Carnegie Mellon University

CoRAST: Towards Foundation Model-Powered Correlated Data Analysis in Resource-Constrained CPS and IoT

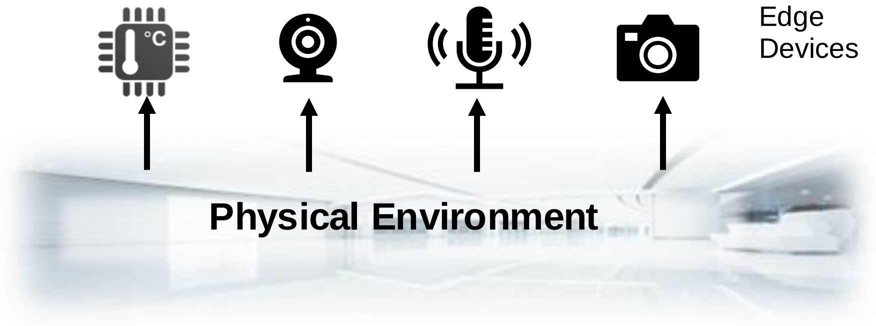
Yi Hu, Jinhang Zuo, Alanis Zhao, Bob Iannucci, Carlee Joe-Wong
CMU, Electrical & Computer Engineering

Overview

- **Correlated Data Analysis**
  - Background & Motivation
- **FM & Challenges**
- **CoRAST Framework**
  - FM for *Analyzing Spatially and Temporally CoRrelated data*
- **Evaluation**
- **Conclusion**

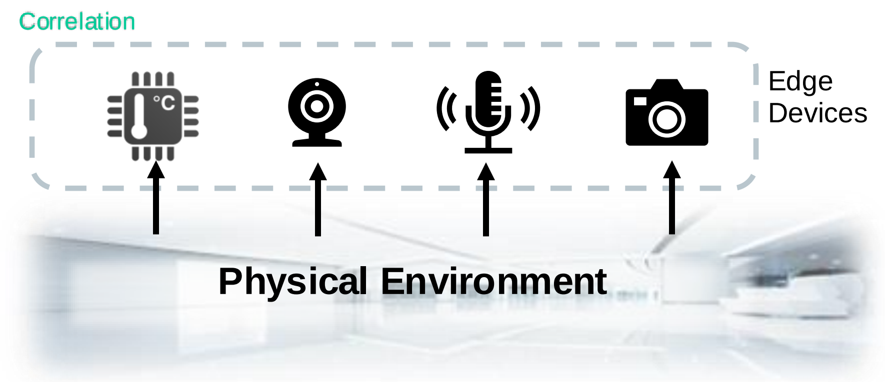
Motivations: Correlated Data in CPS/IoT

- Data collected from *the same physical environment*



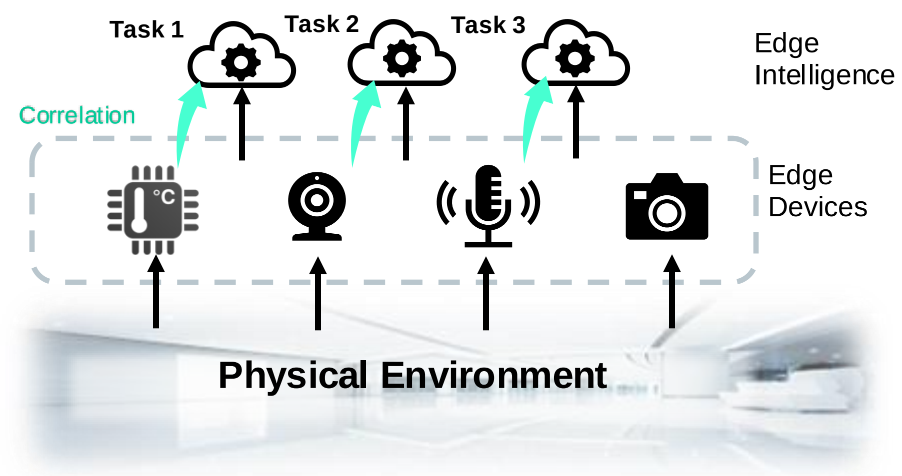
Motivations: Correlated Data in CPS/IoT

- Data collected from *the same physical environment*
- Inherently rich in spatial/temporal/cross-modal correlations
- E.g., visual-supported speech recognition



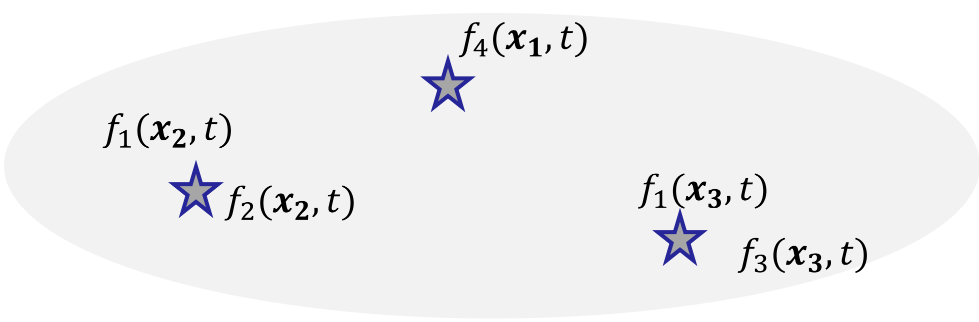
Motivations: Correlated Data in CPS/IoT

- Data collected from *the same physical environment*
- Inherently rich in spatial/temporal/cross-modal correlations
- E.g., visual-supported speech recognition



