ACCEPTED FROM OPEN CALL

# AI-Enabled Next-Generation Communication Networks: Intelligent Agent and AI Router

Chunxiao Jiang, Ning Ge, and Linling Kuang

## Abstract

Nowadays, network users' demands for low latency services and individual data security grow rapidly. In this article, we conceive a new communication network architecture based on the concept of intelligent agents. The proposed intelligent agent serves as a virtual secretary of the corresponding user to request desired service from Internet companies or neighboring agents, which can reduce Internet surfing latency and prevent privacy leakage. Moreover, to improve the forwarding efficiency of data packages, a new mixed routing structure and the AI router are also introduced, to support various on-demand services with low latency. Overall, the proposed communication network architecture is expected to be more efficient than the existing one, and inspire readers to deliberate the revolution of future communications networks.

## Introduction

With the development of mobile terminals and Internet of Things (IoT), massive data imposes new challenges to current communication networks, especially for the security of individual data and efficient services with low latency. First, privacy leakage is a prevailing phenomenon that commonly happens in current electronics commerce [1]. Many people may have encountered such an experience that you browse some commodities using an App on your mobile phone, but another App recommends similar commodities soon. The fundamental reason for the aforementioned situation is because some Internet companies collude with each other and share users' individual private data to earn more profits [2]. Second, a mass of smart mobile terminals frequently visit the remote application clouds, which imposes a heavy burden to the current access networks and inevitably increases service latency. Meanwhile, the router tables in current networks grow exponentially with the increasing number of deployed routers. Naturally, the overlarge scale of router tables deceases the efficiency of routing path selection, as well as bring forth relatively high forwarding latency of data packages, making it difficult to satisfy the quality of service (QoS) of the clients.

Current communication networks consist of the Internet content provider (ICP), the Internet service provider (ISP) and the client [3]. The ICP represents the companies that provide various applications for clients, such as WeChat, Taobao, Amazon, Facebook, etc., while the ISP refers to the telecommunication operators, such as China Mobile Communications Corporation, American Telephone & Telegraph, WorldNet, and and so on. The client refers to the users who request desired service such as E-mail, FTP, information retrieval, and etc., with the aid of the ISP and ICP. On top of that, the current network routing structure, that is, the network control plane, is in a distributed manner, where each router collects local routing information and selects the relatively appropriate path to forward data packages.

Generally, there are two shortcomings in the current network. First, the status of the ISP and the ICP are relatively unbalanced. The ISP is only in charge of data transmission between clients and ICPs, such that ICPs can collect lots of individual data when clients are enjoying the service provided by ICPs. In order to obtain the desired service, users have to surrender their individual data to the ICP, which inevitably leads to the situation that users loss access and control rights of their individual data. Meanwhile, the service access latency in the network inevitably becomes the bottleneck for improving the QoS of users, especially with a large number of active clients. Second, the network control layer in the current network architecture is also worthy of improvement. The current distributed routing strategy cannot adapt to the requirement of the service packets when processing the forwarding operation at routers, such that various service demands are difficult to be satisfied. Although the Software-Defined Network (SDN) [4] has built a centralized control plane, the complexity issue has bothered it a lot when the network scale explodes.

Considering the aforementioned issues, in this article, we introduce a new network architecture based on intelligent agents and artificial intelligence (AI) routers, to improve the security of clients' private data and support the on-demand network services with low latency. The rest of the article is organized as follows. We first present the next-generation network architecture in the following section, and then discuss the functions of the intelligent agents conceived in the new network. Following that, a new mixed routing structure and the AI router for the new network architecture are introduced. Finally, we give the conclusion.

