Adaptive Crowdsourcing Task Generation and Workflow Control for Human Feedback Data Collection

Liqing Wang
School of Information Science and Engineering, Yunnan University, Kunming

Wanjin Chen
School of Information Science and Engineering, Yunnan University, Kunming

Yongyue Xu
School of Information Science and Engineering, Yunnan University, Kunming

Juan Wang
School of Information Science and Engineering, Yunnan University, Kunming

ABSTRACT
Due to the dire scarcity of corpora, the quality of low-resource language translation falls short of the public expectation. However, adopting Reinforcement Learning from Human Feedback (RLHF) can remarkably improve model quality. Nonetheless, obtaining human feedback data is typically time-consuming, costly, or plagued by severe inconsistency. Thus, this paper develops and implements a self-generating crowdsourcing workflow tailored for low-resource translation to address the above-mentioned issues. Under the consideration of quality and cost, this workflow will operate automatically and continually until it obtains the final results, according to a process in which the generated options are filled in the blanks for evaluating and selecting crowdsourcing tasks in various formats, then managing their iterative execution. This method enables the acquisition of various feedback data with varied requirements in diverse forms—including ranking, scoring, comparative judgments, and error correction—at low cost and with high efficiency. These data can then be used to train reward models with the consequence of enhancing reinforcement learning performance. This paper's experimental results testify to the effectiveness of this approach.

Keywords
RLHF; low-resource translation; adaptive; iterative workflow; crowdsourcing

1. INTRODUCTION
In machine translation, current Neural Machine Translation (NMT) systems stand out in translation between lingua franca, including English, Chinese, French, and German, by leveraging deep learning and large-scale corpus resources. However, in the case of non-lingua franca, the extreme scarcity of corpora fails to provide large-scale, high-quality data for model training, resulting in machine translation outputs that often fall short of practical needs [1].

Along with the release of ChatGPT, Reinforcement Learning from Human Feedback (RLHF) has attracted widespread attention and demonstrated robust applicability in fields like text conversation and AI-generated text [2]. RLHF optimizes and fine-tunes translation models according to human feedback, which can compensate for data deficiencies in resource-scarce scenarios when data is extremely scanty. However, achieving desirable results also requires high quality and consistency of feedback data [3].

As for large-scale feedback in low-resource translation, crowdsourcing presents a lower-cost approach to access a broader pool of contributors. Crowdsourcing is a collaborative mode that employs network technology to disintegrate largescale tasks and distribute them to a multitude of non-specific online users for completion flexibly and voluntarily. Crowdsourcing is frequently utilized to address problems that are hardly resolved by mere computer algorithms, encompassing data annotation, quality assessment, and translation correction of low-resource languages [4,5]. Due to the limitations of crowdsourcing workers, the critical challenge lies in devising a targeted, automated, or semi-automated workflow to achieve consistent and quality-assured outcomes within cost constraints [6].

For this purpose, taking the manual evaluation, correction, and feedback of the results of Chinese-Lao low-resource language machine translation as an example, this paper devises and implements a method for generating crowdsourcing collaborative translation iterative tasks based on mixed question types and adaptive control of the workflow.

2. RELATED WORKS
2.1 RLHF
RLHF is a form of reinforcement learning (RL). Reinforcement learning enables agents to make optimal decisions through continuous interaction and trial-and-error with the environment, and the reward model determines whether the decision is optimal. This operation requires objective and effective assessment and measurement of the agent's performance to train the reward model to provide correct reward signals. However, this is susceptible to environmental noise and data bias, subsequently causing the agent to generate meaningless or adverse results through the loopholes in the reward mechanism, thereby giving rise to the problem of reward abuse [7].

To tackle this issue, RLHF introduces the feedback signals of human evaluators into the standard reinforcement learning process, providing more accurate and human value cognition-oriented feedback for the reward model. This measure not only helps overcome the limitations and problems of traditional RL methods but also prompts the learning goals of the agent to be closer to human values, thereby achieving better alignment.