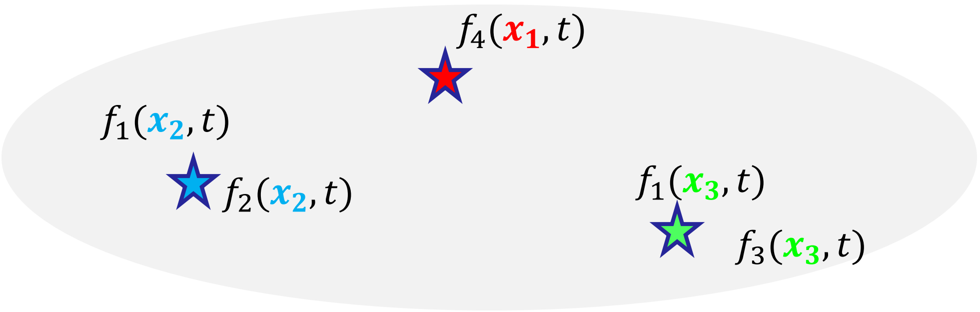
Motivations: Correlated Data Analysis

- Spatiotemporal Correlation



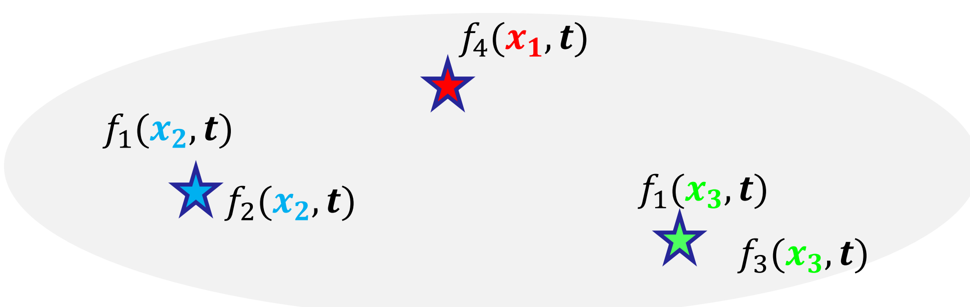
Motivations: Correlated Data Analysis

- Spatiotemporal Correlation
  - Data collected at different locations



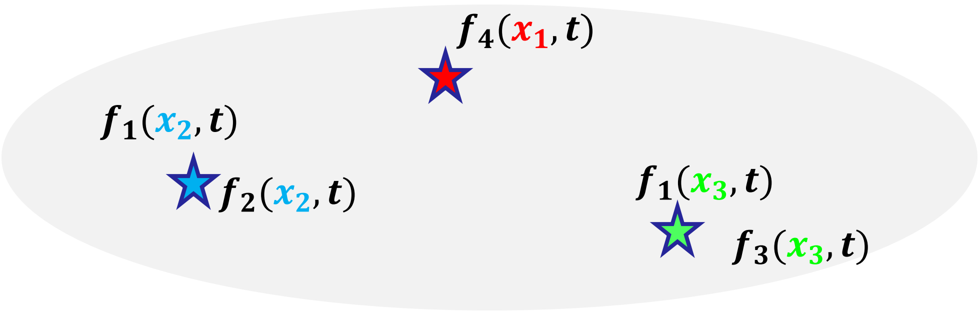
Motivations: Correlated Data Analysis

- Spatiotemporal Correlation
  - Data collected at different locations
  - Temporal evolution/dynamics



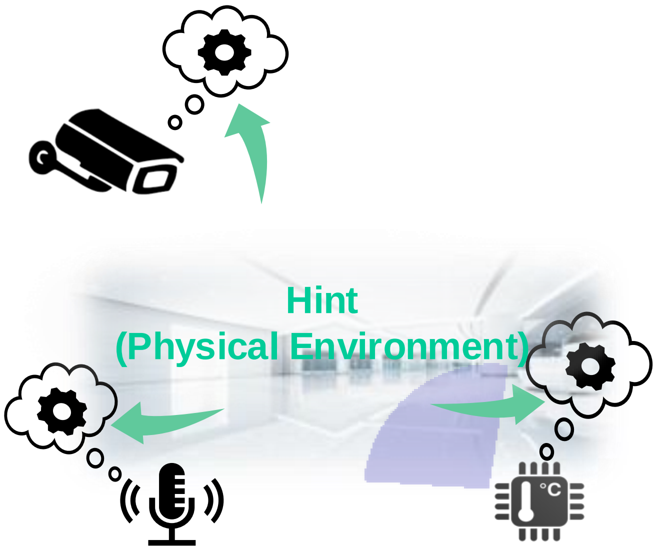
Motivations: Correlated Data Analysis

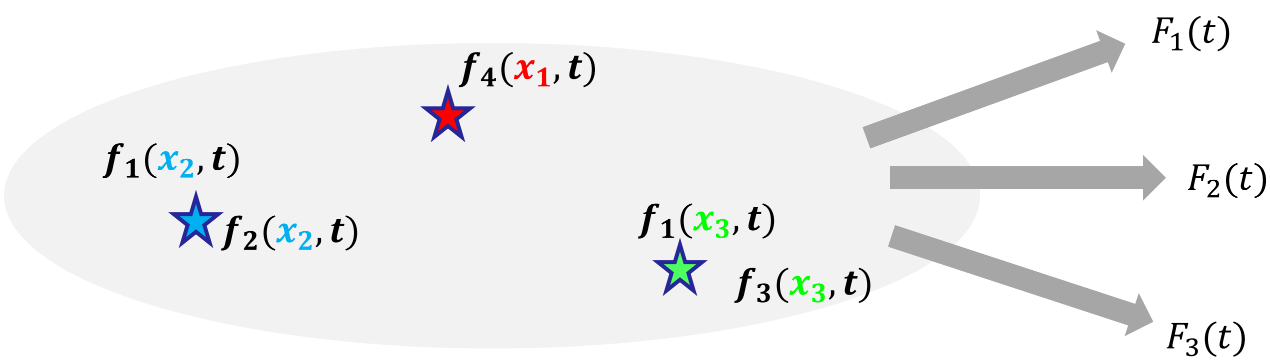
- Spatiotemporal Correlation
  - Data collected at different locations
  - Temporal evolution/dynamics
- Multi-Modality & Data Heterogeneity
  - Time-series, audio, video streams…



Motivations: Correlated Data Analysis

- Spatiotemporal Correlation
  - Data collected at different locations
  - Temporal evolution/dynamics
- Multi-Modality & Data Heterogeneity
  - Time-series, audio, video streams…
- Data Fusion & Learning
  - Facilitate the extraction of useful aspects from distributed data





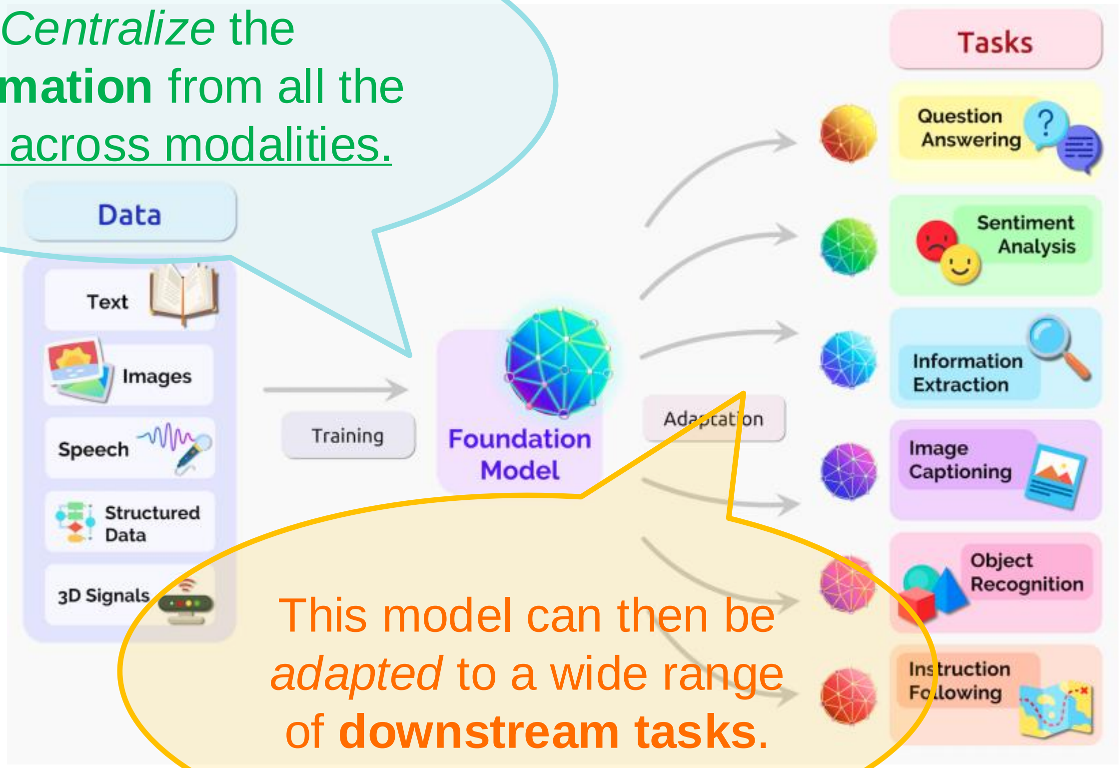
Foundation Models

- Trained on broad data, generally in a self-supervised manner
- Can be adapted (fine-tuned) to a wide range of downstream tasks [1]

Foundation Models Can…

Transfer learning

- ***Transformers***: convert diverse inputs into embeddings
- ***Self-supervised learning***: pretraining from unannotated data



