AI Conferences Made Easy by AI - A Case Study at MIWAI 2025

Tai Q. To[1*], Long S. T. Nguyen[1*], Hung C. Nguyen[1], Nguyen B. Le[1], Tam M. Nguyen[1], Tung T. Nguyen[1], Chattrakul Sombattheera[2], Tho T. Quan[1(✉)]

[1]URA Research Group, Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Vietnam
[2]Mahasarakham University, Thailand

2025 RIVF International Conference on Computing and Communication Technologies (RIVF) | 979-8-3315-7790-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/RIVF68649.2025.11365171

***Abstract*—Assigning suitable reviewers to submitted papers is critical for maintaining the quality and fairness of academic conferences. The Reviewer Assignment Problem (RAP) is inherently multi-objective and uncertain, requiring a balance between expertise, workload distribution, conflicts of interest, and fairness. In practice, both reviewer expertise and paper topics are difficult to specify precisely, as they are dynamic, context-dependent, and often only partially overlapping. To address these challenges, we present a novel approach to RAP that combines the Non-dominated Sorting Genetic Algorithm II (NSGA-II) with fuzzy set theory. Fuzzy sets are employed to model subjective and imprecise factors such as reviewer expertise and workload tolerance, while NSGA-II provides an efficient search mechanism to optimize conflicting objectives simultaneously. Our method yields a set of Pareto-optimal reviewer assignment solutions, allowing program chairs to choose from well-balanced trade-offs. Experimental results on submission data from the 18th International Conference on Multi-disciplinary Trends in Artificial Intelligence (MIWAI 2025) demonstrate that our approach improves assignment quality and fairness compared to baseline strategies, while also reducing manual effort in reviewer selection.**

***Index Terms*—reviewer assignment problem, multi-objective optimization, fuzzy set theory.**

I. INTRODUCTION

Peer review is a cornerstone of scientific progress, ensuring the quality, fairness, and credibility of scholarly communication [1]. At the core of this process lies the *Reviewer Assignment Problem* (RAP), which requires conference organizers to allocate submitted papers to suitable reviewers while balancing expertise, workload, conflicts of interest, and fairness. The challenge is amplified in international AI conferences, where submissions continue to rise while reviewer pools remain limited [2]. Although platforms such as *EasyChair*[1] provide automated assignment functionalities, they often struggle with accurately capturing expertise similarity, resolving lexical mismatches in topic descriptions, and maintaining balanced workloads, which frequently requires manual intervention by program chairs.

Historically, several approaches have been proposed to address RAP. Early methods relied on heuristic or rule-based systems based on explicit keyword matching between submissions and reviewers [3]. Probabilistic and topic-modeling approaches, such as Latent Dirichlet Allocation (LDA), were later employed to provide more flexible expertise modeling [4]. More recent work has explored optimization-based formulations, including network flow theory [5], integer programming [6], and bipartite matching frameworks [7]. These approaches offer formal guarantees and efficiency but still struggle to represent the nuanced and high-dimensional structure of reviewer expertise and paper content.

Multi-objective approaches have also emerged as promising solutions for RAP. Goldsmith and Sloan [8] introduced fairness-aware constraint satisfaction, and Kobren et al. [9] developed large-scale matching algorithms to improve balance and equity. Evolutionary algorithms, particularly the *Non-dominated Sorting Genetic Algorithm II* (NSGA-II), are naturally suited for RAP because they can simultaneously optimize multiple conflicting objectives such as expertise matching, workload balance, conflict avoidance, and fairness [10]. In our approach, similarity scores derived from embedding models form one objective, fuzzy membership values for workload and fairness form others, and NSGA-II identifies a diverse Pareto front of high-quality assignment solutions. This allows program chairs to choose from multiple trade-offs instead of a single aggregated outcome.

