Learning from data streams for automation and orchestration of 6G industrial IoT: toward a semantic communication framework

Shiva Raj Pokhrel[1]

Abstract

Established methods of communication are based mainly on Shannon's theory of information, which purposefully overlooks semantic elements of communication. The future wireless technology should promise to facilitate many services, based on content, needs, and semantics, precisely customized to network capabilities. This gave rise to significant concern for *Semantic Communication* (SC), a novel paradigm considering the message's meaning during transmission. Federated learning (FL) and Asynchronous Advantage Actor Critic (A3C) are the two emerging distributed and artificially intelligent approaches that provide diverse and possibly massive network coverage for data-driven SC solutions of industry 4.0 automation. Although SC is still in an early development stage, FL-empowered architecture has been recognized as one of the most promising solutions to meet the ubiquitous intelligence in the anticipated sixth-generation (6G) networks. This paper identifies industry 4.0 automation needs that drive the convergence of artificial intelligence and 6G for learning from data streams. We *develop a novel SC framework based on the FL and A3C networks* and discuss its potential along with *transfer learning* to address most of the new difficulties anticipated in 6G for industrial communication networks. Our proposed framework has been evaluated with extensive simulation results.

1 Introduction

Fifth generation (5G) wireless communication networks are rapidly being launched on a large scale worldwide [1]. The academic and industry research community has already started investigating the advancements of the cellular technologies for the next decade, i.e., the sixth generation (6G) [1, 2]. The semantic communication (SC) paradigm has received much attention for next-generation advancements [3–5]. SC has a great potential to disrupt current technical problems in existing communication systems based on the possibility of improved performance and effectiveness with the communication from the second level (complementing the impeding issues at the first level communication identified by Shannon and Weaver [6]) in 6G.

6G networks are anticipated to follow a much more comprehensive strategy, catalyzing creative technologies and intelligent infrastructure while conducting prompt and highly effective/productive data collection, transmission, learning, and analyzing everywhere at any moment in time [7]. 6G should concentrate in specific on a modern concept of ubiquitous artificial intelligence for goal-oriented communication, an ultra-flexible infrastructure that facilitates human-level knowledge, semantic processing, and intelligence into all facets of communication networking.

With SC in 6G, the mechanism underlying SC can be the semantic information considering the significance and veracity of the original content (since it may be both instructive and factual). But defining such semantic information or representing semantic characteristics with a specific mathematical model is a highly non-trivial task, making the direct coupling with the first level (Shannon/Weaver communication framework) infeasible.

✉ Shiva Raj Pokhrel
shiva.pokhrel@deakin.edu.au

[1] Deakin University, Geelong, Australia

Table 1 Summary of the survey works on edge intelligence (c.f. Fig 1)

Refs.	Theories	Remarks	Perspective
[18]	Edge intelligence	Training/inference systems, DL for edge optimization and application	Macro-level
[19]	Edge intelligence	Device-cloud-edge coordination, Optimization in training/inference	Macro-level
[20]	Edge intelligence	Offloading, Edge caching/training/inference	Micro-level
[21]	Federated intelligence	Training/inference systems, DL for edge optimization and application	Macro-level
[22]	Deep learning	Device-cloud-edge coordination, Optimization in training/inference	Micro-level
[23]	IoT big data	Offloading, Edge caching/training/inference	Micro-level
[24]	Intelligent networking	Offloading, Edge caching/training/inference	Micro-level
This work	Asynchronous actor-critic	Comprehensive edge intelligence with federated and transfer learning	Micro & Macro

One of the main challenges when developing an AI-based SC network design for a realistic future 6G system is;

How to integrate distributed data processing and collaborative learning for semantic communication through an extensive range of heterogeneous wireless devices?

To this end, federated learning (FL) [8, 9] and Asynchronous Advantage Actor Critic (A3C) [10] are now two evolving distributed AI approaches. They allow data-driven AI and machine learning (ML) over a vast amount of dispersed data residing on mobile devices and collaborative learning via knowledge sharing. Thanks to their capability to share knowledge bases, perform model training and learning on diverse and potentially huge scale networks, FL has already drawn tremendous attention while maintaining all relevant information localized [11–13].

One of the representative industry verticals considered in this paper is smart manufacturing in industrial IoT networks. In particular, we focus on a cellular-based smart factory characterized by IoT integration and related services in intelligent manufacturing. We believe that advancements of *ultra-reliable low-latency communication* (URLLC) and *massive machine-type communications* (mMTC) proposed in 5G by adopting the SC paradigm will be the two key enablers for future Cellular-based factories automation [14].

Smart manufacturing of *Industry 4.0*, marks the transition from traditional to highly connected digital technologies in industrial settings as shown in Figs. 1 and 2. As demonstrated in Figs. 1 and 2, this may include the use of Industrial IoT machines, intelligent computation, distributed learning, self-curing networks, and automated industry for smart manufacturing [15]. Such use cases demand highly reliable and minimal latency goal-oriented communication, which the standard 5G approaches cannot support [2].

This research work has been motivated by the issue of how to make industry machines and robots, shown in Fig. 2, (i) learn (understand and adapt) efficiently in a new environment, (ii) perform human-level learning based on solid comprehension (*semantic learning*) and (iii) transfer their experience or share knowledge-base between sender and receiver to enable all intelligent devices to use prior knowledge and semantics effectively.

We present an FL design for edge intelligence in industry 4.0 systems to tackle the challenges mentioned above thru *Continuous Federated Reinforcement Learning* (CFRL) approach (extending relevant insights from [10, 16]). We suggest an information fusion-based approach to update a standard global model and knowledge-base implemented at the network edge. Then, effective methods of transfer learning (TL) [17] are applied for consistent human-level cognitive intelligence and semantic learning for the best fitting in the industry 4.0 systems.

