Incremental Learning for Mobile Encrypted Traffic Classification

Yige Chen∗†, Tianning Zang∗†, Yongzheng Zhang∗†,
Yuan Zhou ‡, Linshu Ouyang∗†, Peng Yang‡

∗ Institute of Information Engineering, Chinese Academy of Sciences

†School of Cyber Security, University of Chinese Academy of Sciences

‡National Computer Network Emergency Response Technical Team/Coordination Center of China

中国科学院信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING,CAS

中国科学院大学
University of Chinese Academy of Sciences

IEEE ICC®

IEEE International Conference on Communications
14-23 June 2021 // Virtual / Montreal

Closed-world mobile encrypted classification

- Classify encrypted traffic into its belonging application

Closed-world mobile encrypted classification

- Classify encrypted traffic into its belonging application

Open-world mobile encrypted classification

- Breaks the closed-world assumption

Closed-world mobile encrypted classification

- Classify encrypted traffic into its belonging application

Open-world mobile encrypted classification

- Breaks the closed-world assumption
- Deal with the unseen applications

Closed-world Encrypted Traffic Classification

Mobile Encrypted Traffic

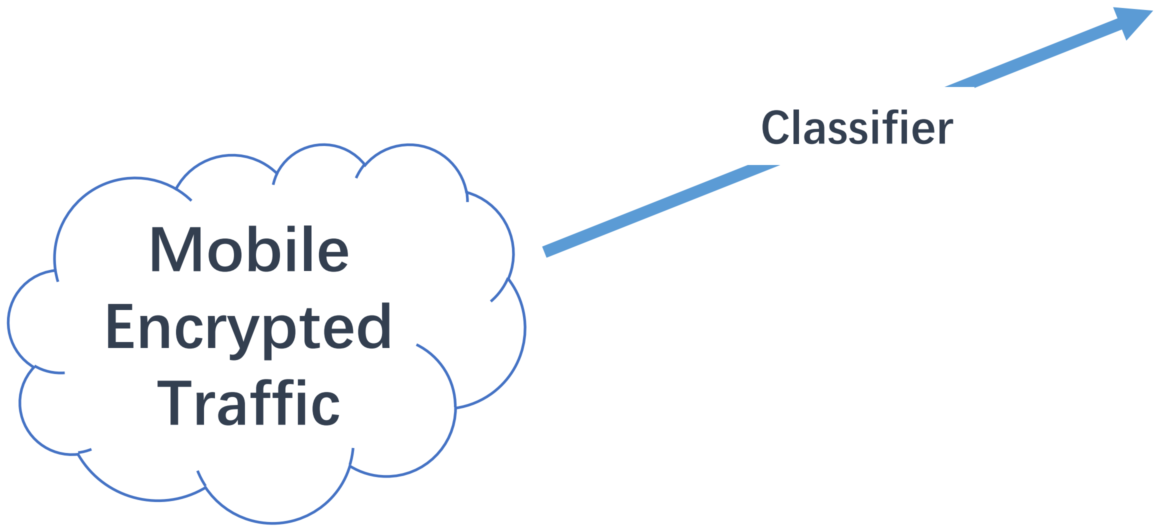
Classifier

Existing Applications

Existing state-of-the-art methods		
SoB [1]	91.22% TPR	0.08% FPR
MaMPF [2]	96.40% TPR	0.20% FPR
FS-Net [3]	99.14% TPR	0.05% FPR
MAAF [4]	98.69% Acc	98.64% F1

[1] M. Shen, M. Wei, L. Zhu, and M. Wang, "Classification of encrypted traffic with second-order markov chains and application attribute bigrams," IEEE Transactions on Information Forensics and Security, vol. 12, no. 8, pp. 1830–1843, 2017.
[2] C. Liu, Z. Cao, G. Xiong, G. Gou, S.-M. Yiu, and L. He, "Mampf: Encrypted traffic classification based on multi-attribute markov probability fingerprints," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–10.
[3] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "Fs-net: A flow sequence network for encrypted traffic classification," in 2019 IEEE International Conference on Computer Communications (Infocom). IEEE, 2019, pp. 1–9.
[4] Y. Chen, T. Zang, Y. Zhang, Y. Zhou, and Y. Wang, "Rethinking encrypted traffic classification: A multi-attribute associated fingerprint approach," in 2019 IEEE 27th International Conference on Network Protocols (ICNP). IEEE, 2019, pp. 1–11.

Closed-world Encrypted Traffic Classification

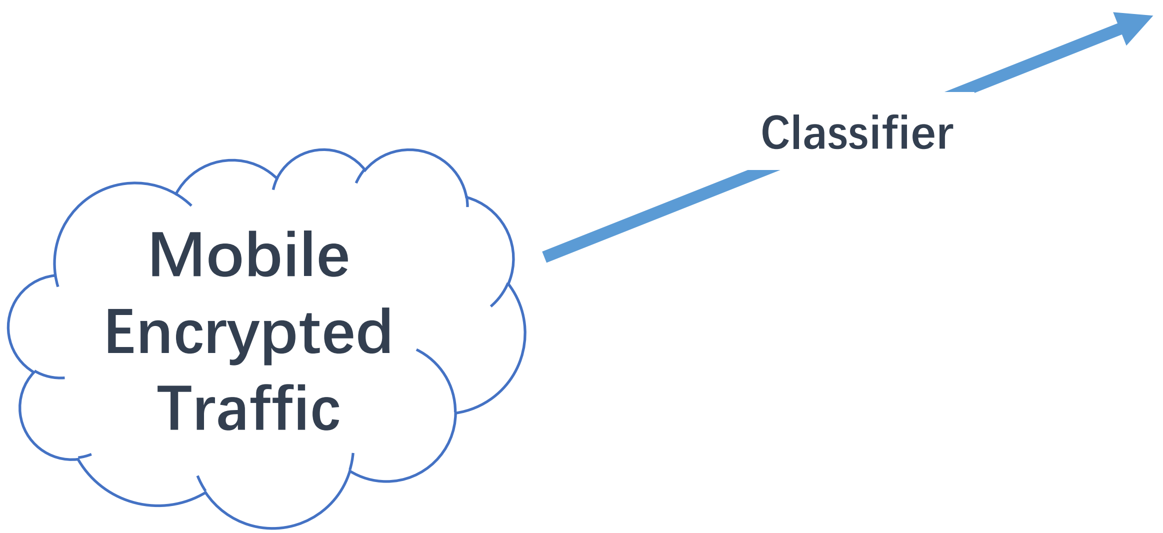


Existing Applications

Existing state-of-the-art methods		
SoB [1]	91.22% TPR	0.08% FPR
MaMPF [2]	96.40% TPR	0.20% FPR
FS-Net [3]	99.14% TPR	0.05% FPR
MAAF [4]	98.69% Acc	98.64% F1

[1] M. Shen, M. Wei, L. Zhu, and M. Wang, "Classification of encrypted traffic with second-order markov chains and application attribute bigrams," IEEE Transactions on Information Forensics and Security, vol. 12, no. 8, pp. 1830–1843, 2017.
[2] C. Liu, Z. Cao, G. Xiong, G. Gou, S.-M. Yiu, and L. He, "Mampf: Encrypted traffic classification based on multi-attribute markov probability fingerprints," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–10.
[3] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "Fs-net: A flow sequence network for encrypted traffic classification," in 2019 IEEE International Conference on Computer Communications (Infocom). IEEE, 2019, pp. 1–9.
[4] Y. Chen, T. Zang, Y. Zhang, Y. Zhou, and Y. Wang, "Rethinking encrypted traffic classification: A multi-attribute associated fingerprint approach," in 2019 IEEE 27th International Conference on Network Protocols (ICNP). IEEE, 2019, pp. 1–11.

Closed-world Encrypted Traffic Classification

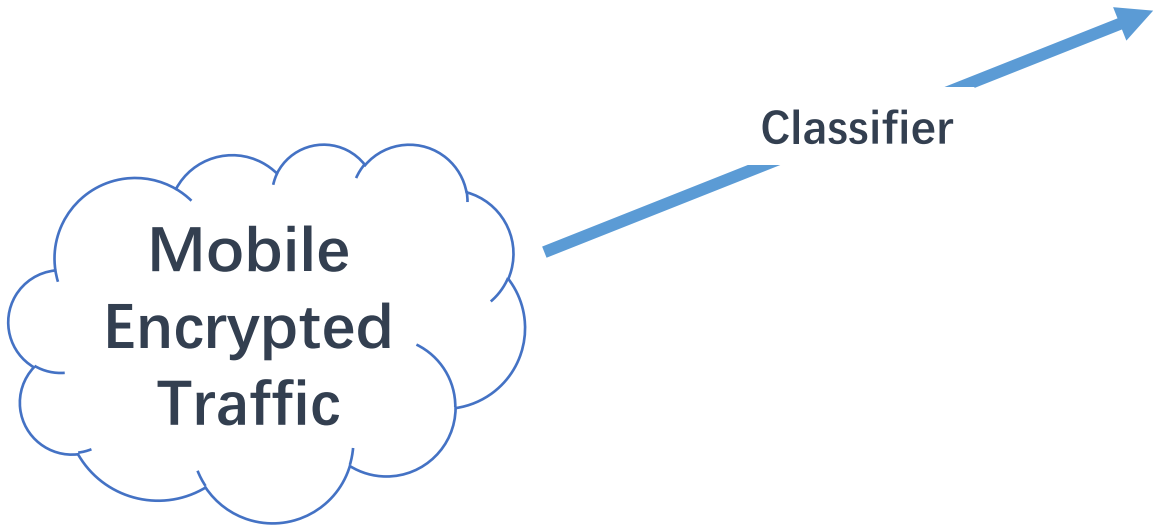


Existing Applications

Existing state-of-the-art methods		
SoB [1]	91.22% TPR	0.08% FPR
MaMPF [2]	96.40% TPR	0.20% FPR
FS-Net [3]	99.14% TPR	0.05% FPR
MAAF [4]	98.69% Acc	98.64% F1

[1] M. Shen, M. Wei, L. Zhu, and M. Wang, "Classification of encrypted traffic with second-order markov chains and application attribute bigrams," IEEE Transactions on Information Forensics and Security, vol. 12, no. 8, pp. 1830–1843, 2017.
[2] C. Liu, Z. Cao, G. Xiong, G. Gou, S.-M. Yiu, and L. He, "Mampf: Encrypted traffic classification based on multi-attribute markov probability fingerprints," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–10.
[3] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "Fs-net: A flow sequence network for encrypted traffic classification," in 2019 IEEE International Conference on Computer Communications (Infocom). IEEE, 2019, pp. 1–9.
[4] Y. Chen, T. Zang, Y. Zhang, Y. Zhou, and Y. Wang, "Rethinking encrypted traffic classification: A multi-attribute associated fingerprint approach," in 2019 IEEE 27th International Conference on Network Protocols (ICNP). IEEE, 2019, pp. 1–11.