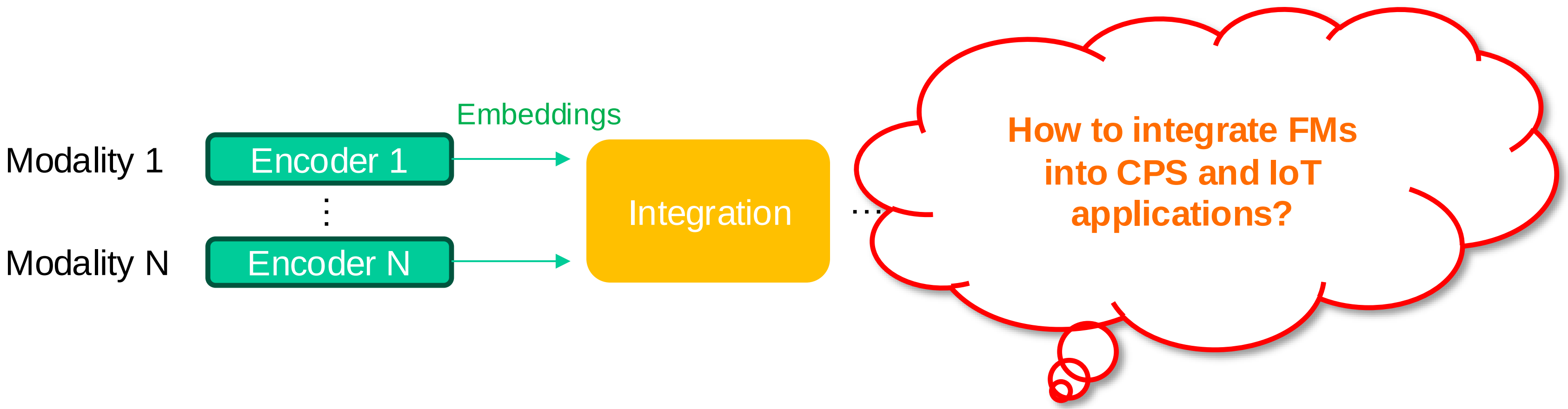
Centralize the **information** from all the data across modalities.

This model can then be *adapted* to a wide range of **downstream tasks**.

Source: [1]

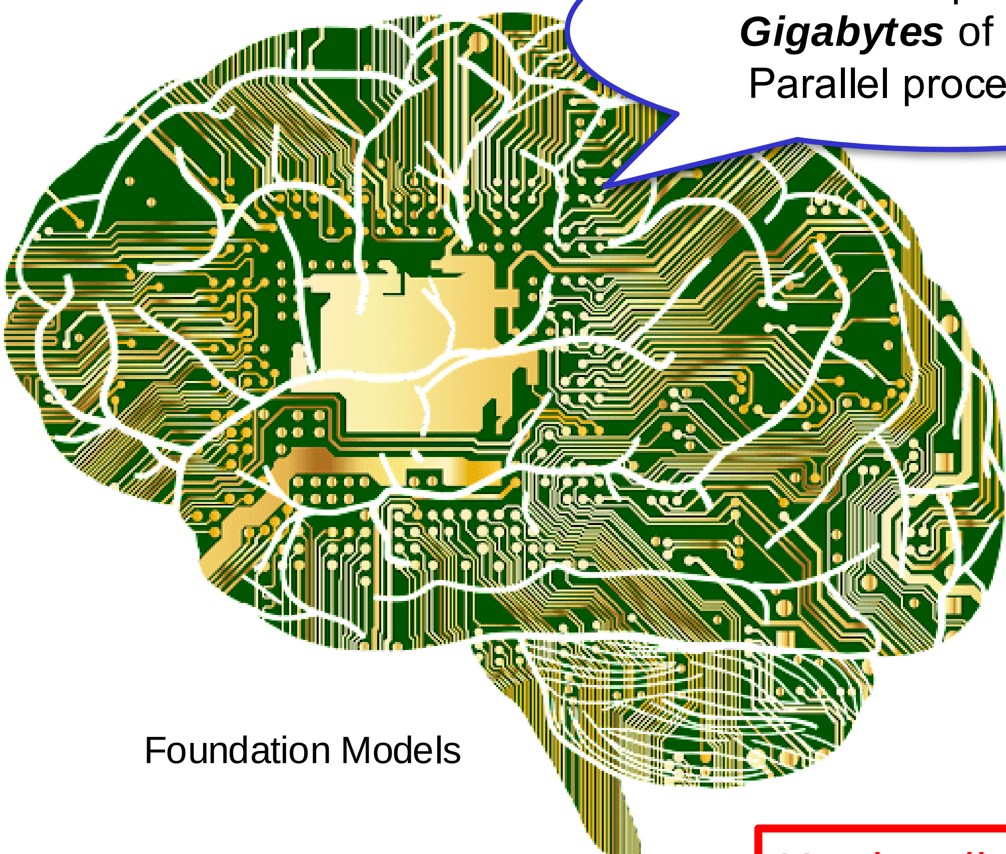
Foundation Models are Promising…

- LLM for time series analysis (**adaptation**)
  - GPT4TS [2], LLM4TS [3] for forecasting, anomaly detection
- Multi-modal models (**representation learning**):
  - CLIP [4]: text-image translation (text/image encoders + contrastive learning)
  - Macaw-LLM [5]: integrate image, audio, video, and text
  - CreamFL [6]: federated learning of modality-specific representations



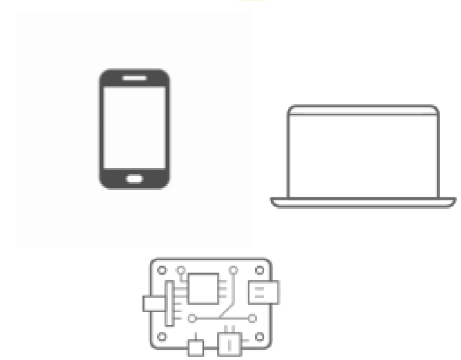
Challenge 1: Resource Constraints

Millions of parameters
Gigabytes of memory
Parallel processing…



Foundation Models

Battery-powered
Limited memory
Limited compute



Edge Devices

Hard to directly host FMs on edge devices

Source: https://ghostmountainllc.com/foundation-models/

Challenge 2: Correlated Local Tasks

Distributed decision-making → localized tasks/models

- Spatiotemporal Correlation
  - Anomaly detection at different points
  - Intersection traffic light control