Industry 4.0 relies entirely on seamless real-time communication. Therefore, we consider the forthcoming 6G-based factory automation and orchestration using edge intelligence and modern machine learning for contexts.
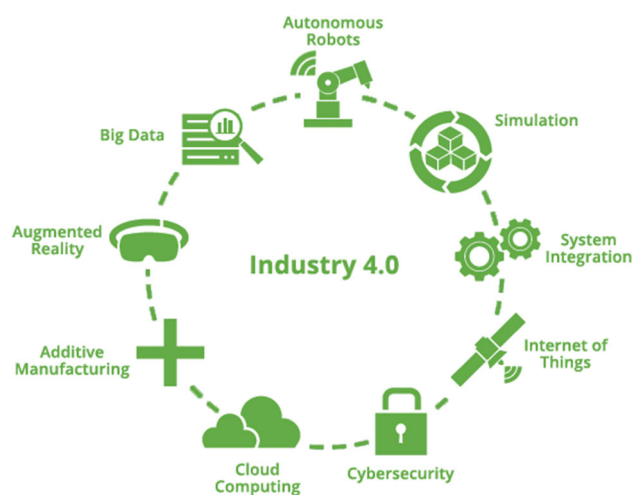

Fig. 1 A high-level view of industry 4.0 evolution and applications. In this work, we focus on the need for semantic communication to automate and orchestrate machines and robots. See the comparison of survey works later in Table 1
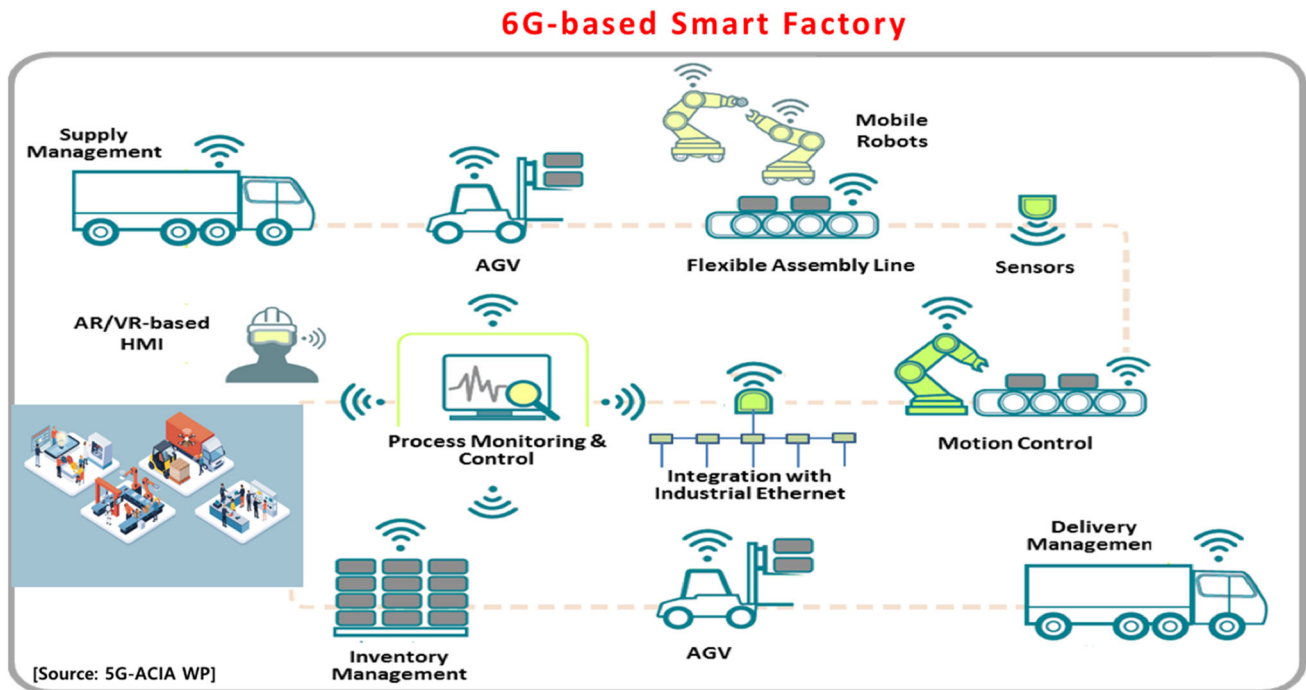
Fig. 2 Industry 4.0 automation and orchestration is relying on URLLC and mMTC. Without an ultrareliable and low latency link, industrial networks cannot collect or process data on the floor, created by IoT sensors or on edge. This knowledge could not be used for surveillance or regular maintenance. Technicians are unable to see issues occurring in real-time, so both software-defined networking and automation-based systems cannot work correctly. To ease the communication issues, we consider a 6G-based smart factory and focus only on the learning and automation of the industry 4.0 system using edge intelligence

Without an ultra-reliable and low latency link, industrial networks cannot collect or process data on the floor, created by IoT sensors or on edge. This incomplete knowledge base (without background information, environments, and semantics) could not be used for surveillance or regular maintenance. Technicians cannot see issues occurring in real-time, so both SDN and automation-based systems often fail to work correctly. To ease the communication issues, we consider a novel SC-based smart factory envisioning 6G and focus only on the learning and automation/orchestration aspect of the industry 4.0 system with edge intelligence.

1.1 Contributions

Our main contributions in this work are outlined as follows:

- We develop a Semantic Communication Framework with continuous federated reinforcement learning (CFRL) capabilities considering successful features of *human cognitive neuroscience* and semantic learning.

It enables industrial machines and robots to carry out continuous learning from data streams in the edge-intelligent industry 4.0 process automation/orchestration.

- We are designing a new knowledge fusion method for learning from data streams aimed at IIoT automation and orchestration. It can integrate previous experience, environmental factors, and expertise of continuously learning machines (for revealing the semantic and pragmatic meaning) and create a standard paradigm for SC-based edge-intelligent manufacturing systems.
- Two successful methods to learning transition are proposed and evaluated for transfer learning to allow machines to respond efficiently to new settings. The proposed methods and CFRL framework have been assessed with extensive experiment results.

A novel distributed, asynchronous reinforcement learning (recommender-like) system is developed in this paper using the asynchronous advantage actor-critic approach, which combines ideas from TL, A3C, and FL. Our core idea is to keep the machine preferences and/or interactions as local knowledge/learning and adopt integrated local, on-machine, and complementary global models. For example, the training procedure for the global model could only be based on the loss gradients of the local learnings.