Despite their potential, evolutionary approaches remain underexplored in reviewer assignment, especially in combination with embedding-based expertise modeling and fuzzy set theory. Reviewer expertise and paper domains are inherently uncertain and dynamic, and crisp representations often fail to capture partial alignment or semantic ambiguity. Fuzzy set theory addresses this challenge by modeling vagueness and gradual degrees of suitability, while embedding models provide dense semantic representations enriched by generative AI. However, embeddings alone only model semantic similarity and do not handle the combinatorial nature of RAP. To date, no widely adopted system integrates semantic embeddings, fuzzy uncertainty modeling, and evolutionary multi-objective optimization in real conference settings. Existing methods either prioritize computational efficiency at the expense of fairness or rely on oversimplified expertise representations that fail to capture uncertainty and semantic richness.

*These authors contributed equally to this work.
[1]https://easychair.org/

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.

To address these limitations, we propose a hybrid framework that integrates embedding-based expertise modeling, fuzzy set theory, and the NSGA-II genetic algorithm to achieve robust and fair reviewer assignment. The embedding component generates semantically rich representations of papers and reviewers [11], enhanced by generative AI to expand latent topics and relevant keywords. The fuzzy component models uncertainty in expertise levels and workload tolerance, while NSGA-II jointly optimizes expertise matching, workload balance, conflict avoidance, and fairness, producing a diverse set of Pareto-optimal solutions. Applied to the *18th International Conference on Multi-disciplinary Trends in Artificial Intelligence* (MIWAI 2025), this hybrid embedding, fuzzy, and evolutionary approach demonstrates its feasibility and practical benefits. Beyond improving assignment quality and fairness, it significantly reduces manual effort for organizers and provides a scalable and adaptive solution for future conferences.

In summary, our contributions are as follows.

- **Reviewer information pipeline.** We develop an automated mechanism to retrieve and process reviewer expertise from prior publications, offering a richer and more accurate basis for reviewer–paper matching.
- **Integrated hybrid framework.** We design a unified system that combines generative AI for semantic enrichment with fuzzy set theory and NSGA-II, enabling uncertainty modeling and simultaneous optimization of multiple conflicting objectives.
- **Deployment at MIWAI 2025.** We deploy our framework in the MIWAI 2025 conference, where it demonstrates substantial improvements in assignment quality and fairness while reducing manual workload for program chairs.

II. RELATED WORKS

The RAP has been widely studied due to its central role in ensuring fair and effective peer review. A common approach is to compute the similarity between reviewers and papers, and then generate an assignment that maximizes the total similarity across all reviewer–paper pairs. This similarity-driven strategy is employed in conference management systems such as EasyChair, *HotCRP*[2], and the *Toronto Paper Matching System*[3].

The first research direction focuses on defining reviewer–paper similarity. Early studies relied on heuristic or keyword-based methods. [12] applied TF-IDF to estimate reviewer suitability. [4] used topic models to represent reviewer expertise, while [13] combined *Linear Discriminant Analysis* (LDA), linear regression, and collaborative filtering to compute suitability scores. [14] assumed reviewer expertise was known beforehand and defined matching rules accordingly. More recent work has introduced neural embedding methods. [15] calculated reviewer–paper similarity using Word2Vec embeddings. [16] proposed the *Word Mover's Distance Constructive Covering Algorithm* (WMD–CCA), which treats RAP as a classification task based on research field labels. [17] further advanced this line of research by using sentence-pair modeling and leveraging associations between paper titles and abstracts as supervision signals.

The second direction focuses on constructing the assignment itself, which is often formulated as an optimization problem with multiple objectives and constraints. *Integer Linear Programming* (ILP) has been one of the most widely used methods. [18] demonstrated that reviewer assignment criteria can be modeled through ILP, and [14] extended this approach under the assumption of known reviewer suitability. *Mixed Integer Programming* (MIP) was later introduced to incorporate additional continuous variables that help relax hard constraints [18]. Beyond ILP, metaheuristic approaches have also been explored. [19] employed a combination of the *Greedy Randomized Adaptive Search Procedure* (GRASP) and genetic algorithms for large-scale reviewer assignment. [15] integrated genetic algorithms with ant colony optimization and reported improvements in assignment quality at the cost of increased computational complexity. Alternative learning-based approaches have also been proposed. [10] evaluated several models, including language models, linear regression, and Bayesian probabilistic matrix factorization, for predicting reviewer suitability.