Digital Object Identifier:
10.1109/MWC.001.2000100

*The authors are with Tsinghua University. Ning Ge is the corresponding author.*



1536-1284/20/$25.00 © 2020 IEEE

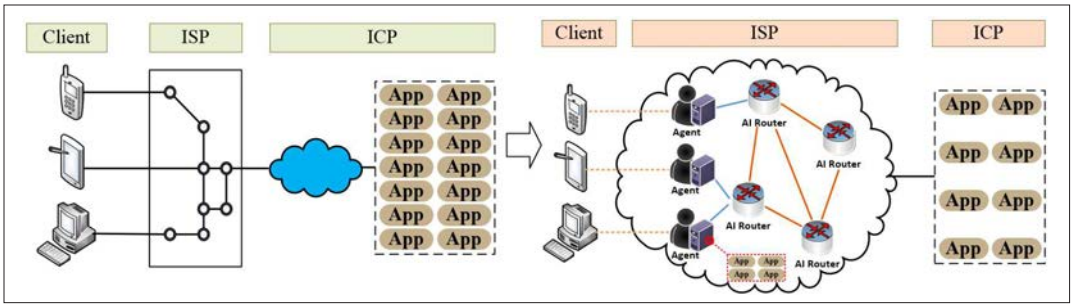


**FIGURE 1**. The next generation communication network architecture.

## Next-Generation Communication Network Architecture

As shown in Fig. 1, the current network and the new network architecture both consist of the ICP, the ISP and the clients. While different from the current network architecture, the roles of these three parts would be changed greatly. In this section, we articulate the new network architecture, and then introduce the transformation of the ICP, as well as the network control policy in detail.

### Description of the New Architecture

As shown in the right part of Fig. 1, in the new network architecture, a series of intelligent agents are deployed either at the edge of the ISP in a distributed manner or at the cloud of the ISP in a centralized manner. The intelligent agent represents the virtual secretary for the corresponding client, which can help the specific user request various service from ICP. Essentially, the intelligent agents would serve as intermediaries between clients and ICPs, as well as among clients, which avoids the direct interaction between clients and ICP and achieve the purpose of protecting clients' private data. On top of that, intelligent agents return the control right of the clients' individual private data to users themselves. It is worth mentioning that wherever the intelligent agent is deployed and whatever the intelligent agent is designed, only the corresponding client has the right of control over it. In addition to the improvement related to data security, the new network can also reduce the service access latency to a large extent, since lightweight applications can be deployed at the edge of the network, such that users can obtain desired service from their agents directly without sending a request to the ICPs, which can improve the QoS of the clients and eases the network congestion to a large extent.

Different from the current network control schemes, the new network structure adopts a new mixed routing structure based on AI routers, with the purpose of reducing service acquirement latency and offering various on-demand services for the clients. On one hand, the distributed AI routers can learn and classify the requirement of various service packets to allocate appropriate network resources based on machine learning [5]. On the other hand, the routers can collect the local network state information and delivery collected information to the network centralized controller timely. By means of machine learning algorithms, the network centralized controller can analyze various network state information, then dispatch long-term control policies to the AI routers to perform.

In addition, to cope with various kinds of malicious network attacks, the dynamic game based security system for intelligent agents is expected to be applied in this new network structure. Based on the dynamic game, the interactions between the attacker and the intelligent agent can be viewed as a two-player dynamic game. Therefore, the game model can be constructed to compute best response strategies for intelligent agents to cope with malicious network attacks, which can enhance the security of this new network structure.

### Transformation of ISP and ICP

With the deployment of intelligent agents, roles of the ISP and the ICP have changed greatly. The former not only serves as data transfer tunnels just like in the current networks, but also stores and manages users' individual data using intelligent agents. In other words, each client controls uniquely corresponding agents to store and manage their individual data. By means of intelligent agents, when requesting desired service, clients only need to inform corresponding agents what service, and then intelligent agents interact with ICPs to obtain users' desired service in an anonymous manner. In addition, for the sake of reducing service requesting latency, the ISP also supports to provide some lightweight applications for clients organized by the corresponding intelligent agents through the mobile edge computing technology (MEC) [6]. Different from the role played in current networks, the ICP only provides the requested service content to corresponding intelligent agents without requiring users' individual private data. It is expected that with the aid of intelligent agents, the capacity of the network itself can be enhanced, while the privilege of ICP in the management of users' individual data can be weakened.

