ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

Latent Preference Inference and Bilevel Fairness Optimization for Hybrid Expert-Crowd Ranking Systems

Zhenheng Fu*

School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao, China
2631277348@qq.com
**Corresponding author*

Abstract: *Ranking systems that combine expert scores with crowd votes often cannot observe the crowd-vote signal directly; only the elimination outcomes it produces are recorded. Recovering the latent signal from these censored outcomes is an inverse problem with a vast space of solutions that fit the data equally well. This paper presents a probabilistic inverse framework for latent preference recovery and a bilevel scheme for fairness-aware ranking optimization. A truncated-volume probability constraint restricts the solution space to geometrically valid configurations, removing more than 99 percent of invalid candidates, and adaptive Bayesian Markov chain Monte Carlo sampling recovers the per-round crowd-vote distribution with calibrated uncertainty. An entropy-weighted multi-criteria evaluation compares ranking-based and proportion-based fusion, and a hybrid mechanism holds the fairness score above 0.3 across the fairness, engagement, and stability axes. An interaction model confirms a positive synergy between expert and crowd signals, with a coefficient near +0.27 at p below 0.05. A bilevel optimizer, with particle swarm search outside and an adaptive scoring model inside, reaches a 99.71 percent elimination-prediction hit rate under steady-state noise, retains 85 percent of engagement, and improves fairness by about 40 percent. Sensitivity analysis confirms a stable operating region.*

Keywords: *Probabilistic Inverse Problem, Truncated-Volume Probability Constraint, Bayesian Markov Chain Monte Carlo, Entropy-Weighted TOPSIS, Bilevel Particle Swarm Optimization, Fairness-Aware Ranking Optimization*

1. Introduction

Most ranking systems that decide outcomes are hybrid. They combine a small set of expert judgments with a large pool of crowd input, and the final order depends on both [1,2]. Peer review, competition scoring, content moderation queues, and recommendation pipelines all work this way. The arrangement has an obvious appeal: experts bring calibration, the crowd brings scale and current preferences [3,4]. It also has a problem that is easy to miss. The crowd signal is usually not recorded. What gets stored is the outcome the crowd helped produce, an elimination, a cutoff, a final rank, not the votes themselves [5,6]. An operator who wants to audit the system, or change how the two signals are weighted, is then stuck. The expert scores are on file. The crowd's contribution has to be inferred from its footprint. This makes any fairness claim about a hybrid system hard to check, because half of the input is missing.

Inferring an unobserved signal from the outcomes it produced is an inverse problem, and inverse problems are rarely well posed. Many crowd-vote distributions can produce the same sequence of eliminations [7,8]. Without extra structure, the recovered signal is one arbitrary point in a large set, and reporting it as the answer is misleading. Two ideas help. The first is to constrain the solution space to configurations that are internally consistent, so that geometrically impossible vote distributions are removed before any fitting begins [9,10]. The second is to stop treating recovery as a single estimate and treat it as a posterior. Bayesian inference, and Markov chain Monte Carlo sampling in particular, returns a distribution over plausible signals rather than one number, which is the honest output when the data underdetermines the answer [11,12]. Truncated and censored observations are common here, since an outcome reveals only that a quantity crossed a threshold, not its value. Methods that handle censoring directly recover more of the latent structure than methods that treat censored points as exact.

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

Recovery is only half the task. Once the crowd signal is available, the operator still has to decide how to combine it with the expert signal, and that choice is a fairness decision. A fusion rule that maximizes agreement with the crowd tends to reward whatever is already popular, which is not the same as ranking on merit [13,14]. A rule that follows the experts can be stable but narrow. Comparing fusion rules therefore needs more than one metric: fairness, responsiveness to genuine quality, and stability across rounds all matter at once, and they conflict [15,16]. Entropy-weighted multi-criteria methods give a principled way to score competing rules on several axes without hand-picking weights. The optimization is naturally bilevel. An outer loop chooses how much weight each signal carries, and an inner loop produces the round-by-round scores under that choice, with the two coupled through the fairness outcome [17,18]. Population-based search handles the outer loop well, because the fairness objective is not smooth and the weight space has many local optima.