RLHF has demonstrated its effectiveness in multiple domains with the etcetera of fine-tuning large language models (LLMs) [8,9], continuous control [10], and robotics [11]. In the context of machine translation, RLHF contributes to generating more localized translations that can accurately express semantic and emotional connotations. However, RLHF still confronts numerous issues and challenges, among which is how to acquire high quality human feedback annotation data efficiently and economically.

Currently, the commonly employed methods for collecting feedback data consist of human annotation, crowdsourcing, and automata. The data obtained through Human annotation is of high quality and aligns better with human preferences. However, this approach is costly and time-consuming, especially for tasks in low-resource settings like non-general language translation. By crowdsourcing, it is possible to lower costs and obtain data with certain reliability [12]. Nevertheless, this method might lead to more data noise due to the possible discrepancy in subjective preferences among annotators and its inherent limitations. In recent years, due to the rapid progress of large models, training a machine learning model or constructing a specific automaton to evaluate results quality has become a feasible approach [13]. However, to acquire feedback data that are more in line with human preferences, a considerable amount of high-quality data is necessary for attaining a high-performance evaluation model. For low-resource translation situations, how to obtain a high-performance evaluation model with an inherent scarcity of data?

2.2 Crowdsourcing employed for the collection of feedback data

Based on the web's openness, crowdsourcing offers scope for addressing issues that are challenging to solve by relying solely on computer algorithms.

According to the distinct time points at which crowdsourcing is integrated into the processing procedures, existent crowdsourcing collaborative workflows can be classified into two types. One involves utilizing crowdsourcing workflows to accomplish the pre-processing steps of the machine model. For instance, Georgescu et al. proposed employing crowdsourced data for active learning training of machine translation models and investigated how to strike a balance between cost and effect [14]. Simpson et al. employed the relationship between labels and data features to set labels automatically. They experimentally demonstrated that this approach could significantly reduce the requirement for crowdsourced labels without compromising the final model accuracy of the machine learning system [15]. The second category involves employing crowdsourcing approaches for the post-processing of machine algorithm outcomes after the model training yields results. For instance, the crowdsourcing workflow is utilized to address the problem of coreference resolution [16]; the crowdsourcing collaborative workflow is employed for Chinese translation, and a ranking model based on graphs is applied for filtering redundant results and controlling the quality of the results within the workflow [17].

On the other hand, according to the varying priority weights assigned to cost and quality, it can be classified into two types: cost-constraint priority and quality priority. Regarding the workflow design with cost-constraint priority, the cost expenditure of crowdsourcing tasks is given priority and controlled through methods such as precise task decomposition, task reward design, and workflow optimization [18,19].

In the aspect of quality-priority workflow design, the main objective is to obtain high-quality task results. To this end, researchers have put forward multiple approaches, encompassing quality control [20], worker selection and task allocation [21]. In crowdsourcing, task iteration is also frequently employed to enhance quality, particularly for crowdsourcing tasks like translation and quality assessment which are hard to obtain consistent results within a single round. Nevertheless, iteration inevitably increases economic and time costs, and when the number of iterations reaches a certain threshold, it becomes difficult to achieve further quality improvement [22].

Therefore, the key to designing the optimal cost-quality balance of the crowdsourcing collaborative, iterative workflow lies in three aspects: the crowdsourcing collaborative mechanism, the generation and allocation of collaborative tasks, and the execution of collaborative tasks. Concerning the task requirements of Chinese-Lao translation investigated in this paper, a workflow with higher adaptability and control capacity needs to be designed to reduce costs and guarantee efficiency.

3. THE METHOD

3.1 The task

The objective of this research is to collect human-generated feedback for the refinement and evaluation of Chinese-Lao subtitle translations through a crowdsourcing-driven collaborative workflow. Preliminary Lao translations have been generated via model training. The focus of this study is to gather human revisions and evaluations. These final results will be utilized as feedback for subsequent reinforcement learning processes.