Closed-world Encrypted Traffic Classification

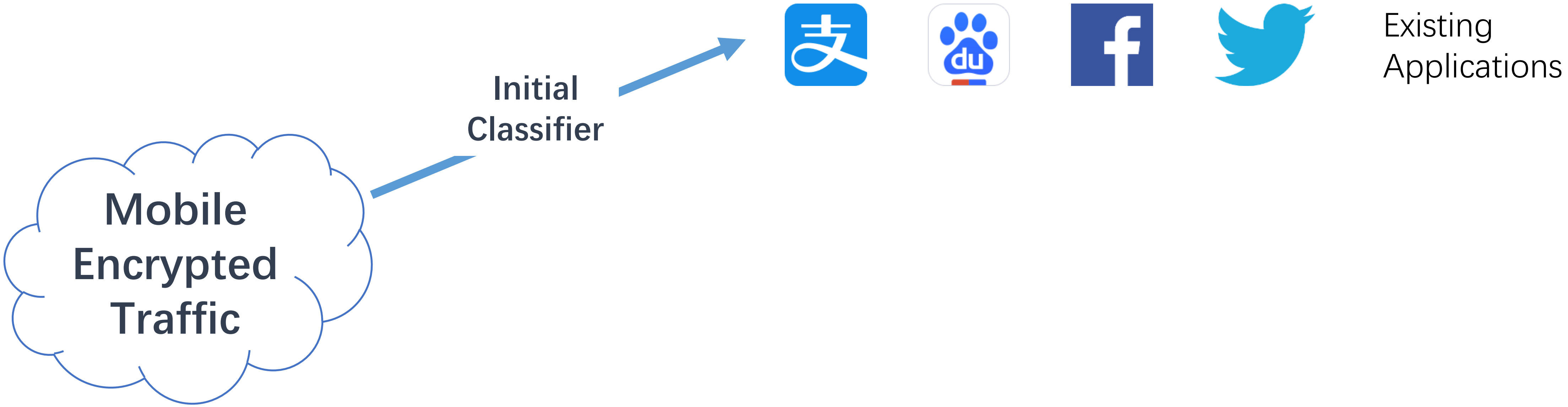


Existing Applications

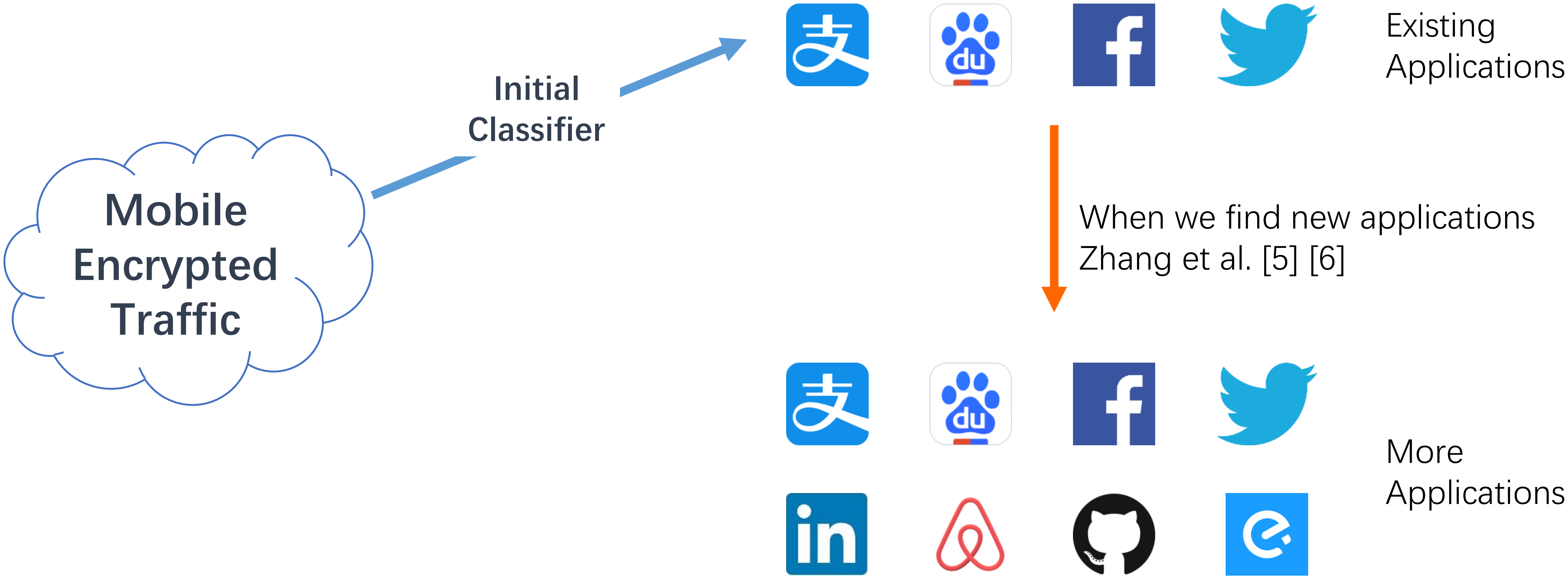
Existing state-of-the-art methods		
SoB [1]	91.22% TPR	0.08% FPR
MaMPF [2]	96.40% TPR	0.20% FPR
FS-Net [3]	99.14% TPR	0.05% FPR
MAAF [4]	98.69% Acc	98.64% F1

[1] M. Shen, M. Wei, L. Zhu, and M. Wang, "Classification of encrypted traffic with second-order markov chains and application attribute bigrams," IEEE Transactions on Information Forensics and Security, vol. 12, no. 8, pp. 1830–1843, 2017.
[2] C. Liu, Z. Cao, G. Xiong, G. Gou, S.-M. Yiu, and L. He, "Mampf: Encrypted traffic classification based on multi-attribute markov probability fingerprints," in 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). IEEE, 2018, pp. 1–10.
[3] C. Liu, L. He, G. Xiong, Z. Cao, and Z. Li, "Fs-net: A flow sequence network for encrypted traffic classification," in 2019 IEEE International Conference on Computer Communications (Infocom). IEEE, 2019, pp. 1–9.
[4] Y. Chen, T. Zang, Y. Zhang, Y. Zhou, and Y. Wang, "Rethinking encrypted traffic classification: A multi-attribute associated fingerprint approach," in 2019 IEEE 27th International Conference on Network Protocols (ICNP). IEEE, 2019, pp. 1–11.

Open-world Encrypted Traffic Classification

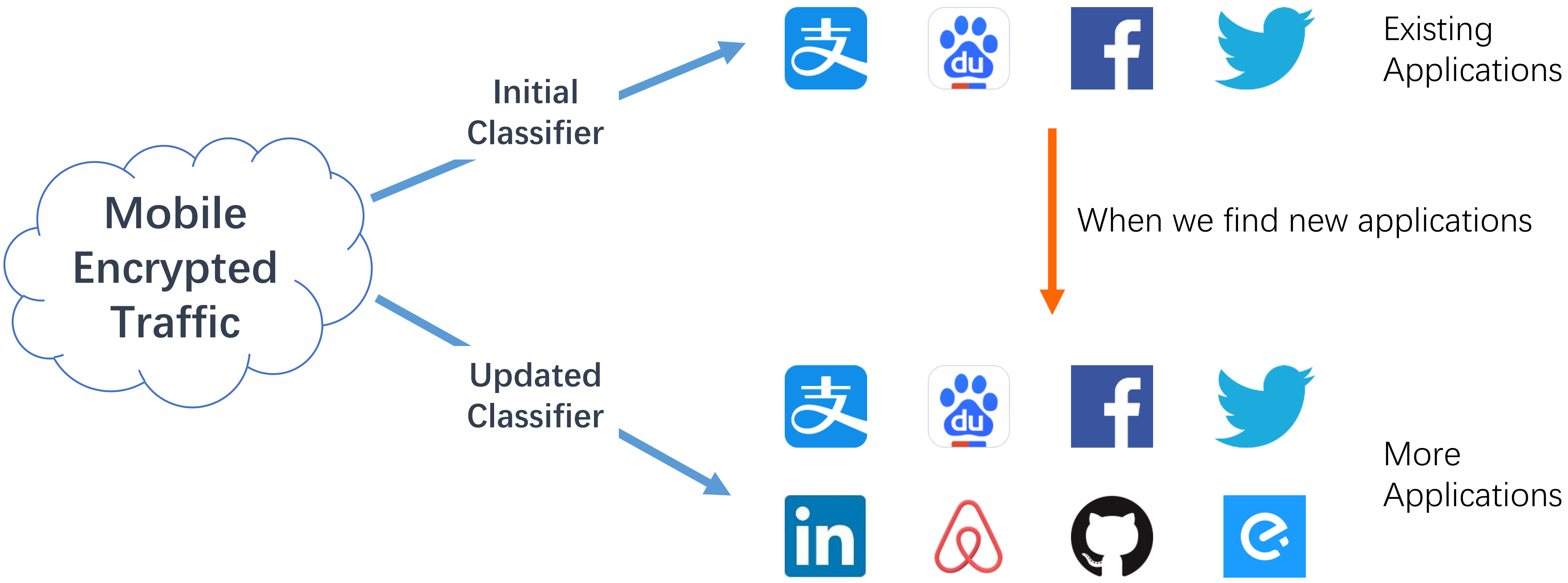


Open-world Encrypted Traffic Classification

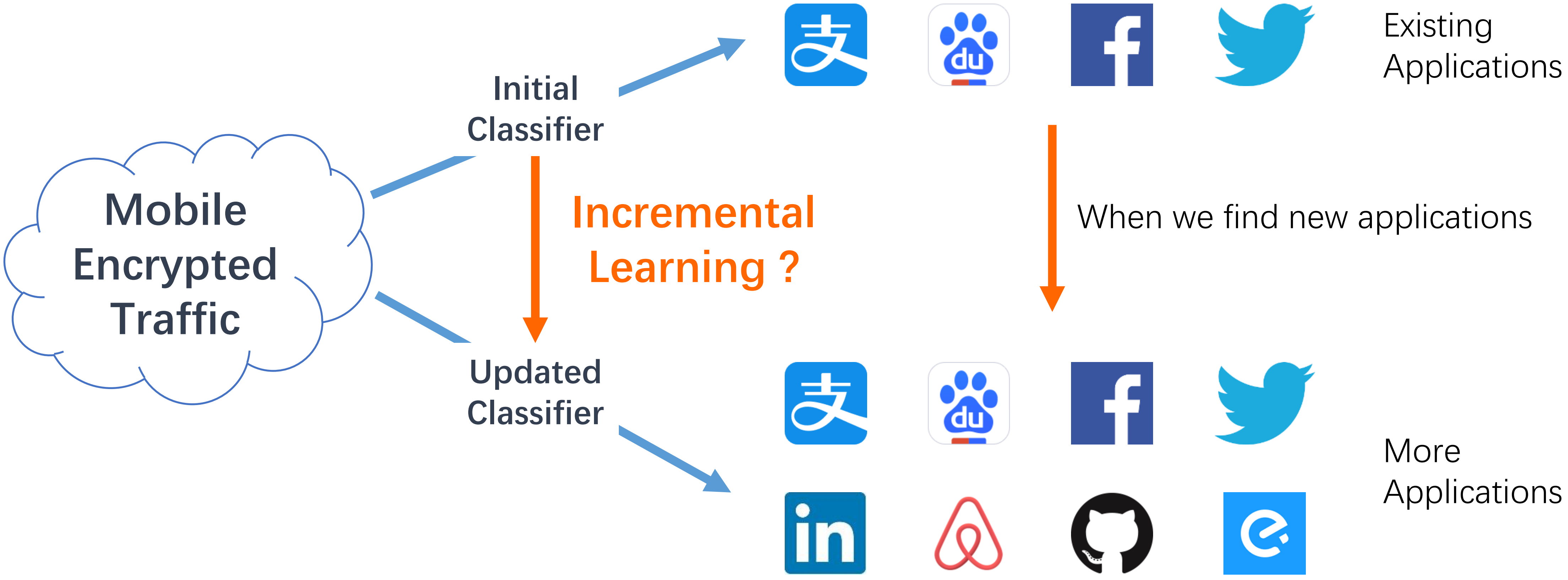


[5] J. Zhang, F. Li, H. Wu, and F. Ye, "Autonomous model update scheme for deep learning based network traffic classifiers," in 2019 IEEE Global Communications Conference (GLOBECOM). IEEE, 2019, pp. 1–6.
[6] J. Zhang, F. Li, F. Ye, and H. Wu, "Autonomous unknown-application filtering and labeling for dl-based traffic classifier update," in 2020 IEEE International Conference on Computer Communications (Infocom). IEEE, 2020, pp. 1–9.