2 Related works and theories

Shannon and Weaver theory for reliable communication is essentially mature and established by now [25]. With 5G and advances in technology, all wired and wireless communication can be ultra-reliable. The communication performance at the first level can be enhanced quantitatively but improving the performance qualitatively seems extremely challenging and close to impossible simply by adjusting first level communication parameters [26, 27]. Modern communication systems have smart endpoints, which implies intelligence, capability, and diversity. It is known that diversity suggests (potential) misunderstandings [26] and therefore, we shall now start investing in schemes/algorithms to make the endpoints more reliable for detecting/correcting misconceptions—the tenets of *semantic communication*.

One may consider a general model of semantic communication where two intelligent interacting agents, sender and receiver, communicate with each other, and the sender wants to accomplish some goals while the receiver is trying to help the sender. We can easily generalize this process at time t in terms of interacting agents where *Agent: States (t) × Inputs (t) = States (t+1) × Outputs (t)*. However, there is an uncertainty associated with the receiver, where the sender does not know the receiver and vice-versa. This needs more attention as it creates a new class of problems and new challenges.

To this end, along the lines of other researchers [5, 6, 26], we believe that the focus should be toward goal-oriented communication, where the sender always attempts to reach the predefined goal with the help of the receiver. However, "*how shall we model and quantify this mechanism?*" is an open problem in the field. Classical approaches of modelling/capturing the dynamics by combining some functions of the state of the sender/receiver with their interaction tasks fail in such a semantic-uncertainty setup, which mandates the need for any advice/suggestion from the third agent (such as Knowledge base). The knowledge base, for instance, can pose the tasks for the sender/receiver interaction and judges the success. It can act as another referee agent in keeping track and determining whether the state evolution reflects goal-oriented achievements or not, thus incrementally building more experiences for the knowledge base. In this way, we can see that semantic communication can deal with uncertainly at the endpoints, which is impractical simply by using information theory fundamentals.

We attempt to bridge the gap using recent findings and the current state-of-the-art edge intelligence with relevant insights from the above discussions. Edge intelligence is exceptionally rich in literature (which (potentially) helps to build the knowledge base for the proposed semantic communication); see the summary of survey works in Table 1 and the references therein (viz. [18–24]). Of particular relevance to this work are the three closest theories—TL, FL, and A3C [10]. FL differs from usual distributed learning as it uses non-IID and unbalanced, massively distributed data over limited communication resources.

A3C determines error using the rewards produced in the state transition, which is different from FL that determines error based on the error in the output (like general NN)—details are discussed as follows.

2.1 Asynchronous advantage actor critic (A3C)

The A3C algorithm developed by Google's DeepMind community [10] has splashed by obsoleting the standard deep reinforcement learning (DRL) algorithms. It was faster, simpler, more robust, and attained much better scores on the standard DRL tasks. In contrast to other DRL approaches, A3C functions well in both continuous and discrete spaces of operation. Therefore, it has now become the most reliable DRL algorithm for new challenging issues with complex spaces of states and actions.

A high-level architecture of A3C is as shown in Fig. 3 and the three **A**'s of A3C are discussed as follows:

Asynchronous. In contrast to conventional DRL techniques, in which a single agent represented by a single neural network interacts with a single environment, A3C uses multiple incarnations (see Fig. 3) to facilitate the learning process. A global network in A3C and several worker agents have their own set of network parameters. Each of these agents communicates with their own copy of the world, whereas the other agents communicate with their surroundings. This works better than having a single agent (beyond the speed of getting more work done) because each agent's experience is independent of the others' experience. It is worth noting that the overall expertise available for training becomes more diverse and asynchronous.

Advantage. From our Policy Gradient implementation perspective, the updated policy used the discounted returns from a series of interactions to inform the agent which of their acts were 'great' and which were 'negative'. The system can then be modified with an advantage to promote better and deter actions adequately. The perspective to use advantage estimates rather than merely discounted returns is to enable the agent to determine not about how decent their actions have been but how much smarter they were than expected. Rationally, this allows the system to concentrate on where the network's predictions have been sorely missing.

Actor-Critic. It provides the best of both policy gradient and Q-learning approaches. In the scenario of A3C, both a
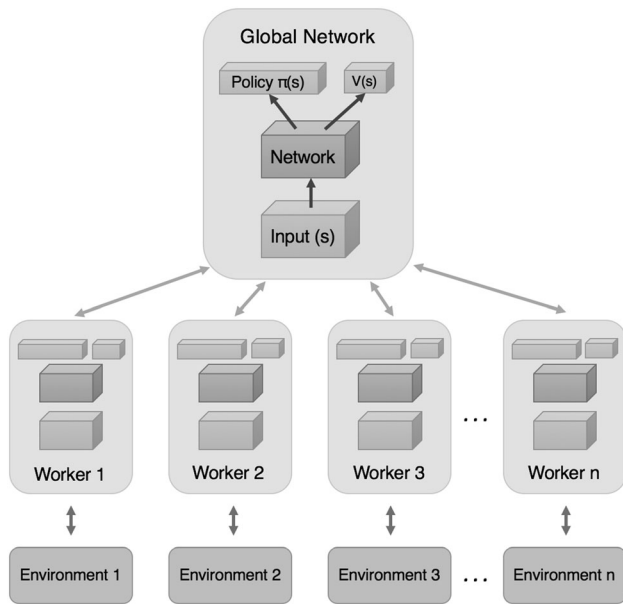
Fig. 3 An abstract view of the overall architectural design diagram for A3C

value function $V(s)$ (how good a particular state is to be in) and a policy $\pi(s)$ (a set of probability outputs for action) will be estimated. These will each be distinct layers sitting at the top of the network, fully connected. Pragmatically, the agent uses the value estimate (the critic) to more intelligently update the policy (the actor) than conventional methods of policy gradients.

Our framework implements federated multi-machine learning through the fusion of knowledge. Google first proposed federated learning, which demonstrated its effectiveness through experiments on different datasets [8]. The datasets are collected and stored at multiple network end nodes in federated learning systems. A learning model is trained from the decentralized datasets at a centralized global server location [28]. Unlike the traditional standard learning method, where multiple edges simultaneously learn, our framework adopts A3C and the first training method, then fuses knowledge to reduce dependence on the communication quality.