3.2 General design

In Figure 1, Chinese-Lao translation sentence pairs requiring feedback data serve as the initial task content to form the Task Sets. The **Task Generation** module creates tasks based on feedback requirements, including fill-in-the-blank, scoring, and selection tasks. These tasks are then distributed through the **Task Assignment** module, which leverages the **Worker Evaluation & Monitoring** module to identify crowd worker characteristics and monitor their real-time performance, ensuring tasks are allocated to the most suitable workers. Feedback results are collected via the **Results Feedback** module and evaluated by the **Quality & Cost Control** module.

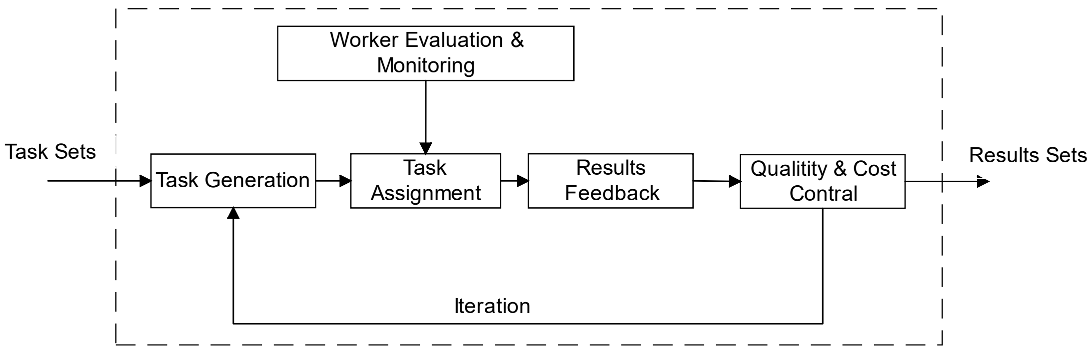

Figure 1: Adaptive collaborative workflow design for the task

If quality requirements are satisfied, the results are finalized and output. Otherwise, if cost constraints permit, an iterative task loop is initiated to achieve better outcomes.

The primary objective is to ensure and enhance quality within cost constraints. To achieve this, it is crucial to accurately model the attributes of crowd workers and implement comprehensive controls across multiple dimensions, including task content, interaction formats, task assignment strategies, worker incentives, evaluation mechanisms, and status monitoring.

In terms of cost management, key measures include the efficient decomposition of hu-man-machine tasks, precise workflow control to minimize iteration cycles, and the strategic use of incentives to boost both efficiency and quality.

In terms of quality assurance and enhancement include the following designs:

- Real-time monitoring and quantification of worker performance, with task assignment controlled accordingly.
- To mitigate the risk of malicious behavior and declining worker performance, the entire task execution process is monitored, including task cost and completion quality.
- Given the challenges in ensuring translation quality with non-professional translators, strategies such as adaptive workflow iterations, collaborative correction, and cross-evaluation are employed to improve result quality while maintaining cost control.

In summary, task content is dynamically generated based on demand, and the workflow's iterative execution is governed by cost and quality thresholds. Worker performance is continuously monitored in real time to ensure optimal task-worker matching at all times. As such, task generation, workflow control, and worker selection constitute the core modules.

3.3 Task generation

In this study, two types of tasks are generated. The first type, referred to as the golden detection task, is derived from a parallel corpus of known, corresponding translations, referred to as the Golden Corpus. These tasks are designed to be similar to the sentences requiring translation and are restricted to a maximum length of 30 words. Their primary purpose is to monitor worker performance and status.

The second type involves tasks that require translation, modification, or evaluation. Task Sets are generated from both types of tasks.

For the golden detection task, tasks are generated either before a worker's initial task assignment or after they have been working for several hours. These tasks may involve activities such as fill-in-the-blank exercises or judgment tasks, with no more than five sentences extracted for each. The primary goal of this task is to evaluate both the worker's abilities and their work status.

