Autonomous Network Provisioning for Digital Transformation Era

Intent oriented service provisioning assisted by Machine Learning

Taro Ogawa
IoT & Cloud Services Business Division
Hitachi, Ltd.
Tokyo, Japan
e-mail: taro.ogawa.tg@hitachi.com

Tomokazu Makino
IoT & Cloud Services Business Division
Hitachi, Ltd.
Tokyo, Japan
e-mail: tomokazu.makino.ax@hitachi.com

Kenji Arai
IoT & Cloud Services Business Division
Hitachi, Ltd.
Tokyo, Japan
e-mail: kenji.arai.jo@hitachi.com

***Abstract*— Systems engineering, especially in requirements engineering, has become a more complex process because of the diversification in both application and infrastructure aspects. The needs of application users are diversifying industry by industry and are not easy to formulate in a conventional mass-optimization way. Infrastructure technologies are also evolving endlessly. Mitigating these complicated and changing gaps between the requirements and infrastructures should be crucial for the business process optimization in digital transformation. This contribution provides the machine learning assisted transformation of ambiguous user intents to network service specifications conforming to underlying network infrastructures. The proposed system utilizes ensemble learning with several decision-tree based algorithms stacked. The vertical-industry classification process is also implemented as a feature space reduction methodology to exploit each category's knowledge of domain experts. The preliminary evaluation of the prediction performance achieves an accuracy of around 80%.**

Keywords-service provisioning; intent-based networking; machine learning; digital transformation; vertical indutry.

I. Introduction

Digital Transformation as refinement in any business procedure context has become relevant for every industry: Industry 4.0 in manufacturing, autonomous cars in transportation, remote diagnosis, operations in medical care, etc.

Networking as an enabler for these systems advancements can not be a “one-size-fits-all” type solution, however. In other words, just the cloud-smartphone infrastructure is not sufficient for the wide range of requirements spectrum from various vertical industries. Manufacturing industries demand low-latency and high-availability, while some medical care facilities necessitate higher bandwidth for high-definition diagnostic image transfer, for example.

Since such diversification of networking requirements causes the customized networking-capability provisioning for each network service user, accurate comprehension of an individual's network service requirements must be essential. An accurate understanding of network service requirements, however, can not be straightforward because of the multifaceted business situations of users, as well as the diversification of technologies and available services in networking or cloud infrastructure.

Additionally, network service users may not be experts in networking technologies and just aim at their business operation efficiency. Usually, their requirements are expressed as various “intent” levels, namely, business-, service-, or resource-related ones [5]. Moreover, since each user's business situation will be changing dynamically, the intents themselves will also change swiftly.

Network service provisioning should be adapting to these circumstances, preferably in an autonomous manners, where Machine Learning technologies will come in.

Applying machine learning technologies to communication networks has been mainstream in research communities. However, almost all the efforts are focusing on network resource efficiencies [1]. Further, there has been hard to find case studies with communication networks composed of multiple technology- or administrative-domains infrastructures. Only several vision articles mentioned machine learning and artificial intelligence aspects on the network service provisioning, such as vertical industries or service orchestration related themes [2][3][4].

Our main contributions are:

- Presenting the overall procedure of network-service requirements engineering and provisioning for vertical industries.
- Introducing the machine learning architecture for network service requirements engineering, which transforms the user's ambiguous intents into a dedicated network service specification conforming to underlying network infrastructures.

 ISBN: 978-1-61208-838-9

- Initial experimental evaluation of the proposed architecture, combining with various industries' expert knowledge.

The rest of this paper is organized as follows. Section II describes conventional network provisioning procedures and problems related to the current diversified situations of application usage and infrastructure technologies. In Section III, a detailed description of the proposed system architecture with an ensemble learning approach is presented. Section IV provides the evaluation results of the proposed architecture for various vertical industries, including transportation, medical care, and e-commerce. Section V concludes the article with our considerations for future enhancements.

II. NETWORK SERVICE PROVISIONING

Usually, requirements engineering for network service provisioning has been relying on the knowledge of experts of each industry domain. The experienced domain experts interact with each user to derive the network service specifications from the user's intents, mainly business and service level ones. Although some rule-based approaches might be possible in the past business environments, the more advanced and expeditious ways must be necessary for the digital transformation era depicted above.