Although these studies provide valuable insights, most existing methods rely on deterministic representations of reviewer expertise, workload, and suitability. Such representations struggle to capture uncertainty, vagueness, and partial overlap between reviewer interests and paper topics. This limitation motivates our work, which integrates NSGA-II with fuzzy set theory to provide a practical and adaptive framework for reviewer assignment that explicitly models uncertainty while optimizing multiple conflicting objectives.

III. PROBLEM FORMULATION

We consider the Reviewer Assignment Problem (RAP) in the context of MIWAI 2025. Let $P = \{p_1, \ldots, p_m\}$ denote the set of submitted papers, and let $R = \{r_1, \ldots, r_n\}$ denote the set of available reviewers. Each paper is required to receive exactly λ reviews, while each reviewer r_j has a maximum capacity L_j that specifies the number of papers they can reasonably be expected to review. The relationship between papers and reviewers is further constrained by conflicts of interest, represented as a set $C \subseteq P \times R$, which prohibits certain paper–reviewer pairs from being assigned. We introduce binary decision variables a_{ij}, where $a_{ij} = 1$ if reviewer r_j is assigned to paper p_i and $a_{ij} = 0$ otherwise. The assignment matrix $A = (a_{ij})$ must satisfy the following feasibility constraints: (i) each paper receives exactly λ reviewers, (ii) no reviewer is assigned more than L_j papers, and (iii) no assignment is allowed for any paper–reviewer pair that appears in the conflict set C.

While these hard constraints determine the space of feasible assignments, the quality of an assignment is governed by several soft objectives that often conflict with one another, as summarized in Table I. The first objective is to maximize overall expertise by assigning each paper to reviewers with

[2]https://hotcrp.com/
[3]https://torontopapermatching.org/

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.

TABLE I
HARD CONSTRAINTS AND OBJECTIVES USED IN THE REVIEWER ASSIGNMENT FORMULATION

ID	Description and Formulation
H1	Each paper must receive exactly λ reviews: $\sum_{j=1}^{n} a_{ij} = \lambda, \quad \forall i \in \{1, \ldots, m\}$
H2	Reviewer capacity must not be exceeded: $\sum_{i=1}^{m} a_{ij} \leq L_j, \quad \forall j \in \{1, \ldots, n\}$
H3	Assignments involving conflicts of interest are prohibited: $a_{ij} = 0, \quad \forall (p_i, r_j) \in C$
O1	Maximize total reviewer expertise: $f_1(A) = \sum_{i=1}^{m} \sum_{j=1}^{n} s_{ij}\, a_{ij}$
O2	Balance reviewer workloads by minimizing load variance: $f_2(A) = -\mathrm{Var}\left(\sum_{i=1}^{m} a_{ij}\right)_{j=1}^{n}$
O3	Ensure fairness by maximizing the minimum similarity per paper: $f_3(A) = \max_{A} \min_{i} \sum_{j=1}^{n} s_{ij}\, a_{ij}$

the highest topical similarity. Let $s_{ij} \in [0, 1]$ denote the similarity score between paper p_i and reviewer r_j, computed using a combination of fuzzy expertise modeling and text similarity methods. The second objective is to balance reviewer workloads. Because highly experienced reviewers tend to attract many suitable submissions, balancing workload ensures that the reviewing burden is distributed fairly rather than concentrated among a few individuals. The third objective focuses on fairness across papers by protecting weaker submissions. In particular, we aim to ensure that even papers with relatively low similarity profiles receive reviewers with adequate expertise. This objective is expressed by maximizing the minimum aggregate similarity across all papers.

IV. PROPOSED APPROACH

The overall workflow is illustrated in Figure 1. The input consists of submitted papers and reviewers' profiles, including their past publications. Reviewer expertise and submission content (titles and abstracts) are semantically encoded using a Transformer-based model [11]. Cosine similarity scores are then computed to capture topical alignment between reviewers and papers. These similarity scores, together with workload statistics, bidding information, and conflict of interest data, are transformed into fuzzy satisfaction values through the fuzzy modeling module.