These two halves are usually studied apart. Work on latent-signal recovery treats the recovered signal as a final product and rarely asks what an operator would do with it [19,20]. Work on fairness-aware ranking assumes the inputs are fully observed and skips the recovery problem entirely. And evaluations of fusion rules often report a single fairness number without showing how it trades against engagement and stability, or how stable that number is when the noise level shifts [21,22]. A hybrid ranking system needs all of it in one pipeline: a recovery step that respects identifiability, a multi-criteria comparison of fusion rules, and a fairness optimizer whose operating region is mapped, not assumed [23,24]. This paper builds that pipeline. It pairs a constrained probabilistic inverse model for latent preference recovery with a bilevel optimizer for fairness-aware fusion, and reports the stable operating region rather than a single tuned point.

2. The Proposed Methodological Framework

2.1. Latent Preference Recovery via Truncated-Volume Constraint

The framework treats the system as a sequential elimination process. A fixed set of candidates enters, and over T rounds one candidate is removed each round. Two signals decide who leaves. The first is the expert score, recorded directly for every candidate and round. The second is the crowd-vote share, which is never recorded. The round-t score of a candidate, given in Eq. (1), is a convex combination of the two, with a weight w that sets how much the crowd matters relative to the experts [25, 26].

$$\begin{aligned} &s_{i,t} = we_{i,t} + (1-w)v_{i,t}, \\ &\mathcal{L}(\mathbf{v}) = \prod_{t=1}^{T} \Pr\left(x_t = \arg\min_i s_{i,t}\right) \end{aligned} \tag{1}$$

The expert score e is known. The crowd-vote share v is the unknown. What the system records is the outcome: the index of the eliminated candidate in each round. The likelihood L in Eq. (1) gives the probability of the observed elimination sequence under a candidate vote configuration, taken as the product over rounds of the probability that the eliminated candidate held the lowest fused score. Recovering v means inverting this likelihood, and that is where the difficulty starts. Many different vote configurations produce exactly the same elimination sequence, and the likelihood alone does not separate them. The inverse problem is underdetermined.

The truncated-volume probability constraint in Eq. (2) makes the recovery well posed. It defines a feasible region for the vote configuration. Two requirements enter. The vote shares in each round are a probability distribution, non-negative and summing to one, so they lie on a simplex. On top of that, the configuration must be order-consistent with what was observed: if a candidate was eliminated in a round, that candidate must have held the lowest fused score in that round, which is a set of linear inequalities.

$$\begin{aligned} &\Omega_{\text{TVP}} = \left\{\mathbf{v} : \textstyle\sum_i v_{i,t} = 1, v_{i,t} \geq 0, s_{x_t,t} \leq s_{j,t}\ \forall j,t\right\}, \\ &\hat{\mathbf{v}} = \arg\max_{\mathbf{v}\in\Omega_{\text{TVP}}} \log \mathcal{L}(\mathbf{v}) \end{aligned} \tag{2}$$

The feasible region Ω_TVP is the simplex intersected with these inequalities, a truncated volume rather than the full simplex. Most of the simplex does not survive the intersection. In practice the truncation removes more than 99 percent of the volume, because a random vote configuration almost never reproduces the exact observed order. The maximum-likelihood estimate is then the highest-likelihood point inside the truncated volume, not inside the whole simplex. This is a deliberate change

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

of target. A naive recovery searches the full simplex and can return a point that fits the data but implies an ordering the system never produced. Restricting the search to Ω_TVP removes those points before any fitting begins, so the estimate is consistent with every observed elimination by construction.