2.2 Transfer learning

As discussed above, upon identifying an expert machine, the learning agent utilizes the transferred DQN model from the expert machine [29] and its current native DQN model to produce an aggregate DQN model [30]. Given the transfer rate $r \in [0, 1]$, the new Q-learning agent can therefore be mathematically interpreted as[1]

---

[1] After each learning stage, r will slowly be decreased to minimize the influence from the expert machine model on the new DQN model.

$$Q^{\text{new}}(s,a) = rQ^{\text{transfer}}(s,a) + (1-r)Q^{\text{current}}(s,a). \qquad (1)$$

In a distributed cooperative 6G empowered IIoT multiple machines system, the policy vector of all agents can be updated by:

$$\Pi_{t+1}(s_t) = \begin{pmatrix} \pi_{t+1}^1 \\ \vdots \\ \pi_{t+1}^i \\ \vdots \\ \pi_{t+1}^M \end{pmatrix} = \begin{pmatrix} \arg\max_{a^1}\{Q_{t+1}^1(s_t^1, a_t^1)\} \\ \vdots \\ \arg\max_{a^i}\{Q_{t+1}^i(s_t^i, a_t^i)\} \\ \vdots \\ \arg\max_{a^M}\{Q_{t+1}^M(s_t^M, a_t^M)\} \end{pmatrix}.$$

3 Our proposed methodology

3.1 CFRL framework for semantic communication with extended A3C

CFRL will reduce the training period without losing accuracy of decision functionality on industrial edge computing systems. CFRL allows the use of the edge-computing configuration to understand the policies and goals desired. It comprises an edge server consisting of a knowledge base, groups of settings, and one or more robots/machines. Moreover, we are developing a federated learning framework using relevant insights from A3C for the fusion of local models into a global model shared at the knowledge base of an edge server. The edge server maintains local fusion models into a mutual environment, and then, the mutual environment develops the aforementioned semantic communication capabilities with changes in the settings.

As illustrated in Fig. 4, CFRL is a lifetime learning-cum-communication framework of industry network automation/orchestration systems. Compared to the standard A3C that simultaneously updates policy network parameters, the recommended information fusion solution is more appropriate for the proposed SC framework using the federated architecture. The presented technique will be capable of fusing asynchronous evolutionary models for goal-oriented communication. At the very same time, the method of modifying parameters has some environmental requirements. At the same time, the new knowledge fusion idea considers the embodiment of the semantics and therefore has very little or no environmental impact. Using a generative network and dynamic weighting approach, knowledge incorporation and building can be done evolutionarily in lieu of A3C (which only produces a decision model during learning and provide recommendations).
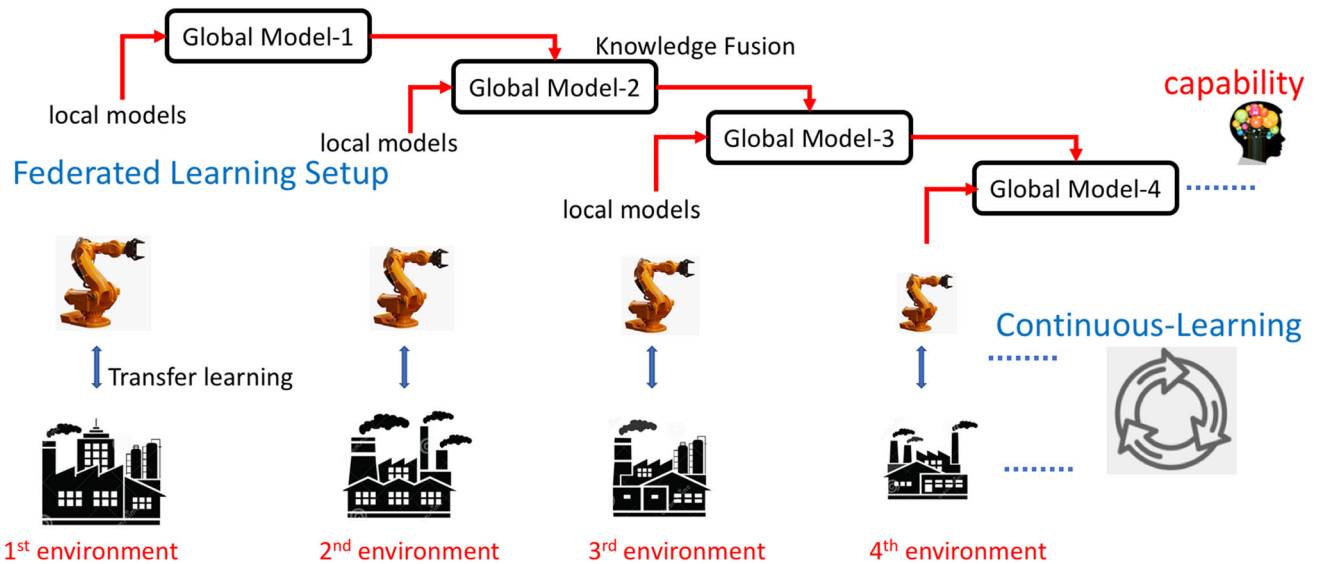
Fig. 4 A high-level view of the proposed CFRL framework. In an industrial environment, through reinforcement learning, the robots/machines learn to automate specific new tasks in the new environment and obtain the private local model. Local models may also be the product of several robots, not only from one robot working in various settings. After that, Robots must transfer the local model to the knowledge base of the edge server. By fusing local models into the global model, the knowledge base of the edge server evolves the global model. Motivated by the transfer learning approach, at the knowledge base of the edge server, the robot/machine uses successor features to transfer the strategy to a new environment

Unlike A3C, the training environment will be highly dynamic in the proposed CFRL framework for the intended goal-oriented communications as shown in Algorithm 2. State of the agents can be considered as uncountable with more uploading of the training environments. The structure of the policy's hidden layers can be in different actors, and even it could be in a different network setting. Typically, training results fuse in the knowledge base of the edge server. Based on the shared model, the robots are continually trained in new environments. In upload and download procedures, robots/machines and the edge server have continuous interactions. Therefore, the proposed SC framework is best suited to edge intelligence for future industry 4.0 automation/orchestration systems where the environment is uncountable. In fact, the proposed framework can be applied from a lifelong learning perspective. However, the answer to a question: "Why does natural (human) communication differ so much from such a designed communication?" requires further investigations. This is beyond the scope of this work.