These tasks are designed as fill-in-the-blank, scoring, or selection tasks. Task difficulty is determined by sentence length and task type, with fill-in-the-blank tasks being the most difficult, followed by multiple-choice and scoring tasks in decreasing order of difficulty.

- **Fill-in-the-blank tasks.** These tasks are generated from a task database using a fill-in-the-blank template. To avoid excessive repetition of the same content, each corpus is extracted at least three times and assigned to different workers. A flag is set to track the number of extractions for each item.
- **Scoring tasks.** Once all fill-in-the-blank results have been collected, scoring tasks are generated based on a predefined scoring template.
- **Selection tasks.** If necessary, the scores for identical content are sorted, and the top three results will be used to generate a selection task.

3.4 Adaptive workflow design and control

The title (Helvetica 18-point bold), authors' names (Helvetica 12-point), and affiliations (Helvetica 10-point) run across the full width of the page – one column wide. This paper also recommends email address (Helvetica 12-point). See the top of this page for three addresses. If only one address is needed, center all address text. For two addresses, use two centered tabs, and so on. For three authors, you may have to improvise.

3.4.1 Adaptive Workflow Design

As shown in Figure 2, the key steps of the workflow are as follows:

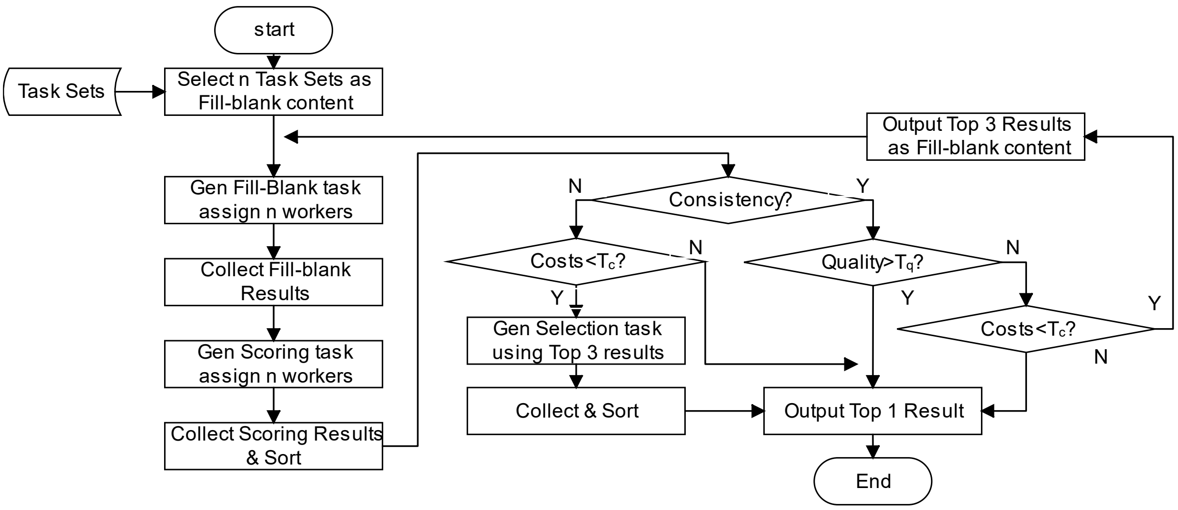

Figure 2: Flowchart of the Adaptive workflow

- **Set control thresholds**: Define two control thresholds, T_c and T_q, for cost and quality constraints.
- **Worker classification:** Workers are categorized into senior and junior based on their professionalism, credibility, and work efficiency. Senior workers are assigned tasks such as fill-in-the-blank or modification and review tasks, while junior workers are assigned simpler tasks, such as judging "good" or "fail."
- **Task generation:** R_1 is constructed by extracting sentences from the Task Sets or results from the previous iteration and assigning them to senior workers as fill-in-the-blank tasks. Each sentence is extracted n times, and the returned results form R_2.
- **Scoring task assignment**: Scoring tasks are derived from the result set R_2 and assigned to m junior workers. The scoring task results are then collected to create the set R_3.
- **Consistency check**: The consistency of the scores for identical content R_3 is calculated. If the scoring results meet the consistency requirement and the quality threshold T_q is satisfied, the results are matched with the original text to form a translation pair, which is added to the final result set R_s. If not, the next iteration is initiated, provided the cost threshold T_c allows.