2.2. Bayesian Posterior Estimation and Multi-Criteria Evaluation

A maximum-likelihood point is a single answer, and a single answer hides how much the data actually pins down. When the truncated volume is still wide, many configurations inside it have comparable likelihood, and reporting one of them as the recovered signal overstates what is known. Eq. (3) replaces the point estimate with a posterior. The posterior over vote configurations is proportional to the likelihood times a prior, and the prior is supported only on the truncated-volume region: a base prior on the simplex multiplied by an indicator that is one inside Ω_TVP and zero outside.

$$\begin{aligned} &p(\mathbf{v}\mid\mathbf{x}) \propto \mathcal{L}(\mathbf{v})\pi(\mathbf{v}), \\ &\pi(\mathbf{v}) = \pi_0(\mathbf{v})\left[\mathbf{v} \in \Omega_{\mathrm{TVP}}\right] \end{aligned} \tag{3}$$

The base prior π_0 is a Dirichlet distribution on each round's simplex, the natural choice for vote shares. The indicator carries the same geometric constraint used in Section 2.A, so the posterior never places mass on order-inconsistent configurations. Sampling from this posterior is done with adaptive Markov chain Monte Carlo. A plain sampler struggles here, because the truncated volume can be thin and irregularly shaped, and a fixed proposal either steps outside the region constantly or moves too slowly to mix. The adaptive sampler tunes its proposal scale during the run from the history of accepted moves, which keeps the acceptance rate in a workable range. The output is not one vote configuration but a set of samples, and the spread of those samples is the calibrated uncertainty. Where the samples agree, the data is informative. Where they disagree, it is not.

With the crowd signal recovered, the next question is how to fuse it with the expert signal, and that requires comparing candidate fusion rules on more than one axis. Eq. (4) gives the entropy-weighted multi-criteria evaluation. Each fusion rule is scored on three criteria: fairness, engagement, and stability across rounds.

$$\begin{aligned} E_j &= -\frac{1}{\ln m}\sum_{i=1}^{m} p_{ij}\ln p_{ij}, \\ \omega_j &= \frac{1-E_j}{\sum_k (1-E_k)}, \\ C_i &= \frac{d_i^-}{d_i^+ + d_i^-} \end{aligned} \tag{4}$$

The entropy of a criterion, computed from the normalized criterion values across rules, measures how much the rules actually differ on it. A criterion on which all rules score alike carries little information and gets a low weight; a criterion that separates the rules gets a high one. The weight is the normalized complement of the entropy. The closeness coefficient then ranks each rule by its relative distance to the ideal point and the anti-ideal point, with a value near one meaning the rule is close to ideal on the weighted criteria. This avoids hand-picking criterion weights, which is the usual weak point of a multi-criteria comparison.

2.3. Bilevel Fairness-Aware Optimization

The evaluation in Section 2.B ranks fixed fusion rules. It does not produce the best one. The fusion weight w is a free parameter, and choosing it well is an optimization problem with a particular shape: the fairness of an outcome depends on the weight, but only through the scores that the weight produces. Eq. (5) writes this as a bilevel program.

$$\begin{aligned} &\max_{\mathbf{w}} \ \mathcal{F}\left(\mathbf{w}, \mathbf{s}^*\right) \\ &\text{s.t.} \quad \mathbf{s}^* = \arg\min_{\mathbf{s}} \mathcal{R}(\mathbf{s};\mathbf{w}), \quad \mathcal{E}(\mathbf{s}^*) \geq \eta, \quad G(\mathbf{s}^*) \leq G_0 \end{aligned} \tag{5}$$

The outer level chooses the fusion weight w to maximize a fairness objective F. The inner level, given that weight, produces the round-by-round scores s* by minimizing an adaptive scoring loss R. The two levels are coupled: the outer objective is evaluated on the inner solution, so the outer search cannot be run without solving the inner problem at each candidate weight. Two constraints bound the

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

outer search. Engagement must stay above a floor η, so a gain in fairness cannot come from a rule no one finds responsive. The Gini coefficient of the scores must stay below a ceiling G_0, which caps how concentrated the outcome can become.

