PEKING UNIVERSITY

FAMOS: Robust Privacy-Preserving Authentication on Payment Apps via Federated Multi-Modal Contrastive Learning

Yifeng Cai[1,2], Ziqi Zhang[3], Jiaping Gui[4], Bingyan Liu[5],
Xiaoke Zhao[6], Ruoyu Li[6], Zhe Li[6], and Ding Li[1,2]

[1] Key Lab of HCST (PKU), MOE [2] School of Computer Science, Peking University
[3] Department of Computer Science, University of Illinois Urbana-Champaign
[4] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University
[5] School of Computer Science, Beijing University of Posts and Telecommunications
[6] Ant Group

Background: Authentication in Payment Apps

PEKING UNIVERSITY

- As of 2024, one of the modern payment app in China, Alipay, has more than **300 million** daily active users and more than **$20 million** in daily payments.
- All payments need to be carefully protected.
- Traditional authentication methods, such as passwords and biometrics, are vulnerable once a device is compromised.

Background: New Solutions and Limitations

PEKING UNIVERSITY

- AuthentiSense [1] and KedyAuth [2] emerge as the SOTA solution for transparent and user-friendly authentication.
- They utilize build-in sensors to capture distinctive user behavioral patterns in an imperceptible manner, thus prevent unauthorized payments and do not harm the user experience.
- **However, they face two limitations in deployment to the payment apps.**

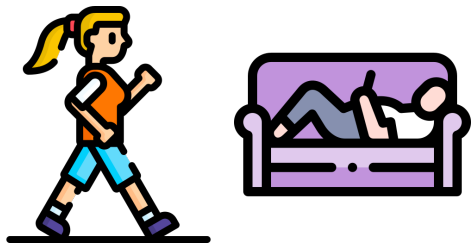
[1] Fereidooni, Hossein, et al. "AuthentiSense: A Scalable Behavioral Biometrics Authentication Scheme using Few-Shot Learning for Mobile Platforms." *NDSS 2023.*

[2] Huh, Jun Ho, et al. "On the long-term effects of continuous keystroke authentication: Keeping user frustration low through behavior adaptation." *ACM IMWUT 2023.*

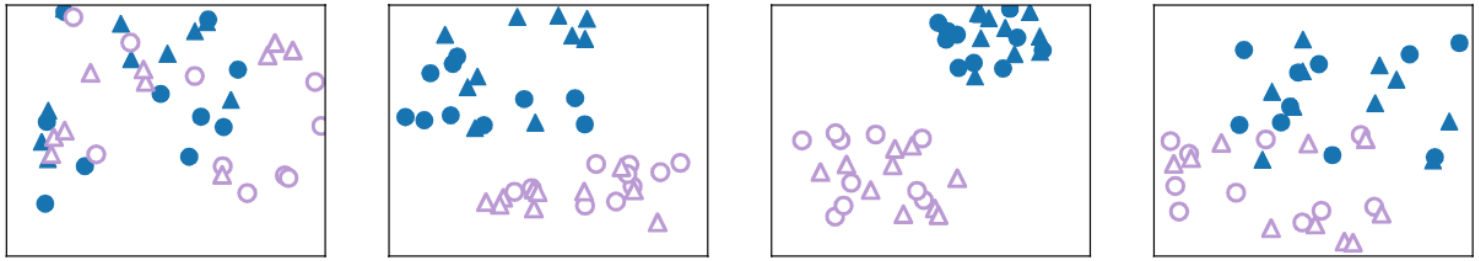
Background: New Solutions and Limitations

PEKING UNIVERSITY

- The first limitation: not considered the negative influence of background activities on sensor readings in real-world scenarios.
  - They assume that users should keep stationary and take similar actions while using mobile apps.
  - Users are not always stationary.