3.4.2 Determination of the threshold value

The cost threshold T_c is determined by a combination of the theoretical cost and the total task cost. The total task cost represents the overall cost required to complete the task, denoted as C_t. The theoretical cost corresponds to the cost necessary to achieve a specified quality level, based on the relationship between the number of iterations and the quality of the results.

If m iterations are required, the theoretical cost is the sum of the product of the cost per task C_n and the number of tasks per iteration N_n, where n ranges from 1 to m. The cost constraint threshold T_c is then set as the minimum of the total task cost and the theoretical cost to ensure that the highest possible quality is achieved within the specified cost constraints.

The calculation formula is shown in equation (1).

$$T_c = Min(C_t, \sum_{n=1}^{m} C_n * N_n) \tag{1}$$

The quality threshold T_q is a score used to assess whether the task result meets the required quality standards, determining whether the task is complete or needs further iteration. The quality threshold is set based on the specific quality requirements of the translation. In this study, the third quartile (Q3) of multiple scores is used as the quality threshold T_q. If the third quartile cannot be obtained, the top-ranked result will be directly output.

3.5 Worker selection and motivation

To assess the initial linguistic proficiency of workers, a golden detection task is assigned, which includes both fill-in-the-blank and scoring tasks. Initially, the fill-in-the-blank task with “Submit” and “Skip” options is used. Based on the correctness of the answers, workers are classified as either Senior or Junior workers. If a worker selects “Skip,” the number of skips will also be considered to determine their classification. These selection methods are designed to maximize the utilization of worker resources while minimizing costs.

In addition, a hybrid incentive system is employed to enhance worker engagement and focus. This system includes the following strategies:

- Incorporating a sense of achievement into tasks, such as through mission completion or challenges against other workers.
- Using indirect incentives, such as TOP N rankings, medals, grading, or bonus points, to improve worker participation and persistence.
- Offering direct incentives through points, virtual goods exchanges, or monetary rewards to encourage active participation.

4. IMPLEMENTATION AND RESULTS ANALYSIS

4.1 Preparation

4.1.1 Experiment Data

The sentence pairs to be translated are subtitles from a Chinese documentary short film. According to the International Standard for Subtitles, the maximum number of characters per line should be between 11 and 16, with no more than three lines. Therefore, the Chinese subtitles to be translated are pre-processed and segmented according to these subtitle requirements, with each sentence containing a minimum of one Chinese character and a maximum of 27 Chinese characters.

For the testing, 200 bilingual parallel sentence pairs, each containing more than six words, were extracted from the expert-translated and reviewed subtitle corpus. Of these, 50 pairs were designated as Golden detection task data.

4.1.2 Crowdsourcing Workers

The volunteers participating in the crowdsourcing task are Laotian students with a basic knowledge of Chinese and Chinese students learning Laotian, all of whom participated voluntarily.

4.2 Evaluation method

This paper uses the BLEU evaluation method to assess the translation results.

BLEU (Bilingual Evaluation Understudy) [23] is one of the most widely used metrics for evaluating machine translation systems. It measures the accuracy of a translated text by calculating the N-gram matching rate between the machine-generated translation and reference translations. The BLEU score ranges from 0 to 1, where a score closer to 1 indicates a high similarity between the machine translation and the reference.

The calculation formula is:

$$BLEU = BP * \exp(\sum_{n=1}^{N} W_n \log P_n) \quad (2)$$

BP is the brevity penalty, which is applied to penalize the machine translation result if its length significantly differs from that of the reference translation. n represents the maximum order of the n-grams (typically up to 4), and W_n is the weight coefficient for each n-gram, usually set to 1/4. P_n is the precision of the n-grams, representing the ratio of matching n-grams to the total number of n-grams in the machine translation result.

