

Recommendation

ITU-T Y.4494 (11/2023)

SERIES Y: Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

Internet of things and smart cities and communities –
Frameworks, architectures and protocols

Reference architecture of collaborative decentralized machine learning for intelligent Internet of things services

ITU-T Y-SERIES RECOMMENDATIONS

Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

GLOBAL INFORMATION INFRASTRUCTURE	Y.100-Y.999
INTERNET PROTOCOL ASPECTS	Y.1000-Y.1999
NEXT GENERATION NETWORKS	Y.2000-Y.2999
FUTURE NETWORKS	Y.3000-Y.3499
CLOUD COMPUTING	Y.3500-Y.3599
BIG DATA	Y.3600-Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800-Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	Y.4000-Y.4999
General	Y.4000-Y.4049
Definitions and terminologies	Y.4050-Y.4099
Requirements and use cases	Y.4100-Y.4249
Infrastructure, connectivity and networks	Y.4250-Y.4399
Frameworks, architectures and protocols	Y.4400-Y.4549
Services, applications, computation and data processing	Y.4550-Y.4699
Management, control and performance	Y.4700-Y.4799
Identification and security	Y.4800-Y.4899
Evaluation and assessment	Y.4900-Y.4999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T Y.4494

Reference architecture of collaborative decentralized machine learning for intelligent Internet of things services

Summary

A collaborative decentralized machine learning (CDML) architecture can support machine learning (ML) model distributed training and inference across highly heterogeneous and resource-constrained Internet of things (IoT) devices, which results in less latency, higher reliability, lower energy consumption and saving bandwidth resources. Using CDML, spare resources across decentralized IoT devices can be fully used to perform computation-intensive ML tasks collaboratively with high performance.

Recommendation ITU-T Y.4494 introduces CDML for intelligent IoT services, and provides the characteristics and reference architecture of CDML for intelligent IoT services.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T Y.4494	2023-11-29	20	11.1002/1000/15688

Keywords

Collaborative decentralized machine learning (CDML), intelligent IoT services, Internet of Things (IoT).

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2024

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope	1
2 References.....	1
3 Definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	1
4 Abbreviations and acronyms	2
5 Conventions	2
6 Overview of CDML for intelligent IoT services	2
7 Characteristics of CDML for intelligent IoT services	4
8 Reference architecture of CDML for intelligent IoT services.....	5
8.1 General architecture.....	5
8.2 Functional components.....	6
9 Common capabilities of CDML for intelligent IoT services.....	13
9.1 Continuous optimization of ML task allocation strategies.....	13
9.2 Publication of ML capabilities	14
9.3 Subscription to ML capabilities.....	14
9.4 Access to ML capabilities	14
Appendix I – Use Case: The implementation of the CDML for intelligent IoT services.....	16
Appendix II – Common procedures of CDML for intelligent IoT services	18
II.1 Publication of ML capabilities	18
II.2 Subscription to ML capabilities.....	18
II.3 Accessing the subscribed ML capabilities in CDML.....	19
Bibliography.....	21

Recommendation ITU-T Y.4494

Reference architecture of collaborative decentralized machine learning for intelligent Internet of things services

1 Scope

This Recommendation aims to propose the reference architecture of collaborative decentralized machine learning for intelligent IoT services.

The scope of this Recommendation includes:

- Overview of collaborative decentralized machine learning for intelligent IoT services;
- Characteristics of collaborative decentralized machine learning for intelligent IoT services;
- Reference architecture and common capabilities of collaborative decentralized machine learning for intelligent IoT services.

Use case and common procedures will be provided in the appendices.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T Y.4000] Recommendation ITU-T Y.4000/2060 (2012), *Overview of the Internet of things*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 application [b-ITU-T Y.2091]: A structured set of capabilities, which provide value-added functionality supported by one or more services, which may be supported by an API interface.

3.1.2 capability [b-ITU-R M.1224-1]: The ability of an item to meet a service demand of given quantitative characteristics under given internal conditions.

3.1.3 device [ITU-T Y.4000]: With regard to the Internet of things, this is a piece of equipment with the mandatory capabilities of communication and the optional capabilities of sensing, actuation, data capture, data storage and data processing.