Different Background Activities



Different Stability between Sensors

Background: New Solutions and Limitations

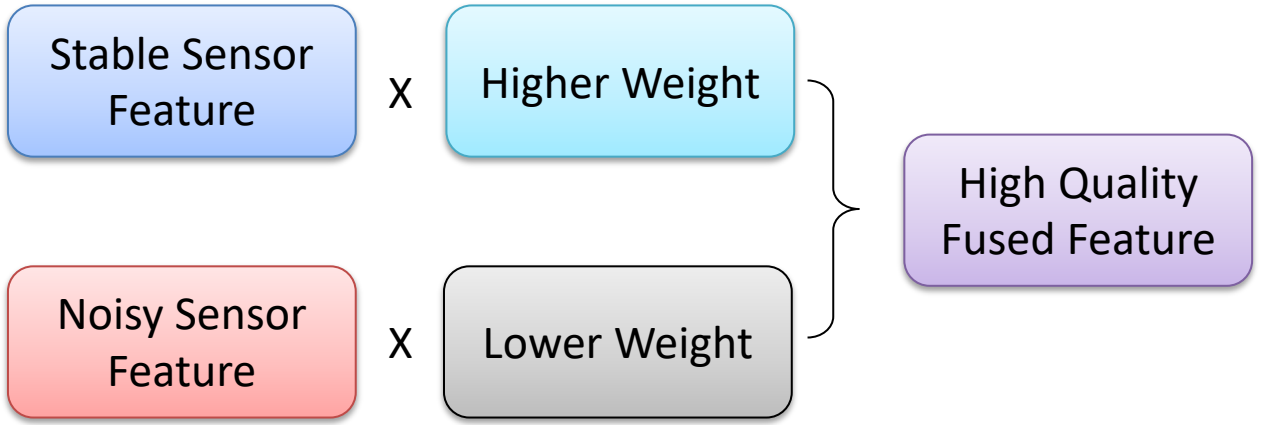
PEKING UNIVERSITY

- The second limitation: the violation of user privacy in the sensor data.
  - Their training strategy requires one user's data as positive sample and other users' data as negative sample.
  - Collecting and sharing such data is naturally against regulations (e.g., GDPR and PIPL).

Key Insight

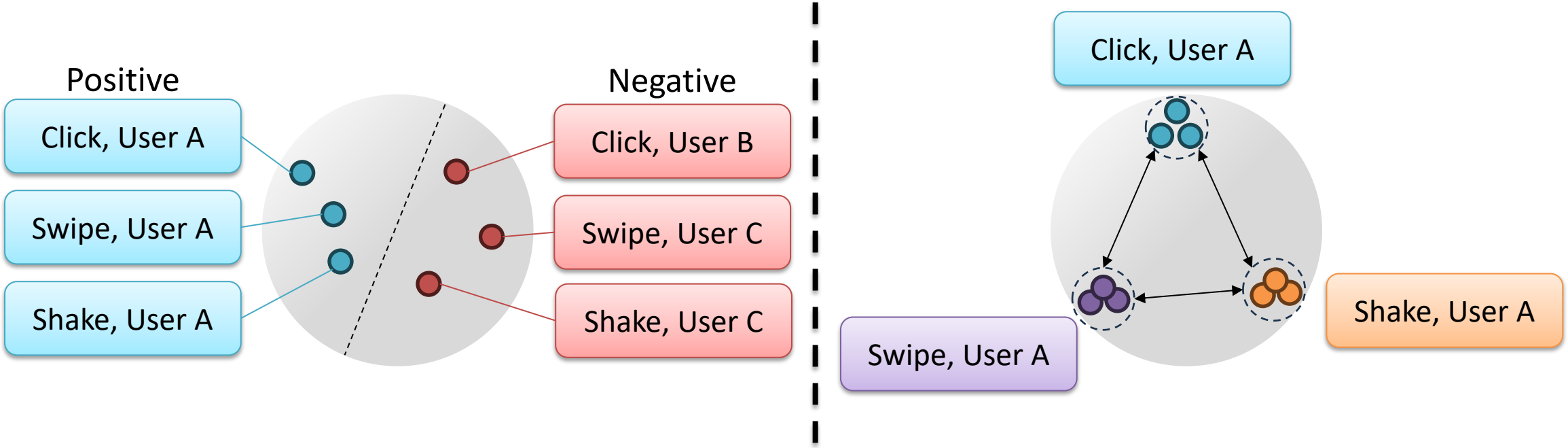
PEKING UNIVERSITY

- *Fusing multi-modal sensors data*, we can effectively eliminate background activities.



