Research on Improving the Efficiency of English Network Assessment using Computer-Aided Design

Xuebing Zhang[1], Ruolin Shi[2], Na Wang[3], and Rui Chen[4]

[1]Cangzhou Normal University, Cangzhou, Hebei 061000, China, srl891216@163.com
[2]Cangzhou Normal University, Cangzhou, Hebei 061000, China, zlj@caztc.edu.cn
[3]Cangzhou Normal University, Cangzhou, Hebei 061000, China, wangnacangzhou@163.com
[4]Cangzhou Normal University, Cangzhou, Hebei 061000, China, chenruihebei@126.com

Corresponding author: Ruolin Shi, zlj@caztc.edu.cn

Abstract. At present, most of the examinations in schools and in the society use traditional examination methods, and the examination efficiency is low. In the context of the rapid development of computer technology, this study built an English network test system based on computer-aided technology. Moreover, the Internet-based college English online examination system was developed based on the campus network and combined with the analysis of the requirements of the actual examination of College English, and the design goals of the system were proposed. In addition, combined with the actual needs, this paper built the functional structure of the system, and combined the existing computer technology to achieve system functions. Finally, the experiment was designed to verify the functionality of the system. Studies have shown that each function of the system meets the expected goals and can provide a theoretical reference for future research.

Keywords: computer aided; English; network assessment; automatic assessment; intelligence.
DOI: https://doi.org/10.14733/cadaps.2020.S1.68-77

1 INTRODUCTION

Question support is an important requirement in the design of an English test system, and the English test system is designed for English exams. Through the analysis of the computer-based college English test system (referred to as the four-six-level computer test system), the author finds that the four-six-level computer-based test system supports more comprehensive questions and covers four aspects of listening, speaking, reading and writing, but the total number of questions is too small. On the other hand, the relationship between the examination system and the question type is very close to the analysis and presentation of the test paper [1]. When the number and types of questions supported by the system increase, how to parse the test paper and present it for students to answer is a new problem.

The research and development of computerized examinations abroad is much earlier than domestic ones. Many new evaluation theories such as CAT (computer adaptive test) and IRT (project reflection theory) are first proposed by foreign researchers and applied to the actual test [2]. The American Educational Testing Service Center (ETS) is currently the world's largest private non-profit education test assessment agency and a leader in educational research, and the TOEFL, GRE and TOEIC tests are developed by ETS. At present, ETS has more than 50 million machine tests per year [3]. ETS is mainly responsible for the development of the test content, and the implementation of the test is the responsibility of its wholly-owned subsidiary Prometric [4]. Prometric is currently the world's largest computerized certification testing services company. Prometric mainly focuses on the implementation of international manufacturer certification exams, international professional certification exams and some foreign entrance exams (such as SAT) [5]. Prometric's test site system is a general-purpose system built on the core of UTD (Unified Test Driver). UTD adopts the architecture based on COM plug-in, and the plug-in is divided into visual plug-in and non-visual plug-in. The visual plugin is used to implement the user interface, and the non-visual plugin is used to implement logic such as timing, navigation, and judgment [6]. The system connects the system, COM plug-in and content through VBScript script, and realizes the customization of the system by means of scripting. Moreover, UTD also encapsulates IE controls into visual plug-ins for displaying test questions, displaying test introductions, and so on. UTD's exam result data is stored in OLE structured form, and each candidate's exam result has only one file, and the test server provides a shared directory for the client to save the test results during the exam. This way of saving data also leads to a relatively heavy network load, so it is not suitable for subjective and operational questions with relatively large answers. In addition, from the perspective of the organization of the exam, Prometric adopts the appointment-based application mode, that is, the candidates themselves schedule the test sites and test time online. This model reduces operating costs, but it also affects the development of its domestic examination business, because many important exams in China use the unified examination model [7].