$\boldsymbol{f_3(x_3, t_3)}$

$\boldsymbol{f_1(x_1, t_1)}$

$\boldsymbol{f_2(x_2, t_2)}$

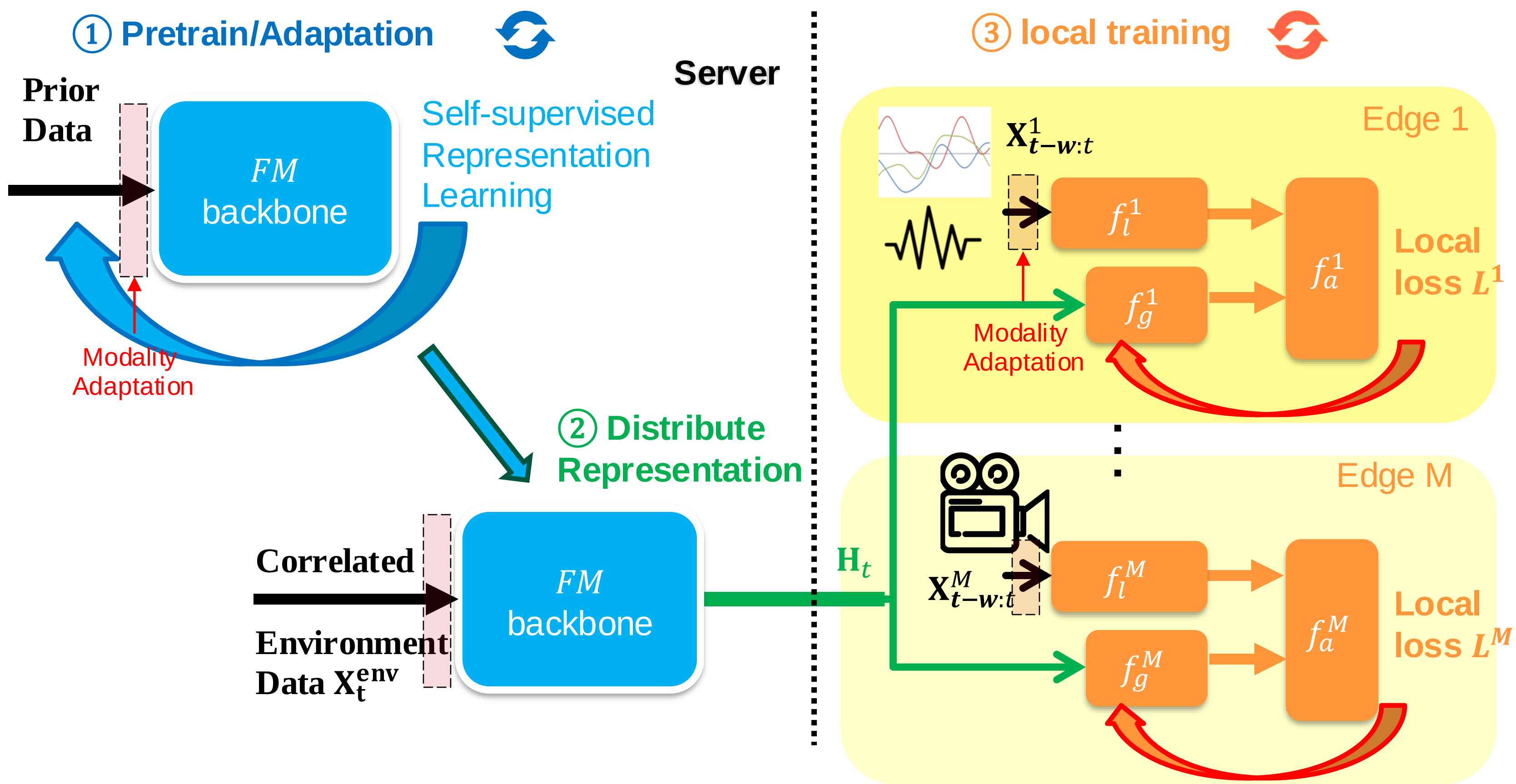
- Cross-modality Correlation
  - Visual (lip reading) and audio (speech recognition)
  - Weather data (pressure, moisture) and forecast

CoRAST:

FM for Analyzing Spatially and Temporally CoRrelated data

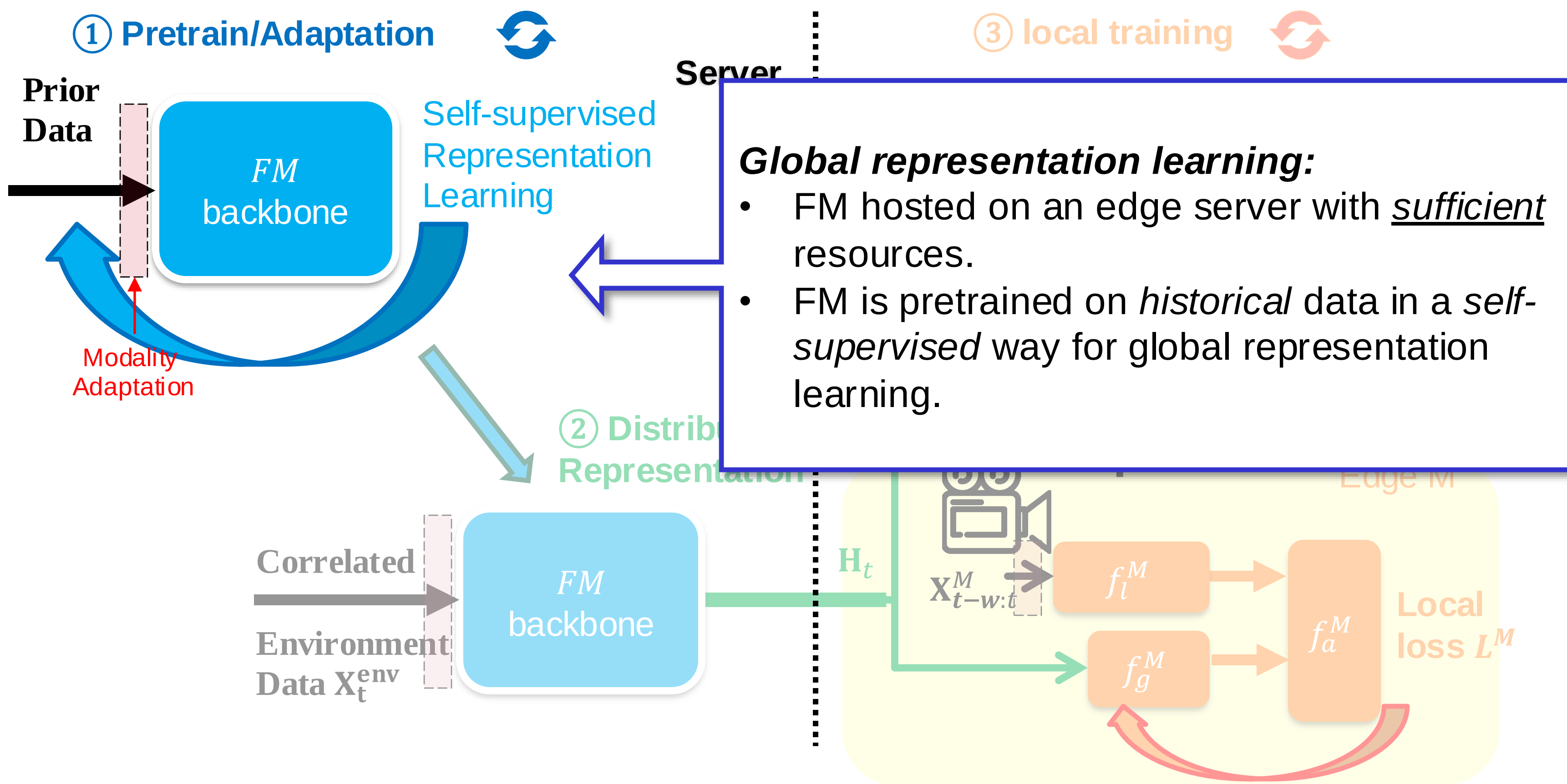
- Server-Based Representation Learning
- Client Local Training

CoRAST



① Pretrain/Adaptation
Server
Prior Data
FM backbone
Self-supervised Representation Learning
Modality Adaptation
② Distribute Representation
Correlated Environment Data X_t^{env}
FM backbone
H_t
③ local training
Edge 1
$X_{t-w:t}^1$
f_l^1
f_g^1
f_a^1
Local loss L^1
Modality Adaptation
Edge M
$X_{t-w:t}^M$
f_l^M
f_g^M
f_a^M
Local loss L^M