As a matter of fact, it is undeniable that the new network architecture also brings the transformation of the service mode of the client. In the current network architecture, the client has to surrender user's individual information to the ICP, so as to obtain desired service, while in the new network architecture, intelligent agents act as clients' virtual secretary to achieve indirect interaction in an anonymous manner. Basically, the significant transformation of the interactive manner protects privacy leakage in turn.

## Intelligent Agents

As aforementioned, the intelligent agents (Fig. 2) play a vital role in individual data management and various kinds of interactions in the new network architecture. In such a multi-agent net-

In the current network architecture, the client has to surrender user's individual information to the ICP, so as to obtain desired service, while in the new network architecture, intelligent agents act as clients' virtual secretary to achieve indirect interaction in an anonymous manner. Basically, the significant transformation of the interactive manner protects privacy leakage in turn.

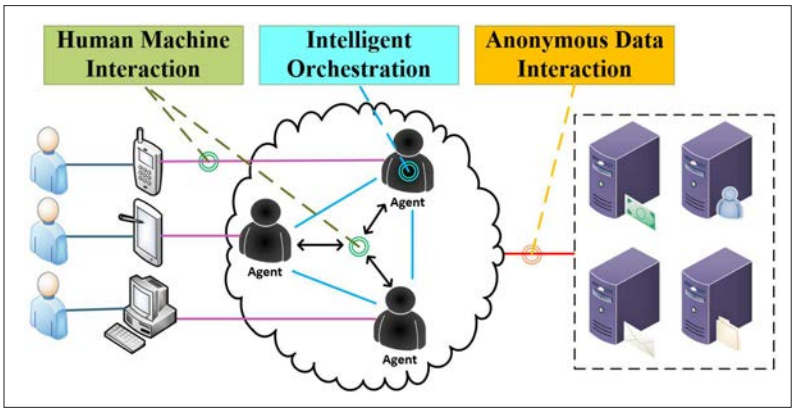


**FIGURE 2**. Intelligent agents.

work system, the intelligent agents are supposed to support a series of functions, such as Human Machine Interaction, Intelligent Orchestration and Anonymous Data Interaction. In this section, we articulate the main functions of intelligent agents in detail.

## Human Machine Interaction

For the sake of providing intelligent services for clients, on one hand, intelligent agents should be able to understand, learn and infer the behavior of clients accurately; on the other hand, to provide more intelligent recommendations for clients, each agent should also exchange the preference information with others to establish the preference database for the corresponding client. In order to realize the aforementioned functions, two problems are supposed to be considered:

- How can intelligent agents understand users' behavior accurately?
- How to obtain and exchange the preference information efficiently?

The former issue mainly refers to the interaction between clients and intelligent agents. Overall, natural language processing (NLP) [7] is one of the most popular technologies in the research area of artificial intelligence in recent years, which can help intelligent agents understand natural language by means of speech recognition, syntactic and semantic analysis based on deep learning. With the aid of the NLP technology, users' behavior information can be more easily processed, so as to help agents provide more intelligent services for users. The latter issue mainly represents the interaction among agents. An efficient approach to analyze users' habits, interest, preferences etc. is the technology of preference information processing [8], which is employed to summarize preference information by analyzing the large amount of user data, and then customize the recommendation database for the specific user. In addition, different from the current collaborative filter based model in preference information processing, deep inference would become a new technical issue to be explored for the sake of making better recommendations.

## Intelligent Orchestration

Based on the intelligent agents, lightweight applications can be deployed at the edge of the network in order to provide desired service for users with less latency. To ensure smooth operation of these applications running in intelligent agents, intelligent agents are suppose to support intelligent orchestration of computing resources and network resources [9]. The former usually refers to CPU resources and memory resources in each lightweight computing unit, while the latter usually represents bandwidth resources and link resources between agents and ICP, as well as among agents.

