

Introduction (Example: Observing multi-labels for images)

- Multi-label feature selection improves the classification performance of multi-label data by removing redundant and irrelevant features

Multi-label Feature Selection

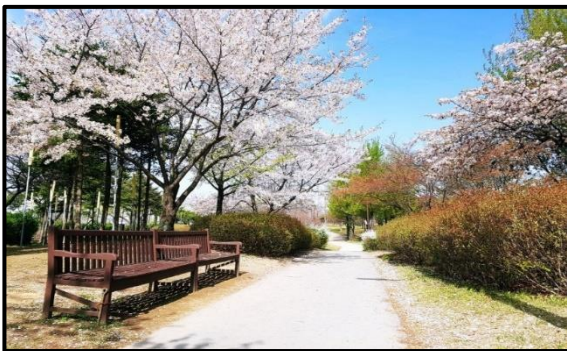


Fig. 1



Fig. 2



Fig. 3

redundant features

: Requires only one

Irrelevant features

: Not applicable
for classification

Fig. 1

Fig. 2

Fig. 3

Street light	Bench	Bird	Cherry Blossom	Snow	Sand
O	O	X	O	X	X
O	O	X	X	O	X
X	X	X	X	O	O

Multi-label = $\{l_1, l_2, \dots, l_{|L|}\}$

Fig. 1 (spring, park)

Fig. 2 (winter, park)

Fig. 3 (winter, sea)

Selected Features

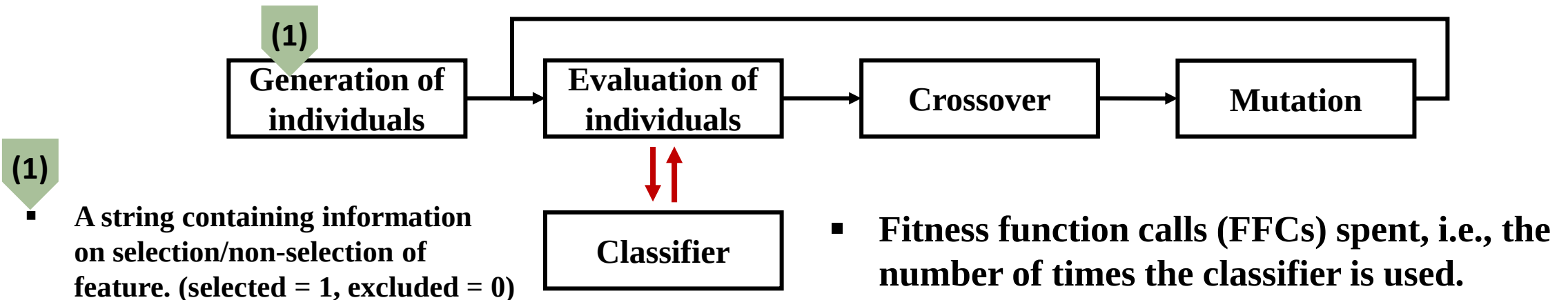
Bench	Cherry Blossom	Snow	Sand
O	O	X	X
O	X	O	X
X	X	O	O

Types of Multi-label Feature Selection

- An individual refers to a character string containing information on the selection/non-selection of a feature. For the earlier example, the selected individuals are expressed as depicted on the right.

Bench	...	Snow	Sand
1	...	0	1

Wrapper method (Example: genetic algorithm) [1, 2, 3]



Filtering method (Example: mutual information measures) [4, 5, 6]