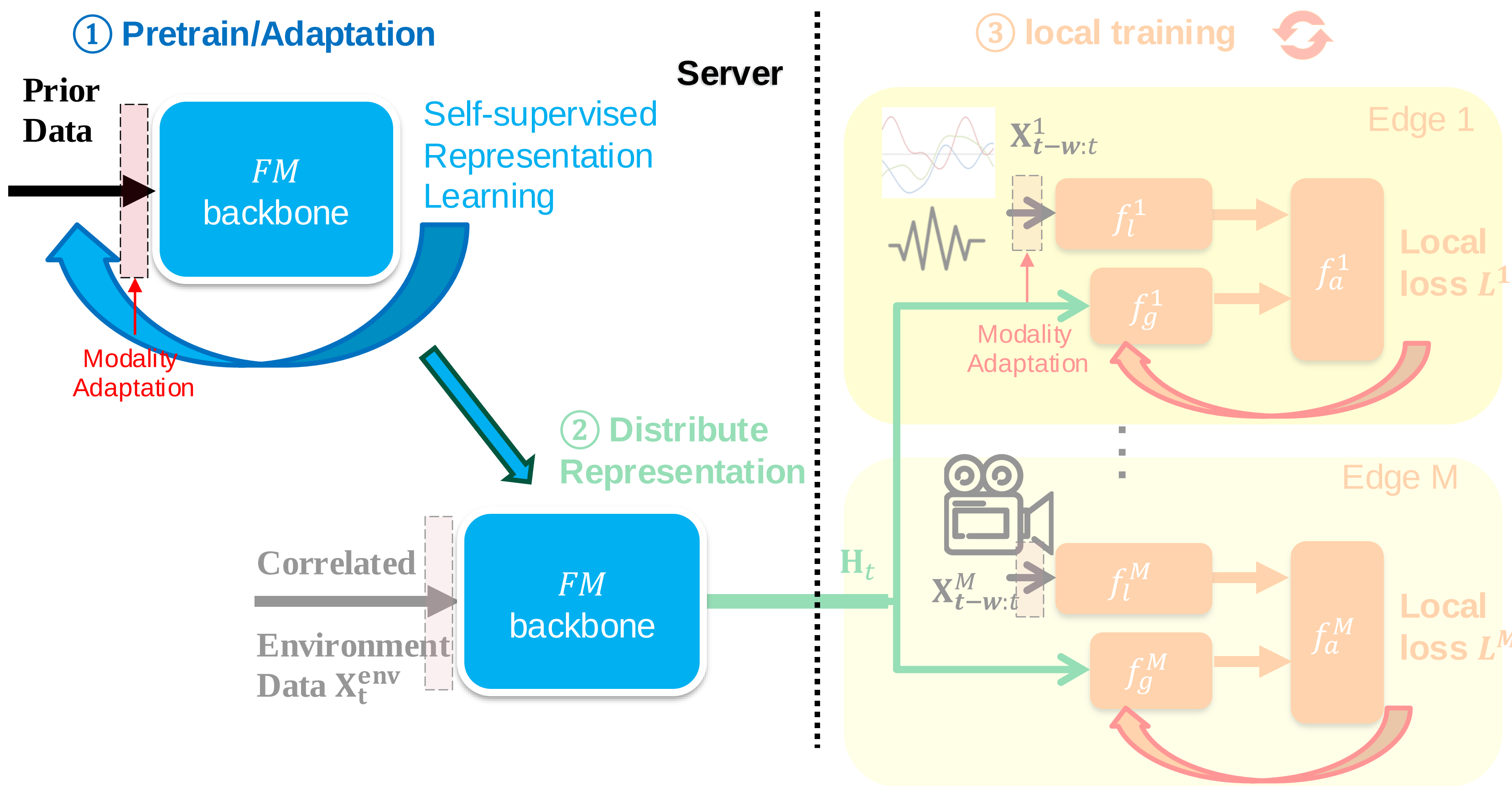
CoRAST: Representation Learning



Global representation learning:

- FM hosted on an edge server with *sufficient* resources.
- FM is pretrained on *historical* data in a *self-supervised* way for global representation learning.