Open-world Encrypted Traffic Classification



Open-world Encrypted Traffic Classification



Naive Incremental Learning Methods

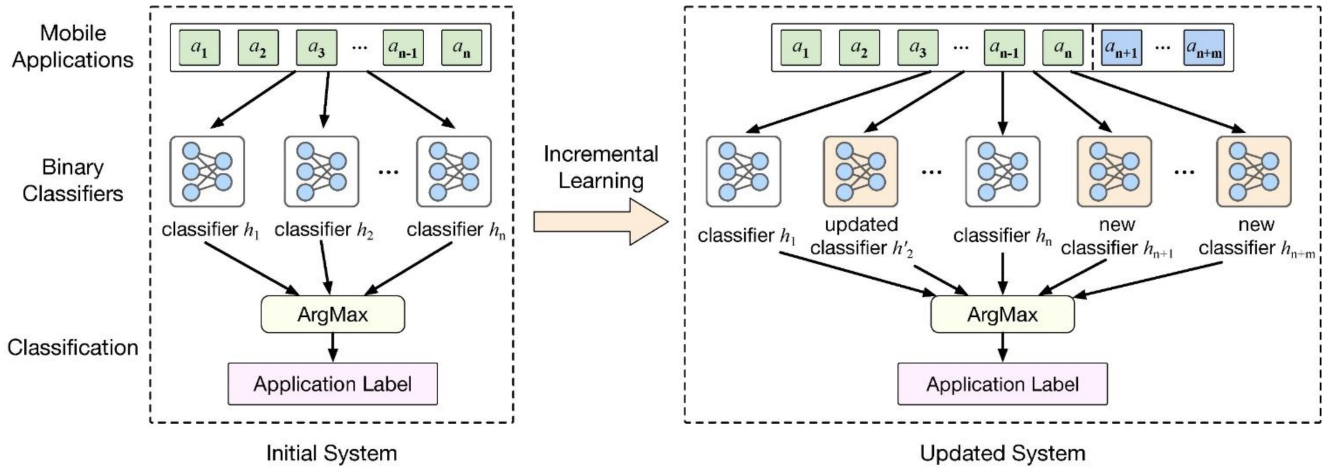
1. Retraining the updated classifier from scratch
   - Considerable training time and effort
   - Expansion of the dataset

Naive Incremental Learning Methods

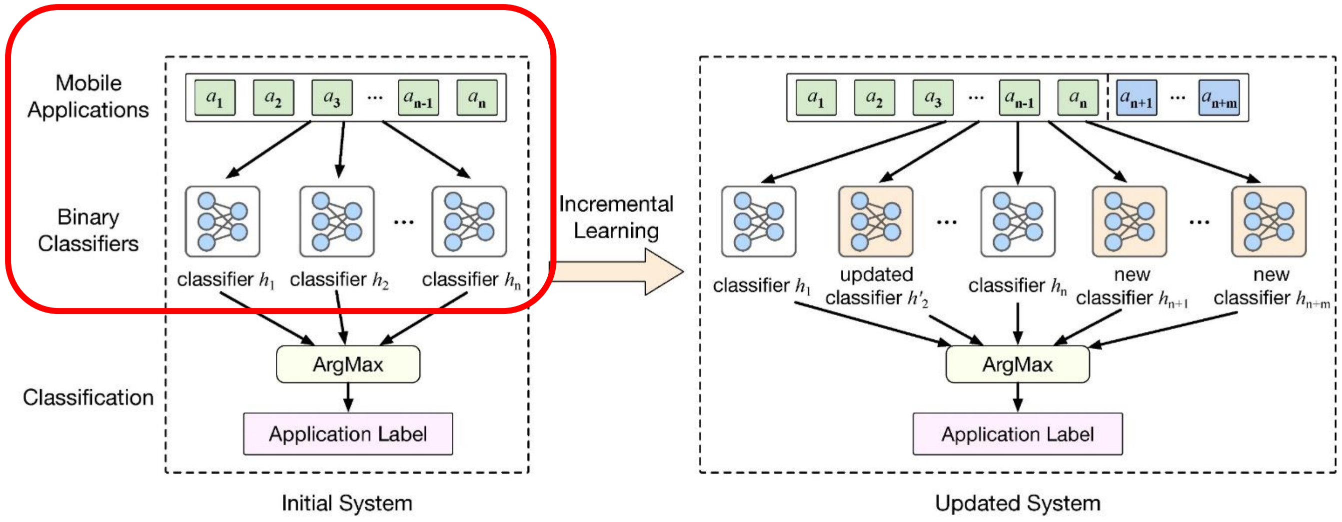
1. Retraining the updated classifier from scratch
    - Considerable training time and effort
    - Expansion of the dataset

2. Fine-tuning the existing classifier
    - catastrophic forgetting problem

Incremental Learning based on (OvR) Strategy



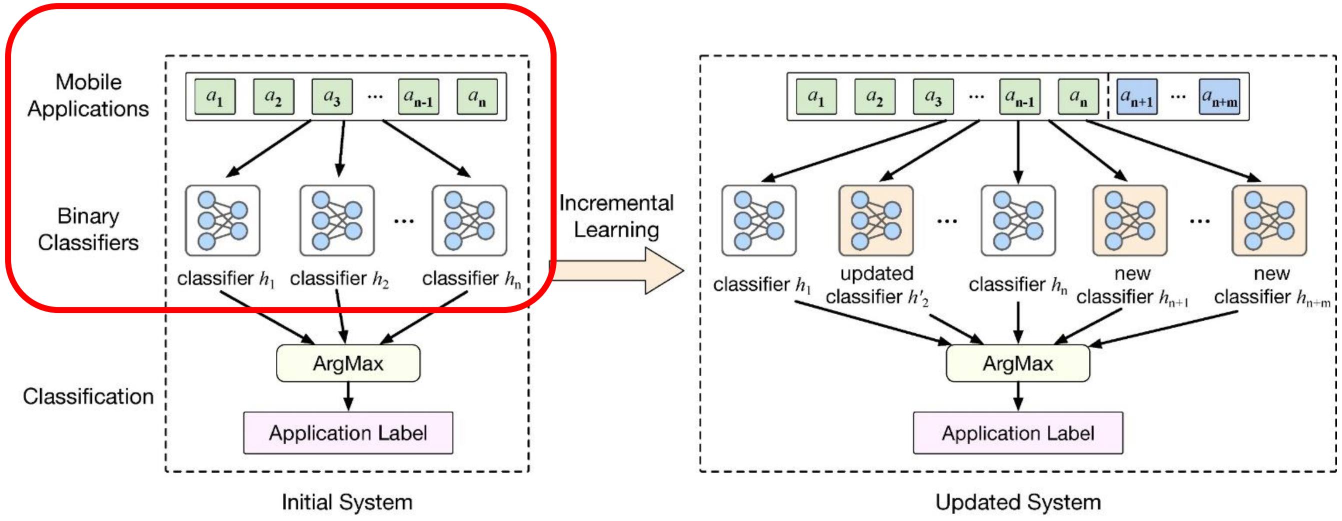
Incremental Learning based on (OvR) Strategy



One vs Rest Strategy

- n binary classifiers. The classifier h_i correspond to i^{th} mobile application.

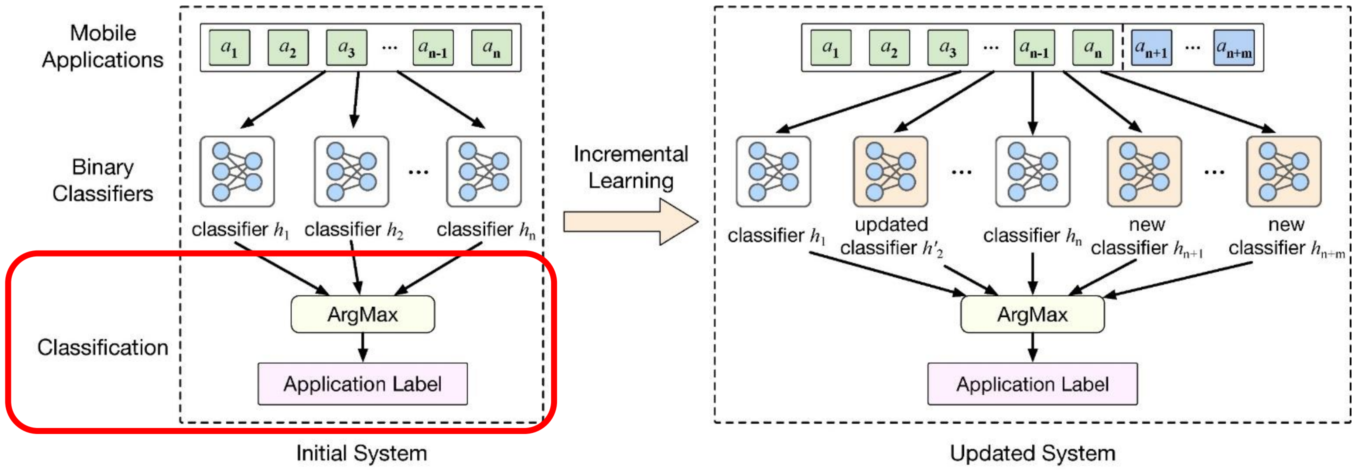
Incremental Learning based on (OvR) Strategy



One vs Rest Strategy

- n binary classifiers. The classifier h_i correspond to i^{th} mobile application.
- The binary classifier h_i considers i^{th} application as positive while other applications as negative.