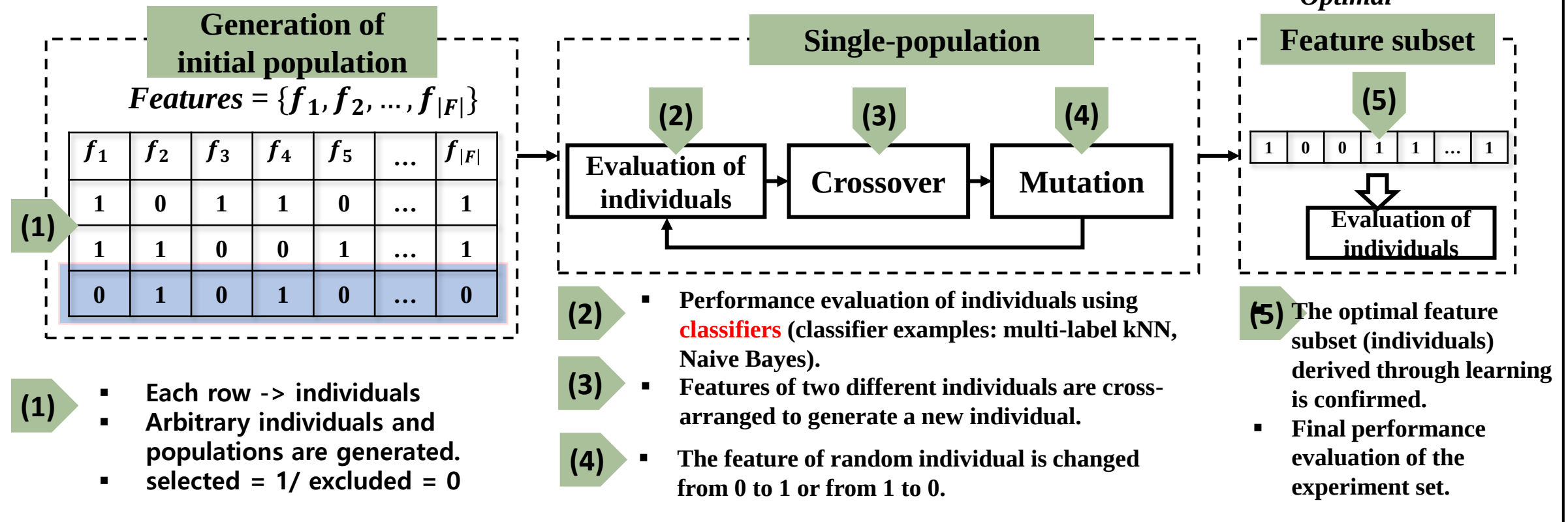
$$I(x, y) = H(x) - H(x|y) = H(y) - H(y|x)$$

- Mutual information measures: a method that measures two random dependent variables.

Multi-label Feature Wrapper (Example: Single-population Genetic Algorithm)

- Genetic algorithm method has been often used for solving feature selection problems as it provides good solutions based on randomness.
- Design issue 1: Prevent premature convergence of population until the optimal solution is obtained
- Design issue 2: A structure that can quickly introduce important features to the population.

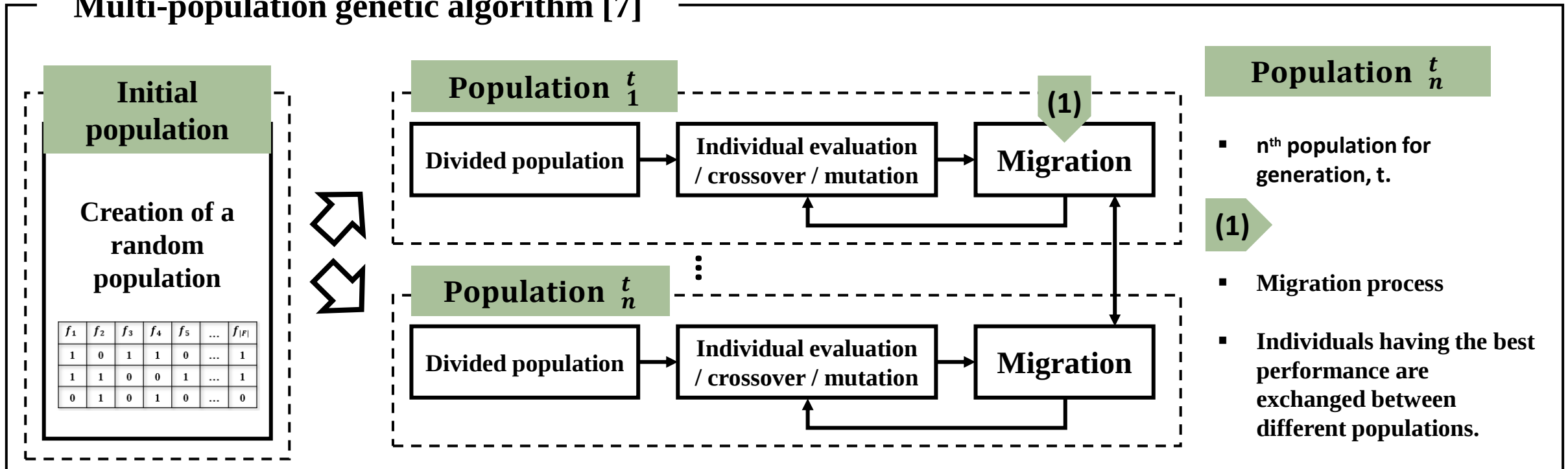
Single-population genetic algorithm



Introduction of Multi-population : Preservation of Population Diversity

- Traditional genetic algorithms are weaker in preserving the diversity of populations due to the characteristics of single-population.
- Multi-population genetic algorithms perform search by dividing the population into several independent populations. The advantage of this is that even if a certain population is converged prematurely, the remaining populations are unaffected.
- Feature set information of different populations are obtained through migration. Consequently, each population is naturally composed of various individuals.