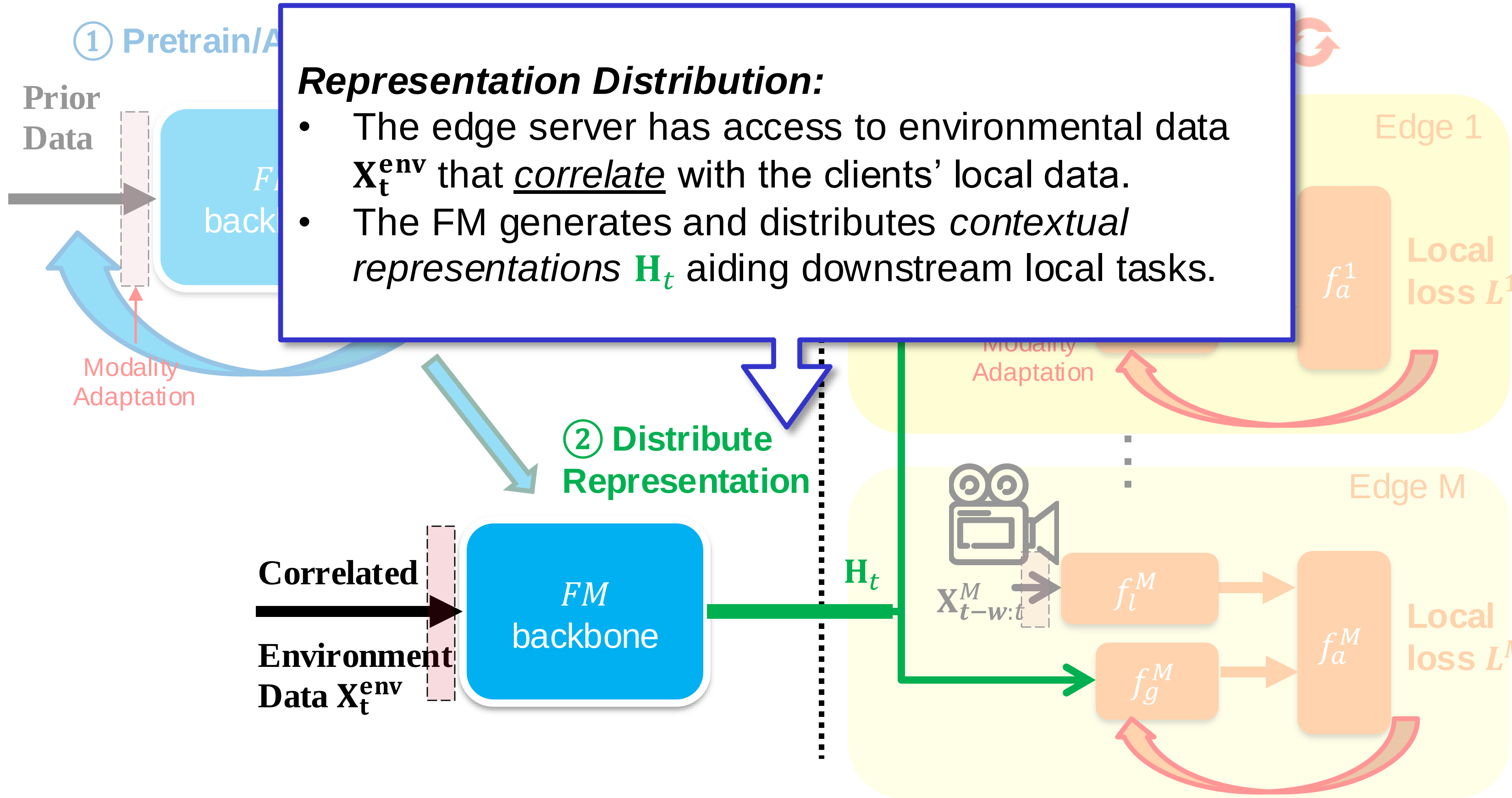
CoRAST: Representation Learning



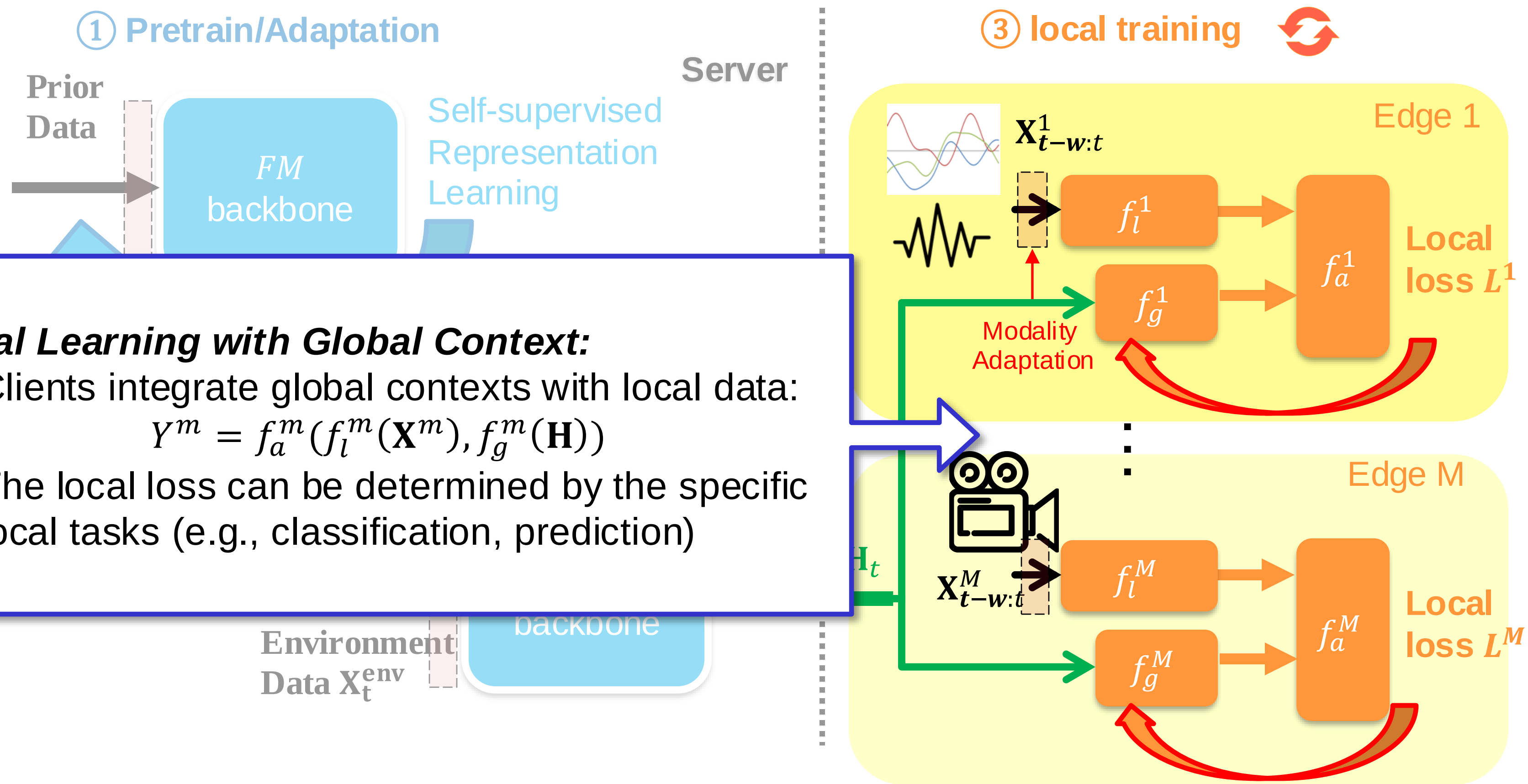
CoRAST: Representation Learning

Representation Distribution:

- The edge server has access to environmental data $\mathbf{X}_\mathbf{t}^{\mathbf{env}}$ that *correlate* with the clients' local data.
- The FM generates and distributes *contextual representations* $\mathbf{H}_t$ aiding downstream local tasks.



CoRAST: Local Training



Local Learning with Global Context:

- Clients integrate global contexts with local data:

$$Y^m = f_a^m(f_l^m(\mathbf{X}^m), f_g^m(\mathbf{H}))$$

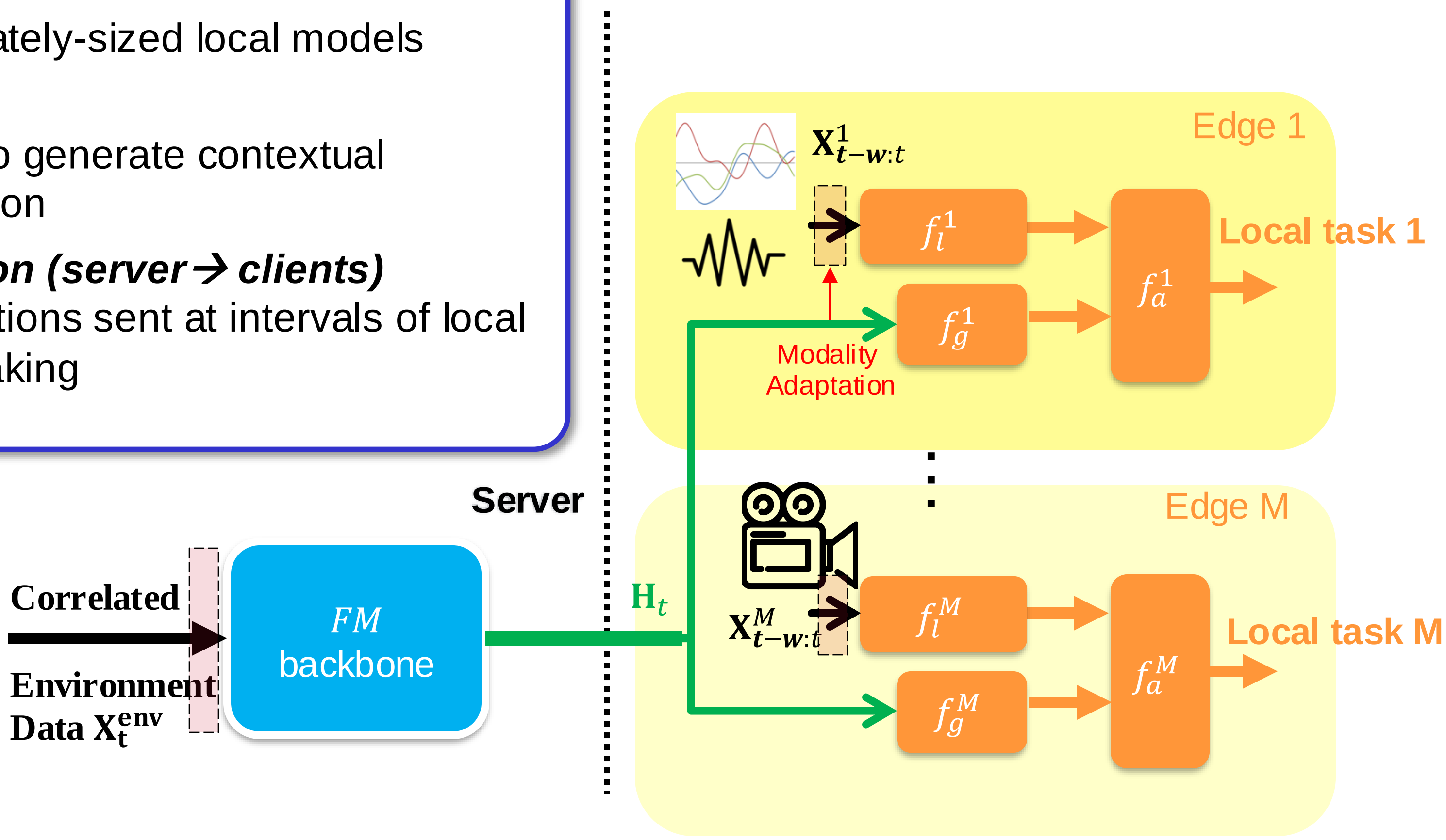
- The local loss can be determined by the specific local tasks (e.g., classification, prediction)