At present, the implementation of the test system mainly includes B/C mode and C/S mode. The B/C mode test system is also known as the online test system. The online test system has the advantage of convenient deployment. There is no need to install any software on the test machine, and candidates can log in to the test through the browser. The online test system is also very flexible in the choice of test location and test time. As long as the machine can access the Internet, the test can be completed [8]. The main disadvantage of the online exam system is the existence of security issues. The online exam system relies on a stable and reliable network. Once a network failure occurs during the exam, the exam cannot continue. In addition, the online test system is difficult to solve how to prevent candidates from cheating in the course of the test. Another shortcoming of the online test system is that it is limited by network bandwidth and server computing power. The general system can only support hundreds of people online at the same time, which limits the application of the online test system in large-scale examinations [9]. Therefore, the online test system is generally used for some low-risk tests, such as: recruitment exams, job skills assessment, or should be aware of the test. Large-scale exams and high-risk exams generally choose the c/s mode exam system. Unlike the online test system, the c/s mode test system has no restrictions on accessing local resources and can achieve stronger anti-cheating methods. Clients can take advantage of native resources for better performance, use multimedia technology, run interactive questions, and support operational questions. At the same time, the c/s mode test system can be independent of the network during the examination process and improve the fault tolerance of the system [10]. In addition, the c/s mode test system can adopt a distributed architecture, and the independent operation of each test site does not affect each other, and the number of test sites can be arbitrarily expanded to meet the system stability and fault tolerance requirements of large-scale examinations. However, the c/s mode test system needs to be installed on the test machine, which is relatively complicated to deploy. The general examination room system uses c/s mode to support daily exams and exams and supports large-scale exams and high-risk exams [11].

This paper studies and implements the research and implementation process of the College English test system based on B/S structure, including key technology analysis, system analysis and

system design, database design, system implementation and system testing. This thesis mainly introduces a network-based college English online examination system based on the campus network and the specific method of designing and implementing the system. In addition to user registration, multi-user concurrency control, real-time online exams, automatic judgment, and question bank maintenance, the system also has dynamic random questions, test time control, user management, score management and other functions.

2 RESEARCH METHOD

2.1 Mathematical model of generating papers

Generally speaking, in the automatic generating papers, the user will put forward various requirements for the quality of the test paper, such as the total amount of questions, the average difficulty, the proportion of the questions, the proportion of the chapters, the proportion of the key chapters, the intersection and integration of the knowledge points, etc. Automatic generating papers should satisfy the user's requirements to the greatest extent [12]. Therefore, before generating papers we first establish the corresponding state space of the control indicator for the process of automatic generating papers.

$$\mathrm{D} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1,m+n} \\ d_{21} & d_{22} & \cdots & d_{2,m+n} \\ \cdots & \cdots & \cdots & \cdots \\ d_{m+n,1} & d_{m+n,2} & \cdots & d_{m+n,m+n} \end{bmatrix} \tag{1}$$

Each line of D consists of control indicators of a test question, such as the total score of the test paper, the scores of each chapter, the score of the question type, the knowledge points, etc., and these attribute indicators are coded and expressed in binary form. Moreover, each column is the full value of an indicator in the question bank.

Genetic algorithms are a simulation of the natural evolution of humans. The natural evolution of humans is an evolutionary process that takes place on chromosomes. Natural selection allows more chromosomes with better fitness than those with poor fitness, and mutations can make the chromosome different from the parent chromosome. A new chromosome can be produced by the combination of two parental chromosomes. These concepts are reflected in mathematics to form the basic concepts and theorems of genetic algorithms [13]. This section will combine the basic theory and basic theorem of genetic algorithm to study the algorithm of automatic generating papers.

Genetic algorithm is a parallel algorithm that can be effectively optimized. Based on Morgan's gene theory and Eldridge and Gould's theory of discontinuous balance, the algorithm combines some of the ideas of Mayr's edge species formation theory and Bertalan's general system theory and simulates Darwin's natural genetics: inheritance (genetic inheritance), evolution (gene mutation) and survival of the fittest (excellent genes are genetically replicated in large numbers, and inferior genes are less genetically replicated). Its essence is a search algorithm that combines the natural selection of the survival of the fittest of the natural organisms, the evolutionary mechanism of the survival of the fittest, and the random information exchange system between individuals and individuals in the same group. When using genetic algorithm to solve the problem, we first need to express the problem of the required solution into binary code, and then perform basic operations according to the environment: selection, crossover, mutation [14]. By doing this so-called "survival selection", we finally converge to an individual who is most suitable for environmental conditions and get the optimal solution to the problem.