Key Insight

- *Clustering one user's data representation by action categories*, we can achieve user authentication without training data from other users.

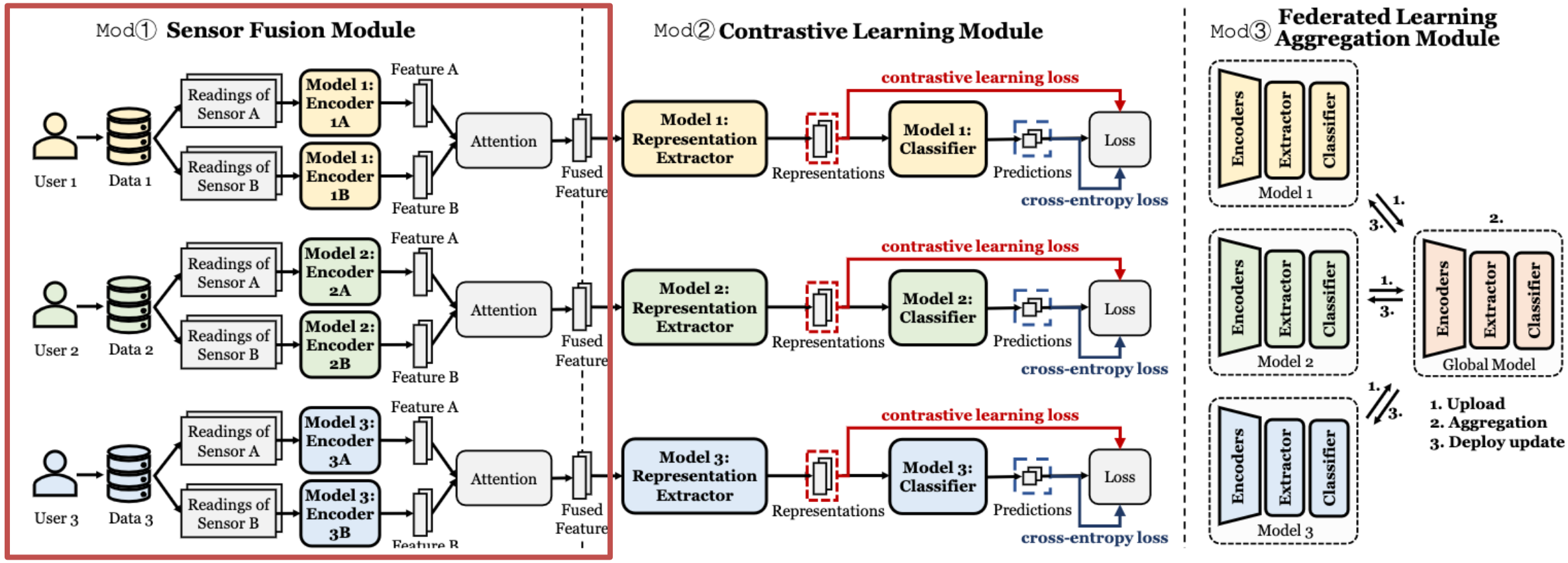


Design of FAMOS: Goals and Solutions

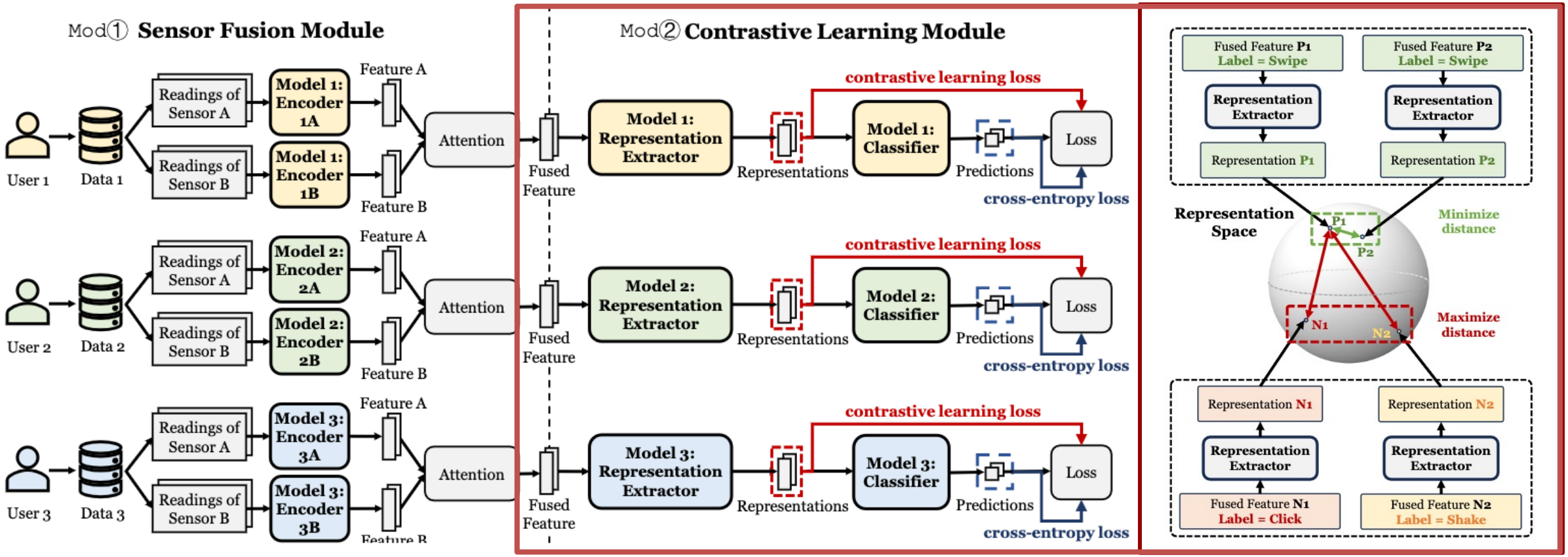
PEKING UNIVERSITY

- **Robustness.**
  - FAMOS should be robust to noises introduced by background activities.
- **Lightweight.**
  - FAMOS should be deployable within users' smartphones.
- **Privacy-Preserving.**
  - FAMOS should prioritize user privacy, minimizing the risk of data leakage.