There will be more diversified network infrastructures, including mobile accesses (4G/5G/6G, WiFi6 and beyond, local- and private-cellular, CBRS, LPWA, etc.), metro accesses, core transport lease lines. There will also be more sophisticated cloud-based networking services, such as SD-WAN and intra-/inter-cloud networking gateway services. Cloud services are also expanding their capabilities, including edge and serverless computing.

Regarding these advancements in both technologies and services, the conventional network service provisioning with human experts might become impossible or inefficient at least. Without accommodating such networking environmental changes, network service users may lose their business opportunities.

However, the necessary procedure for domain experts should remain. Domain experts extract the generic networking requirements from the user's intents expressed by the user, although having said ambiguously. They could use their expert knowledge relating to the domain and the user's business situation. They also classify the requirements into functional and non-functional ones.

The extracted and inferred generic networking requirements are translated into network service specifications alining to the underlying network infrastructures. There might be many choices for selecting the underlying network services in the multiple domains for end-to-end network system configuration. The domain experts also utilize the cost or reliability performance knowledge for such selection.

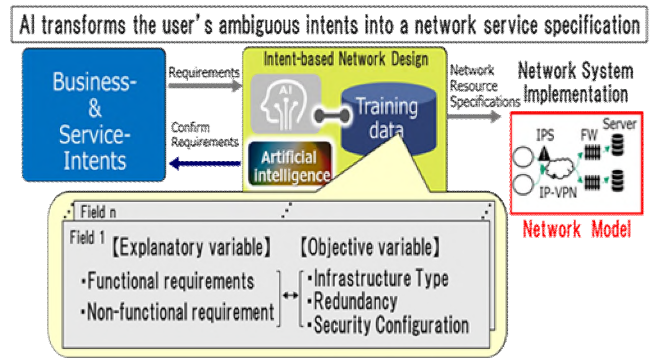

Figure 1. Proposed system workflow.

III. NETWORK REQUIREMENTS ENGINEERING AND MACHINE LEARNING

A. *System Architecture*

The proposed system workflow is depicted in Figure 1. The extractor accepts the corresponding user's intents via a graphical user interface (GUI). It also validates and preprocesses the input data into the appropriate network service requirements. The preprocessing procedure includes not only conventional operation, such as regularization but also pre-classification of the input feature space based on the industry-specific class structure relevant to the user concerned [6].

The analyzer then classifies the network service requirements with several types of machine learning methodology. It derives the candidates of the network configuration with machine learning performance indicators for each classified model.

B. *Model-based engineering*

The most critical part of the system is how to mitigate the gap between the user's intents and network service requirements sufficient to the configuration of the workable end-to-end network service composed by the various underlying network infrastructures.

The extractor functional capabilities are shown as a model-based engineering GUI. Figure 2 depicts a case of a smart-meter distribution network system.

The GUI provides two types of interface: a system model input part (left) and a topology input part (right). The former part accepts the system model, including the information related to network scalability, social impacts of failure (measured by some metric, such as resulted economic loss and affected population). The latter part obtains the parameters, such as the service category (a relevant vertical industry), types of the end system, and the connecting topology between the end systems. The extractor deduces the generic network service requirements, both functional, such as required bandwidth and non-functional, such as resiliency and availability.

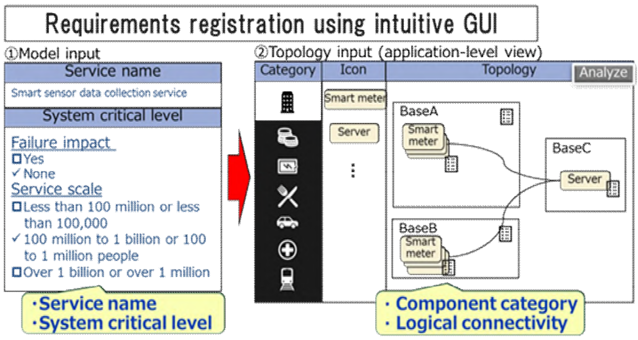

Figure 2. Model-based engineering GUI.

Currently, this extrapolation of network service requirements from the input intents is almost rule-based, with the conversion table for the relevant industry classification, which is expert-knowledge based.

C. *Classifier*

The generic network requirements transformed from the user's intents by the extractor are fed to the analyzer; its functional configuration is shown in Figure 3.