4.3 Results and analysis

FirstBLEU refers to the BLEU score of the translation results generated by the model using the experimental data. These results need to obtain feedback through this workflow and serve as the initial values for the fill-in-the-blank tasks. FinalBLEU represents the BLEU score of the feedback results obtained through this workflow.

As shown in Figures 3 and 4, the quality of the results from the collaborative translation task improves after the first round of the fill-in-the-blank task, compared to the initial FirstBLEU score. Additionally, Figure 4 illustrates that the results of multiple iterations are significantly better than the initial results from the worker translation task. However, as the number of tasks increases, the quality gap between the iterative results and the initial results gradually narrows. This is mainly attributed to worker fatigue and task monotony. The overall curve exhibits a cyclical pattern, consistent with the cyclical nature of workers' performance.

To further analyze the results FirstBLEU and FinalBLEU, as shown in the boxplot in Figure 5.

The minimum, quartiles, and median values of the boxplot are detailed in Table 1.

Table 1: Boxplot Descriptive Statistics Table

	FirstBLEU	FinalBLEU
Min	**0.096529**	0
Q1	0.308987	**0.318296**
Median	0.428495	**0.453904**
Q3	0.595993	**0.619962**

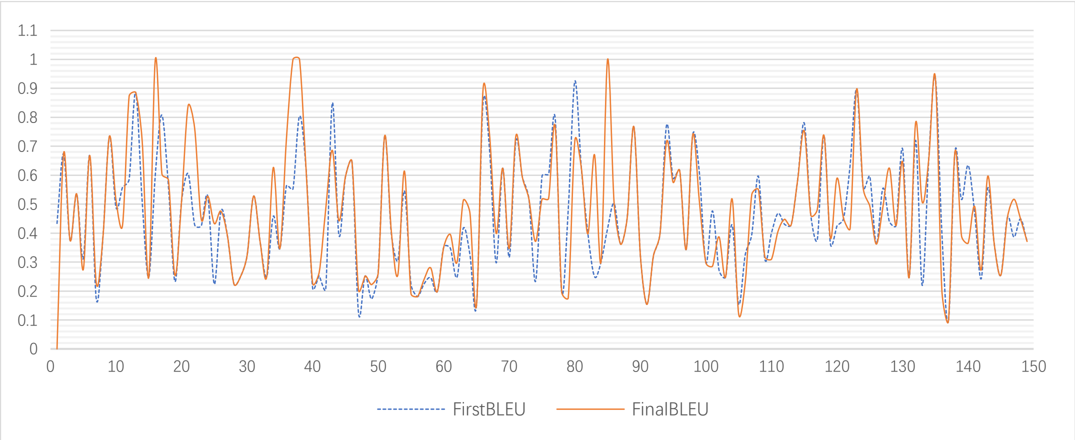

Figure 3: Quality changes before and after workflow iteration

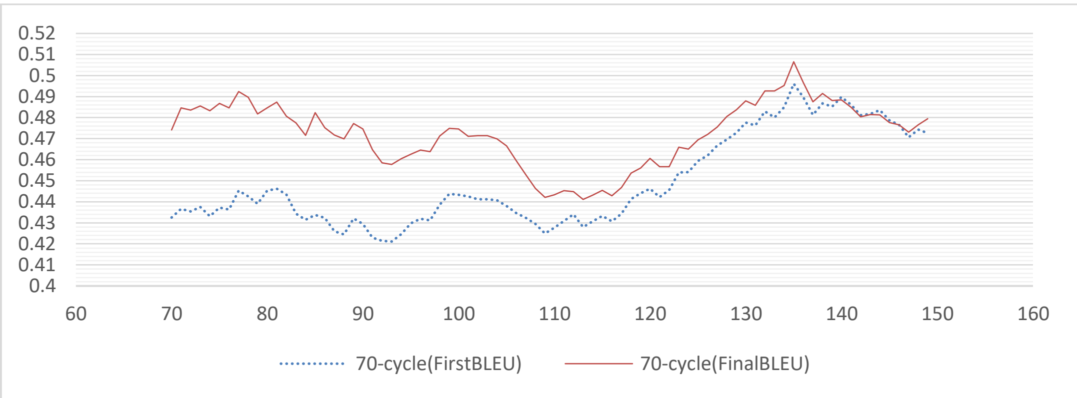

Figure 4: Mean smoothing of result quality change before and after iteration

Based on the median line of the boxplot, the quality of the results before and after the collaborative workflow improves by approximately 5.93%. In Figure 5, the boxplot for the crowdsourcing task exhibits a larger box, reflecting greater variability and dispersion in the data. This discrepancy is due to the inconsistent quality resulting from the fluctuating performance of the crowd workers.

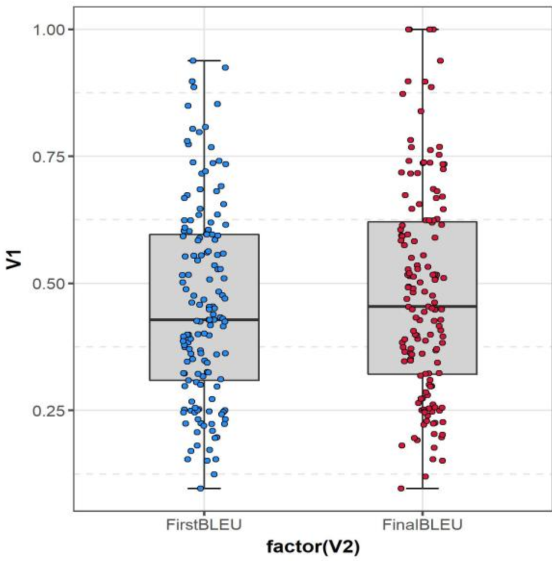

Figure 5: Boxplot of result quality change

For the collaborative workflow proposed in this paper, which is tailored for non-general language translation, strict worker requirements are essential. Designing additional crowdsourcing interaction methods to reduce task difficulty—such as using Thai as a bridge language—could further lower costs and improve translation quality.