Subsequently, the NSGA-II framework is initialized with a population of candidate reviewer–paper assignments, incorporating a bias toward high similarity scores to ensure strong initialization. The population evolves iteratively through crossover and mutation operators, while strictly enforcing all hard constraints. Each candidate solution is evaluated using the three-objective fitness function introduced in Section III. Guided by non-dominated sorting, elitism, and diversity preservation, NSGA-II searches for high-quality solutions. Upon termination, the algorithm produces a set of Pareto-optimal assignments, from which conference organizers can select the most appropriate trade-off between expertise, workload balance, and fairness.

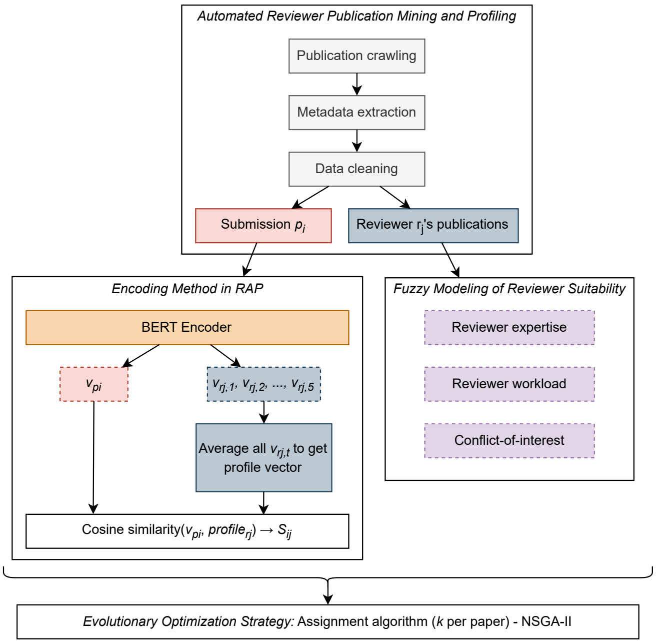

Fig. 1. Overview of the proposed reviewer assignment framework

A. Automated Reviewer Publication Mining and Profiling

To establish a reliable basis for reviewer–paper matching, our framework automatically collects and organizes reviewer-related data. Specifically, an automated crawler retrieves publicly available information from *Google Scholar*[4], including recent publications and research keywords. This ensures that reviewer expertise is captured in a timely and unbiased manner.

The automated pipeline consists of three main stages:

- **Publication Crawling:** Automatically retrieving titles, abstracts, and keywords of top publications within the last five years.
- **Metadata Extraction:** Systematically collecting co-authorship and affiliation data to support conflict of interest detection and workload balancing.
- **Data Cleaning:** Performing automated deduplication and text normalization to guarantee consistent, high-quality semantic representations.

By automating these steps, the tool minimizes manual effort while providing a scalable reviewer knowledge base for embedding and optimization. The pipeline operates exclusively on publicly available metadata (e.g., titles, abstracts, keywords,

[4]https://scholar.google.com/

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.

and co-authorship) and avoids any personal or sensitive information, thereby ensuring compliance with established research ethics standards on transparency, privacy, and anonymization.

B. Encoding Method in Reviewer Assignment Problem

To represent candidate solutions for the RAP, we employ a fixed-length integer vector encoding that effectively maps reviewer-paper assignments into a structure compatible with evolutionary algorithms. Let $P = \{p_1, p_2, \ldots, p_m\}$ denote the set of submitted papers and $R = \{r_1, r_2, \ldots, r_k\}$ represent the available reviewers. Each paper $p_i \in P$ requires exactly q reviewers (with $q = 4$ in our implementation).

The encoding represents a complete assignment as a concatenated vector of blocks:

$$x = [b_1 \| b_2 \| \ldots \| b_m] \in \mathbb{Z}^{m \times q},$$

where each block $b_i = [r_{i1}, r_{i2}, \ldots, r_{iq}]$ corresponds to the reviewer indices assigned to paper p_i, with each gene $r_{ij} \in [0, k-1]$ referencing a specific reviewer in R. This encoding facilitates the application of evolutionary operators while maintaining the problem constraints.