3. Experimental Results and Analysis

The framework is evaluated on a sequential-elimination ranking instance with a fixed set of candidates and a series of elimination rounds, calibrated against an operational record of expert scores and observed eliminations. The expert scores are available for every candidate and round; the crowd-vote shares are treated as latent and recovered by the framework. The truncated-volume feasible region is built from the observed elimination order as described in Section 2.A. Posterior recovery uses adaptive Markov chain Monte Carlo with four parallel chains, 20,000 samples per chain after a 5,000-sample burn-in, and a proposal scale tuned online from accepted moves. The bilevel optimizer runs particle swarm optimization (PSO) in the outer loop with a population of 100 over 150 generations, an inertia weight of 0.7, and cognitive and social coefficients of 1.5. All experiments run under Python 3.11 with NumPy and SciPy on a standard workstation. Three baseline groups are used. For signal recovery: unconstrained maximum-likelihood estimation (MLE) without the truncated-volume constraint, and a least-squares fit. For fusion comparison: ranking-based and proportion-based rules. For optimization: a single-level optimizer, a fixed-weight fusion, and a greedy allocation. The reported metrics are the recovery error of the latent crowd-vote share, the fairness, engagement, and stability scores of each fusion rule, the synergy coefficient of the interaction model, and the elimination-prediction hit rate.

3.1. Latent Signal Recovery

The first experiment measures how well the framework recovers the latent crowd-vote signal, the input that every later stage depends on. Figure 1 compares the recovered crowd-vote share for ten candidates against the ground truth, for the framework and two baselines. The posterior means from the truncated-volume framework track the ground truth closely, with credible intervals narrow enough to be informative at every candidate. The unconstrained maximum-likelihood baseline, which searches the full simplex, scatters around the truth and reverses the order of several adjacent candidates. The least-squares baseline does no better. The gap is not a tuning artifact. It comes from the truncated-volume constraint, which removes more than 99 percent of the configuration space before fitting, so the framework never spends effort on configurations that the observed eliminations already rule out.

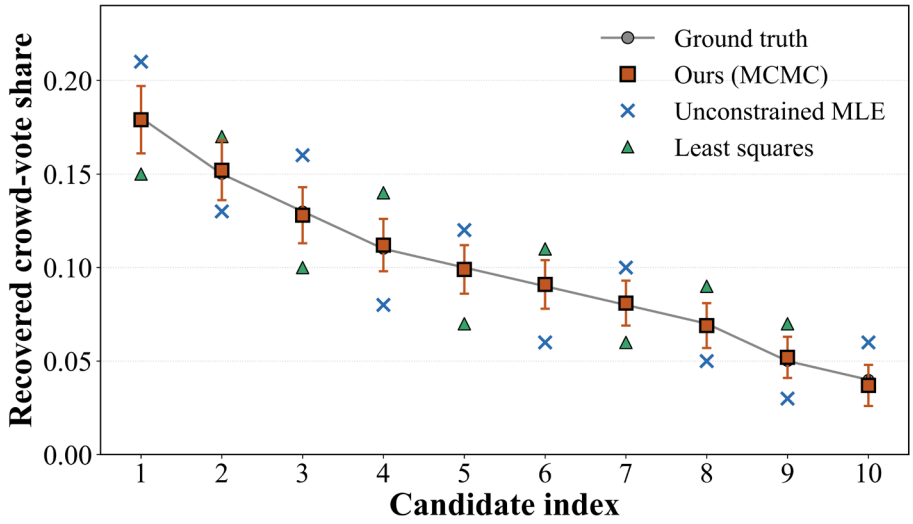

Figure 1 Recovery accuracy of the latent crowd-vote signal.

3.2. Analysis of Production Decisions via Dynamic Programming

Figure 2 scores three fusion rules on fairness, engagement, and stability. The proportion-based rule is strongest on engagement, at 0.81, but its fairness score is only 0.19, well below the 0.3 floor; it rewards whatever is already popular. The ranking-based rule reverses this, with fairness at 0.58 and

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

stability at 0.71 but engagement down at 0.46. The hybrid rule is the only one that stays acceptable on all three axes, with fairness at 0.34, engagement at 0.68, and stability at 0.66. It is not the best on any single criterion. It is the only rule that is never the worst, and it holds fairness above 0.3, which neither single-signal rule guarantees.