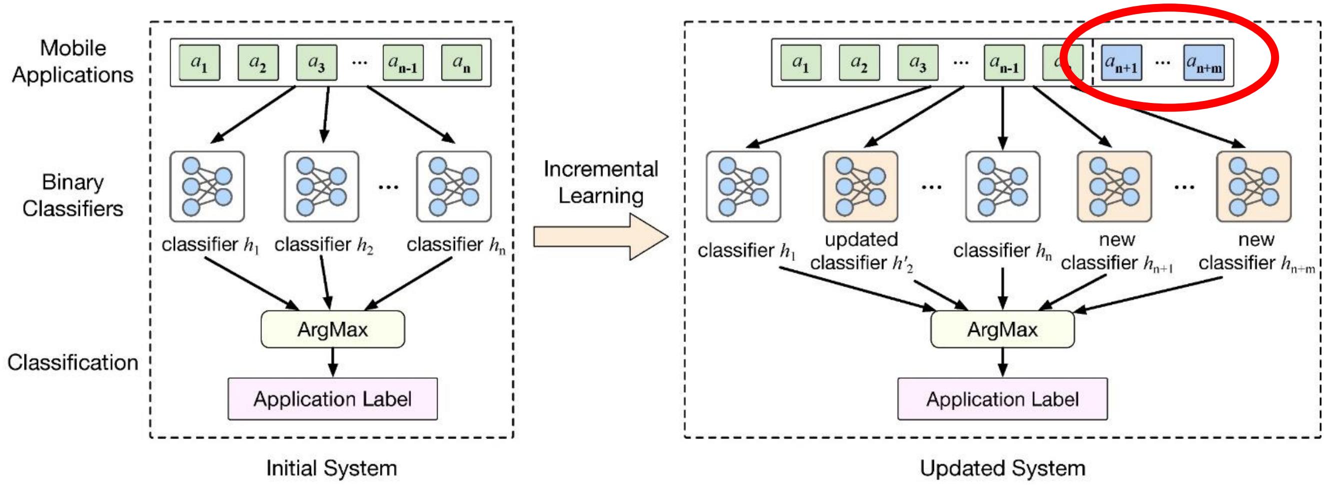
Incremental Learning based on (OvR) Strategy



One vs Rest Strategy

- n binary classifiers. The classifier h_i correspond to i^{th} mobile application.
- The binary classifier h_i considers i^{th} application as positive while other applications as negative.
- The system integrates all binary classifiers to make classification.

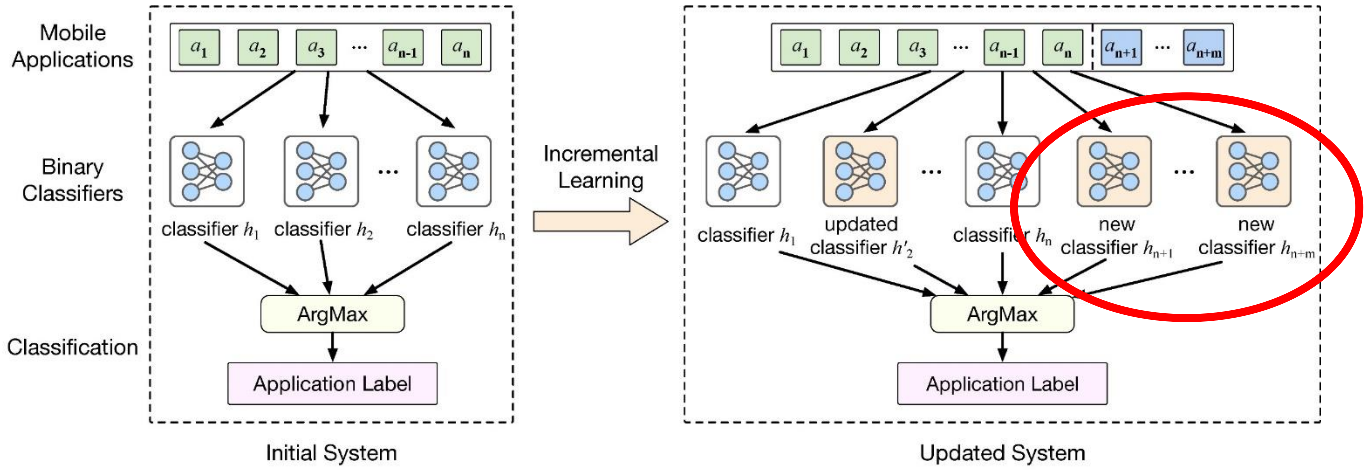
Incremental Learning based on (OvR) Strategy



Incremental Learning

- Collect the dataset of new applications.

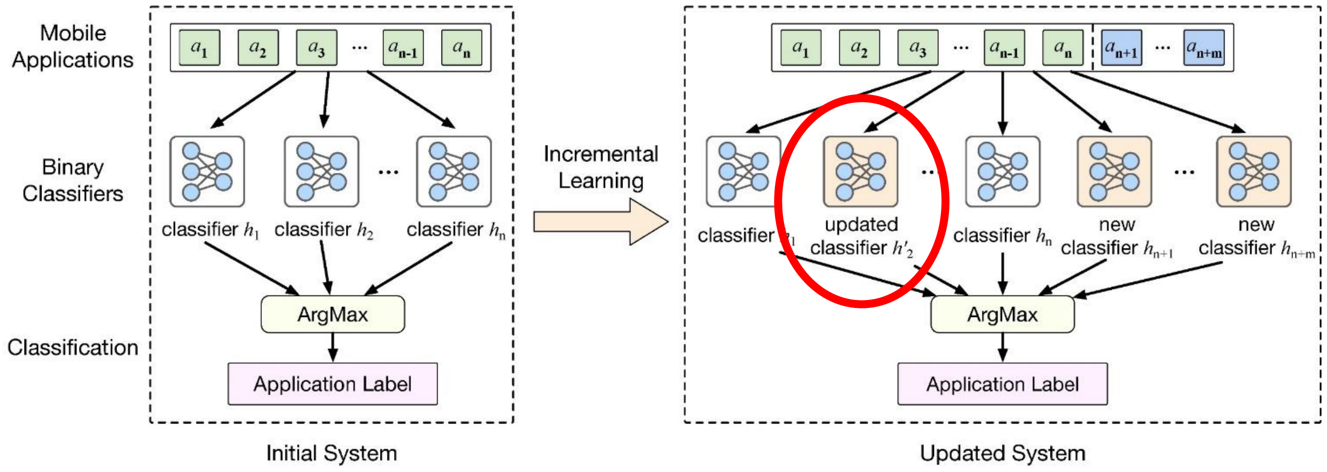
Incremental Learning based on (OvR) Strategy



Incremental Learning

- Collect the dataset of new applications.
- Build extra new binary classifiers for the new applications.

Incremental Learning based on (OvR) Strategy

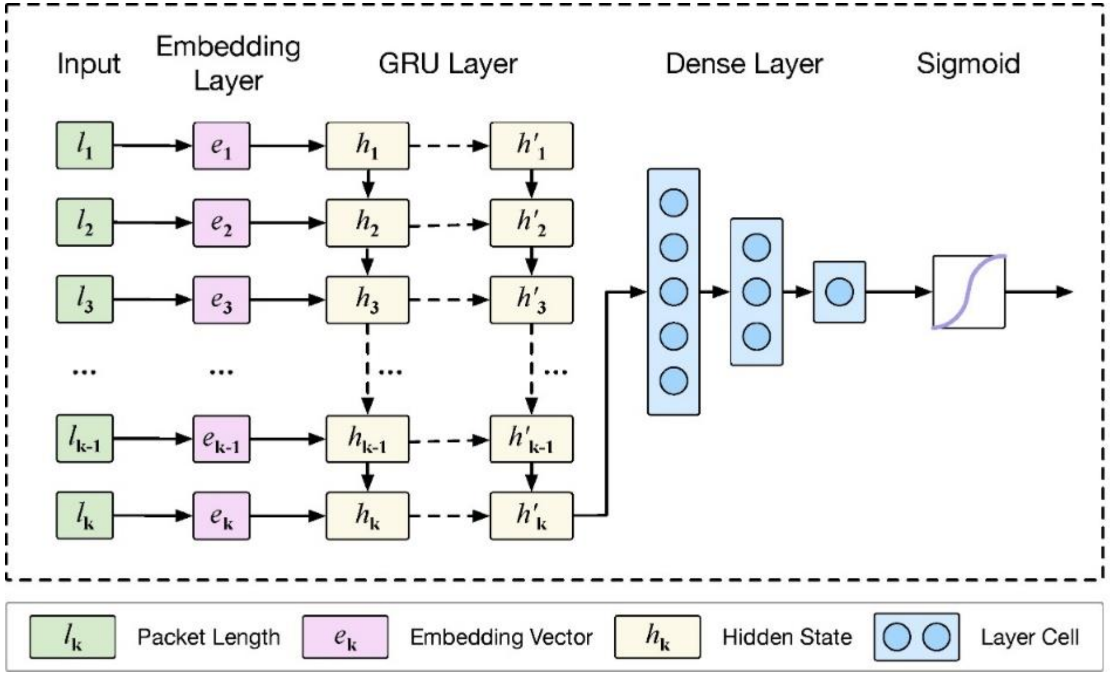


Incremental Learning

- Collect the dataset of new applications.
- Build extra new binary classifiers for the new applications.
- Retrain the outdated classifier that accept more than the retraining threshold τ of the new applications' traffic

Binary Classifier

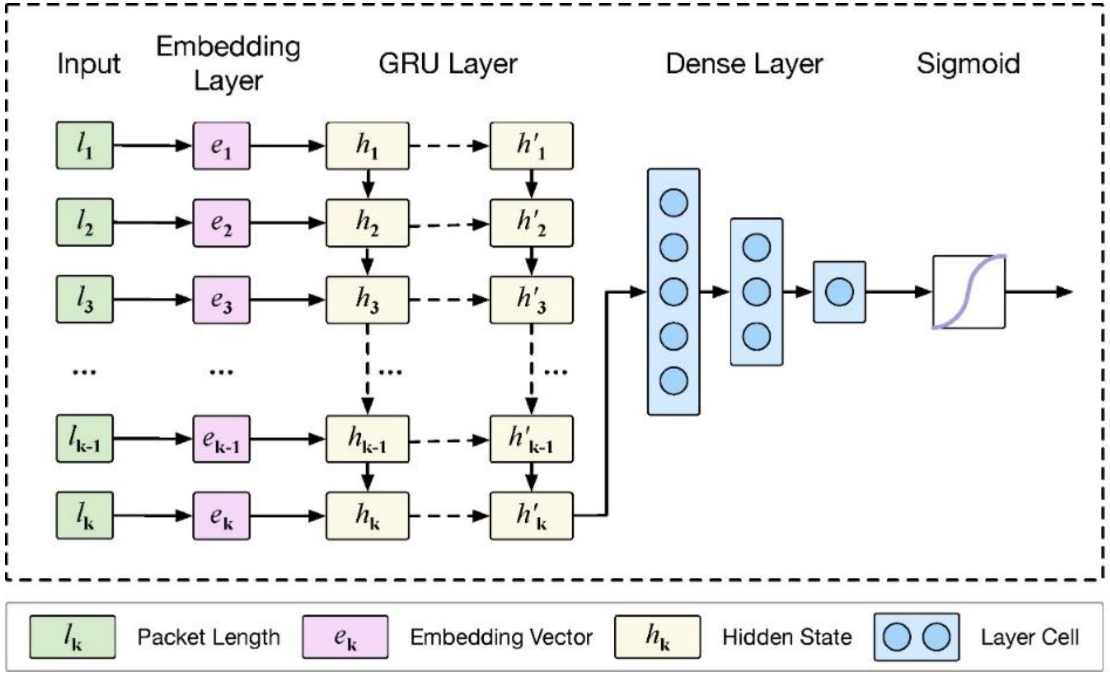
1. A neural network-based Implementation.