To ensure topical relevance in assignments, we use a semantic similarity framework based on BERTopic embeddings [20]. Each paper p_i and reviewer r_j (represented through their publication history) are encoded into vector representations $\mathbf{v}_{p_i}$ and $\mathbf{v}_{r_j}$. The similarity score S_{ij}, computed using cosine similarity between these vectors,

$$S_{ij} = \cos(\theta) = \frac{\mathbf{v}_{p_i} \cdot \mathbf{v}_{r_j}}{\|\mathbf{v}_{p_i}\| \|\mathbf{v}_{r_j}\|},$$

serves as a key metric in our fitness function to optimize for topical coherence alongside other constraints. This approach enables efficient evolutionary optimization while ensuring semantically meaningful reviewer–paper matches.

C. Fuzzy Modeling of Reviewer Suitability

To evaluate the suitability of reviewer assignments as soft constraints, we employ fuzzy set theory to capture the inherent vagueness in expertise, workload, and conflict levels. Each assignment (p_i, r_j) of paper p_i to reviewer r_j is associated with a satisfaction degree $s(p_i, r_j) \in [0, 1]$, where $s = 1$ denotes a highly suitable match and $s = 0$ indicates an unsuitable one. Intermediate values reflect partial suitability, balancing multiple factors.

Three fuzzy membership functions are defined to model the core aspects of reviewer assignment. First, reviewer expertise is measured using a triangular fuzzy set, where higher similarity between paper content and reviewer interests yields higher membership values. Second, reviewer workload is modeled by a trapezoidal fuzzy set, in which light to moderate workloads achieve high membership, while excessive load reduces suitability. Third, conflict of interest is handled using an inverted triangular fuzzy set, where low conflict levels are mapped to higher membership values. The combined base satisfaction is computed as a weighted sum of these three components, with greater emphasis placed on expertise, followed by workload and conflict.

In addition to these intrinsic factors, reviewer bidding data is incorporated as an external modifier. When a reviewer explicitly bids on a paper (e.g., "yes" or "maybe"), the bid score is blended with the fuzzy-based satisfaction score to reflect self-declared interest or willingness. This integration ensures that the model captures both implicit expertise derived from text similarity and explicit reviewer preferences expressed during bidding.

Formally, the overall satisfaction function is given by:

$$s(p_i, r_j) = \alpha \cdot \Big(w_1 \cdot \mu_{\text{exp}} + w_2 \cdot \mu_{\text{wl}} + w_3 \cdot \mu_{\text{cf}} \Big) + (1-\alpha) \cdot \text{bid}(p_i, r_j),$$

where μ_{exp}, μ_{wl}, and μ_{cf} denote the fuzzy membership values of expertise, workload, and conflict, respectively. The weights w_1, w_2, w_3 capture their relative importance, while α balance-fuzzy evaluation and bidding data.

D. Evolutionary Optimization Strategy

In our context, fitness evaluation balances three competing objectives: maximizing reviewer–paper satisfaction, minimizing workload imbalance, and ensuring fairness across reviewers. All hard constraints (each paper receives the required number of reviewers, no conflicts, reviewer capacity limits) are treated as mandatory; any violation produces an infeasible solution that is discarded. Among feasible solutions, fitness is computed from the fuzzy satisfaction model introduced in Section IV-C together with simple load statistics.

For clarity, let P_{r_j} denote the set of papers assigned to reviewer r_j under solution s, and let

$$N = \big| \{ (p_i, r_j) \in s \} \big|$$

be the total number of reviewer–paper assignments. The workload of reviewer r_j is defined as

$$w_j(s) = |P_{r_j}|,$$

and the workload vector is $w(s) = (w_1(s), \ldots, w_n(s))$. We use the standard deviation of $w(s)$ as the imbalance measure.