Based on the idea of the least-squares method in curve fitting, the fitness function is set to [15].

$$\mathrm{f}(x) = \sum_{i=1}^{10} \left(e^i\right)^2 \tag{2}$$

Among them, e^i is the error between the proportion of the test score and the actual proportion of the difficulty less than or equal to x_i. In order to speed up the optimization and avoid the premature phenomenon of the algorithm, and make the algorithm converge to the better solution as soon as

possible, we adopt the method of setting the penalty function to adjust the fitness function value of the prepared chromosome according to formula (3) after each generation is completed [16].

$$f(x) = ae^{bf(x)} \quad (3)$$

In a new generation of individuals, if a suitable individual does not appear, the value of a is set to 0.2. If 80% of the individuals satisfy the condition, the value of a is set to 0.8. In other cases, the value of a is set to 1. The parameter b is set to an empirical value of about 0.15.

Cross-operation is a must, it is a guarantee of global search. The crossover operator design requires two practical codes for the intersection of any two codes, and the characteristics of the parent are inherited as much as possible. The specific design adopts the segmentation single point crossover strategy. For the two randomly selected chromosomes, the set crossover probability is randomly selected in each sub-coding segment by a random point, and the number of questions in the segment is determined to be unchanged after the intersection. If the number of questions in the segment changes, the intersection is reselected.

Variation is a means of achieving group diversity and is also a guarantee of local search. Here, the segmentation and one-time mutation strategy is adopted: the intra-segment mutation operation is sequentially performed on each sub-coding segment of the chromosome, and the mutation operators of each segment are independent of each other. First, according to the probability of mutation, a certain point in the segment is randomly inverted, and then within the segment where the point is located, it is randomly determined to look for a point in the backward or forward direction, and its value is inverted. The purpose is to eliminate illegal genes and meet the constraints.

2.2 System structure

Based on the principles of abstraction, modularity, information hiding and consistency of software engineering, this paper adopted a modular design for the system. The modular design requires high cohesion within the module and loose coupling between modules. By following the above principles, the complexity of the software is reduced, and the efficiency of software development is improved. First, in the JSP solution, each functional module in the system is encapsulated and stored in a file of a different project in the form of a class file. Among them, the Product MANAGE SYTEM of the experimental test management system is the total entrance of the system.

Product MANAGE SYTEM's main role is to initialize the system when the lab test management system starts working or when the system fails. Moreover, the design of the system interface of the experimental test management system is also realized by the module, and the coordination work between the modules of the system is also realized by the module. These functions of the module correspond to the APPLICATION layer in the pattern stack. The definition of each business class of the system is defined in the file of the ProductManageSystem.BLL project. The function of the system module also includes the logical processing of the experimental test management system.

This corresponds to the business layer in the three-tier architecture. The class definitions for database-related operations are contained in the ProductManageSystem.DAL project file, which is used to complete the system's operations on the database and corresponds to the data access layer. General classes for some systems, such as general classes that implement functions such as data conversion or data validation, are provided by project files in ProductManageSystem.common. In addition, the definition of entity attributes, entity class attributes and their methods in the data model of the lab test management system is done by the ProductManageSystem.Model project file. The entity object that processes logic in the business layer is imple nmmented by the entity layer composed of the latter two. The processing and delivery process of data between different layers in the three-tier architecture of the system is shown in Figure 1.

The figure above represents the processing and direction of the system data is the solid arrow, the dotted arrow represents the feedback of the corresponding layer of the system on the data processing results. The logical processing and access layer of the business layer of the data in the database in the system respectively introduce the entity class to process the relevant experimental data and services, so the implementation of the above two layers of functions involves the analysis and construction of the corresponding entity object.