The analyzer predicts the appropriate network configuration for the user concerned utilizing the learned data accumulated through related industry cases' design and operation.

The diversification of business intents and technologies, however, makes the prediction difficult. The classifier optimization might not be straight forward like the usual machine learning process. We adopt a sort of ensemble learning approach and human intervention inside the prediction and selection pipeline to alleviate this issue.

D. *Ensemble Learning and Confidence*

Table I shows the consideration of classifier selection, comparing the techniques, such as Cosine Similarity, Support Vector Machine, and Random Forest, from the viewpoints of computation and accuracy. Considering the current problem space depicted in Figure 2, Random Forest can be concluded as a fundamental approach to the present purpose of classification from the comparison.

Furthermore, several tree-based techniques are stacked to achieve a broader range of application of this approach. Other than Random Forest, we apply Gradient Boosting Decision Tree (GBDT) and Light Gradient Boosting Machine (LightGBM). These classifiers are stacked and independently predict the optimized network configurations using the same learned data. The performance of the predictions is figured up as confidence of each classifier's prediction.

If the particular classifier's confidence exceeds the threshold set, the analyzer recommends the classifier's prediction. In case that none of the prediction confidence passes the threshold, the analyzer requests human interventions, and takes the human decisions as learning data. Even if the prediction exceeds the threshold, the analyzer provides each classifier's confidence to human considerations to enhance the learning process.

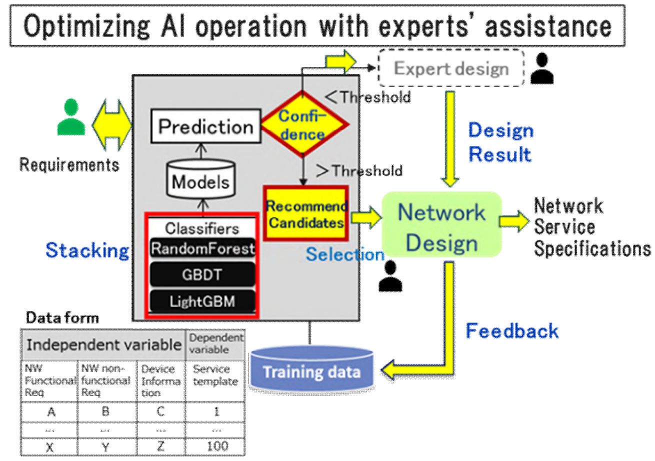

Figure 3. Architecture.

TABLE I. CLASSIFIERS COMPARISON

ML Method	Computational Complexity	Accuracy
Cosine Similarity	Low	Low
Support Vector Machine	High	Medium to high
Radom Forest	Medium	Medium to high
Stacking	Acceptable	High

E. *Networking Components and Configuration Model*

The derived generic network configuration should be translated to network service specifications for underlying infrastructures. Such an arrangement can be modelled by networking components and their setup as a system. Figure 4 depicts an example of the network model as a type of client-server system, which is the prevalent cloud system architecture. Figure 4 also includes the public networking apparatus, such as access and core.

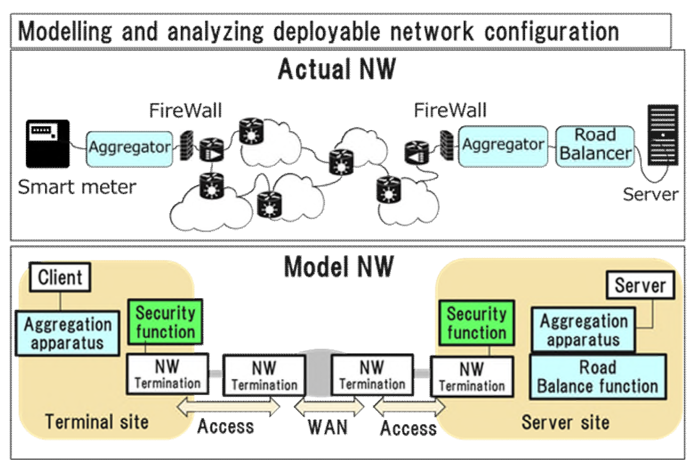

Figure 4. Network model.