Specifically, intelligent orchestration is the joint dispatch of computing resources and network resources. On one hand, considering the demand of computing resources for different applications are various, intelligent agents should be able to allocate computing resources for different applications reasonably and intelligently, so as to perform multiple computing tasks in parallel efficiently. On the other hand, some specific network applications also need to invoke necessary information from other places over the network to achieve specific functions, such that intelligent agents are supposed to configure the network resources reasonably according to the demands of different applications with the purpose of reducing service access latency, which in turn can ensure smooth operation of the applications.

## Anonymous Data Interaction

Anonymous data interaction mainly refers to the interaction between agents and ICPs, as well as among agents. Specifically, to request desired service from ICPs and exchange the preference information with other agents without loss of individual private information, intelligent agents are supposed to hide the user's individual data when delivering interactive information [10]. When it comes to anonymous interaction, anonymous access technology would be appropriate in intelligent agents to protect the confidentiality and integrity of users' individual data. Specifically, the double-locked box protocol [11] is a feasible method to achieve anonymous data interaction. It consists of an encrypted out box, an encrypted inner box and a destination address, which enables the exchange of information between agents and ICPs, as well as among agents, without knowing the identity of users. By means of the double-locked box protocol, the identity of the user is never directly to the ICP and the intelligent agent is unaware of the identities of other users, which can protect users' private data efficiently.

Generally, for the sake of preventing individual data from leakage in the anonymous interaction, intelligent agents are also in charge of storing and managing the clients' private data. Considering that the management of clients' private data should be in a more secure way than usual, the technology of distributed data base and advanced encryption technology are appropriate for intelligent agents. The former technology can be used to develop the multi-agent distributed data storage system [12] to separate users' data, and store individual data and access data respectively [13], in order to further prevent the users' privacy from leakage even if intelligent agents are breached, while for the latter, to prevent privacy leakage from internal attacks, with the aid of advanced encryption technology, clients' individual private data cannot be available and disclosed to unauthorized individuals including

network administrators, which further improves the security of individual data.

## Mixed Centralized And Distributed Routing

In addition to the intelligent agents analyzed above, the network control plane is also of importance to consider in new network architecture, so as to further reduce the service latency. In this section, we first discuss the deficiencies of the current routing structure, then introduce a new hybrid routing structure and the AI router.

### Shortcomings of Existing Routing Structures

Generally, the current routing structure can be divided into two types: centralized routing and distributed routing. For the former, as the growth of network scale, the amount of network state information increases exponentially, and the dynamic network structure changes frequently. Those would inevitably lead to the situation that the centralized routing structure cannot collect all the state information from the whole network, such that the network would become malfunctioned because of the belated update of network state information. Different from the former, the latter distributed one only collects local routing information, which can incur the difficulty of the selection of global optimal routing path and global optimal resource configurations.

Recently, intelligent routing is proposed to improve forwarding efficiency of data packages. An important thing is that whatever the router is designed, the intelligent routing is currently relying on the SDN architecture, which only focuses on the optimization of the control plane without considering the improvement of the forwarding plane. Generally, the intelligent routing structure consists of the forwarding plane, the knowledge plane and the intelligent control plane. In a word, the forwarding plane is in charge of forwarding data packages according to the routing forwarding table and instruction sent by the control plane, while the knowledge plane is responsible for collecting various network state information, and providing decision basis for the intelligent control plane. Basically, the intelligent control plane is the core of the intelligent routing, based on the machine learning technology [14], the control plane would synthesize network state information collected by the knowledge plane and delivery forwarding strategies to the forwarding plane. However, constrained by the centralized SDN architecture, centralized intelligence also encountered the complexity issue, especially for the knowledge plane to abstract global network knowledge.

### New Routing Structure and AI Router

Combining the concept of centralized routing, distributed routing and AI, we consider a new mixed centralized and distributed routing structure to improve the forwarding capacity of data packages, so as to reduce data transfer latency. Different from current routers, in the new mixed routing structure, each AI router has the right to process routing state information and adjust the routing strategy according to the network environment. Meanwhile, the entire network has a network brain in the aforementioned intelligent control plane, which is utilized for designing and distributing some learning policies trained by relatively long-term data of network states, in order to facilitate the distributed AI router to approach the global optimality. Generally, the brain of the entire network is equivalent to the "network master brain," and each AI router is equivalent to the "network collaboration brain."