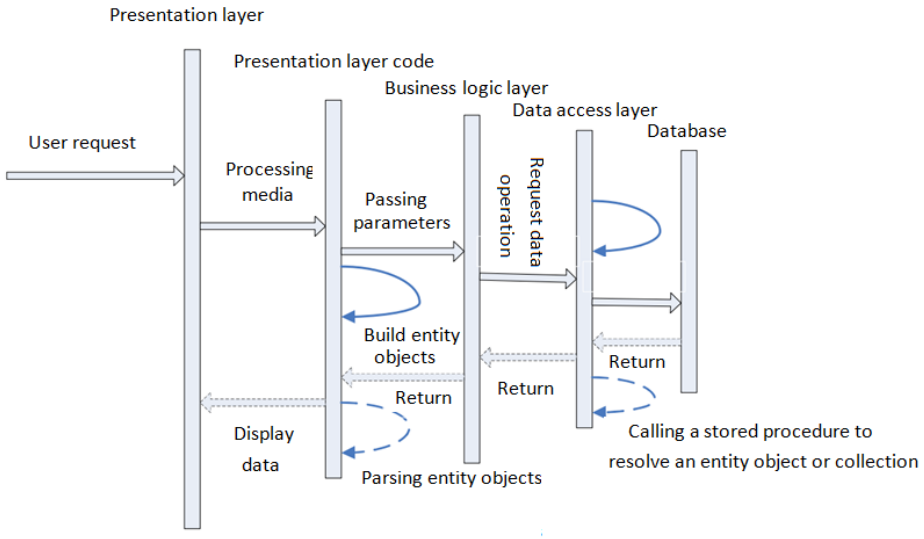

Figure 1: The processing and delivery process of data in system.

3 SYSTEM CONSTRUCTION

The system architecture adopts C/S architecture and is divided into client and server. The users of the client mainly include candidates, teachers, and administrators. Candidates log in to the client for identity verification, information check, device check, and login candidates. The teacher logs in to the client for identity verification, test update, and deletion. The administrator logs in to the client for identity verification, system maintenance, and personnel management. The server side is mainly used to store data, and various data operations of the client are finally stored on the server side.

Figure 3 is a window of the login module. In the candidate examination subsystem, it needs to deal with the direct use of the system, so the design of this part of the human-machine interface is particularly important. At present, the design of human-machine interface in system design has been as important as data design, structure design and process design. Especially in the human-computer interaction system, the workload of human-machine interface design accounts for more than half of the system design workload. The problems encountered in the design of the human-machine interface are: system response time, error handling, user assistance, and so on. In the human-computer interaction, if the response time of the system is too long, the waiting time of the user becomes long, the user feels nervous and frustrated when the system is not clear, and the user may speed up the operation and cause the system to respond more slowly, thereby entering a vicious circle. User response time is especially important in the exam system. Once the system response time is too long, the candidates will be nervous. Therefore, in this system, it is necessary to pay special attention to the system response time and minimize the response time as much as possible.

The current system software is very complicated, and it is obviously impossible for users to explore it themselves. Therefore, the system should bring its own user assistance facilities so that users can familiarize themselves with the system as soon as possible and solve the problem by themselves without leaving the user interface. This system is an examination system, so user assistance is essential, and there is an introduction to user assistance in the next chapter. Error message processing. The process of using the system by the user will inevitably lead to errors. On the one hand, the user input data error causes the occurrence of error information, and on the other hand, the system itself has a bug. Regardless of the error, the system needs to report the error to the user

to make the user to know the error so that the error is not repeated. In the human-computer interaction interface of the system, the user will be prompted for errors that may occur in the system.

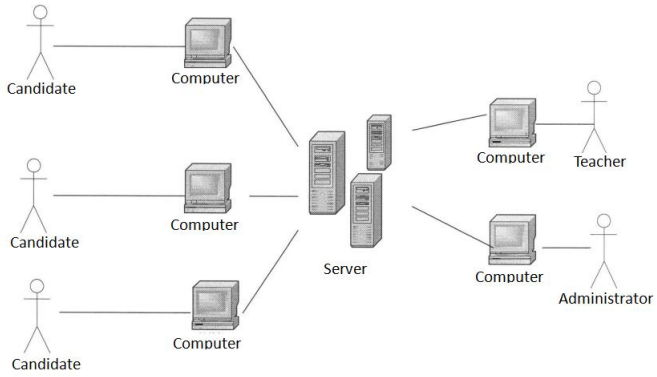

Figure 2: Overall architecture of the system.