CoRAST: Runtime

Clients
- Run moderately-sized local models

Server
- Utilize FM to generate contextual representation

Communication (server→ clients)
- Representations sent at intervals of local decision-making



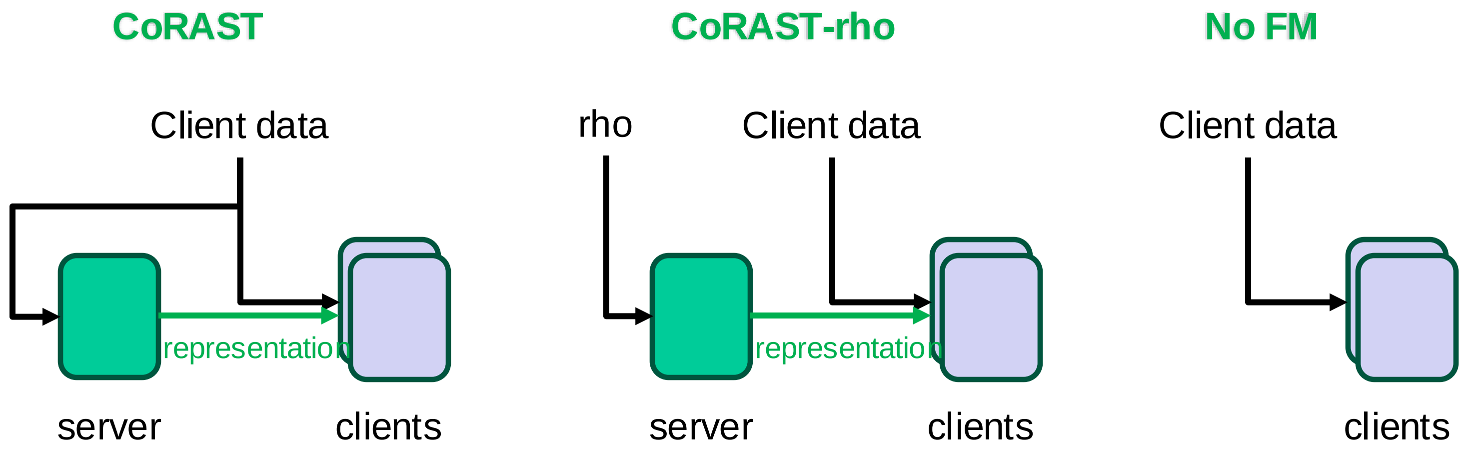
Evaluation

- Aligned Objective
- Diverse Local Tasks

Experiment Setup

Weather dataset: Temperature, pressure (p), relative humidity (rh)…

	Server	Clients	Representation
Model	TS2Vec	TCN	
# of parameters	Over 33.7k	13~15.2k	256

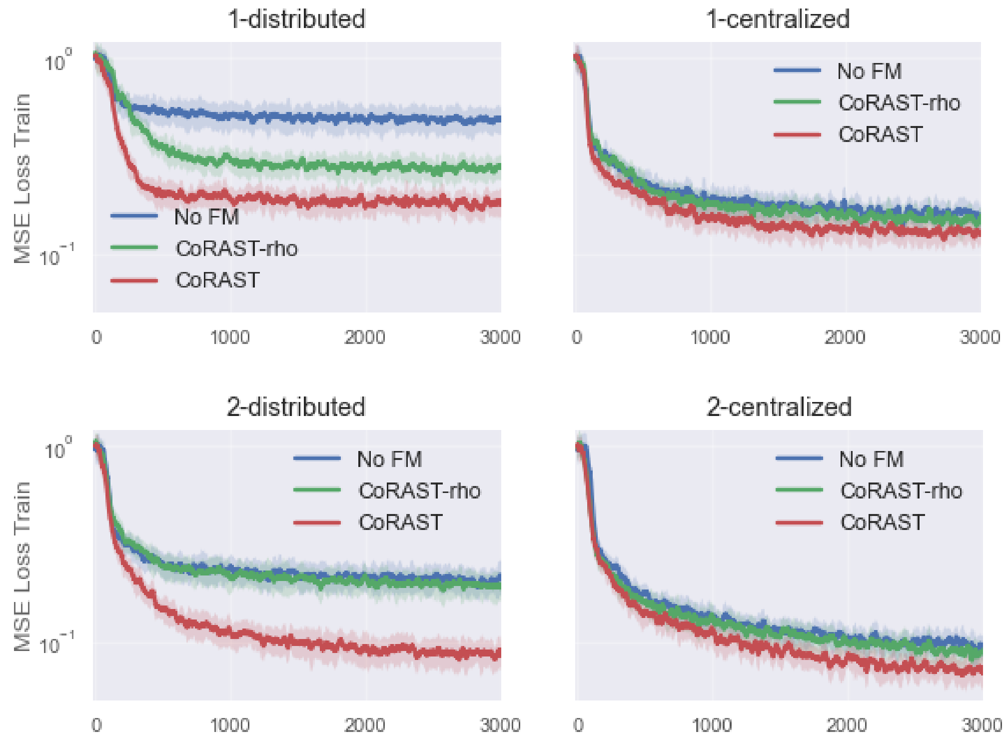


H2CO Forecast

Setting 1	Local data	Task
Centralized	Tdew, rh, sh	→ H2CO
Distributed	Tdew	→ H2CO
	rh	→ H2CO
	sh	→ H2CO

Setting 2	Local data	Task
Centralized	Tdew, rh, sh, Tpot, p	→ H2CO
Distributed	Tdew, Tpot	→ H2CO
	rh, p	→ H2CO
	sh	→ H2CO

Train MSE Loss



Test MSE

Setting	No FM	CoRAST-rho	CoRAST
1-centralized	0.195	0.182	**0.171**
1-distributed	0.391	0.303	**0.201**
2-centralized	0.075	**0.055**	0.061
2-distributed	0.151	0.146	**0.063**