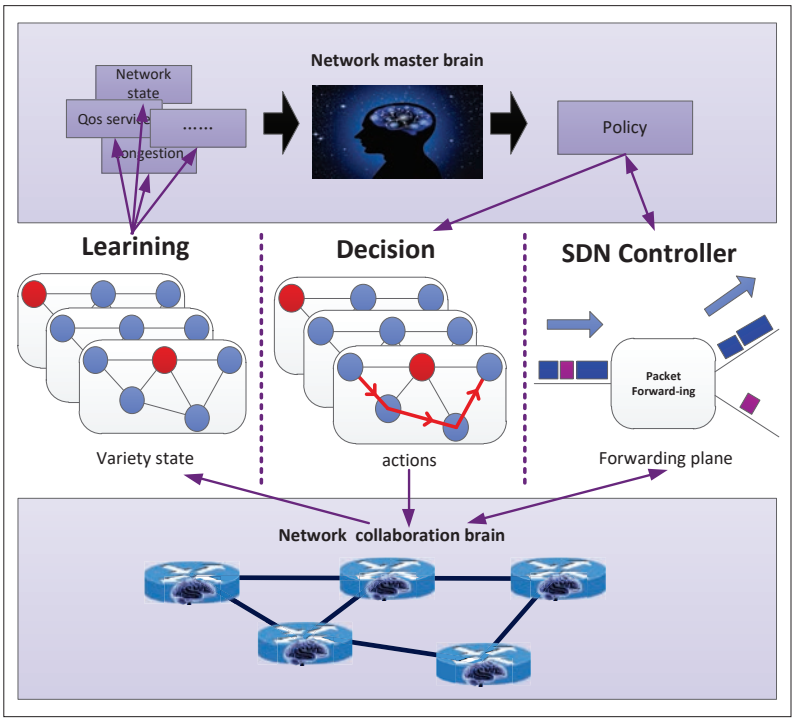


**FIGURE 3.** Time varying graph.

This mixed network routing architecture is shown in Fig. 3, from which we can see that the centralized routing acting as the network master brain only provide global policies to configure the AI router according to the various state information of the network, while each AI router router performs its own decision based on its surrounding environment and the configured policy. As a matter of fact, it is also equivalent to a mixed Q-learning architecture or a multi-agent Q-learning architecture [15]. The AI routers, that is, the network collaboration brain, can distributively collect the fine-grained network state information and adjust the routing policy in a short-term period according to the QoS requirement of the data packets. Meanwhile, through the process of collecting information from the network, the network master brain can obtain coarse-grained global network state information, such as the router state, the QoS of users, the congestion information etc., to adjust long-term routing policies, so as to avoid the network overload caused by the excessive granularity in the dynamic scenarios.

## Conclusion

In this article, we first analyzed the disadvantages of the current network architecture. To solve the issues of efficient service and privacy protection, we introduced a new intelligent agents based next-generation network architecture, which deploys intelligent agents acting as the clients' secretary at the edge of the ISP in a distributed manner to reduce service latency and

Different from current routers, in the new mixed routing structure, each AI router has the right to process routing state information and adjust the routing strategy according to the network environment. Meanwhile, the entire network has a network brain in the intelligent control plane.

prevent privacy leakage. Moveover, for the sake of improving the QoS of clients further, a new mixed routing structure and the AI router for the new network architecture were also presented, which combined the concept of centralized routing, distributed routing and artificial intelligence. Compared with traditional network architectures, our proposed architecture was expected to be better than current network architectures in terms of improving the QoS of clients and protecting their individual private data.

## Acknowledgment

This work was supported by the National Key Research and Development Program No. 2018YFA0701601; and by the National Natural Science Foundation of China under Grant 61922050 and 91638205. The corresponding author is Ning Ge.