 ISBN: 978-1-61208-838-9

The analyzer's output for the translation based on the network model is expressed as a parameter set up on a network template. The network template includes the network components to achieve the networking connectivities between the end systems and servers. An example of the template is depicted in Figure 5, including concentrators aggregating the signals from IoT devices (smart meters), middleboxes for security, load balancers at the server-side.

It also defines the access lines and WANs. Each networking components can have various capabilities depending on the networking service requirements. The capability ranges and examples of the networking components are also shown respectively in Figure 5.

F. Data representations and constraint

Appropriate networking configurations for various industry categories have been generated by each vertical industry's domain experts and validated through machine learning processes. These learned configurations are associated with the network templates as a network menu. An example of the network menu is shown in Figure 6.

The network menu also contains the restrictions for the combination of network components and configuration. For example, a smart factory network configuration requires real-time, low-latency, and availability. Connected cars necessitate edge computing capabilities in addition to that. High-definition image transfer should be essential for remote medical cares. These category-specific requirements can be accomplished through specific network configurations, and no versatile layout should be existing.

Figure 7 depicts the examples of the output of the network configuration for the smart-meter networking case.

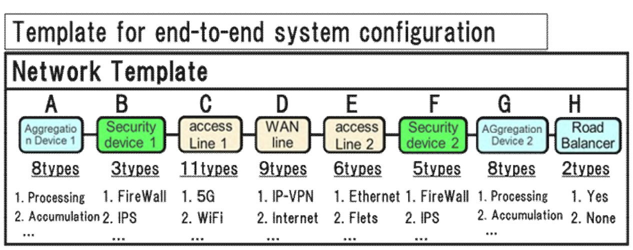

Figure 5. Network template example.

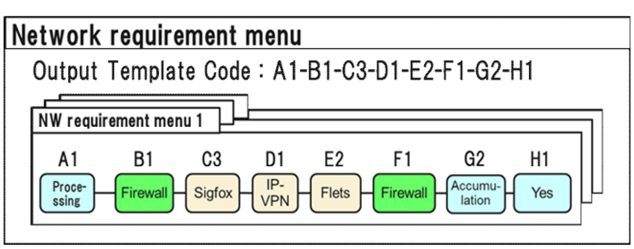

Figure 6. Network menu example.

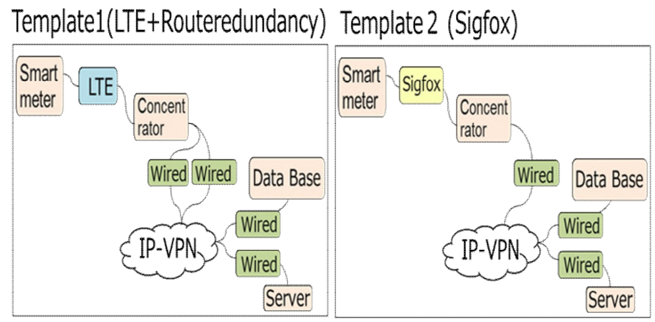

Figure 7. Examples of the outputs.

Two candidates for such a system are selected by the analyzer, one for LTE-based and another using Sigfox. The analyzer recommends the optimized one considering the various possibilities, including the cost related viewpoints.

IV. PERFORMANCE EVALUATION

Even though usual data-shortage problems for this investigation exist, the initial preliminary performance evaluation has been carried out. A base data set is generated to bootstrap the machine learning system through the procedure as follows.

Table II lists several reference sites of the actual use cases incorporated into the evaluation. 52 IoT related system configurations are picked up as basic patterns from the examples in these sites.

Furthermore, a data augmentation approach is deployed to generate more training data from the samples mentioned above. Some parameters, such as the number of end-devices or branches, can be varied in a reasonable range for each use case. It is to be noted that the parameters are not independent of each other but correlated, as depicted in Figure 8. Additionally, the correlation strength should differ parameter by parameters.

According to the industry categories, the generated data is pre-classified, such as transportation, medical care, and e-commerce.

An example result of the process with hold-out validation is shown in Table III. The accuracies of the classification for the categories are about 80%.