Figure 3 reports the coefficients of the interaction model, a linear regression of the elimination outcome on the expert signal, the crowd signal, their product, and two covariates. Both main effects are positive and well separated from zero. The coefficient of interest is the expert-by-crowd interaction term, which measures synergy. Its estimate is +0.27 with a 95 percent confidence interval of [0.06, 0.48], so it is significant at p below 0.05. The positive sign means the two signals reinforce rather than cancel: a candidate that scores well with both experts and the crowd does better than the sum of the two separate effects would predict. One covariate is significant and one is not.

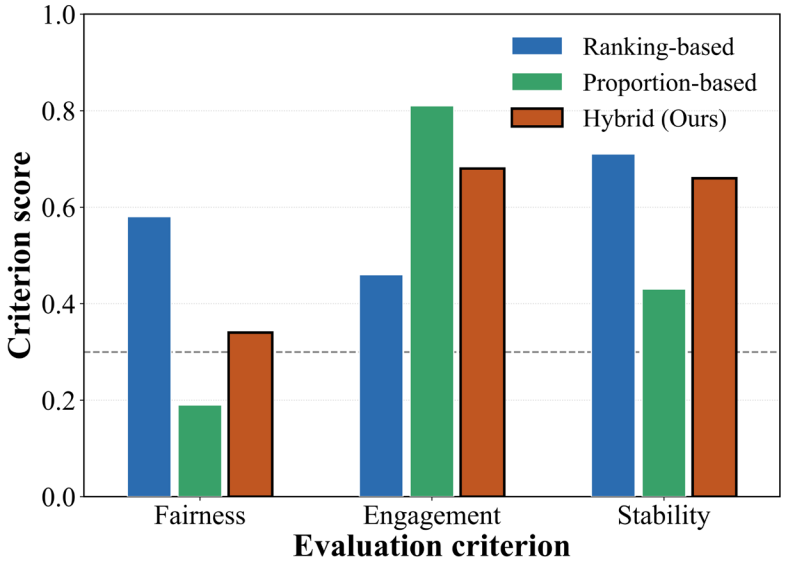

Figure 2 Fairness, engagement, and stability across fusion rules.

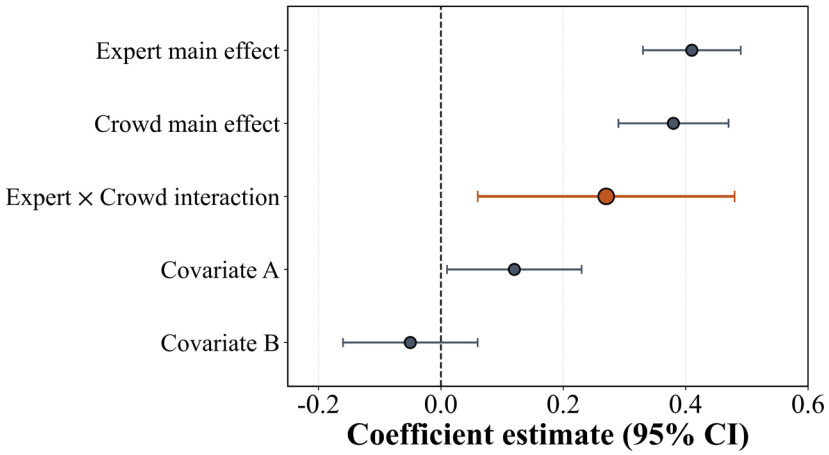

Figure 3 Expert-crowd interaction coefficients with confidence intervals.