## References

[1] L. Xu *et al.*, "Information Security in Big Data: Privacy and Data Mining," *IEEE Access*, vol. 2, no. 1, 2014, pp. 1149–76.
[2] L. Xu *et al.*, " Privacy or Utility in Data Collection? A Contract Theoretic Approach," *IEEE J. Selected Topics in Signal Processing*, vol. 9, no. 7, 2015, pp. 1256–69.
[3] D. Lee *et al.*, " Network Neutrality and Difference in Efficiency Among Internet Application Service Providers: A Meta-Frontier Analysis," *Telecommunications Policy*, vol. 35, no. 8, 2011, pp. 764–72.
[4] A. Gumaste *et al.*, " SDN Use Cases for Service Provider Networks: Part 2," *IEEE Commun. Mag.*, vol. 55, no. 4, 2017, pp. 62–63.
[5] J. Wang *et al.*, "Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks," *IEEE Commun. Surveys & Tutorials*, 2020.
[6] J. Ren *et al.*, "Collaborative Cloud and Edge Computing for Latency Minimization," *IEEE Trans. Vehicular Technology*, vol. 68, no. 5, May 2019, pp. 5031–44.
[7] T. Young *et al.*, "Recent Trends in Deep Learning Based Natural Language Processing," *IEEE Computational Intelligence Mag.*, vol. 13, no. 3, Aug. 2018, pp. 55–75.
[8] J. Fürnkranz *et al.*, *Preference Learning: An Introduction*, Springer Berlin Heidelberg, 2010.
[9] T. Taleb *et al.*, "On Multi-Access Edge Computing: A Survey of the Emerging 5G Network Edge Cloud Architecture and Orchestration," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 3, 2017, pp. 1657–81.
[10] C. Jiang *et al.*, "Information Credibility Modeling in Cooperative Networks: Equilibrium and Mechanism Design," *IEEE JSAC*, vol. 35, no. 2, Feb. 2017, pp. 432–48.
[11] E. S. Peters and N. Maxemchuk, "Privacy in the Cloud: Anonymous Tax Preparation," *Proc. IEEE Int'l. Conf. Intelligent Data and Security (IDS)*, New York, 2016, pp. 177–82.
[12] H. Zhuge, " Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning," *IEEE Trans. Knowledge and Data Engineering*, vol. 21, no. 6, 2009, pp. 785–99.
[13] S. Yu *et al.*, " Achieving Secure, Scalable, and Fine-Grained Data Access Control in Cloud Computing," *IEEE Infocom*, 2010.
[14] C. Jiang *et al.*, "Machine Learning Paradigms for Next-Generation Wireless Networks," *IEEE Wireless Commun.*, vol. 24, no. 2, 2017, pp. 98–105.
[15] J. Wang *et al.*, "Distributed Q-Learning Aided Heterogeneous Network Association for Energy-Efficient IIoT," *IEEE Trans. Industrial Informatics*, vol. 16, no. 4, Apr. 2020, pp. 2756–64.

## Biographies

Chunxiao Jiang (SM) is an associate professor in School of Information Science and Technology, Tsinghua University. He received the B.S. degree in information engineering from Beihang University, Beijing in 2008, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing in 2013, both with the highest honors. His research interests include application of game theory, optimization, and statistical theories to communication, networking, and resource allocation problems, in particular space networks and heterogeneous networks.

Ning Ge (M) received his B.S. degree in 1993, and his Ph.D. in 1997, both from Tsinghua University, China. From 1998 to 2000, he worked on the development of ATM switch fabric ASIC at ADC Telecommunications, Dallas, TX. Since 2000, he has been with the Department of Electronics Engineering at Tsinghua University. His current interests are in the areas of communication ASIC design, wireless communications, and communication networks.

Linling Kuang (M) received the B.S. and M.S. degrees from the National University of Defense Technology, Changsha, China, in 1995 and 1998, respectively, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004, where she has been with the Tsinghua Space Center since 2007. Her research interests include wireless broadband communications, signal processing, and satellite communication.