3.1.4 Internet of things (IoT) [ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

3.1.5 machine learning (ML) [b-ITU-T Y.3172]: Processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AI	Artificial Intelligence
ARM	Advanced Reduced instruction set computer Machines
CDML	Collaborative Decentralized Machine Learning
CGC-FC	Computation Graph Construction Functional Component
CGO-FC	Computation Graph Optimization Functional Component
CL-FC	Continuous Learning Functional Component
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DNN	Deep Neural Network
DSCA-FC	Device Service Capability Abstraction Functional Component
FE-FC	Feature Extraction Functional Component
FPGA	Field Programmable Gate Array
GCN	Graph Convolutional Network
GPU	Graphics Processing Unit
ICTA	Intelligent Computing Task Allocation
IoT	Internet of Things
ML	Machine Learning
OS	Operating System
QoS	Quality of Service
RAM	Random Access Memory
RKGC-FC	Resource Knowledge Graph Construction Functional Component
RSC-FC	Resource Subgraphs Construction Functional Component
SPP-FC	System Performance Prediction Functional Component
TPU	Tensor Processing Unit

5 Conventions

In this Recommendation:

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

6 Overview of CDML for intelligent IoT services

Machine learning (ML) has become an important methodology in many artificial intelligence (AI) applications, such as intelligent monitoring, home service robot, VR and intelligent control, etc., and has achieved state-of-the-art performance in various fields including smart elderly care, smart home, Internet of vehicles, smart community, industrial Internet, etc. ML and AI are widely used in intelligent IoT services. However, most ML models are highly computationally intensive, leading to stringent requirements on resources, including computing power, memory and communication

resources. In particular, computing capability provided by powerful hardware (e.g., GPU, FPGA, TPU) is the key to boost the performance of these ML models. Cloud computing with powerful computing capability is a promising solution, which requires Internet of Things (IoT) devices to transfer raw data collected to the cloud for processing. However, uploading massive amounts of data to cloud data centres would consume significant bandwidth resources, and might lead to unbearable delay and energy consumption. So there is a trend to offload computing tasks from the centralized cloud to decentralized devices at the edge of network, referred to as decentralized ML.

Along with the explosive development trend of IoT, as well as the rapid increases in both the number and the computing capability of smart IoT devices, there is a growing demand to deploy intelligent applications to a wide spectrum of IoT devices. The connectivity, communications and intelligence of IoT devices make many of them agents for intelligent services. These connected IoT devices have not only a large amount of user raw data but also slightly idle resources and computing power. Nevertheless, a single IoT device may have resources so limited that it can hardly support ML model towards training/inference on its own. The good news is that many IoT devices have some spare resources most of the time and can communicate with each other over local networks; these distributed devices can potentially be powerful contributors when orchestrated effectively. Therefore, a decentralized ML architecture based on device collaboration, namely collaborative decentralized machine learning (CDML), may be a feasible and practical solution. It can make full use of the spare resources across diverse IoT devices and thus meet the increasing demands from the ML-enabled smart IoT services. The advantage of CDML is that it can concurrently take into account latency, privacy, and energy consumption, and efficiently use limited computing, storage and communication resources at the edge in the meanwhile.

CDML can achieve distributed training and inference of an ML model across diverse IoT devices in a cooperative manner, which will push the computation of ML models from the cloud data centre to the IoT devices near the edge of the network. CDML contains three main functions: resource management, task decomposition and task allocation. Resource management has the ability to perceive, manage, and orchestrate the idle resources on heterogeneous IoT devices, including computing, storage and communication resources. Task decomposition is able to partition a computation-intensive ML model on the operator level so as to allocate certain operators to the proper IoT devices. Task allocation, which is the core of CDML, is responsible for generating multiple task allocation strategies and choosing the optimal strategy as a result, so as to realize the best matching of available resources and computing tasks. In summary, CDML achieves a full-stack optimized decentralized intelligent IoT system design based on device collaboration. CDML is able to support ML model training and inference across heterogeneous and resource-constrained IoT devices by maximizing resource utilization and optimizing task allocation. The overview of CDML is shown in Figure 1.