A Test System for Vehicular Voice Cloud Service

Li Yu, Kailiang Zhang, Jiang Man, Hao Yu, Yuqing Yao, and Lei Chen(✉)

Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology, Xuzhou 221018, China
chenlei@xzit.edu.cn

Abstract. Some test systems for voice cloud services have been developed in recent years. However, the automobile manufacturers, communications equipment merchants and network operator still lacks methods and tools to evaluate the vehicular voice cloud services from the perspective of the end user experience. Considering the user behavior and user experience, a light weight vehicular voice cloud evaluation system is designed in this paper. The system is able to send voice information to voice cloud server according to user habit, and record the user experience indicators, such as accurate, voice quality, service delay, server computation capacity, and so on. The study shows that the vehicle voice cloud evaluation system can avoid complex communication and language processing, evaluate the performance of the service from view of end user.

Keywords: Voice cloud · Vehicular unit · Performance evaluation · Quality of experience

1 Introduction

Some new approaches are proposed to predict the network traffic in the end-to-end network [1–3] and analyze the end user experience [4]. This will help to construct the test case and evaluate the user experience. On the premise of guaranteeing user experience, new routing schemes are proposed to increase the energy efficiency [5, 6]. However, it is difficult to measure the performance of these approaches for the emerging service of vehicle voice cloud because of the fast changing network topology and complex communication process. And the exits simulation tools and evaluation methods cannot evaluate the key quality indicators of vehicle voice cloud service.

Since 2000, BMW and Acura, the Apple Corp and Ferrari, Mercedes Benz and Volvo jointly developed the vehicular voice assistant system. BAIDU and IFLYTEK, have published their vehicular speech service based on cloud platform.

Vehicular voice cloud service, as an emerging multimedia service, has attracted more and more users. The vehicle environment is characterized by frequent handover,

H. Song et al. (Eds.): SIMUtools 2019, LNICST 295, pp. 346–352, 2019.
https://doi.org/10.1007/978-3-030-32216-8_33

network topology change, background noise and Doppler Effect, and the user behavior and the traffic flow also greatly affect the user experience [7]. This is difficult during the test. In past decades, many traditional speech quality assessment models have been proposed to evaluate the speech quality in limited communication bandwidth [8]. Because most of voice cloud service transmit voice data through TCP protocol, the low transmit rate only leads to high latency. The high noise affects the recognition precision. By adopting these models, some new approaches are proposed to evaluate the speech quality in high noise environment [9, 10]. However, the evaluation score can not reflect the recognition precision of machine.

From the mentioned above, it can be seen that the vehicular voice cloud will become a common configuration of the car. However, there is no mature testing tool in the market for the user experience evaluation of vehicular voice cloud service.

A test approach, which can be deployed in a car, is proposed in this paper to test the user experience in real scenario. By adopting typical user behavior, our proposed approach is also can simulate the complicated scenarios in the lab environment.

2 Design of the Evaluation System

The impacts of protocol and jitter have been evaluated in laboratory environment [11]. However, there are much more factors that can affect the end user experience. Therefore, the test system should be able to evaluate the appreciable indicators in real vehicular environment or in complicated simulated environment.

According to the general architecture of voice cloud service, we proposed a system structure for simulating the user behavior and evaluate the performance indicators of user experience.

In the real vehicular environment, the terminal sends selected typical pre-recorded voice data to cloud. The controller is responsible for issuing test instructions and test scripts to multiple service initiation units and simulation servers. The packets are captured in terminal. The controller indentifies these packets and computes appreciable indicators of the end user.

2.1 Recognition Precision Test

For evaluating recognition precision, the terminal calls the API of voice cloud service. First, the test system imports test case of voice. The vehicular terminal preprocesses the voice and sends voice data to cloud. Then the returned data is compared with expected text. This processing is shown in Fig. 1.

2.2 Latency Test

All the network packets between terminal and cloud are captured and saved as PCAP file by the test software. The evaluation program opens PCAP file and takes key fields. Through checking the key fields, the first request packet and the last response packet are indentified. This processing is shown in Fig. 2. The latency can be obtained by the timestamps of these packets. However, because different voice cloud services have different protocol, it is impossible to build a common packet inspection program for all voice cloud services. An approach is to build three-level packet inspection templates [12]. The algorithm process is shown in Table 1.