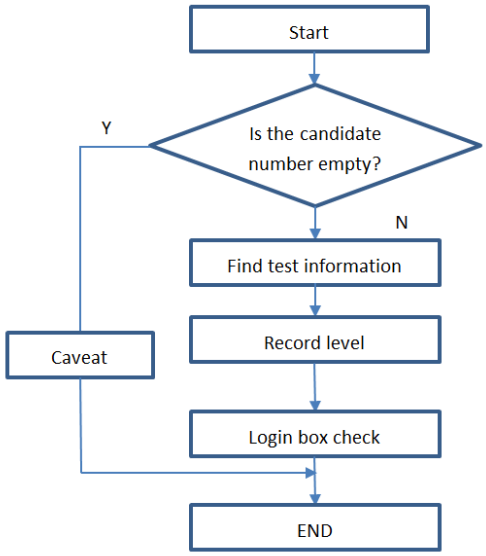

Figure 3: Login module flow chart.

Database design is the foundation for building a database, the foundation of an application, and the basis for software system development. In other words, the database design is to adapt to the existing application environment by constructing an optimal database schema, and to meet the needs of the application, and to ensure data integrity and consistency, which can improve data access efficiency and reduce data redundancy. Therefore, the optimized database structure design is the basis of application development. In general, the design principles and steps of the database are as follows:

The goal of database concept design is to reflect the data relationship between organizational structures through the model, and form the database conceptual structure of information needs, that is, the conceptual model. When we analyze the organizational structure, we should truly reflect the real things and the laws between them according to the general design principles of the database, instead of relying on subjective assumptions, taking it for granted, "what should it be?" or "how

should it not be". In addition, the design should be simple and straightforward, easy to understand, easy to modify, and pay attention to data redundancy. In the end, it can be easily converted into a database hierarchy and it is easy to generate a logical model of the database. The entity-type relationship model is the most commonly used model in database design. The E-R model refers to the extraction of entities and entities from abstract objects through abstract methods. The E-R model consists of entities, entity associations, and entity attributes. In general, the E-R model is analyzed and generated by the illustrated method. We call it the E-R diagram.

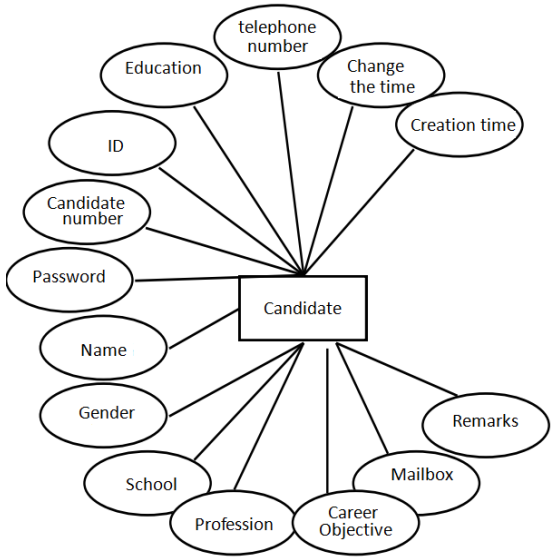

Figure 4: Candidate entity map.

4 SYSTEM TESTS

The software used in this section is HP's automated test software, LoadRunner, which is a professional stress test software and is often seen in software testing. It mainly accesses the system through multiple virtual users to achieve the purpose of system stress testing. The tool that LoadRunner provides for the analysis of stress test results is called Analysis. With this tool, the maximum response time, minimum time and average time of the system during the test can be monitored, and the CPU and memory usage of the server can be detected.

Because the "College English" test involves almost every college student, and there are all kinds of people in various college students, the registration time of the system is almost randomly distributed during the test. However, in general, the login time will not be able to log in half an hour later than the start time, so the design of the business scenario is as follows:

The user opens the "College English" test system home page, click on the "Login System" link, enter the user name and password, click "Login", and exit the system. Moreover, user calculations and startups will be concurrent. The number of concurrent numbers is designed to be 20, and the method of gradually increasing is adopted. First, two concurrent numbers are started, and two are added every five seconds. After reaching the specified number of users, the operation continues for about 5 minutes. The CPU usage is shown in Figure 5. As can be seen from the above figure, during the entire scenario test, the average CPU usage of the test server was 31.3%, and a high peak appeared at about 5 minutes and 30 seconds in the scenario execution, which was about 80.1%. Overall, CPU usage is balanced throughout the testing process. The memory usage is shown in Figure 6.