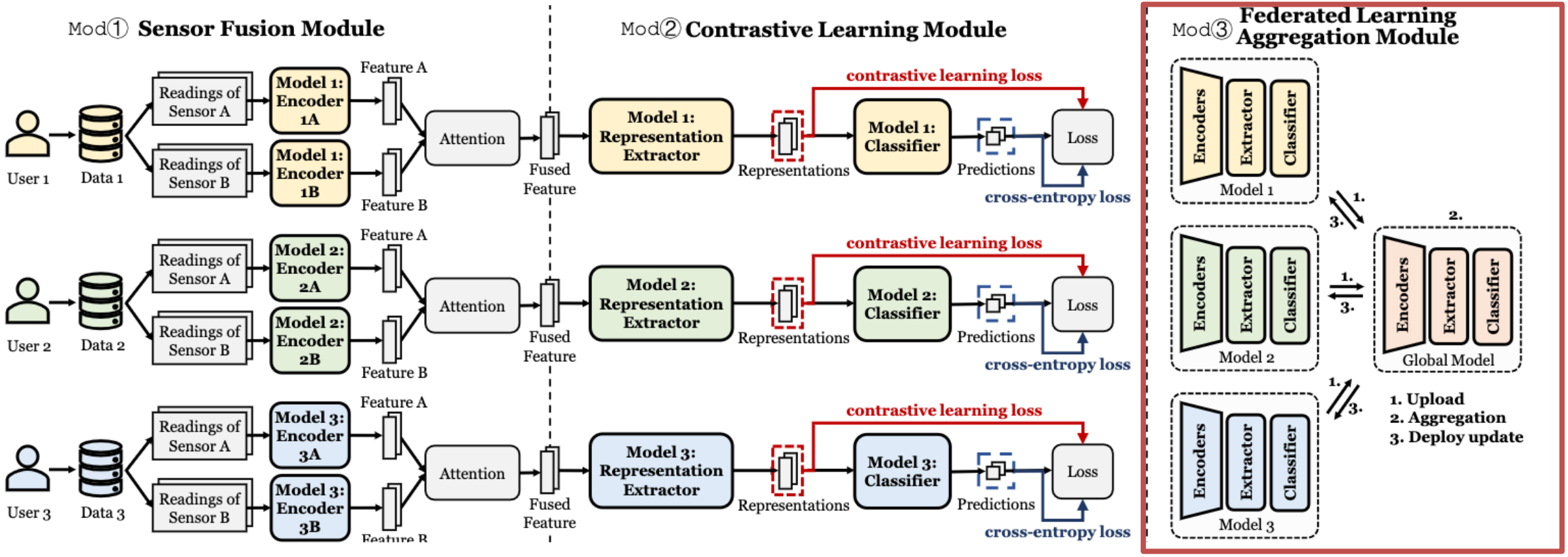
Design of FAMOS: Training Phase



Design of FAMOS: Training Phase



Design of FAMOS: Training Phase



Evaluation: Research Questions

PEKING UNIVERSITY

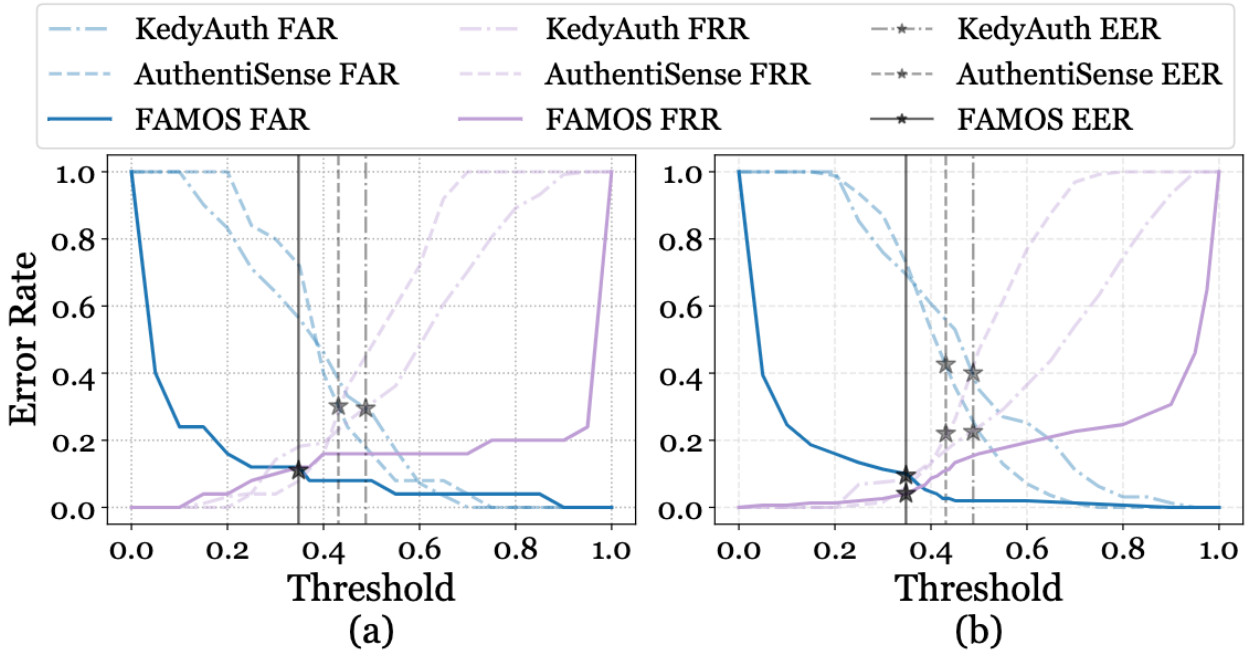
- **RQ1: Overall Effectiveness**
- **RQ2: Mitigating Background Activities**
- **RQ3: Effective of Contrastive Learning**
- **RQ4: Effective of Federated Learning**
- **RQ5: On-device Performance**

