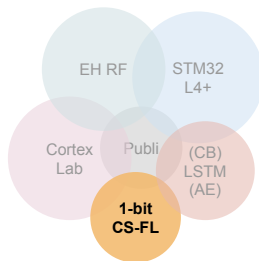
Work plan ... will most likely age poorly



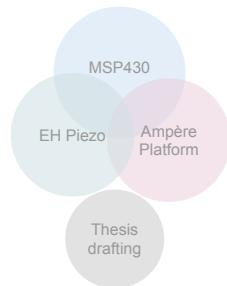
Work plan ... will most likely age poorly



Year 1



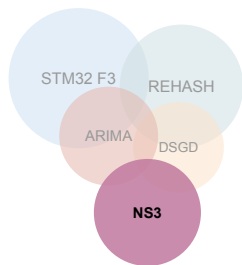
Year 2



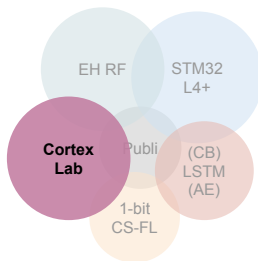
Year 3



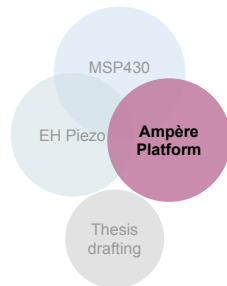
Work plan ... will most likely age poorly



Year 1



Year 2



Year 3



Ultra Low Power Ambient Artificial Intelligence

Antoine Bonneau^{1,2} Frédéric Le Mouël¹ Fabien Mieyeville²

Contact : `antoine.bonneau@insa-lyon.fr`

¹Univ Lyon, INSA Lyon, Inria, CITI, EA3720, 69621 Villeurbanne, France

²Univ Lyon, Université Claude Bernard Lyon 1, INSA Lyon, Ecole Centrale de Lyon, CNRS, Ampère, UMR5005, 69622 Villeurbanne, France

JRAF - Grenoble - Nov. 2022

References I

-  Bakar, Abu et al. (Sept. 2021). "REHASH: A Flexible, Developer Focused, Heuristic Adaptation Platform for Intermittently Powered Computing". In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* DOI: 10.1145/3478077.
-  Beltrán, Enrique Tomás Martínez et al. (Nov. 2022). *Decentralized Federated Learning: Fundamentals, State-of-the-art, Frameworks, Trends, and Challenges*. DOI: 10.48550/arXiv.2211.08413.
-  Lee, Seulki et al. (Dec. 2019). "Intermittent Learning: On-Device Machine Learning on Intermittently Powered System". In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* DOI: 10.1145/3369837.
-  Li, Chengxi et al. (Oct. 2021). "Communication-Efficient Federated Learning Based on Compressed Sensing". In: *IEEE Internet Things J.* DOI: 10.1109/JIOT.2021.3073112.

References II

-  Li, Zexi et al. (Nov. 2022). “Towards Effective Clustered Federated Learning: A Peer-to-peer Framework with Adaptive Neighbor Matching”. In: *IEEE Transactions on Big Data*. DOI: [10.1109/TBDATA.2022.3222971](https://doi.org/10.1109/TBDATA.2022.3222971).
-  Lin, Ji et al. (July 2022). “On-Device Training Under 256KB Memory”. In: DOI: [10.48550/arXiv.2206.15472](https://doi.org/10.48550/arXiv.2206.15472).
-  Ojo, Mike O. et al. (2018). “A Review of Low-End, Middle-End, and High-End IoT Devices”. In: *IEEE Access*. DOI: [10.1109/ACCESS.2018.2879615](https://doi.org/10.1109/ACCESS.2018.2879615).
-  Profentzas, Christos et al. (Sept. 2022). “MiniLearn: On-Device Learning for Low-Power IoT Devices”. In: *International Conference on Embedded Wireless Systems and Networks*. URL: <https://research.chalmers.se/en/publication/531933>.