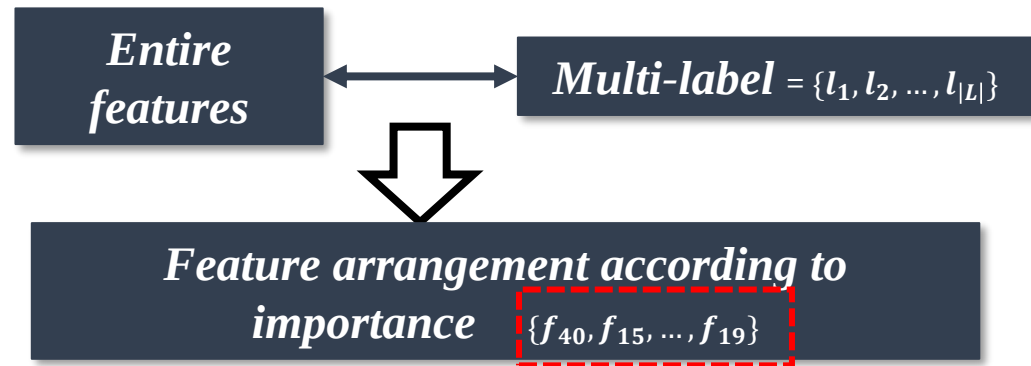
Multi-population genetic algorithm [7]



Introduction of Refinement : Strengthening Search Performance

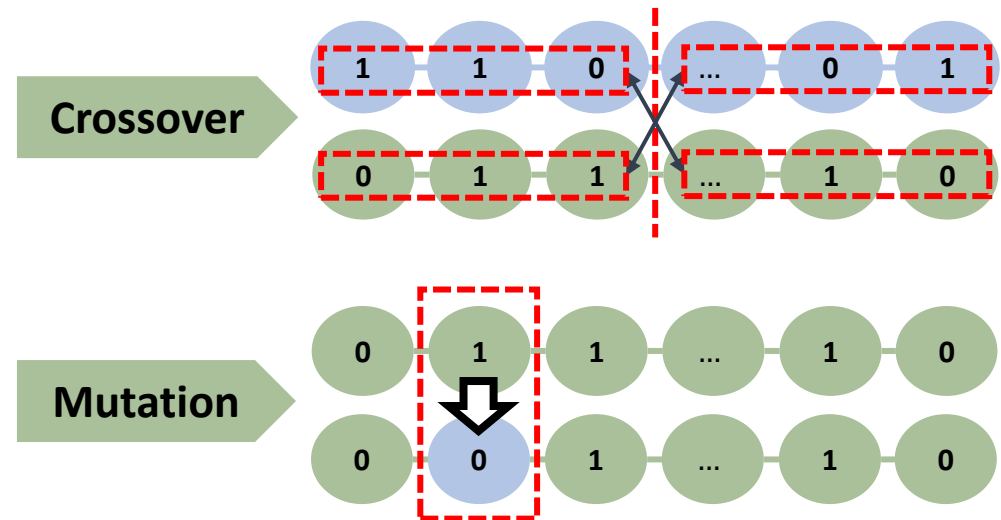
- Traditional genetic algorithms have difficulty introducing features related to multi-labels.
- Refinement calculates the ranking of features by a separate calculation method and applies it to the existing individuals, to create new individuals.
- The refinement is generally applied to individuals having high performance rankings in a population. It displays superior performance compared to simple evaluation/crossover/mutation.

Introduction of Refinement [8, 9]



- Feature arrangement and feature extraction based on relevance with multi-labels.
- Refinement : inflow of features that are highly related to the labels of existing individuals of the population.