Neural Binary Classification Network

Binary Classifier

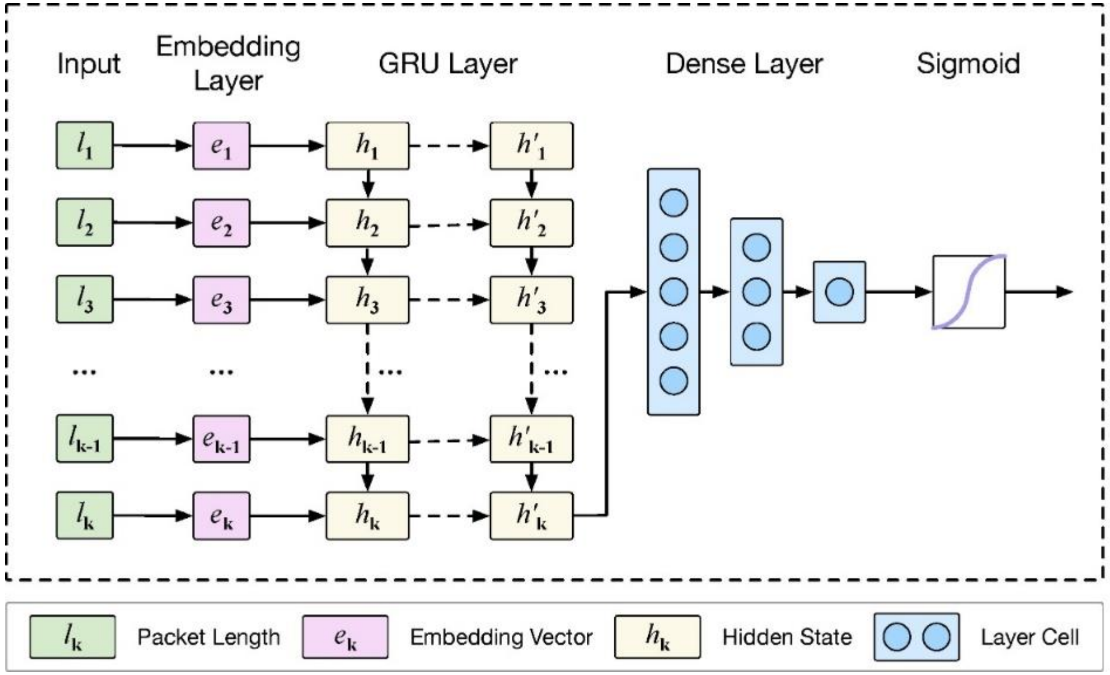
1. A neural network-based Implementation.
2. Take the first k packet lengths of flows as classifier input.



Neural Binary Classification Network

Binary Classifier

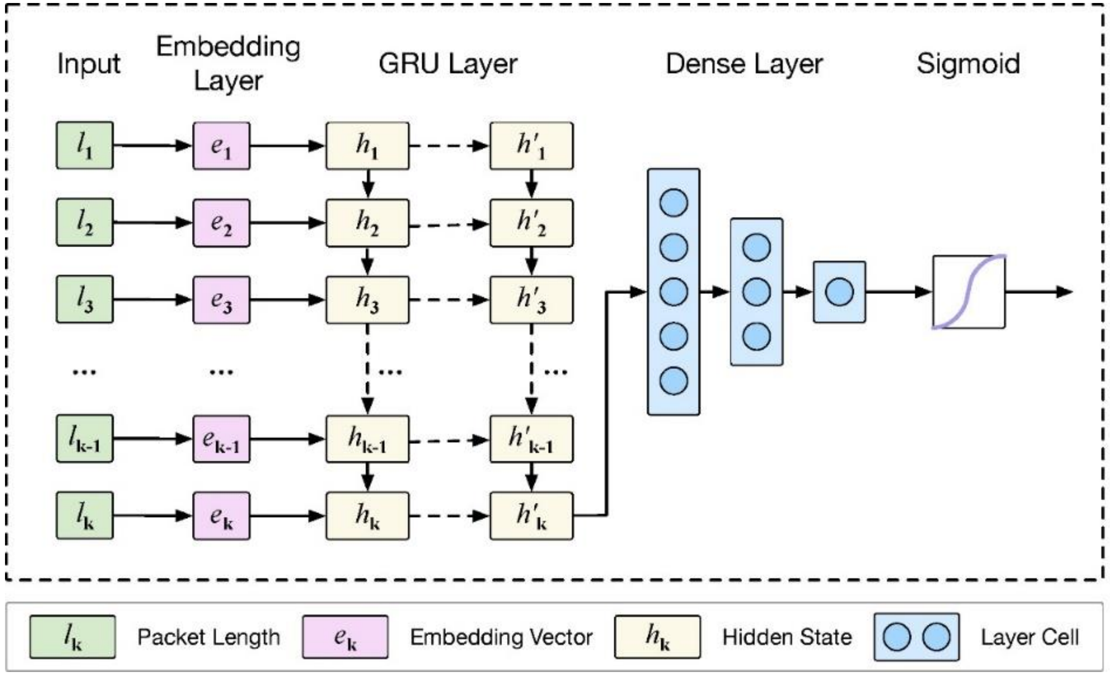
1. A neural network-based Implementation.
2. Take the first k packet lengths of flows as classifier input.
3. Use Gated Recurrent Unit (GRU) to process the sequence inputs.



Neural Binary Classification Network

Binary Classifier

1. A neural network-based Implementation.
2. Take the first k packet lengths of flows as classifier input.
3. Use Gated Recurrent Unit (GRU) to process the sequence inputs.
4. Adopt a sigmoid function to normalize classification probability ∈ (0-1).



Neural Binary Classification Network

Sample Selection

1. Maintain the size of dataset when adding new applications.

Sample Selection

1. Maintain the size of dataset when adding new applications.
2. Select new samples and remove leftover samples through herding selection.

Sample Selection

1. Maintain the size of dataset when adding new applications.
2. Select new samples and remove leftover samples through herding selection.
3. Sequentially selects samples that keeps its vector average nearest to the original vector average.

Evaluation Dataset

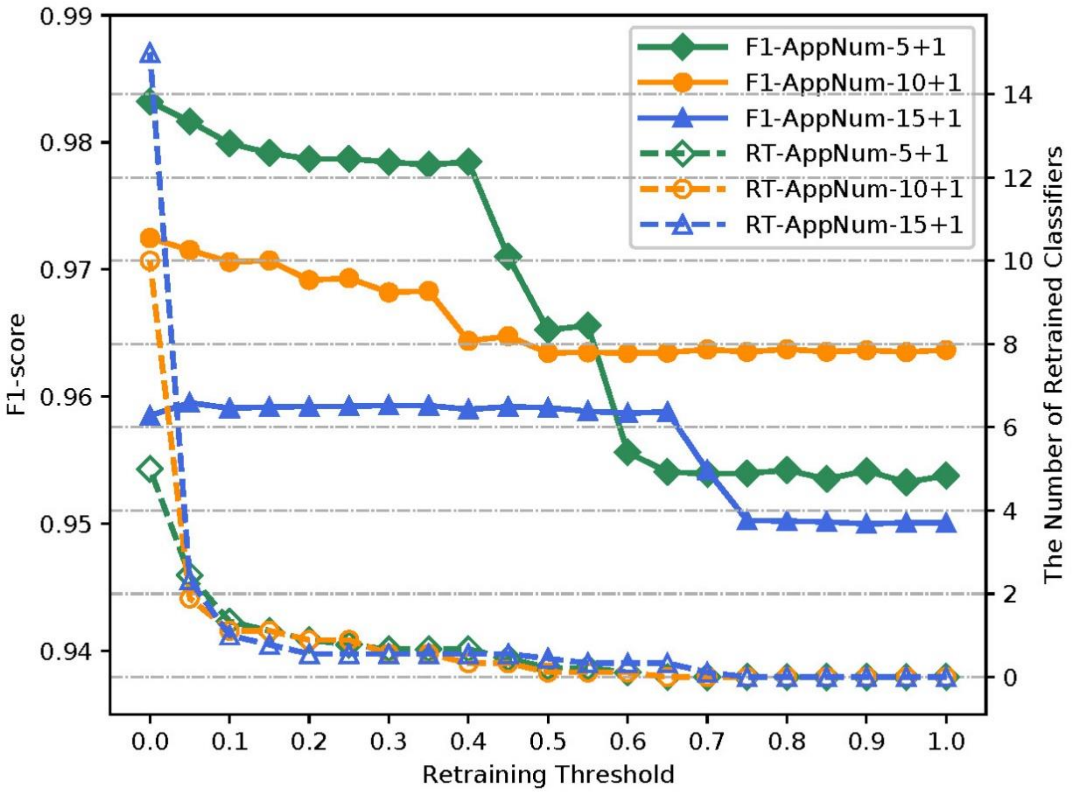
- A manually collected dataset provided by MAAF [1].
- 77,278 real-world encrypted flows of 16 popular mobile applications.

TABLE II
THE STATISTIC OF 16 APPLICATION TRACESETS

Developer	Application	Manually Collected Traceset				
		Flows	Packets	Domain	Cert	Both[1]
Alibaba	Alipay	5201	315234	16.4%	96.3%	97.3%
	Taobao[2]	3231	291348	93.9%	96.8%	99.4%
	AMap[2]	3624	114513	91.7%	98.8%	99.4%
Baidu	Baidu Search	4732	181971	52.5%	90.3%	94.3%
	Baidu Map[2]	5544	215920	40.0%	89.2%	93.8%
Facebook	Facebook	4148	526289	46.3%	82.2%	87.4%
	Instagram	4379	343809	27.0%	5.8%	31.8%
Twitter	Twitter	4463	167166	45.6%	89.7%	93.9%
Sina	Weibo	3817	127057	95.4%	95.2%	99.6%
Airbnb	Airbnb	5843	875837	76.0%	67.7%	82.2%
Linkedin	Linkedin	4203	160614	91.4%	91.8%	98.5%
Evernote	Evernote	7504	202557	98.4%	48.1%	98.5%
Blued	Blued	4833	478467	73.4%	55.6%	73.8%
Ele	Ele	6740	99193	98.9%	98.5%	99.9%
Github	Github	4431	151355	98.6%	96.4%	98.8%
Yirendai	Yirendai	4585	61356	98.1%	97.5%	99.2%
Total		77278	4312686	71.7%	79.9%	90.7%

[1] Y. Chen, T. Zang, Y. Zhang, Y. Zhou, and Y. Wang, "Rethinking encrypted traffic classification: A multi-attribute associated fingerprint approach," in 2019 IEEE 27th International Conference on Network Protocols (ICNP). IEEE, 2019, pp. 1–11.