The three objective components are:

$$f_1(s) = -\frac{1}{N} \sum_{(p_i, r_j) \in s} s(p_i, r_j), \tag{1}$$

$$f_2(s) = \sigma\big(w(s)\big) = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \big(w_j(s) - \overline{w}(s)\big)^2}, \tag{2}$$

$$f_3(s) = - \min_{r_j :\, |P_{r_j}| > 0} \frac{1}{|P_{r_j}|} \sum_{p_i \in P_{r_j}} s(p_i, r_j), \tag{3}$$

where $s(p_i, r_j) \in [0, 1]$ is the fuzzy satisfaction score (Section IV-C), $\overline{w}(s)$ is the mean reviewer load, and the minimum in Equation (3) is taken only over reviewers with at least one assigned paper to avoid division by zero. The negative signs in f_1 and f_3 convert maximization of satisfaction into minimization because our optimization routine minimizes all objectives.

In practice, we normalize each objective component to comparable ranges prior to multi-objective optimization (for

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.

example, dividing f_1 by its maximum possible value and scaling f_2 by the number of papers) to prevent any single objective from dominating the search. NSGA-II then operates on the vector (f_1, f_2, f_3) to approximate the Pareto front of trade-off solutions that satisfy all hard constraints. To evolve the population, we employ genetic operators tailored to the integer-encoded representation of reviewer assignments, where each individual is a fixed-length vector with segments corresponding to the reviewers assigned to each paper. Crossover is implemented as a two-point operator applied at paper boundaries so that complete reviewer sets for selected papers are exchanged between parents, preserving assignment consistency while promoting diversity. Mutation is performed through a reviewer reassignment operator that replaces one reviewer in the vector with another valid candidate selected uniformly at random. Parent selection uses tournament selection to balance exploitation and exploration, and standard probabilities are applied for both crossover and mutation. Any infeasible offspring are discarded immediately to ensure all retained solutions remain valid.

V. Experimental Results

We evaluate the proposed NSGA-II-based framework for reviewer assignment at MIWAI 2025, which integrates semantic embedding and fuzzy modeling. The experiments employ real-world submission data to rigorously assess the effectiveness and robustness of the approach.

A. Dataset and Assigning Environment

For our experiments, the MIWAI conference provided essential information on submissions and reviewers, while additional publication data were automatically crawled to enrich the dataset.

The dataset includes titles and abstracts from the recent publications of 205 potential reviewers, averaging around ten publications each, along with abstracts and keywords from 287 papers submitted to MIWAI 2025. Reviewer assignments are determined primarily by the similarity between each paper and the reviewers' prior work, with bidding preferences from 55 reviewers given high priority. *Conflict of Interest* (COI) data are constructed from institutional affiliation comparisons (as of 2025) and co-authorship records from the preceding three years. Each paper is assigned between two and four reviewers, and no reviewer is assigned more than six papers to maintain balanced workloads. All experiments use the parameter settings summarized in Table II.

TABLE II
GA parameter settings used in all experiments

Parameter	Value
Population size	200
Number of generations	1000
Crossover probability	0.9
Mutation probability	0.35
λ_1	1.0
λ_2	1000.0

B. Results and Analysis

To assess the practical benefits of multi-objective optimization, we compare the proposed NSGA-II framework under three configurations:

- **Two-phased NSGA-II:** Phase 1 prioritizes reviewer bidding preferences, and Phase 2 completes the assignment using semantic similarity.
- **Single-phased NSGA-II:** All reviewers are assigned simultaneously in a single optimization process.
- **Ant Colony Optimization (ACO):** A baseline metaheuristic for comparison.

All methods enforce the same hard constraints and use the three-objective fitness evaluation described in Section III. Each experiment is repeated ten times to mitigate stochastic effects and ensure statistical reliability. Fitness scores are normalized and inverted for consistency with a maximization objective.

The aggregated results in Table III show that ACO yields the lowest fitness range, indicating weaker assignment quality. Both NSGA-II variants achieve higher overall fitness, but with clear differences: the two-phased NSGA-II improves bidding match rates by 19–24 percent, while the single-phased version yields 11–15 percent better semantic similarity. This trade-off allows organizers to weight preference satisfaction or topical relevance depending on internal priorities.