Evolution algorithm: difficulty in fine-tuning

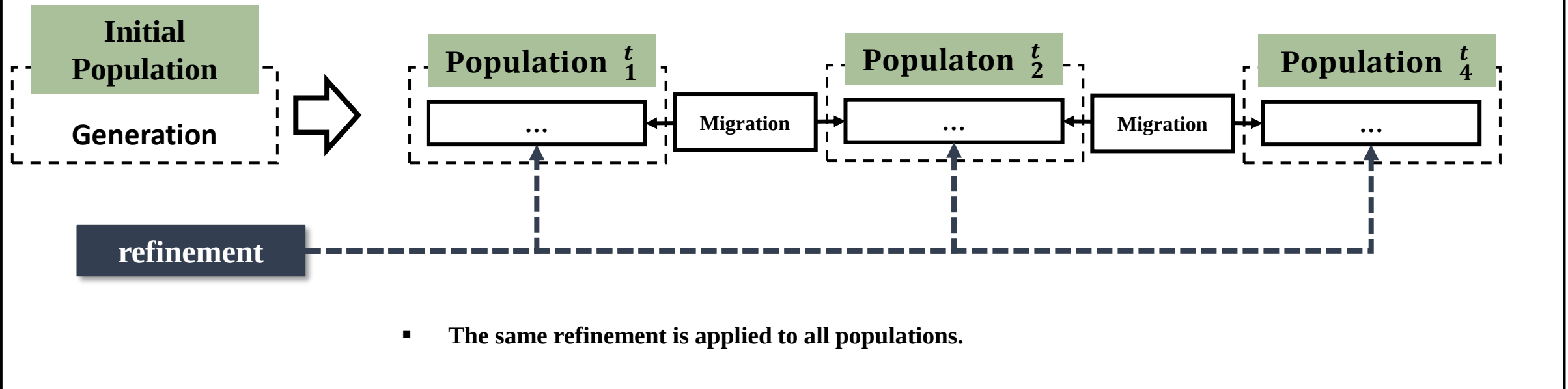


- Crossover & mutation process application: this method selects arbitrary crossover positions and features to create new individuals, and it has trouble fine-tuning based on randomness.

Research Motivation

- The addition of refinement to multi-label feature selection utilizing multi-population genetic algorithm vastly improves performance, despite no related researches.
- **If refinement is applied to each population according to the conventional techniques, it is difficult to quickly preserve the diversity as the generation passes.** Thus, an effective search is difficult.
- Therefore, a method to introduce refinement without losing the ability to preserve diversity is essential, which is an advantage of multi-population genetic algorithms.