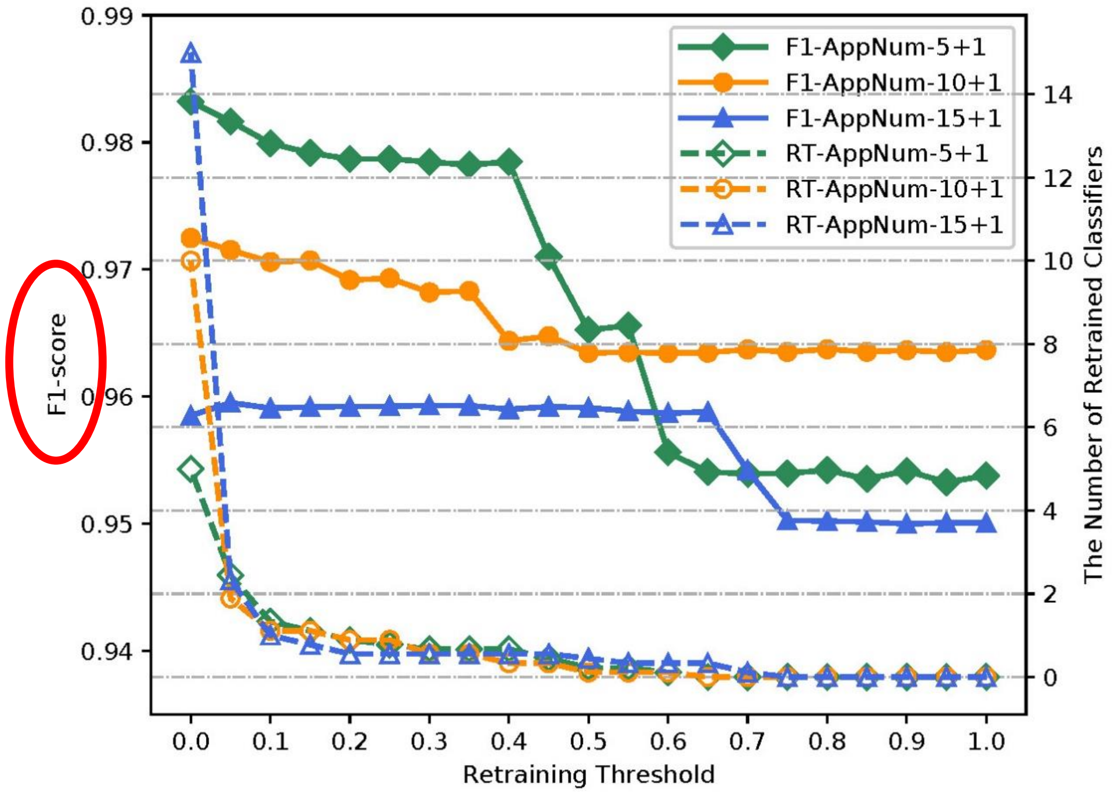
Analysis of the Retraining Threshold



Performance Evaluation on Different Retraining Thresholds

Analysis of the Retraining Threshold

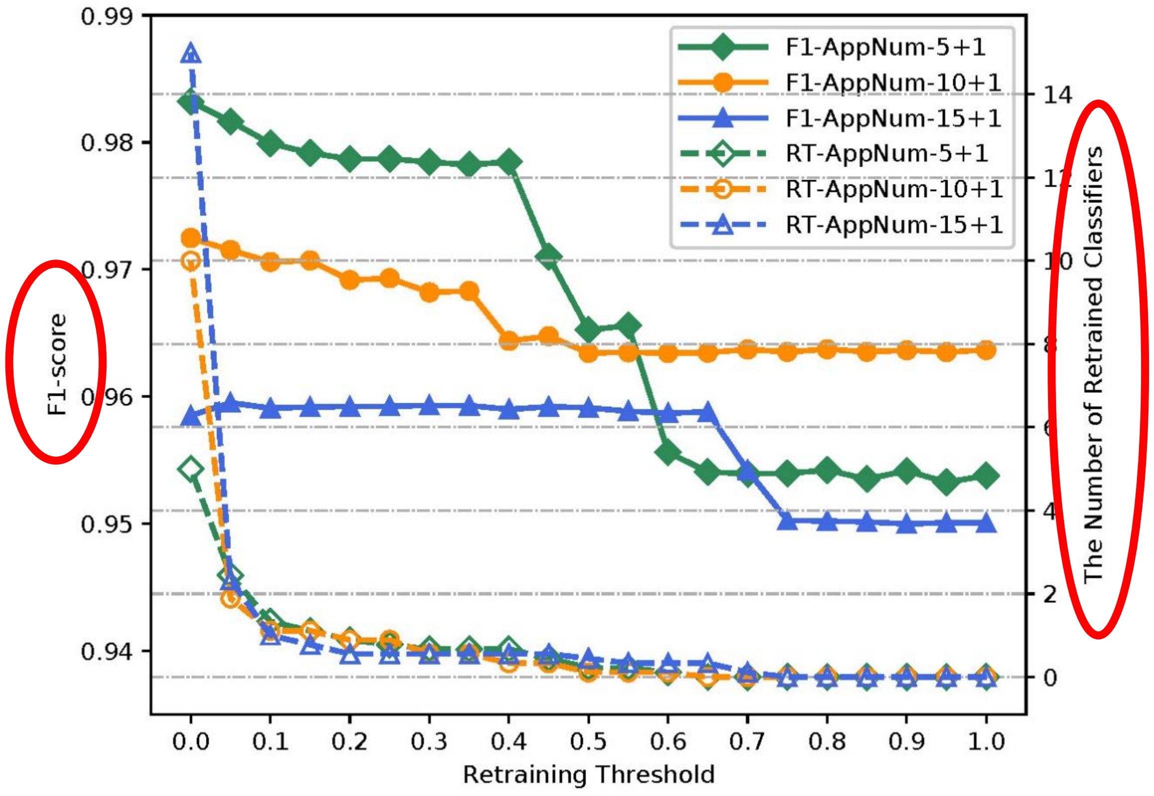
- Both the F1-scores and the number of retrained classifiers declines with the retraining threshold.



Performance Evaluation on Different Retraining Thresholds

Analysis of the Retraining Threshold

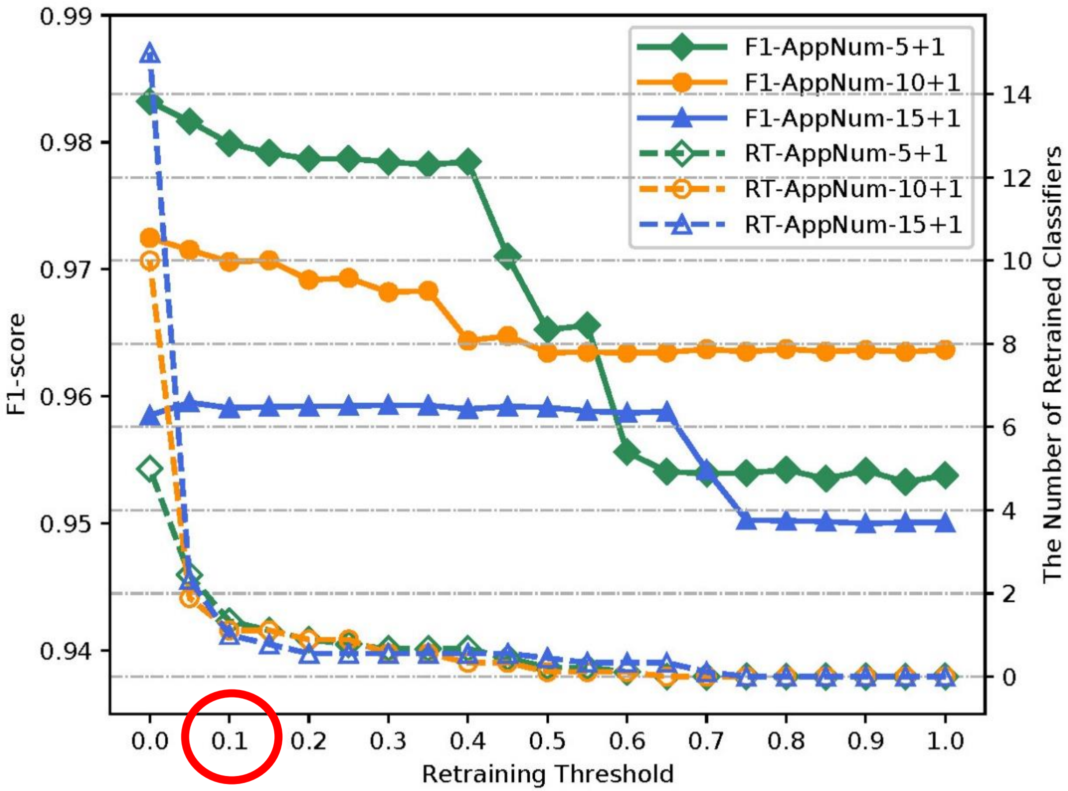
- Both the F1-scores and the number of retrained classifiers declines with the retraining threshold.



Performance Evaluation on Different Retraining Thresholds

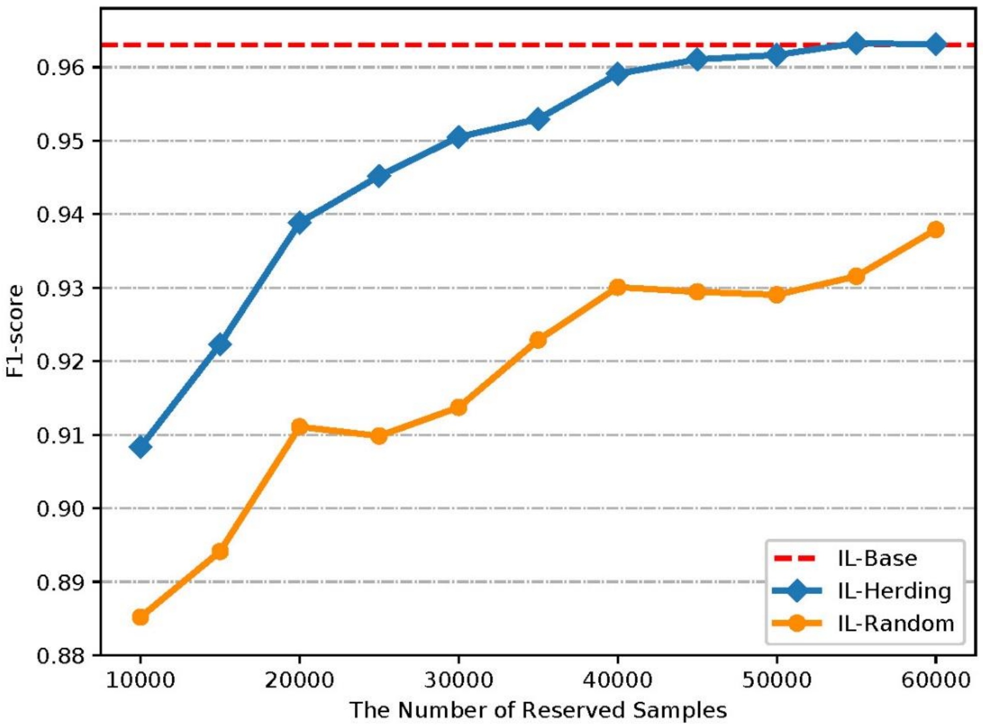
Analysis of the Retraining Threshold

- Both the F1-scores and the number of retrained classifiers declines with the retraining threshold.
- When the retraining threshold increases from 0 to 0.1, the number of retrained classifiers decreases sharply while the classifier only loses a little F1-scores.