Algorithm 1 CFRL Processing Algorithm

1: **procedure** INITIALIZE
2: Action-value Q-network with random weights; Parameters of shared model; Local networks (N)
3: Evolved parameters of the model
4: **end procedure**
5: **procedure** EDGE INTELLIGENCE
6: **While** (Global Network Executing) **do**
7: $\quad\pi \leftarrow$ evolved parameters
8: $\quad$**for** (n=1; n$\leq$ N; n++)
9: $\quad\quad$parameter$_n \leftarrow$ Machine(n) perform DRL with π
10: $\quad\quad$Send parameter$_n$ to Edge
11: $\quad$**end for**
12: **end While**
13: parameter$_{new} \leftarrow$ Fuse (parameter$_1$ ... parameter$_{old}$)
14: **end procedure**

Algorithm 2 CFRL Using A3C Algorithm

1: **procedure** INITIALIZE
2: Globally Shared Parameters, Θ, Θ_n;
3: Shared Counter $C = 0$
4: Specific Parameters, Θ', Θ'_n; Step Counter T=1.
5: **end procedure**
6: **While** (C$\leq$ Max) **do**
7: $\quad$Reset Gradients $d\Theta$ and $d\Theta_n$
8: $\quad$Sync Specific Parameters $\Theta' = \Theta$ and $\Theta'_n = \Theta_n$
9: $\quad$T$_{start}$=T; Get State S_t
10: $\quad$**While** (T$_{start} \leq$ T$_{Max}$ || Final State) **do**
11: $\quad$Take action A$_t$ using policy $\pi(A_t|S_t, \Theta')$
12: $\quad$Get reward and next state S_{t+1}
13: $\quad$C=C+1, T=T+1
14: $\quad$**end While**
15: $\quad$**for** (i=T-1; i$\leq$ T$_{start}$; i++)
16: $\quad$R$\leftarrow$ Reward + γ R
17: $\quad$d$\Theta \leftarrow$ d$\Theta + \nabla_{\Theta'} \log\pi(A_i|S_i, \Theta')\big(\text{R-V}(S_i, \Theta'_n)\big)$
18: $\quad$d$\Theta_n \leftarrow$ d$\Theta_n + \delta \dfrac{\big(R-V(S_i, \Theta'_v)\big)^2}{\delta\Theta'_n}$
19: $\quad$**end for**
20: Update Θ', Θ'_n using $\delta\Theta', \delta\Theta'_n$ asynchronously.
21: **end While**

3.2 Federated learning for knowledge fusion

For a local server in the machine, the initial Q-network must be defined when the machine downloads the global model from the knowledge base of the edge server. Such an initial Q-network is capable of reaching the goal and avoiding specific automation/orchestration tasks. Observe in Fig. 4 and Algorithm 3 that the CFRL will reduce the training time for machines and learn their automation/orchestration tasks. In our experimental setting, the knowledge base of the edge server does not automatically fuse the local model every time a machine uploads it but fuses at a defined rate. So in Algorithms 1, 2 and 3, as shown in Fig. 4, we apply the computing flow of the CFRL algorithm for both knowledge function and model transfer learning [31].

It is known that uncertainty is the embodiment of confidence. Therefore, we use information entropy to define confidence in this work (along the line of that of [16]). In particular, the confidence for a machine i (information entropy) is given by [16, Eq. 1]

$$\text{Conf}_i := -\frac{1}{\log n}\sum_{j=1}^{m}\left\{\frac{\text{Score}_{ji}}{\sum\limits_{j=1}^{m}\text{Score}_{ji}}.\log\frac{\text{Score}_{ji}}{\sum\limits_{j=1}^{m}\text{Score}_{ji}}\right\} \quad (2)$$

where n is the size of local networks and m is the size of actions of a machine. Using Eqn. 2, the memory weightage of the machine i and knowledge fusion can be computed as [16, Eqs. 4–6],

$$\text{d}w_i = \frac{1-\text{Conf}_i}{\sum_{i=1}^{n}1-\text{Conf}_i} \quad (3)$$

$$\text{Label}_j = \text{ Score} \times (\text{Conf}_1, \text{Conf}_2, \ldots, \text{Conf}_n)^T \quad (4)$$

$$\Theta^{\star} = \arg\min_{\Theta}\frac{1}{M}\sum_{i=1}^{M}(y_i - h_{\Theta}(x_i))^2 \quad (5)$$

Using the error $(y_i - h_{\Theta}(x_i))^2$ in the training process, see in Algorithms 2 and 3, Eqs. 4 and 5 are the main aim of the learning process.

Algorithm 3 Knowledge Fusion Algorithm
1: **procedure** INITIALIZE
2: Shared Network Parameters, K:= data samples; N:= private Networks; M:= Action Machines Size
3: **end procedure**
4: **procedure** FUSE KNOWLEDGE
5:　**for** (i=0; i≤ N; i++)
6:　　Calculate indirect features x_i
7:　　**procedure** CALCULATE SCORE
8:　　　**for** (n=1; n≤ K; n++)
9:　　　$Score_{temp} \leftarrow f_n(x_i)$
10:　　　$Score_i \leftarrow$ Append $Score_{temp}$
11:　　**end for**
12:　　**end procedure**
13:　　**procedure** COMPUTE CONFIDENCE
14:　　　**for** (n=1; n≤ K; n++)　　**for** (m=1; m≤ M; m++)
15:　　　Calculate Confidence Values
16:　　　**end for**　　**end for**
17:　　**end procedure**
18:　Calculate and append $Label$
19:　**end for**
20: **end procedure**
21: Train shared network using $(x, Label)$

3.3 Transfer learning for sharing knowledge

Different methods have been proposed to transfer reinforcement learning in the literature. We found that there are two approaches applied in the specific task when a machine learns to automate those tasks. One promising approach is taking the global model as a starting actor in the network, which attains a good score initially but is quite unstable. Another system would be using the global model as a feature extractor in transfer learning. The latter method increases the dimension of the features, and it can improve the effect stably. One problem with the latter approach is to solve in the experiment: the structural difference between the input layer of the global network and local network.