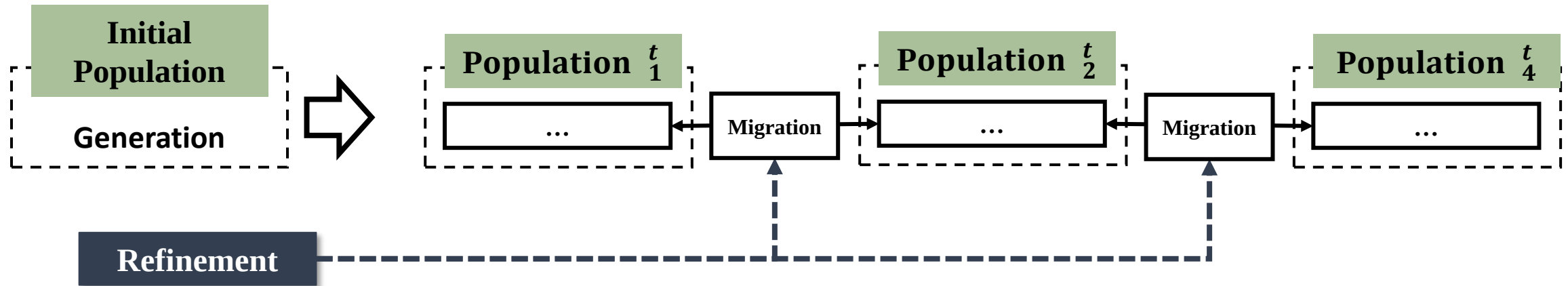
Problem



Proposed Method

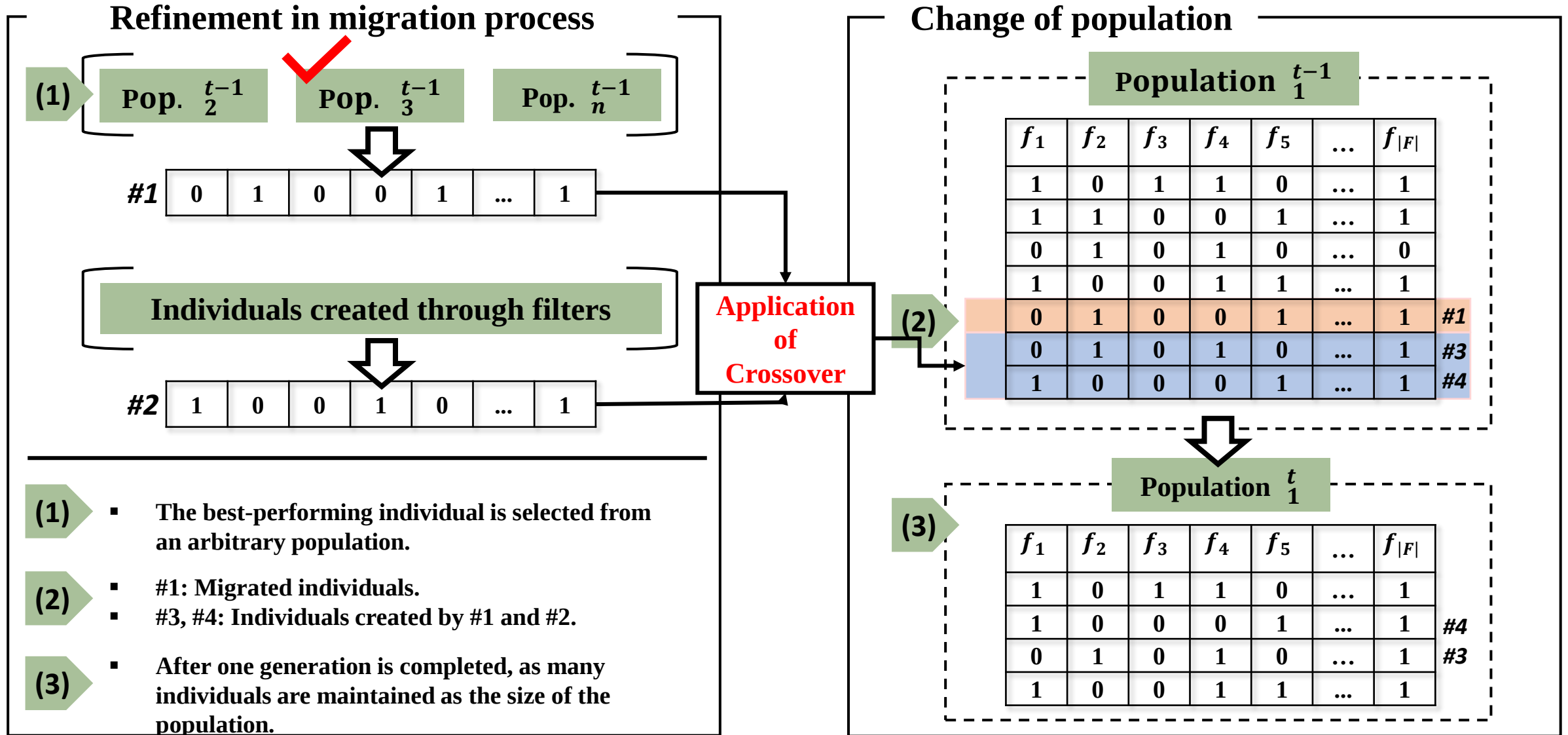
- Refinement for migrated individuals is applied as an effective refinement method for the multi-label genetic algorithm.

Application of refinement to the migrated individuals between populations



- The refinement is applied to the best-performing individual in one of the other population.
- The diversity of populations cannot be preserved as the individuals are not derived by the existing populations but by the independent populations.

The Proposed Algorithm's Operation Process



Experimental Goals and Environment

Experimental goals

- Verify if the proposed algorithm performs better than algorithms of recent studies.
- Experimentally verify if the conventional refinement method reduces the diversity in multi-population genetic algorithm.

Verification method

- Friedman test/critical value ($\alpha = 0.05$) : 2.8451
- Bonferroni-Dunn test/CD refers to critical difference. A mean grade that deviates from the CD value indicates a better performance.