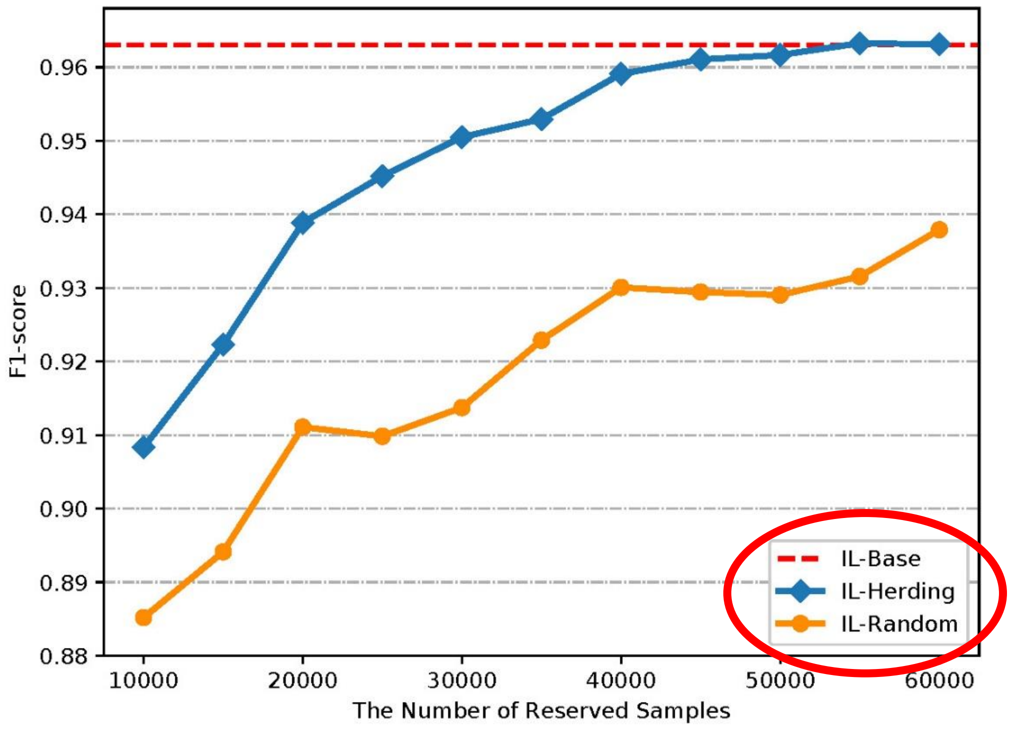
Performance Evaluation on Different Retraining Thresholds

Analysis of the Sample Selection



Comparison Results on Different Sample Numbers

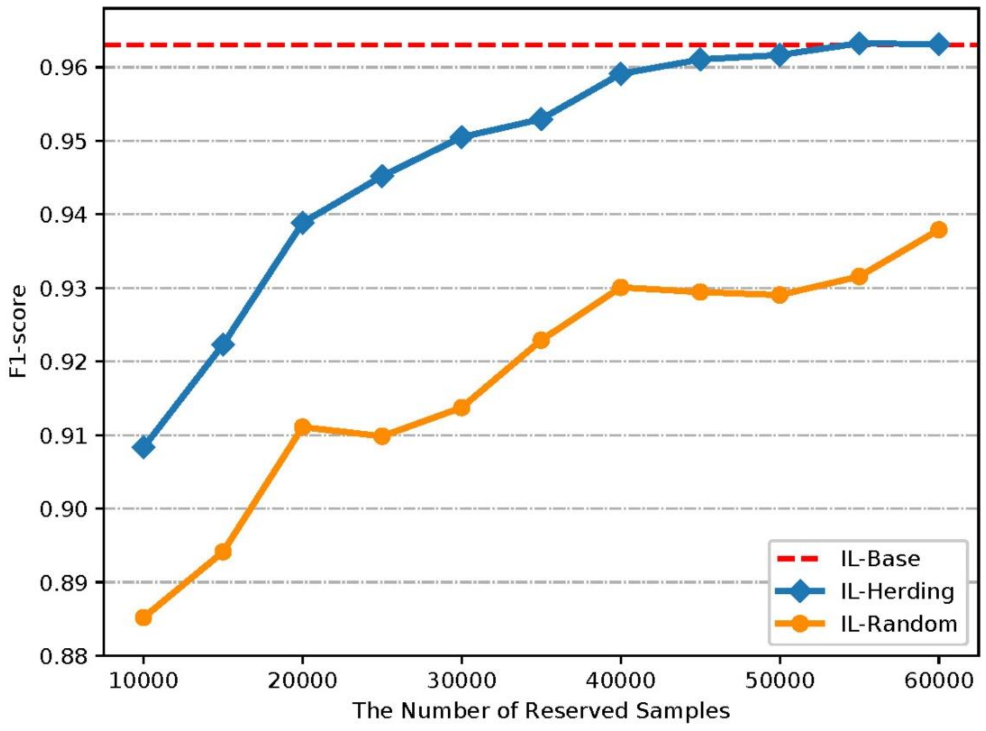
Analysis of the Sample Selection



Comparison Results on Different Sample Numbers

Analysis of the Sample Selection

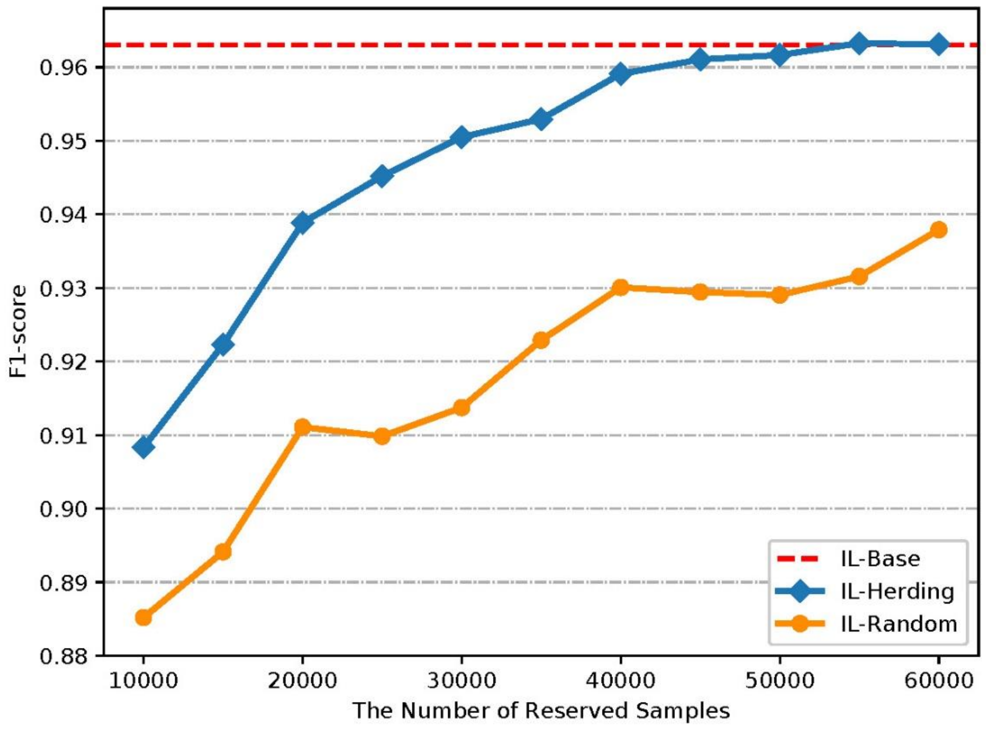
- The F1-score of both IL-Herding and IL-Random increases.



Comparison Results on Different Sample Numbers

Analysis of the Sample Selection

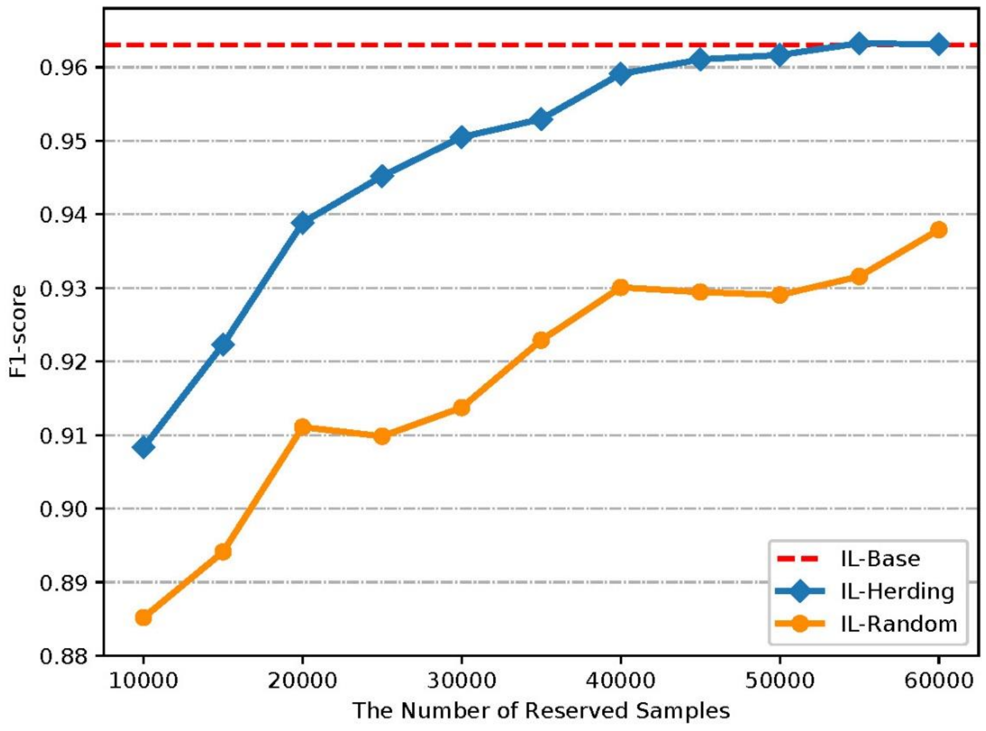
- The F1-score of both IL-Herding and IL-Random increases.
- The IL-Herding gradually approaches IL-Base with the enrichment of reserved samples.