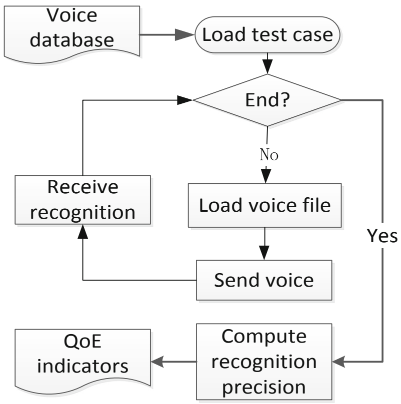

Fig. 1. Diagram of recognition precision evaluation.

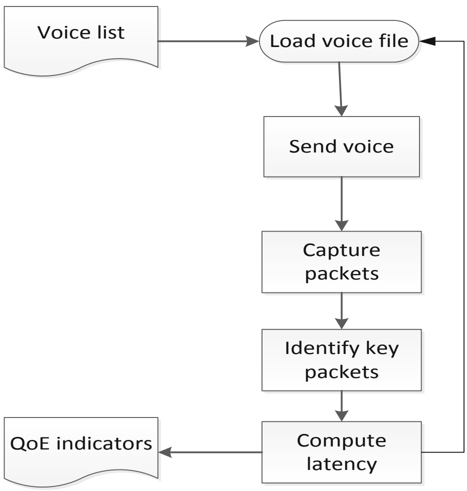

Fig. 2. Diagram of latency evaluation.

Table 1. Latency computation

step	Algorithm
1	Capture network data, get pcap_file;
2	Split the pcap_file into frames;
3	WHILE the set of frames is not empty
4	IF (find DNS datagram) and (find target domain name)
5	Record time_stamp;
6	Flag=open;
7	ELSEIF (find a follow TCP server FIN) and (flag ==open)
8	Compute latency;
9	Flag=close;
10	ENDIF
11	ENDIF
12	ENDWHILE

3 System Test

The authors test the proposed system in vehicular environment. By analyzing the usage ratio of common sentences, six common sentences in vehicular environment are selected to evaluate the recognition precision and latency of voice cloud service. The most common requests in vehicle voice cloud service are shown in Table 2.

Table 2. Statement usage ratio

Test case	Weather	Scenic spot	Gas	Music	Hotel	Introduce	Other
Proportion	29.1%	20.2%	10.4%	23.3%	13.55%	7.41%	9.27%

The sentences include request for forecast, navigation to scenic spot, navigation to gas station, navigation to cafe, music and introduction.

The authors select 10 road spots with various level of background noise to send the voice recognition request. From the first spot to the last one, the background noise of each spot rose by 5%–10%. The precision of each group is shown in Table 3.

Table 3. Recognition precision of test cases under each background noise.

Test case	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Recognition precision	100%	80%	60%	60%	60%	40%	60%	10%	0%	0%

The system automatically transmits the voice to cloud and receive the recognition results. The recognition precision is showed in Fig. 3. The background noise gradually increases from spot 1 to spot 10. As the noise increasing, the voice recognition precision experiences an obvious decline.

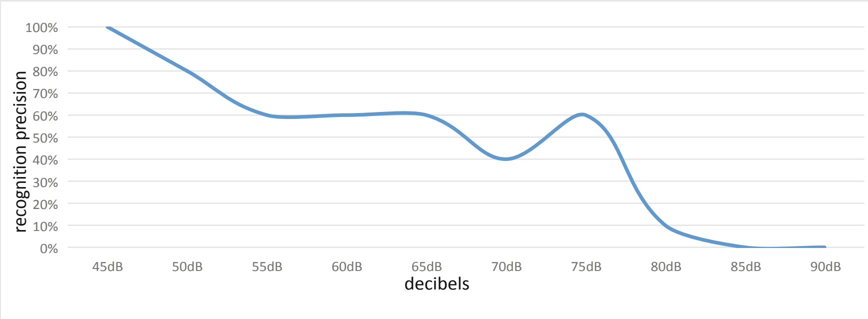

Fig. 3. Recognition precision with various level of noise.

The authors select 7 sites which near restaurant, cyber cafe, street, residential area, urban area, park, and shopping mall to evaluate the latency of voice cloud service. These locations have obvious differences in signal-to-noise ratio, channel interference and network congestion. The average delay of test cases at each test point is shown in Table 4.