TABLE II. USE-CASE REFERENCE

Source	URL
MIC Japan	http://www.soumu.go.jp/main_sosiki/joho_tsusin/top/local_support/ict/index.html
Hitachi Lumada	https://www.hitachi.co.jp/products/it/lumada/usecase/index.html
KDDI	https://iot.kddi.com/cases/

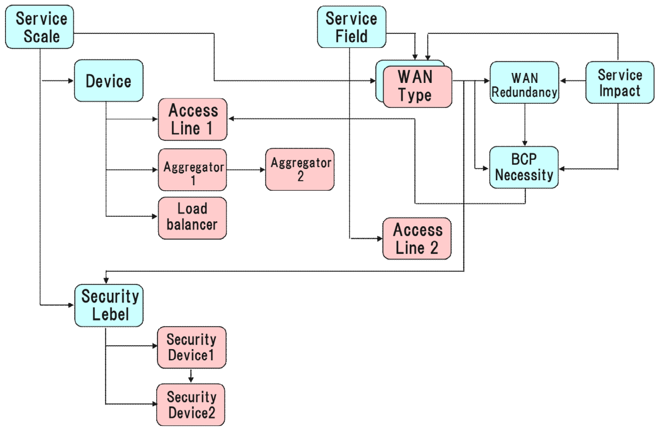

Figure 8. Correlation between the parameters.

TABLE III. EVALUATION RESULTS

Target Industry Category	Accuracy		
	Random Forest	GBDT	LightGBM
Smart city	0.837	0.876	0.855
Construction	0.799	0.858	0.828
Commerce	0.846	0.923	0.928
Manufacturing	0.835	0.877	0.869
Transportation	0.795	0.845	0.831

V. CONCLUSION

This contribution provides the machine learning assisted transformation of ambiguous user intents to network service specifications conforming to underlying network infrastructures. The proposed system utilizes ensemble learning with several decision-tree based algorithms stacked. The vertical-industry classification process is also implemented as a feature space reduction methodology to exploit each category's knowledge of domain experts. The preliminary evaluation of the prediction performance achieves about accuracy of around 80%.

Systems engineering in the digital transformation era, however, may have an intrinsic difficulty of ever-changing conditions, which causes situations of data shortage for the usual statistical machine learning approach. The proposed architecture utilizes not just conventional machine learning techniques but also domain-expert knowledge and other approaches like data augmentation and simulation.

Although the evaluated system is a preliminary one to bootstrap the proposed architecture, the approach should be essential in the diversified and accelerated digital transformation era.

Expanding the applicable industrial categories is a direction of future enhancements. Some hierarchical structure can exist composed of common features to every industry, specific characteristic to each sector, and individually segmented distinction of each user. Network model enrichment is also possible enhancements.

Finally, some standard organizations are starting activities related to service provisioning with machine intelligence [7][8]. The architecture and results presented here should contribute to the advancement in these standardizations.

ACKNOWLEDGEMENT

This work was conducted as a part of the project titled "Research and development for innovative AI network integrated infrastructure technologies (JPMI00316)", supported by the Ministry of Internal Affairs and Communications, Japan.

REFERENCES

[1] M. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine learning for 5G/B5G mobile and wireless communications: potential, limitations, and future directions, " IEEE Access, vol. 7, pp. 137184-137206, 2019

[2] N. Kato, B. Mao, F. Tang, Y. Kawamoto, and J. Liu, "Ten challenges in advancing machine learning technologies toward 6G," IEEE Wireless Communications, vol. 27, No. 3, pp. 96–103, June 2020.

[3] X. Li et al., "Automating Vertical Services Deployments over the 5GT Platform," IEEE Communications Magazine, vol. 58, No. 7, pp. 44–50, July 2020.

[4] B. Spinnewyn, S. Latré, and J. F. Botero, "Service Orchestration in NFV-Based Traditional and Emerging Cloud Environments: State of the Art and Research Challenges," IEEE Communications Magazine, vol. 58, No.8, pp. 76–81, August 2020.

[5] TM Forum White Paper, "Autonomous Networks: Empowering digital transformation for smart societies and industries," Release 2, October 2020.

[6] Information-technology Promotion Agency Japan (IPA), "Non-Functional Requirements Grades Usage Guide [Description Manual]," April 2010.

[7] ITU-T Draft Recommendation Y.3178, "Functional framework of AI-based network service provisioning in future networks including IMT-2020," March 2021.

[8] TM Forum Insight, "Intent oriented customer engagement: understanding customer purpose," July 2020.

 ISBN: 978-1-61208-838-9