Figure 5: CPU usage.

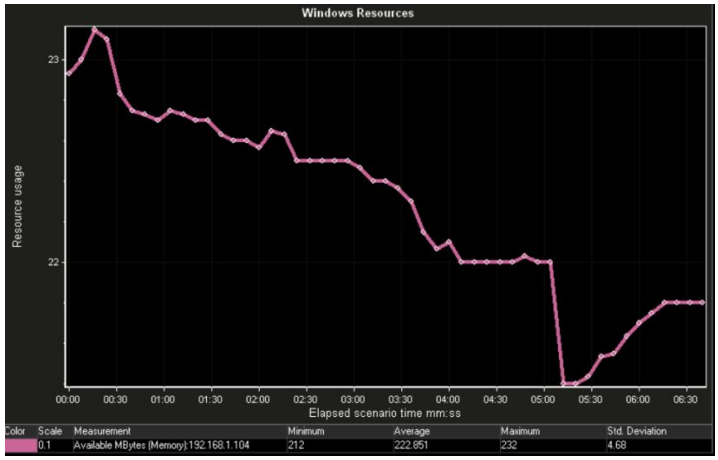

Figure 6: Memory usage.

As can be seen from the above figure, the available physical memory of the system is continuously decreasing. However, from its available maximum (232M) and minimum (212M), it can be seen that during the 5-minute test time, the memory change is not large, the available physical memory is maintained at about 222M, and the memory usage rate is (512M-222M)/512M=57%. This result indicates that there is no significant performance bottleneck. Overall, there is enough memory, but the overall trend of the graph is declining, and if more concurrency is used, there may be an out of memory problem. The transaction response time is shown in Figure 7.

Transaction response time is an important indicator to measure the processing power of the server. As can be seen from the figure, during the entire test, there were two large fluctuations, but in general, the response time of the server processing the submit operation was 3.435 seconds, including the think time of 3 seconds. Excluding thinking time, the server processes the entire login business operation, from opening the home page, logging in, and then exiting. From the above analysis, it can be seen that the server is relatively fast. The test data record table is shown in Table 1.

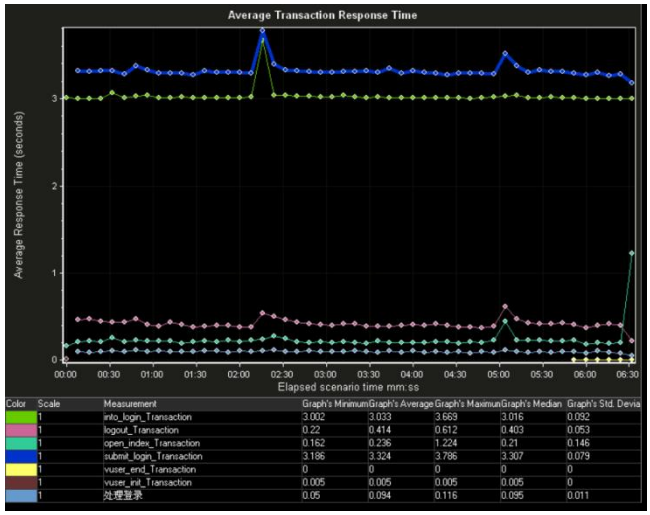

Figure 7: Average transaction response time.

Investigation item	*Target value*	*Actual value*	*Pass (Yes/No)*
Response time	<=5s	1.321	Yes
Business success rate	>99%	100%	Yes
Concurrent number	20	20	Yes
CPU usage	<75%	31.30%	Yes
Memory usage	<75%	57%	Yes

Table 1: Test data record table.

Through the above detailed tests, we have obtained detailed test results. By analyzing the results, we can conclude that the expected indicators of the test met the requirements and met our expected results.