With an underlying Markovian decision cycle consisting of action space, device space, transition matrix, we follow the former approach and consider the DRL problem in factory automation/orchestration. After taking action, a reward is received at any state for the given policy and discount factor. The Q-function gives appropriate action for any given policy π,

$$Q^{\pi}(s,a) = \mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t R(s_t, a_t)\Big| s_0 = s, a_0 = a, \pi\right] \quad (6)$$

where $R(s, a)$ is the reward perceived at state $s \in \mathcal{S}$ when action $a \in \mathcal{A}$ is taken, and $0 < \gamma < 1$ is the discount factor. Assuming that the reward satisfies $0 < R(s,a) < 1$, the optimal policy $\pi^{\star}$ has the Q-function $Q^{\star}(s,a)$

$$Q^{\star}(s,a) = R(s,a) + \gamma \mathop{\mathbb{E}}_{\substack{s' \sim T(s'|s,a) \\ a' \sim \pi(a|s)}}\left[Q^{\star}(s',a')|s_t = s\right] \quad (7)$$

where T is the transition matrix $T = \{T_{s,s'}^a; s \in \mathcal{S}, a \in \mathcal{A}\}$ and $T_{s,s'}^a$ is the transition probability from s to s'. Eq. (7) satisfies optimized conditions of the Bellman equation.

It is worth noting that it is a contraction mapping for the Q-function above and has a special fixed-point structure for optimum Q-function. We use the *max-square error metric*, $M_{Sq}(Q)$, which quantifies the excellence of Q-learning as follows:

$$M_{Sq}(Q) := \max_{s,a}\left(Q^{\star}(s,a) - Q(s,a)\right)^2 \text{ where}$$

$$Q(s,a) = R(s,a) + \gamma \mathop{\mathbb{E}}_{\substack{s' \sim T(s'|s,a) \\ a' \sim \pi(a|s)}}\left[Q(s',a')|s_t = s\right] \quad (8)$$

We have the following intuition underneath Algorithm 4. Whenever a new machine trained offline is installed into a new IIoT network under similar conditions, their optimal

Q-function and evolution would always be identical. It is beneficial to use the transferred target, a better target, therefore consistently helpful to accelerate the rate of convergence in the overall learning process.

Algorithm 4 Transfer Learning for Sharing Knowledge

1: **procedure** BEGIN
2:   Initialize Q-learning Q_n^i, known $Q_o^\star$, Target Q_t, $i = 1, \dots N$ iterations
3: **end procedure**
4: **procedure** TRANSFER KNOWLEDGE
5:   **for** (i=0; i≤ N; i++)
6:     **procedure** CONDITION-CHECK
7:       **if** $\left(M_{Sq}(Q_o^\star) \leq M_{Sq}(Q_n^i)\right)$
8:         $Q_{temp}^n = Q_o^\star$
9:       **else** $Q_{temp}^n = Q_t$
10:       **end if**
11:     **end procedure**
12:     **procedure** Q-LEARNING
13:       **for** $s \in \mathcal{S}, a \in \mathcal{A}$
14:         $Q_c^n(s', a') = \frac{1}{N}\left(R(s,a) + \gamma \max_{a'} Q_{temp}(s', a')\right)$
15:         $Q_n^{i+1}(s,a) = (1 - \frac{1}{N})Q_n^n(s,a) + Q_c^n(s', a')$
16:       **end for**
17:     **end procedure**
18:   **end for**
19: **end procedure**

---

4 Performance evaluation

In this section, we plan to address the following three critical questions.

- How can our SC framework reduces training time substantially despite losing automation/orchestration accuracy and effectiveness in 6G-enabled edge-computing systems?
- How successful is the knowledge fusion algorithm in stimulating the global model and goal-oriented evolution?
- How successful is the transfer learning strategies for SC in exploiting typical pattern to a particular automation/orchestration task?

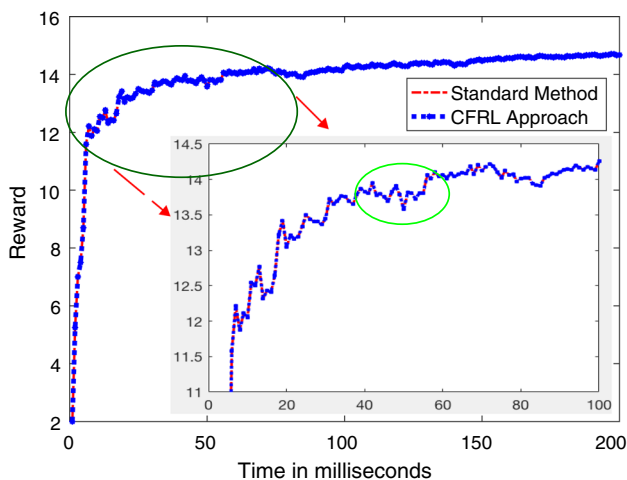To answer the first question, we are conducting experiments to compare the performance of the standard

Fig. 5 With the training approach of environment 1, CFRL has an exactly similar result as that of the standard method. This is because there is no sharing of the learning models from the old machines to new machines for the anticipated industry automation/orchestration
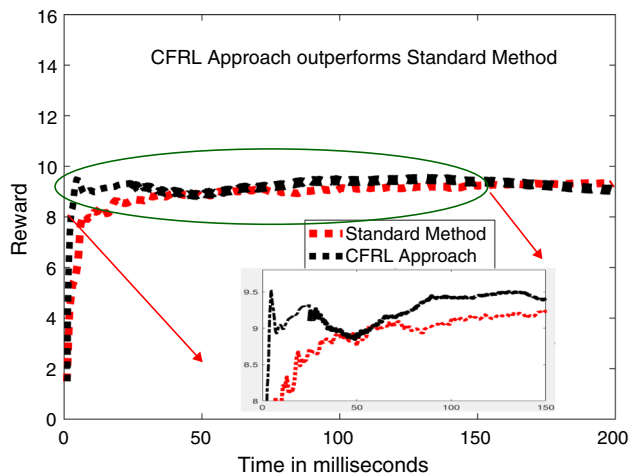