Comparison Results on Different Sample Numbers

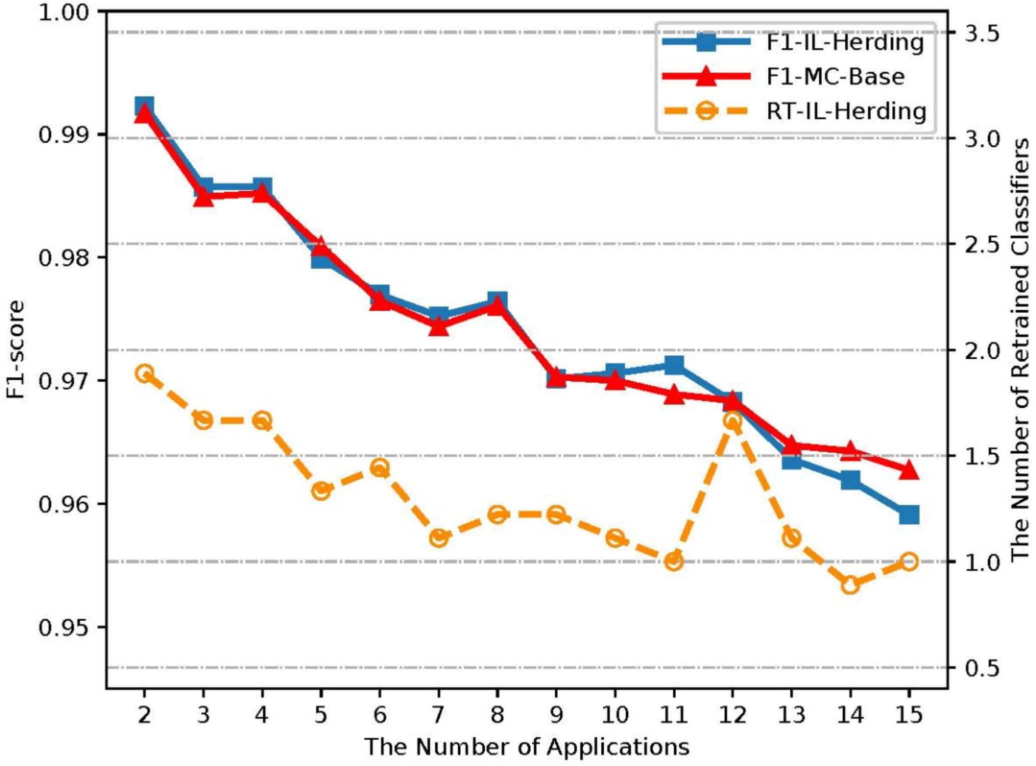
Analysis of the Sample Selection

- The F1-score of both IL-Herding and IL-Random increases.
- The IL-Herding gradually approaches IL-Base with the enrichment of reserved samples.
- The IL-Herding shows overall higher classification accuracy than the IL-Random.



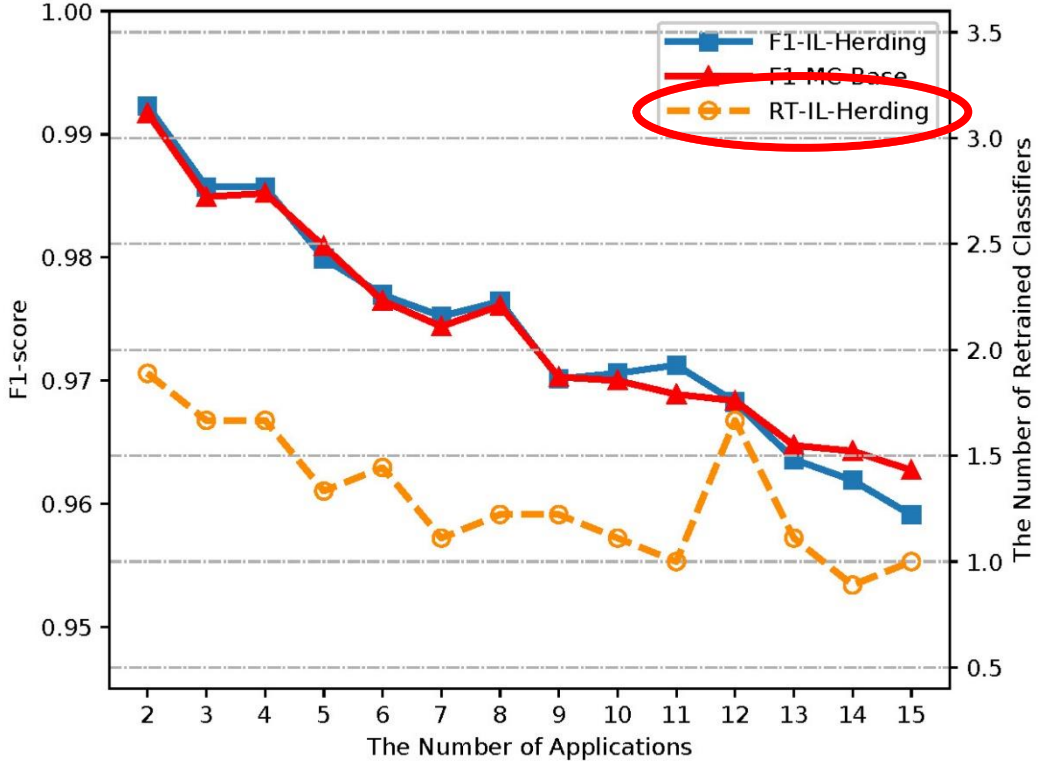
Comparison Results on Different Sample Numbers

Analysis of the Number of Applications



Comparison Results on Different Numbers of Applications

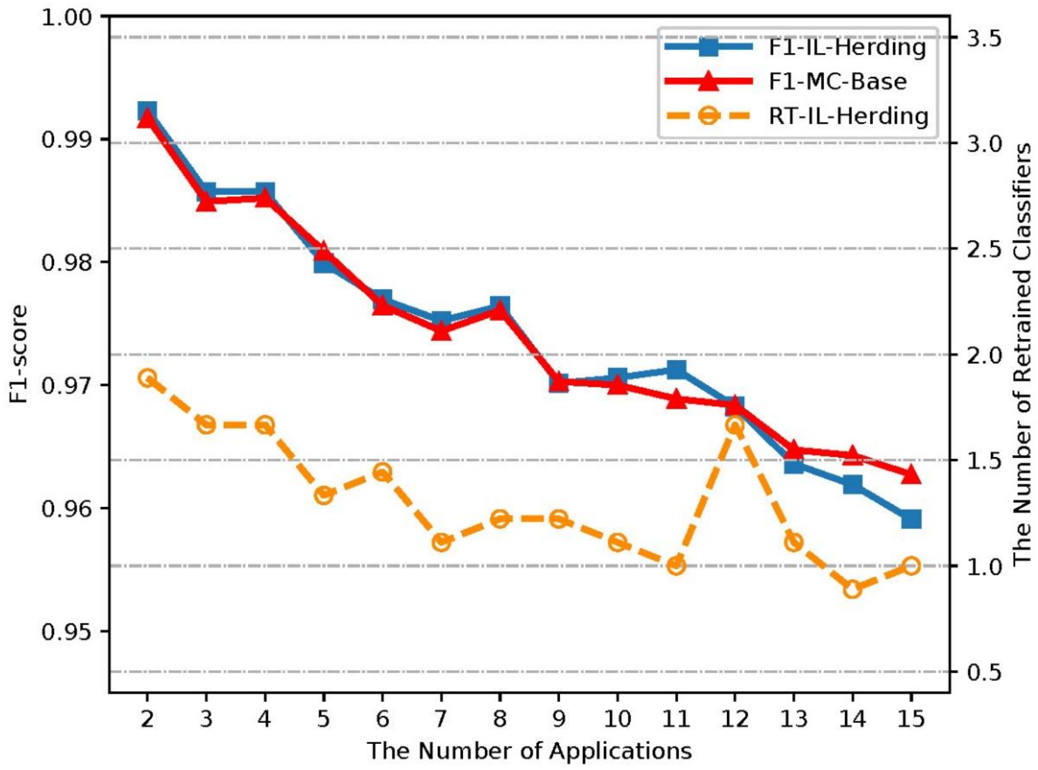
Analysis of the Number of Applications



Comparison Results on Different Numbers of Applications

Analysis of the Number of Applications

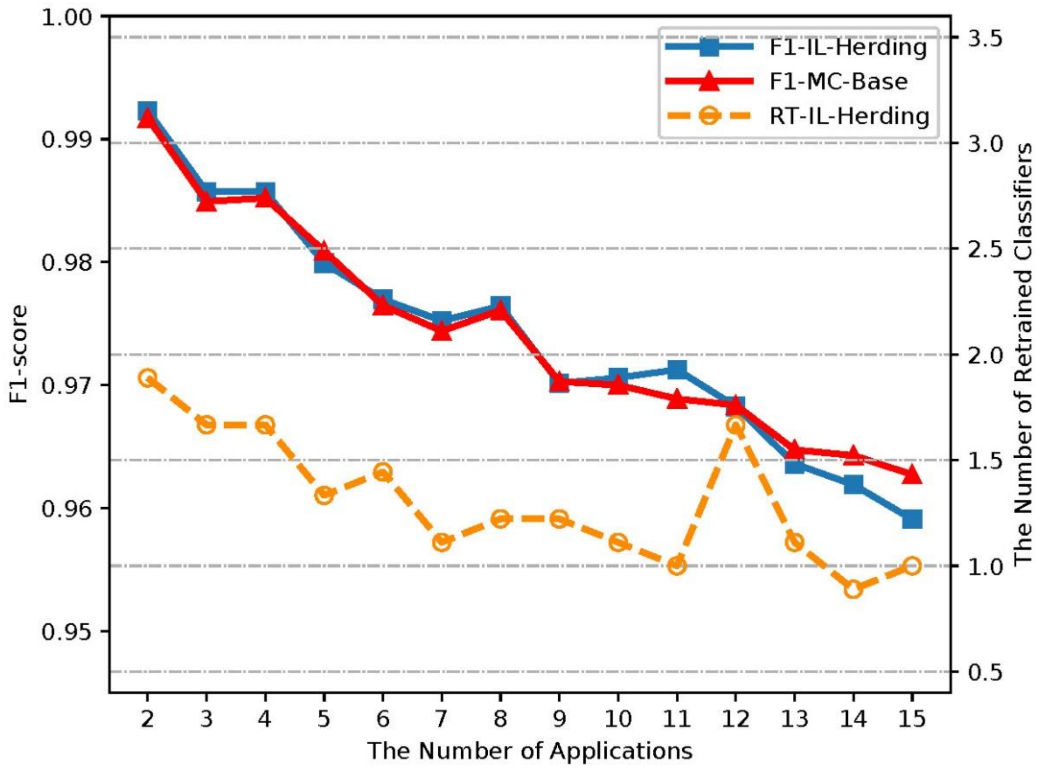
- The F1-scores of both the IL-Herding and MC-Herding slowly decrease with the increase of the number of applications.



Comparison Results on Different Numbers of Applications

Analysis of the Number of Applications

- The F1-scores of both the IL-Herding and MC-Herding slowly decrease with the increase of the number of applications.
- The number of retrained classifiers generally declines from 2.0 to 1.0 with the increase in the number of applications.



Comparison Results on Different Numbers of Applications

- For more details, please contact chenyige@iie.ac.cn
- Questions & Answers