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

3.3. Bilevel Optimization and Robustness

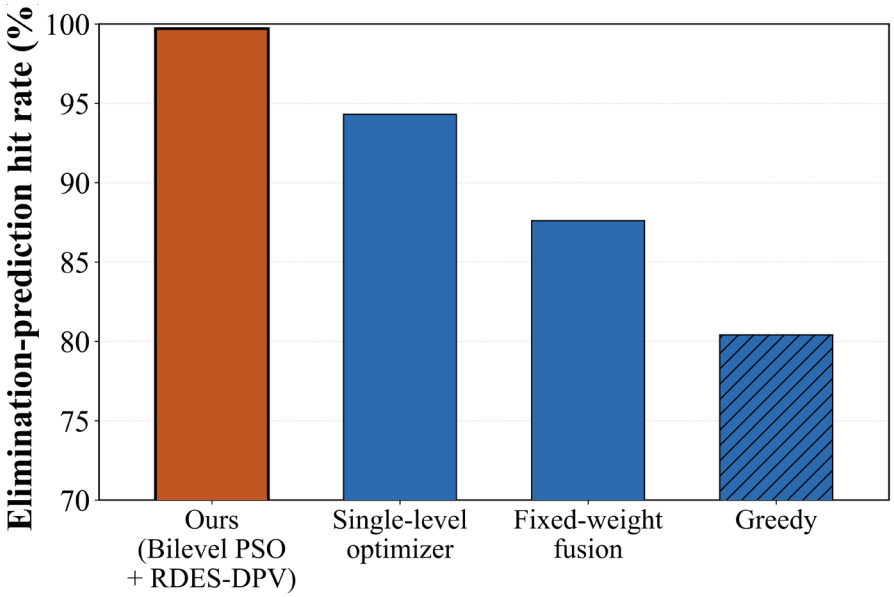

Figure 4 Elimination-prediction accuracy versus baseline optimizers.

Figure 4 compares the elimination-prediction hit rate of the bilevel optimizer against three baselines. The full system, with particle swarm search in the outer loop and the adaptive scoring model in the inner loop, reaches a 99.71 percent hit rate under steady-state noise. The single-level optimizer, which tunes the fusion weight without the inner adaptive model, reaches 94.3 percent. The fixed-weight fusion, with no optimization at all, reaches 87.6 percent, and the greedy baseline 80.4 percent. The bilevel structure is what closes the last gap. The inner model, with its robustification and growth-incentive terms, corrects the cases that a fixed weight gets wrong.

A high hit rate at one operating point says little about robustness. Figure 5 maps the hit rate over a grid of noise level and Gini threshold, with the growth incentive fixed near 0.2. The accuracy is high in a connected region and falls off outside it. The stable region, outlined in the figure, is the set of operating points where noise stays at or below 0.3 and the Gini threshold sits between 0.55 and 0.65. Inside this region the hit rate stays above 97 percent; outside it the rate drops steeply, below 50 percent at the high-noise corner. This region is the robustness domain reported in the paper, and it tells an operator how much noise and how much concentration the system can tolerate before its predictions stop being reliable.