5. CONCLUSION AND FUTURE WORK

To address the high cost of obtaining feedback data in reinforcement learning based on human feedback, an adaptive crowdsourcing workflow, controlled by cost and quality thresholds, was designed.

This method regards the results of machine translation of Chinese-Lao as the data to be evaluated, introduces the crowdsourcing workflow, and distributes the feedback tasks to the crowdsourcing workers through fill-in-the-blank modification questions, judgment questions, and multiple-choice questions. By integrating worker selection, dynamic monitoring, and incentives, under the cost constraint, consistent evaluation and feedback data are ultimately obtained. After testing, compared with the initial input data, the BLEU increases by 5.93% in this method's error correction and modification results. They can be used as feedback data for input into the reward model training. Additionally, diverse question types such as judgment and selection can provide rich feedback data forms, including sorting, scoring, and comparison for reinforcement learning, demonstrating strong universality and adaptability.

This approach allows for the acquisition of evaluation and optimized translation feedback at a reduced cost. With the advancement of large language models (LLMs), this paper plans to utilize LLMs to handle more evaluation or comparison tasks in the future, with manual intervention reserved only for uncertain results. This will further enhance efficiency and reduce costs. A key focus of this study is determining the optimal timing for manual intervention.

6. ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundations of China under Grants 61862064.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Liqing Wang; data collection and Experiment: Wanjin Chen, Juan Wang, Liqing Wang; analysis and interpretation of results: Liqing Wang, Juan Wang; draft manuscript preparation: Liqing Wang, Yongyue Xu, Wanjin Chen. All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

7. REFERENCES

[1] Yazar, B. K., Şahın, D. Ö., & Kiliç, E. 2023. Low-resource neural machine translation: A systematic literature review. IEEE Access, 11, 131775-131813.