Table 4. Recognition time delay of test cases at each test point.

Test case	1st	2nd	3rd	4th	5th	6th	7th
Weather	0.617	0.787	0.486	0.682	0.763	0.986	0.687
Scenic spot	0.518	0.637	0.556	0.543	0.816	1.124	0.567
Gas station	0.588	0.697	0.781	0.764	0.832	1.564	0.772
Music	0.724	0.651	0.634	0.831	0.912	1.245	0.762
Hotel	0.689	0.812	0.682	0.654	0.811	0.965	0.731
Introduce	0.867	0.821	0.614	0.627	0.849	1.265	0.639

As showed in Fig. 4, the experiment results demonstrate that the scenario strongly affects the latency of voice cloud service, which affects the end user experience.

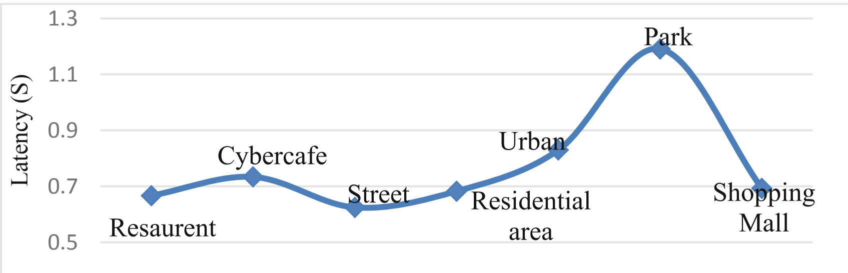

Fig. 4. Latency of voice cloud service.

4 Conclusions

In this paper, the authors analyze the key factors affecting the user experience of vehicular voice cloud service, and propose a light weight test system framework to evaluate the key indicators from view of end user. The results show that the light weight test system can be deployed in the common terminal of commercial network. And an approach to build complicated test scenarios in laboratory environment is also proposed.

For further, it is essential to collect typical user habit and analyzes the distribution characteristics of user behavior. To simulate the road communication link state, we can build the simulation test scenarios in laboratory environment.

Acknowledgements. This work is partly supported by the Natural Science Foundation of Jiangsu Province of China (No. BK20161165), the applied fundamental research Foundation of Xuzhou of China (No. KC17072), Ministry of Housing and Urban-Rural Development Science and Technology Planning Project (2016-R2-060), and the Open Fund of the Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology.

References

1. Jiang, D., Wang, W., Shi, L., Song, H.: A compressive sensing-based approach to end-to-end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. (2018) https://doi.org/10.1109/tnse.2018.2877597
2. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **1**(2), 1–12 (2018)
3. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. PLoS ONE **13**(5), 1–23 (2018)
4. Chen, L., Jiang, D., Bao, R., Xiong, J., Liu, F., Bei, L.: MIMO scheduling effectiveness analysis for bursty data service from view of QoE. Chin. J. Electron. **26**(5), 1079–1085 (2017)
5. Jiang, D., Li, W., Lv, H.: An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. Neurocomputing **220**(2017), 160–169 (2017)
6. Jiang, D., Zhang, P., Lv, Z., Song, H.: Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications. IEEE Internet Things J. **3**(6), 1437–1447 (2018)
7. Vegni, A.M., Loscrí, V.: A Survey on Vehicular Social Networks. IEEE Commun. Surv. Tutorials **17**(4), 2397–2419 (2015). Fourthquarter
8. Dubey, R.K., Kumar, A.: Non-intrusive speech quality assessment using multi-resolution auditory model features for degraded narrowband speech. IET Signal Proc. **9**(9), 638–646 (2015)
9. Zhou, W., He, Q.: Non-intrusive speech quality objective evaluation in high-noise environments. In: 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), pp. 50–54. Chengdu (2015)
10. Islam, M.R., et al.: Non-intrusive objective evaluation of speech quality in noisy condition. In: 2016 9th International Conference on Electrical and Computer Engineering (ICECE), pp. 586–589. Dhaka (2016)

11. Assefi, M., Wittie, M., Knight, A.: Impact of network performance on cloud speech recognition. In: International Conference on Computer Communication and Networks, pp. 1–6. IEEE (2015)
12. Chen, L., et al.: A lightweight end-side user experience data collection system for quality evaluation of multimedia communications. IEEE Access **6**, 15408–15419 (2018)