Fig. 6 In the training procedure of Environment-2, CFRL (black curve) demonstrated the global model sharing. CFRL obtains better rewards in less time when compared with the standard approach (red curve)
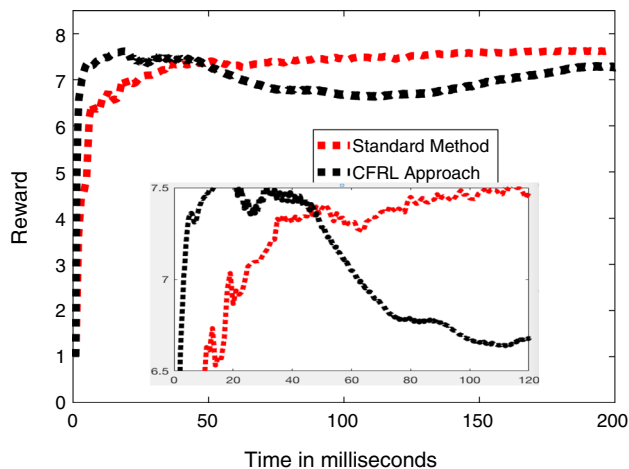

Fig. 7 In Environment 3, CFRL evolves the global model demonstrated excellent results (black curve). Observe that CFRL obtains better rewards in less time when compared with the standard approach (compare red and black curves). However, CFRL is not performing as well as the standard method when learning time is longer than 50 milliseconds

approach with that of the CFRL framework. For the second question, we are performing experiments to compare the performance of generic models and the global model of transfer reinforcement learning.

4.1 Simulation test-bed setup

Four different training settings are considered to illustrate other consequences among the comprehensive approach of the training and the CFRL framework, considering automation/orchestration in the navigation of a moving machine or a robot. There are no obstacles except the walls
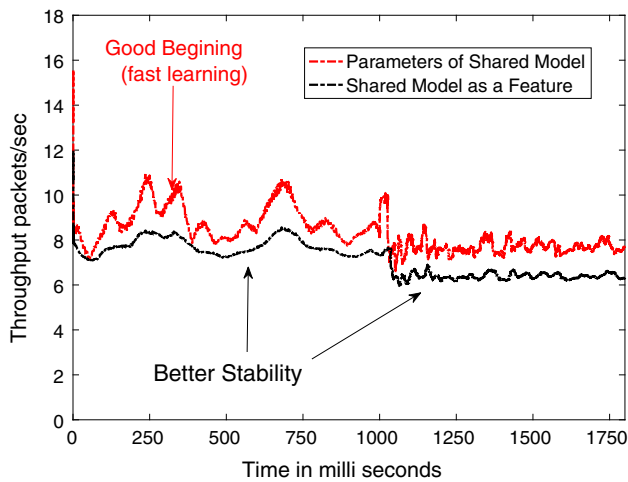
Fig. 8 Comparison of the two transfer learning approaches

Table 2 Results of the extensive experiments to demonstrate the efficiency of knowledge fusion in the CFRL approach

Models	Training time	Average score	Average score of last five
Model 1	1 hr 32 mins	1370	1540
	4 hrs	48	348
	7 hrs	−960	−936
	33 min	2675	3500
Model 2	3.5 hrs	65	790
	6.5 hrs	230	140
	37 mins	2900	3120
Model 3	3 hrs	178	−250
	5hrs	322	1999

in Environment 1. There are four fixed poles as obstacles in Environment 2 and four moving poles as obstacles in Environment 3. In particular, there are both dynamic and static barriers in Environment 4. The observation area is between 0.1 m and 5 m.

4.2 Evaluation of CFRL framework

We have carried out three environments, and the output of CFRL is in Figs. 5, 6, and 7 in accordance with the performance of generic methods. In Environment 2 and Environment 3, CFRL increased the automation/orchestration decision accuracy and reduced training time for the 6G automated edge-intelligent system. The improvement can be observed in Figs. 6 and 7, which is highly efficient compared with the global shared model.

CFRL is very effective in learning policy for all SC constraints and industry automation/orchestration considered. It enhances the ability of our learned model to generalize across commonly faced environments and situations. Experimental results show that CFRL is able to minimize training time without compromising accuracy loss in the automation/orchestration process for the proposed SC-based intelligent industrial network computing.

4.3 Evaluation of knowledge fusion approach

Table 2, we present comparative and quantitative results of our knowledge fusion approach for the SC framework. Observe in Table 2 that the federated learning approach with the shared model reduces training time on a continual learning basis. In fact, it can be found that standard process models are only capable of making excellent decisions in specific settings. In contrast, the CFRL-based model with knowledge infusion is capable of making the best decisions

in a variety of different settings. Overall, the presented information fusion method is therefore successful and highly efficient for the SC framework.

4.4 Evaluation of transfer learning algorithm

We performed a quantitative analysis to validate and equate the two transfer learning methods. The result appears in Fig. 8. From the figure, it can be observed that all strategies to transfer learning will successfully increase the performance of reinforcement learning and hence boosts the performance of the proposed SC framework. The parameter transfer approach has faster learning speed (red curve), and the feature extractor approach has more excellent evolution stability (black curve).

5 Conclusion

We developed a novel edge-intelligence-based semantic communication framework for 6G. It is based on federated-learning-based 6G networking for the automation/orchestration systems of industry 4.0 networks. The proposed intelligent networking approach can effectively use prior knowledge and adapt rapidly to a changing environment and system semantics. We also presented a knowledge fusion algorithm in the framework and provided insights into transfer-based learning. Our approach allows models to fuse as well as share the model to evolve asynchronously using knowledge-base. Using knowledge based on policy tests, we have evaluated our design and the framework extensively. It is worth noting that our method has constrained requirements and limitations for a full-fledged SC of machines/robots. The extension of the developed framework to handle the intended dimensions of SC flexibly requires further investigations and is left for future work. Nevertheless, the developed edge-intelligent 6G framework is an initial innovation to offer a broader range

of industry 4.0 services, and further advancements will be remarkable.