[2] Chen, H., Yuan, K., Huang, Y., Guo, L., Wang, Y., & Chen, J. 2023. Feedback is all you need: from ChatGPT to autonomous driving. Science China. Information Sciences, 66(6), 166201.

[3] Wang, L., & Xiao, Y. 2024. Improving Low-Resource Machine Translation Using Reinforcement Learning from Human Feedback. Intelligent Automation & Soft Computing, 39(4).

[4] Ri, J. 2016. Utilizing crowdsourced translation in translator training—focusing on the practices in Japan—. The Journal of Translation Studies, 17(2), 91-117.

[5] Vaughan, J. W. 2018. Making better use of the crowd: How crowdsourcing can advance machine learning research. Journal of Machine Learning Research, 18(193), 1-46.

[6] Howe, J. 2006. The rise of crowdsourcing. Wired magazine, 14(6), 176-183.

[7] Skalse, J., Howe, N., Krasheninnikov, D., & Krueger, D. 2022. Defining and characterizing reward gaming. Advances in neural information processing systems, 35, 9460-9471.

[8] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Hüllermeier, E. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences, 103, 102274.

[9] OpenAI, T. 2022. Chatgpt: Optimizing language models for dialogue. OpenAI. In.

[10] Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. 2017. Deep reinforcement learning from human preferences. Advances in neural information processing systems, 30.

[11] Hejna III, D. J., & Sadigh, D. 2023. Few-shot preference learning for human-in-the-loop rl. Paper presented at the Conference on Robot Learning.

[12] Krippendorff, K. 2004. Reliability in content analysis: Some common misconceptions and recommendations. Human communication research, 30(3), 411-433.

[13] Gilardi, F., Alizadeh, M., & Kubli, M. 2023. ChatGPT outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences, 120(30), e2305016120.

[14] Georgescu, M., Pham, D. D., Firan, C. S., Gadiraju, U., & Nejdl, W. 2014. When in doubt ask the crowd: Employing crowdsourcing for active learning. Paper presented at the Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14).

[15] Simpson, E. D., Venanzi, M., Reece, S., Kohli, P., Guiver, J., Roberts, S. J., & Jennings, N. R. 2015. Language understanding in the wild: Combining crowdsourcing and machine learning. Paper presented at the Proceedings of the 24th international conference on world wide web.

[16] Fuge, M., Tee, K., Agogino, A., & Maton, N. 2014. Analysis of collaborative design networks: A case study of openideo. Journal of Computing and Information Science in Engineering, 14(2), 021009.

[17] Feng, J. H., & Li, G. L. 2015. A survey on crowdsourcing. Jisuanji Xuebao/Chinese Journal of Computers, 38, 1713-1726. doi:10.11897/SP.J.1016.2015.01713

[18] Bhatti, S. S., Gao, X., & Chen, G. 2020. General framework, opportunities and challenges for crowdsourcing techniques: A comprehensive survey. Journal of Systems and Software, 167, 110611.

[19] Kang, Q., & Tay, W. P. 2020. Task recommendation in crowdsourcing based on learning preferences and reliabilities. IEEE Transactions on Services Computing, 15(4), 1785-1798.

[20] Y Bo, Wang L. 2020. Task assignment method based on worker multi-attribute characteristic model. Computer Engineering and Design.;41(5):1398-403.

[21] Chen B, Wang L, Jiang X, Yao H. 2021. Survey of task assignment for crowd-based cooperative computing. Computer Engineering and Applications.;57(20):1-2.

[22] Carvalho, A., Dimitrov, S., & Larson, K. 2016. How many crowdsourced workers should a requester hire? Annals of Mathematics and Artificial Intelligence, 78, 45-72.

[23] Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. 2002. Bleu: a method for automatic evaluation of machine translation. Paper presented at the Proceedings of the 40th annual meeting of the Association for Computational Linguistics.