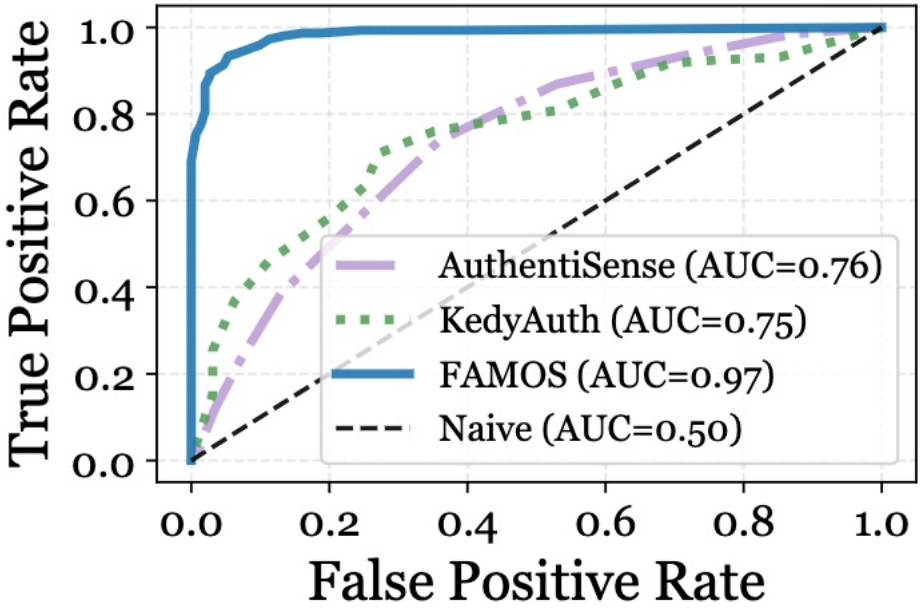
Evaluation: Protocol

PEKING UNIVERSITY

- **Dataset**
  - **Real-world Dataset:** 70 real-world Alipay users (20victim/50attacker)
  - **In-Lab Dataset:** 24 hired volunteers in Ant Group (4victim/20attacker)
  - **Format:**
    - Touch Screen
    - IMU (Accelerometer, Gyroscope, Magnetometer)
- **Device:** Huawei Mate X3, Xiaomi 13 Pro, VIVO X100, Honor Magic 6
- **Baselines:** AuthentiSense, KedyAuth
- **Metrics:** FAR, FRR, ERR, TPR, F1-Score, AUC, etc.