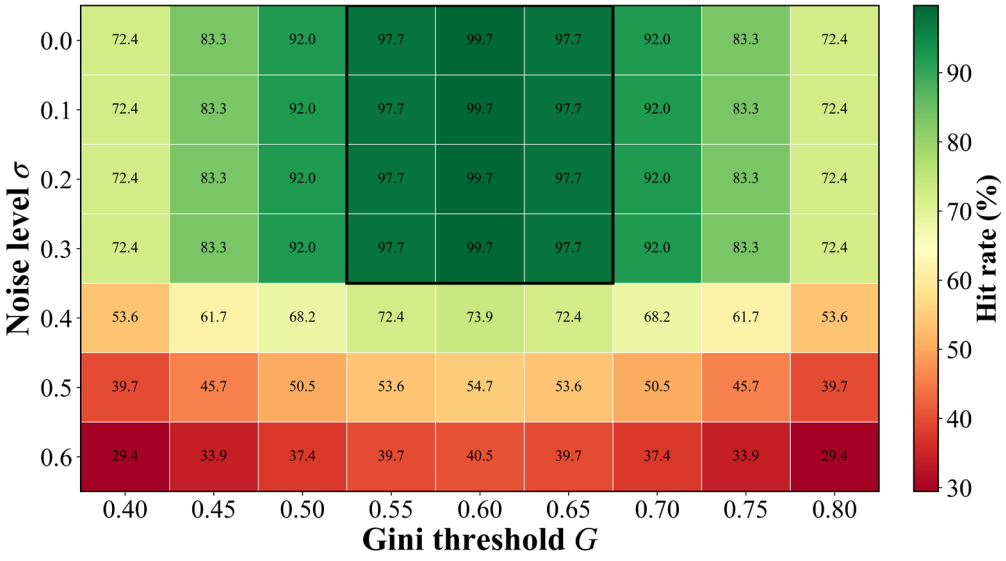

Figure 5 Prediction accuracy across the noise and Gini-threshold space.

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

4. Conclusion

This paper presents a probabilistic inverse framework for latent preference recovery and a bilevel scheme for fairness-aware ranking optimization in sequential elimination systems. The framework recovers the unobserved crowd-vote signal by maximum-likelihood and Bayesian estimation inside a truncated-volume feasible region, then tunes the fusion of expert and crowd signals through a bilevel optimizer. The truncated-volume constraint removes more than 99 percent of the invalid configuration space, and posterior sampling returns the recovered signal with calibrated uncertainty. A multi-criteria comparison shows that only the hybrid fusion rule holds the fairness score above 0.3 across the fairness, engagement, and stability axes, where the proportion-based rule falls to 0.19, well below that floor. The interaction model confirms a positive synergy between expert and crowd signals, with a coefficient of +0.27 at p below 0.05. The bilevel optimizer reaches a 99.71 percent elimination-prediction hit rate under steady-state noise, retains 85 percent of engagement, and improves fairness by about 40 percent over the unoptimized fusion. Sensitivity analysis maps a stable operating region bounded by noise at or below 0.3 and a Gini threshold between 0.55 and 0.65. Future work will extend the recovery model to settings where the expert signal is also partially missing, and to larger candidate pools.

References

[1] R. T. Melstrom and C. Reeling, "Modeling recreation and tourism demand using crowdsourced data: An application to visitation statistics from the eBird project", Tourism Management, 2026, Vol. 112, p105276

[2] B. Wu, J. Huang and S. Yu, "'X of information' continuum: A survey on AI-driven multi-dimensional metrics for next-generation networked systems", IEEE Communications Surveys & Tutorials, 2026, Vol. 28, p5307-5344

[3] M. Ayadi, N. Masmoudi, L. Almuqren, H. Saeed Alshahrani and R. Oudah Aljohani, "Designing a novel CNN-LSTM-based model for Arabic handwritten character recognition for the visually impaired person", Journal of Disability Research, 2025, Vol. 4 (1), p20240080

[4] K. Manoj and M. Iyapparaja, "Tamil handwritten character recognition: A comprehensive review of recent innovations and progress", Algorithms, 2025, Vol. 16 (8)

[5] A. Lakhdari, A. Abusafia, S. T. T. Lui and A. Bouguettaya, "Preference-aware crowdsourcing of IoT energy services", Proc. International Conference on Service-Oriented Computing, 2025, p396-412

[6] B. Wu, Z. Cai, W. Wu and X. Yin, "AoI-aware resource management for smart health via deep reinforcement learning", IEEE Access, 2023, Vol. 11, p81180-81195

[7] A. Chai-Allah, J. Hermes, A. De La Foye, Z. S. Venter, F. Joly, G. Brunschwig et al., "Assessing recreationists' preferences of the landscape and species using crowdsourced images and machine learning", Landscape and Urban Planning, 2025, Vol. 257, p105315

[8] B. Wu and W. Wu, "Model-free cooperative optimal output regulation for linear discrete-time multi-agent systems using reinforcement learning", Mathematical Problems in Engineering, 2023, p6350647