TABLE III
Normalized fitness score ranges across 10 runs

Algorithm	Min	Max
Two-phased NSGA-II	0.826	0.917
Single-phased NSGA-II	0.812	0.879
Ant Colony Opt.	0.693	0.772

Figure 2 shows the convergence behavior of the three methods. The two-phased NSGA-II consistently reaches the highest fitness and converges rapidly in early generations, indicating that incorporating bidding information provides a stronger initialization for the evolutionary search. The single-phased NSGA-II also improves steadily but stabilizes at a slightly lower fitness because it must optimize all objectives simultaneously. The ACO baseline converges more slowly and remains well below both NSGA-II variants, highlighting its limited ability to handle multi-objective reviewer assignment.

We also evaluate our framework against the reviewer assignment procedure used in MIWAI 2024, which relied on EasyChair's automated functions followed by manual adjustments from program chairs. For consistency, similarity scores for MIWAI 2024 are recomputed using the same method described earlier. As shown in Table IV, our approach provides substantially higher total similarity and improved per-paper similarity, demonstrating the practical advantages of integrating semantic modeling, fuzzy evaluation, and evolutionary optimization.

VI. Conclusion and Future Direction

In this paper, we presented a fuzzy multi-objective optimization approach based on NSGA-II to address the reviewer assignment problem at MIWAI 2025. The formulation

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.

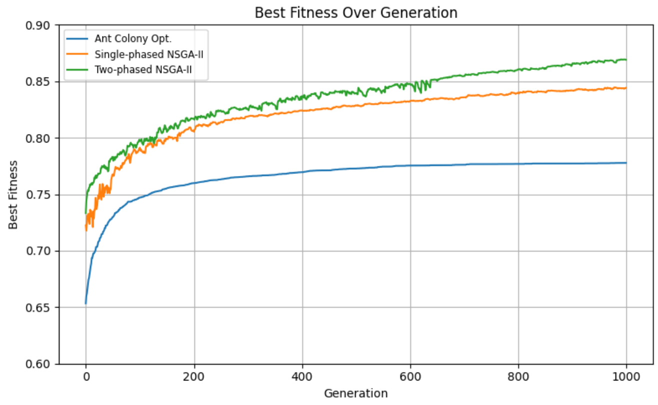

Fig. 2. Best fitness across generations for the three methods

TABLE IV
COMPARISON WITH THE MIWAI 2024 ASSIGNMENT PROCEDURE

	Our method	**Manual process**
Total similarity	120.436	56.483
Avg. similarity per paper	0.441	0.187

integrates hard constraints to ensure feasible assignments and soft objectives to improve expertise matching, workload balance, and fairness. Fuzzy set modeling is used to capture uncertainty in reviewer–paper suitability, while its integration within NSGA-II enables the generation of diverse trade-off solutions that program chairs can select according to their priorities. Experiments on the MIWAI 2025 dataset show that our method outperforms greedy and single-objective baselines, especially in improving semantic alignment and maintaining equitable workloads. The fuzzy representation also enhances realism in cases where expertise cannot be precisely quantified. Future extensions include scaling the framework to larger conferences, supporting dynamic or online assignment, and incorporating reviewer feedback or historical performance data to further improve assignment quality and fairness.

ACKNOWLEDGEMENT

We would like to express our gratitude to Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study. We are also sincerely grateful to Assoc. Prof. Tran Ngoc Thinh, Assoc. Prof. Pham Hoang Anh, and Dr. Panich Sudkhot for their assistance in providing the MIWAI data that made this research possible.

REFERENCES

[1] S. Price and P. A. Flach, "Computational support for academic peer review: a perspective from artificial intelligence," *Commun. ACM*, vol. 60, no. 3, p. 70–79, Feb. 2017. [Online]. Available: https://doi.org/10.1145/2979672

[2] D. Sculley, J. Snoek, and A. Wiltschko, "Avoiding a tragedy of the commons in the peer review process," 2018.