Declarations

Conflicts of interest The authors declare no conflict of interest.

References

1. Huda MN, Hirley A , Alcaraz LO, Mohammad S, Moya ODP, Matti L-A (2019) Six key enablers for machine type communication in 6G." arXiv preprint arXiv:1903.05406
2. Ella P , et al (2020) 6G White paper on edge intelligence arXiv:2004.14850.2020
3. Elif U, Onur K, Anthony E, James G, Marian C, Petar P, Mohamad A et al (2021) Semantic communications in networked systems. arXiv preprint arXiv:2103.05391
4. Kountouris M, Pappas N (2021) Semantics-empowered communication for networked intelligent systems. IEEE Commun Mag 59(6):96–102
5. Emilio CS , Barbarossa S (2021) 6G networks: Beyond Shannon towards semantic and goal-oriented communications. Comput Netw 190:107930
6. Shannon CE, Weaver W (1949) The mathematical theory of communication. The University of Illinois Press, Champaign
7. Samad A et al. (2020) 6G White paper on machine learning in wireless communication networks. arXiv preprint arXiv:2004.13875
8. McMahan HB, Moore E, Daniel R, Seth H (2016) Communication-efficient learning of deep networks from decentralized data. arXiv preprint arXiv:1602.05629
9. Pokhrel SR, Choi J (2020) Federated learning with blockchain for autonomous vehicles: analysis and design challenges. IEEE Trans Commun 68(8):4734–4746
10. Volodymyr M et al (2016) Asynchronous methods for deep reinforcement learning. In: International conference on machine learning, pp. 1928–1937
11. Tariq F et al. (2019) A speculative study on 6G. IEEE Wireless Commun. 27(4):118–25. arXiv preprint arXiv:1902.06700 (2019)
12. Pokhrel SR, Garg S (2020) Multipath communication with deep Q-network for industry 4.0 automation and orchestration. IEEE Trans Ind Inf. https://doi.org/10.1109/TII.2020.3000502
13. Pokhrel SR, Singh S (2020) Compound-TCP performance for industry 4.0 WiFi: a cognitive federated learning approach. IEEE Trans Ind Inf. https://doi.org/10.1109/TII.2020.2985033
14. Changyang S, Rui D, Zhouyou G, Zhanwei H, Yonghui L, Wibowo H, Chenyang Y, Lingyang S, Branka V (2020) Deep Learning for ultra-reliable and low-latency communications in 6G Networks. arXiv preprint arXiv:2002.11045
15. Xiaomin Li, Jiafu Wan, Hong-Ning Dai, Muhammad Imran, Min Xia, Antonio Celesti (2019) A hybrid computing solution and resource scheduling strategy for edge computing in smart manufacturing. IEEE Trans Ind Inf 15(7):4225–4234
16. Boyi Liu et al (2019) Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems. IEEE Robot Autom Lett 4(4):4555–4562
17. Pokhrel SR et al (2021) Multipath TCP meets transfer learning: a novel edge-based learning for industrial IoT. IEEE Internet Things J 8(13):10299–307
18. Zhou Z, Chen X, Li E, Zeng L, Luo K, Zhang J (2019) Edge intelligence: paving the last mile of artificial intelligence with edge computing. In: Proceedings of the IEEE, vol. 107(8): pp. 1738–1762
19. Wang X, Han Y, Leung VC, Niyato D, Yan X, Chen X (2020) Convergence of edge computing and deep learning: a comprehensive survey. IEEE Commun Surv & Tut 22(2):869–904
20. Xu D, Li T, Li Y, Su X, Tarkoma S, Jiang T, Crowcroft J, Hui P (2020) Edge intelligence: architectures, challenges, and applications. arXiv e-prints (2020): arXiv-2020
21. Yang Q, Liu Y, Chen T, Tong Y (2019) Federated machine learning: concept and applications. ACM Trans Intell Syst Technol (TIST) 10(2):12
22. Wang J, Cao B , Yu P, Sun L, Bao W, Zhu X (2018) Deep-learning towards mobile applications. In: 2018 IEEE 38th International conference on distributed computing systems (ICDCS). IEEE, pp. 1385–1393
23. Mohammadi M, Al-Fuqaha A, Sorour S, Guizani M (2018) Deep-learning for IoT big data and streaming analytics: a survey. IEEE Commun Surv & Tut 20(4):2923–2960
24. Patras ZP, Haddadi H (2019) Deep learning in mobile and wireless networking: a survey. IEEE Commun Surv & Tutor 21(3):2224–87
25. Goldreich O, Juba B, Sudan M (2012) A theory of goal-oriented communication. J ACM (JACM) 59(2):1–65
26. Brendan J, Madhu S (2011) Efficient semantic communication via compatible beliefs. In: Proceedings of the second symposium on innovations in computer science-ICS, pp. 7-9. Institute for Computer Science, Tsinghua University
27. Brendan J, Madhu S (2008) Universal semantic communication I. In: Proceedings of the fortieth annual ACM symposium on theory of computing, pp. 123-132
28. Liu Y, Yuan X, Xiong Z, Kang J, Wang X, Niyato D (2020) Federated learning for 6G communications: challenges, methods, and future directions. arXiv preprint arXiv:2006.02931
29. Li J, Huang R, He G, Wang S, Li G, Li W (2020) A deep adversarial transfer learning network for machinery emerging fault detection. IEEE Sens J 20(15):8413–22
30. Moiseeva A, Trautmann D, Schütze H (2020) Multipurpose intelligent process automation via conversational assistant. arXiv preprint arXiv:2001.02284
31. Diez-Olivan A, Del Ser J, Diego G, Basilio S (2019) Data fusion and machine learning for industrial prognosis: trends and perspectives towards industry 4.0. Inf Fusion 50:92–111