5 CONCLUSION

This paper studied and implemented the implementation process of English test system based on C/S structure, including related technical analysis, system requirements analysis and system function design, database design and implementation, and specific implementation methods of each sub-module. This paper mainly introduces a specific method for the development and implementation of the Internet-based college English online examination system and he specific method of design and implementation of this system. The English test system can perform a variety of functions, such as user login and registration, multi-user real-time online exam, question bank management, test paper management, test management, candidate management, report management, automatic timing scoring and system maintenance, etc. In addition, it also has dynamic random questions, test time countdown control, user management, automatic score analysis and other functions. In addition, this article has obtained detailed test results through detailed tests and tested the system under existing hardware and software conditions. Finally, through testing, it was found that the system is in line with the overall objectives of the system, and all expected indicators met the requirements and met our expected results.

6 ORCID

Xuebing Zhang, https://orcid.org/0000-0003-0021-9267
Ruolin Shi, https://orcid.org/0000-0002-2781-2979
Na Wang, https://orcid.org/0000-0001-5767-0732
Rui Chen, https://orcid.org/0000-0001-7555-3521

REFERENCES

[1] Wang, L.; et al: Exploiting cross-sentence context for neural machine translation, 2017, https://doi.org/10.18653/v1/D17-1301
[2] Nguyen, Q. P.; Shin, J. C.; Ock, C. Y: Korean morphological analysis for Korean-Vietnamese statistical machine translation, Journal of Electronic Science and Technology, 15(4), 2017, https://doi.org/10.11989/JEST.1674-862X.61005104
[3] Sharaf, A.; et al: The UMD Neural Machine Translation Systems at WMT17 Bandit Learning Task, 2017, https://doi.org/10.18653/v1/W17-4778
[4] Han, L.: LEPOR: An augmented machine translation evaluation metric, 2017, https://doi.org/10.13140/RG.2.1.3149.9128
[5] Niehues, J.; Cho, E.: Exploiting linguistic resources for neural machine translation using multi-task Learning, 2017, https://doi.org/10.18653/v1/W17-4708
[6] Toral, A.; et al: Crawl and crowd to bring machine translation to under-resourced languages, Language Resources and Evaluation, 2016, https://doi.org/10.1007/s10579-016-9363-6
[7] Zhang, B.; Xiong, D.; Su, J.: Neural machine translation with deep attention, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 1-1, https://doi.org/10.1109/TPAMI.2018.2876404
[8] Zhang, H.; et al: Understanding subtitles by character-level sequence-to-sequence learning, IEEE Transactions on Industrial Informatics, 13(2), 2017, 616-624, https://doi.org/10.1109/TII.2016.2601521
[9] Stahlberg, F.; Byrne, B.: Unfolding and shrinking neural machine translation ensembles, 2017, https://doi.org/10.18653/v1/D17-1208
[10] Klubika, F.; Toral, A.; Sánchez-Cartagena, V. M.: Quantitative fine-grained human evaluation of machine translation systems: a case study on English to Croatian, Machine Translation, 2018, https://doi.org/10.1007/s10590-018-9214-x
[11] Arcan, M.; et al: Leveraging bilingual terminology to improve machine translation in a CAT environment, Natural Language Engineering, 23(5), 2017, 26, https://doi.org/10.1017/S1351324917000195
[12] Chua, C. C.; et al: Meaning preservation in example-based machine translation with structural semantics, Expert Systems with Applications, 78, 2017, 242-258, https://doi.org/10.1016/j.eswa.2017.02.021
[13] Liu, Y.; Vong, C. M.; Wong, P. K.: Extreme learning machine for huge hypotheses re-ranking in statistical machine translation, Cognitive Computation, 9(2), 2017, 285-294, https://doi.org/10.1007/s12559-017-9452-x
[14] Quang-Phuoc, N.; et al: Effect of word sense disambiguation on neural machine translation: A case study in Korean, IEEE Access, 2018, 1-1, https://doi.org/10.1109/ACCESS.2018.2851281
[15] Li, Q.; et al: Research on improved corpus-level and phrase-level pivot language based methods in low-resource machine translation, Chinese Journal of Computers, 2017, https://doi.org/10.11897/SP.J.1016.2017.00925
[16] Isabelle, P.; Cherry, C.; Foster, G.: A challenge set approach to evaluating machine translation, 2017, https://doi.org/10.18653/v1/D17-1263