[3] L. Charlin and R. S. Zemel, "The Toronto Paper Matching System: An automated paper-reviewer assignment system," 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:680003

[4] D. Mimno and A. McCallum, "Expertise modeling for matching papers with reviewers," in *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 500–509. [Online]. Available: https://doi.org/10.1145/1281192.1281247

[5] M. Karimzadehgan and C. Zhai, "Constrained multi-aspect expertise matching for committee review assignment," in *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, ser. CIKM '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 1697–1700. [Online]. Available: https://doi.org/10.1145/1645953.1646207

[6] J. Jin, B. Niu, P. Ji, and Q. Geng, "An integer linear programming model of reviewer assignment with research interest considerations," *Annals of Operations Research*, vol. 291, no. 1, pp. 409–433, Aug 2020. [Online]. Available: https://doi.org/10.1007/s10479-018-2919-7

[7] I. Stelmakh, N. Shah, and A. Singh, "PeerReview4All: fair and accurate reviewer assignment in peer review," *J. Mach. Learn. Res.*, vol. 22, no. 1, Jan. 2021.

[8] J. Goldsmith and R. Sloan, "The AI Conference Paper Assignment Problem," *AAAI Workshop - Technical Report*, 01 2007.

[9] A. Kobren, B. Saha, and A. McCallum, "Paper Matching with Local Fairness Constraints," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. KDD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1247–1257. [Online]. Available: https://doi.org/10.1145/3292500.3330899

[10] Y. Kalmukov, "An algorithm for automatic assignment of reviewers to papers," *Scientometrics*, vol. 124, no. 3, pp. 1811–1850, Sep 2020. [Online]. Available: https://doi.org/10.1007/s11192-020-03519-0

[11] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 3982–3992. [Online]. Available: https://aclanthology.org/D19-1410/

[12] S. Hettich and M. J. Pazzani, "Mining for proposal reviewers: lessons learned at the national science foundation," in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 862–871. [Online]. Available: https://doi.org/10.1145/1150402.1150521

[13] L. Charlin, R. Zemel, and C. Boutilier, "A framework for optimizing paper matching," in *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, ser. UAI'11. Arlington, Virginia, USA: AUAI Press, 2011, p. 86–95.

[14] W. Tang, J. Tang, T. Lei, C. Tan, B. Gao, and T. Li, "On optimization of expertise matching with various constraints," *Neurocomput.*, vol. 76, no. 1, p. 71–83, Jan. 2012. [Online]. Available: https://doi.org/10.1016/j.neucom.2011.04.039

[15] X. Zhao and Y. Zhang, "Reviewer assignment algorithms for peer review automation: A survey," *Information Processing & Management*, vol. 59, no. 5, p. 103028, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306457322001388

[16] S. Zhao, D. Zhang, Z. Duan, J. Chen, Y.-p. Zhang, and J. Tang, "A novel classification method for paper-reviewer recommendation," *Scientometrics*, vol. 115, no. 3, pp. 1293–1313, Jun 2018. [Online]. Available: https://doi.org/10.1007/s11192-018-2726-6

[17] Z. Duan, S. Tan, S. Zhao, Q. Wang, J. Chen, and Y. Zhang, "Reviewer assignment based on sentence pair modeling," *Neurocomputing*, vol. 366, pp. 97–108, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231219309828

[18] K. Leyton-Brown, Mausam, Y. Nandwani, H. Zarkoob, C. Cameron, N. Newman, and D. Raghu, "Matching papers and reviewers at large conferences," *Artif. Intell.*, vol. 331, no. C, Jun. 2024. [Online]. Available: https://doi.org/10.1016/j.artint.2024.104119

[19] Y. Xu, J. Ma, Y. Sun, G. Hao, W. Xu, and D. Zhao, "A decision support approach for assigning reviewers to proposals," *Expert Systems with Applications*, vol. 37, no. 10, pp. 6948–6956, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417410002083

[20] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022.

Authorized licensed use limited to: RMIT University Library. Downloaded on February 21,2026 at 12:18:57 UTC from IEEE Xplore. Restrictions apply.