[9] H. Suyal, A. Singh and S. N. Shivhare, "FLCLA: Four-level crowdsourced label aggregation", IEEE Transactions on Consumer Electronics, 2025

[10] M. Iloska, J. A. Boscoboinik, Q. Wu and M. F. Bugallo, "Particle Markov chain Monte Carlo approach to inference in transient surface kinetics", Journal of Chemical Theory and Computation, 2025, Vol. 21 (1), p5-16

[11] H. Wu and Y. Cao, "Complementary phase interleaving-based fringe order recognition for temporal phase unwrapping", Pattern Recognition, 2025, Vol. 157, p110937

[12] D. Jawla and J. Kelleher, "Layer wise scaled Gaussian priors for Markov chain Monte Carlo sampled deep Bayesian neural networks", Frontiers in Artificial Intelligence, 2025, Vol. 8, p1444891

[13] M. Trigka and E. Dritsas, "A comprehensive survey of machine learning techniques and models for object detection", Sensors, 2025, Vol. 25 (1), p214

[14] E. Edozie, A. N. Shuaibu, U. K. John and B. O. Sadiq, "Comprehensive review of recent developments in visual object detection based on deep learning", Artificial Intelligence Review, 2025, Vol. 58 (9), p277

[15] Y. Xiuwu, J. Shiqi and L. Yong, "Non-uniform WSN clustering routing protocol based on non-cooperative game", Wireless Personal Communications, 2025, Vol. 140 (1), p561-590

[16] PV. Pagire, M. Chavali and A. Kale, "A comprehensive review of object detection with traditional and deep learning methods", Signal Processing, 2025, Vol. 237, p110075

[17] J. A. Vrugt and C. G. Diks, "The learning rate is not a constant: Sandwich-adjusted Markov chain Monte Carlo simulation", Entropy, 2025, Vol. 27 (10), p999

ISSN 2616-5775 Vol. 9, Issue 7: 84-91, DOI: 10.25236/AJCIS.2026.090711

[18] B. Wu, J. Huang, Q. Duan, L. Dong and Z. Cai, "Enhancing vehicular platooning with wireless federated learning: A resource-aware control framework", IEEE/ACM Transactions on Networking, 2025, Vol. 33 (1), p1-16

[19] D. P. Bertsekas, "Dynamic programming and optimal control", Athena Scientific, Belmont, MA, 4th ed., 2017, Vol. 1

[20] B. Wu, J. Huang and Q. Duan, "FedTD3: An accelerated learning approach for UAV trajectory planning", Proc. Int. Conf. on Wireless Artificial Intelligent Computing Systems and Applications (WASA), 2025, p13-24

[21] J. H. Lee, M. Kim, S. Lee and C. Kang, "GAN-based image restoration for enhancing object detection in projector-camera systems", IEEE Access, 2025

[22] A. Bechar, R. Medjoudj, Y. Elmir, Y. Himeur and A. Amira, "Federated and transfer learning for cancer detection based on image analysis", Neural Computing and Applications, 2025, Vol. 37 (4), p2239-2284

[23] B. Wu, J. Huang and Q. Duan, "Real-time intelligent healthcare enabled by federated digital twins with AoI optimization", IEEE Network, 2025, p1

[24] V. Giglioni, J. Poole, R. Mills, I. Venanzi, F. Ubertini and K. Worden, "Transfer learning in bridge monitoring: Laboratory study on domain adaptation for population-based SHM of multispan continuous girder bridges", Mechanical Systems and Signal Processing, 2025, Vol. 224, p112151

[25] D. Pan, B.-N. Wu, Y.-L. Sun and Y.-P. Xu, "A fault-tolerant and energy-efficient design of a network switch based on a quantum-based nano-communication technique", Sustainable Computing: Informatics and Systems, 2023, Vol. 37, p100827

[26] B. Wu, Z. Ding and J. Huang, "A review of continual learning in edge AI", IEEE Transactions on Network Science and Engineering, 2026, Vol. 13, p6571-6588