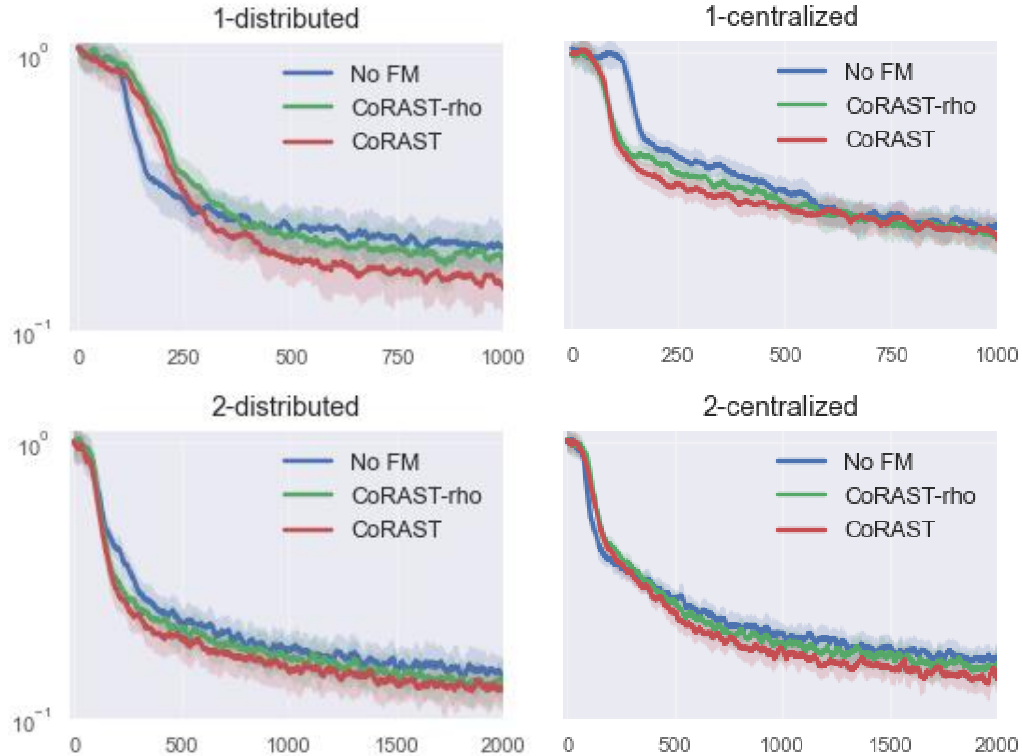
- CoRAST significantly improves distributed learning by adding additional environmental insights.
- CoRAST facilitates a more effective aggregation of client data.

Local Forecast

Setting 1	Local data	Task
Centralized	Tdew, rh, sh	→ Tdew, rh, sh
Distributed	Tdew	→ Tdew
	rh	→ rh
	sh	→ sh

Setting 2	Local data	Task
Centralized	Tdew, rh, sh, Tpot, p	→ Tdew, rh, sh, Tpot, p
Distributed	Tdew, Tpot	→ Tdew, Tpot
	rh, p	→ rh, p
	sh	→ sh

Train MSE Loss



Test MSE

Setting	Variable	No FM	CoRAST-rho	CoRAST
1-distributed	Tdew	0.072	**0.056**	0.075
	rh	0.297	0.318	**0.276**
	sh	0.077	0.072	**0.059**
2-distributed	Tdew	1.953	**1.796**	1.853
	rh	0.223	0.216	**0.211**
	sh	1.945	**1.912**	1.930
	p	0.151	0.159	**0.145**
	Tpot	2.633	2.544	**2.438**

CoRAST can enhance the distributed learning of interrelated tasks.

Conclusion

- **CoRAST Framework**
  - The first FM-based learning framework for analyzing correlated heterogeneous data that support diverse downstream tasks
- **Proof-of-concept Evaluation**
  - CoRAST improves distributed learning on a real-world weather dataset, reducing forecasting errors with its FM-based global learning approach.

References

[1] Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

[2] Zhou, T., Niu, P., Sun, L., & Jin, R. (2024). One fits all: Power general time series analysis by pretrained lm. *Advances in neural information processing systems*, *36*.

[3] Chang, C., Peng, W. C., & Chen, T. F. (2023). Llm4ts: Two-stage fine-tuning for time-series forecasting with pre-trained llms. arXiv preprint arXiv:2308.08469.

[4] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PMLR.

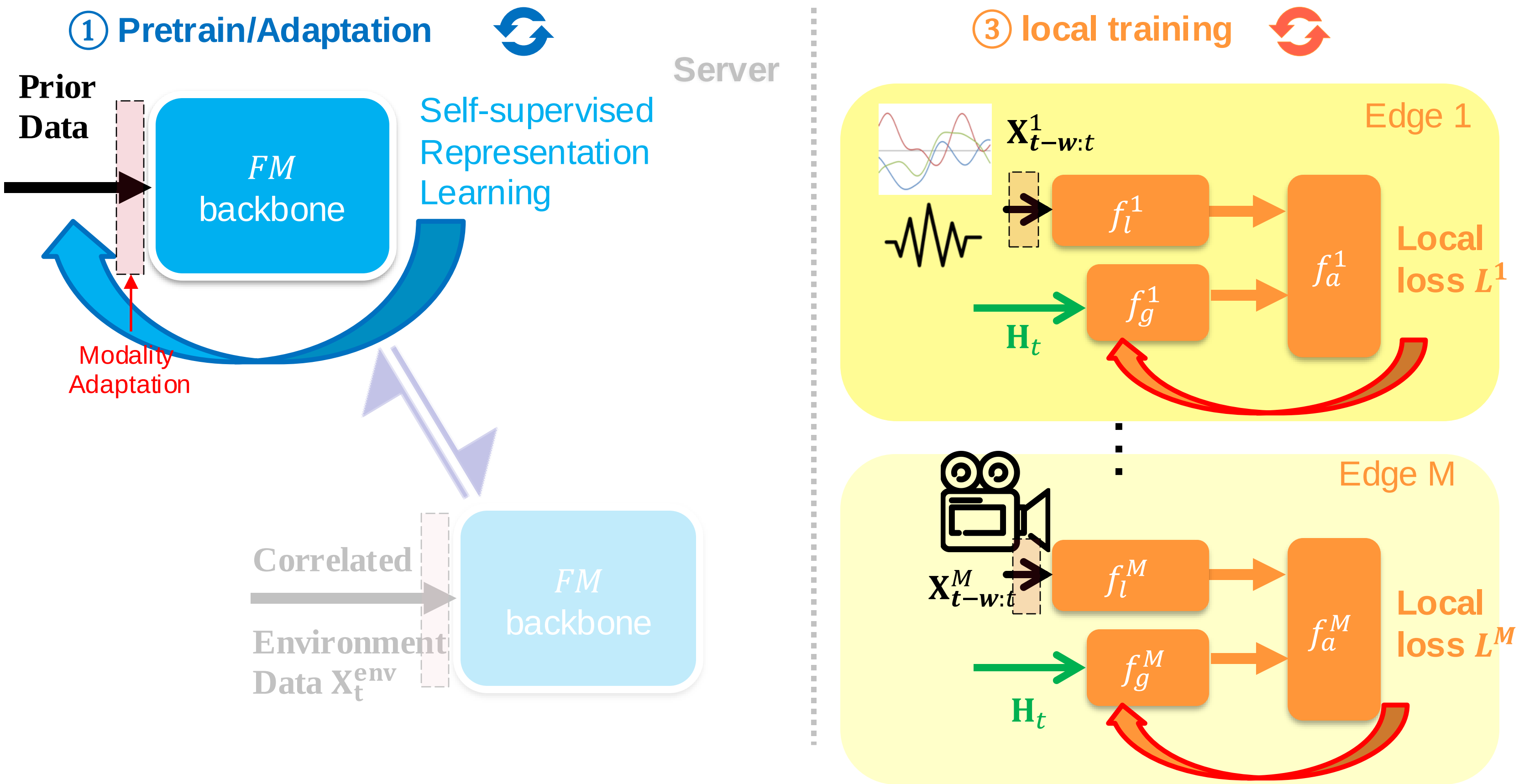
[5] Lyu, C., Wu, M., Wang, L., Huang, X., Liu, B., Du, Z., ... & Tu, Z. (2023). Macaw-llm: Multi-modal language modeling with image, audio, video, and text integration. arXiv preprint arXiv:2306.09093.

[6] Yu, Q., Liu, Y., Wang, Y., Xu, K., & Liu, J. (2023). Multimodal federated learning via contrastive representation ensemble. arXiv preprint arXiv:2302.08888.

Thanks!

- Contact: yihu@andrew.cmu.edu

CoRAST: Continual Learning



Future Work: Architecture Design