RQ1: Overall Effectiveness



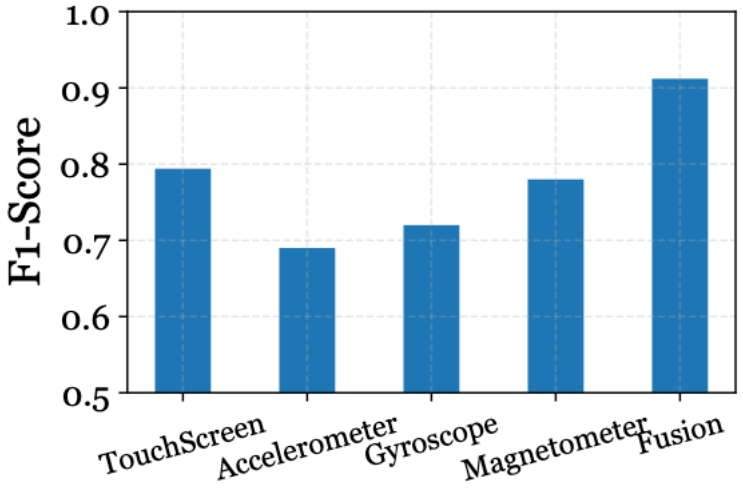


FAR: 4.4X lower / FRR: 5.5X lower / AUC: 27.7% higher / F1-Score: 42.2% higher

RQ2: Mitigating Background Activities

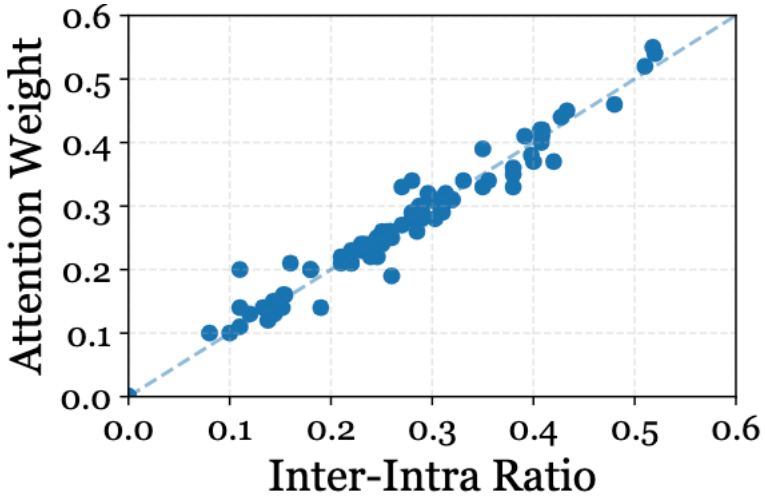
PEKING UNIVERSITY

Improvement on Accuracy by Sensor Fusion



13.2% - 31.9% improvement

Effectiveness of Attention



Different sensors have different stability under different background activities

RQ3 - RQ5

PEKING UNIVERSITY

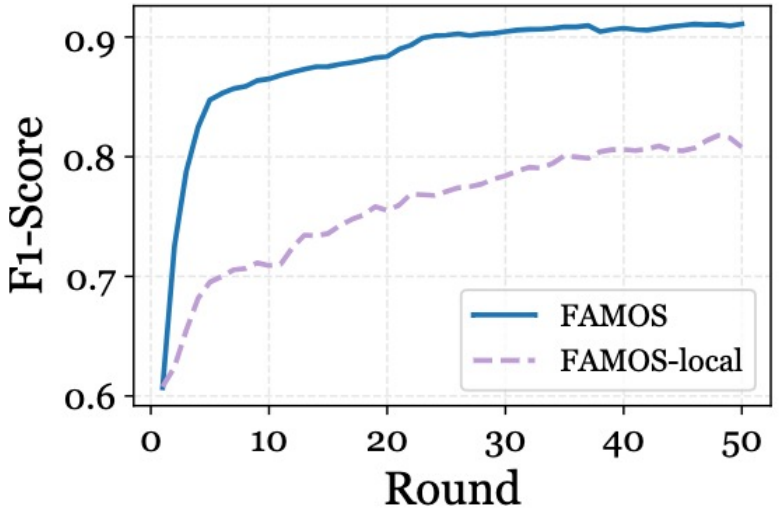
- **RQ3: Effective of Contrastive Learning**

Table 4: Compared of the averaged distance of fused features and representation vectors.

Action type	Features	Representations
Victim Samples (Different actions)	0.54	0.83
Victim Samples (Same action)	0.49	0.31
Victim & Attacker Samples (Same action)	0.52	0.50

1.4X increasement in difference

- **RQ4: Effective of Federated Learning**



FAMOS-local: 15.1% Lower

- **RQ5: On-device Performance**

Table 5: The overhead measurement of four devices.

Device	Memory (MB)		CPU (%)		Battery (%/h)		Time (ms)	
	Train	Infer	Train	Infer	Train	Infer	Train	Infer
Huawei Mate X3	8.42	3.94	23.83	16.21	1.34	0.52	1590K	132
Xiaomi 13 Pro	8.51	3.83	21.28	14.72	1.18	0.45	1542K	109
VIVO X100	8.74	4.11	25.11	17.61	1.22	0.51	1644K	143
Honor Magic 6	8.63	3.90	24.85	16.90	1.33	0.42	1698K	155
Average	8.58	3.95	23.77	16.36	1.32	0.48	1618K	135

Conclusion

PEKING UNIVERSITY

- We identify two practical limitations, the influence of background noises and privacy violations
- We propose a novel authentication framework, FAMOS, based on federated multi-modal contrastive learning.
- We comprehensively evaluate FAMOS using real-world datasets and devices.