Compared algorithms

- Bacterial Colony Optimization method with Multi-Dimensional Population (BCO_MDP) [12] : Swarm and Evolutionary Computation (2019)
- Tribe competition-based genetic algorithm (TCbGA) [10] : Applied Soft Computing (2017)
- Enhanced multi-population niche genetic algorithm (EMPNGA) [11] : Expert Systems with Applications (2019)
- Classifier: Multi-Label Naïve Bayes (MLNB) [13]

Performance evaluation method

- Multi-label Accuracy: The overall effect of the classifier is evaluated.
- Subset Accuracy: The proportion of accurately classified examples is evaluated.
- Hamming Loss: The proportion of incorrectly classified “instance-label” pairs is evaluated.
- One-error: The proportion of incorrectly classified “instance-label” pairs is evaluated for the top-level labels.

Experimental Environment & Data

Experimental environment & data

- 14 multi-label datasets
- Cross validation (8:2), 10 repeated experiments
- Algorithm termination condition: terminate when the number of evaluations of individual FFC exceeds 300.
- refinement method: Generalized Information-Theoretic Criterion (GICS) for MLFS [10]

Dataset	Patterns	Features	Data type	Labels	Cardinality	Density	Distinct	Domain
Arts	7,484	1,157	Numeric	26	1.654	0.064	599	Text
Business	11,214	1,096	Numeric	30	1.599	0.053	233	Text
Computers	12,444	1,705	Numeric	33	1.507	0.046	428	Text
Education	12,030	1,377	Numeric	33	1.463	0.044	511	Text
Entertainment	12,730	1,600	Numeric	21	1.414	0.067	337	Text
Genbase	662	1,185	Numeric	27	1.252	0.046	32	Biology
Health	9,205	1,530	Numeric	32	1.644	0.051	335	Text
Recreation	12,828	1,516	Numeric	22	1.429	0.065	530	Text
Reference	8,027	1,984	Numeric	33	1.174	0.036	275	Text
Scene	2,407	294	Numeric	6	1.074	0.179	15	Image
Science	6,428	1,859	Numeric	40	1.450	0.036	457	Text
Social	12,111	2,618	Numeric	29	1.279	0.033	361	Text
Society	14,512	1,590	Numeric	27	1.670	0.062	1,054	Text
Yeast	2,417	103	Numeric	14	4.237	0.303	198	Biology

Experimental Results (Evaluation Method: Mlacc & Setacc)

Multi-label accuracy

Dataset	Proposed		BCO_MDP		TCbGA		EMPNGA	
	mean	std	mean	std	mean	std	mean	std
arts	0.1951	± 0.004	0.1007	± 0.018	0.0853	± 0.014	0.0939	± 0.020
business	0.5442	± 0.009	0.5256	± 0.013	0.5259	± 0.013	0.5252	± 0.011
computers	0.3726	± 0.006	0.3205	± 0.009	0.3230	± 0.011	0.3204	± 0.012
education	0.1486	± 0.006	0.0809	± 0.011	0.0660	± 0.012	0.0709	± 0.014
entertainment	0.3196	± 0.007	0.1935	± 0.019	0.1701	± 0.022	0.1926	± 0.016
ganebase	0.9106	± 0.024	0.5098	± 0.123	0.2576	± 0.092	0.4288	± 0.070
health	0.4473	± 0.010	0.3576	± 0.026	0.3431	± 0.012	0.3425	± 0.019
recreation	0.2773	± 0.005	0.1435	± 0.019	0.0929	± 0.015	0.1204	± 0.019
reference	0.4192	± 0.008	0.3302	± 0.012	0.3472	± 0.011	0.3253	± 0.011
scene	0.4050	± 0.033	0.3840	± 0.027	0.3518	± 0.031	0.3293	± 0.033
science	0.1696	± 0.011	0.0815	± 0.011	0.0689	± 0.018	0.0730	± 0.010
social	0.5309	± 0.009	0.4340	± 0.049	0.4147	± 0.027	0.4398	± 0.020
society	0.3340	± 0.012	0.2074	± 0.049	0.2184	± 0.046	0.1931	± 0.036
yeast	0.1203	± 0.019	0.0988	± 0.016	0.0969	± 0.014	0.1085	± 0.066
rank	1.0		2.43		3.29		3.29	

Subset accuracy

Dataset	Proposed		BCO_MDP		TCbGA		EMPNGA	
	mean	std	mean	std	mean	std	mean	std
arts	0.1951	± 0.004	0.1007	± 0.018	0.0853	± 0.014	0.0939	± 0.020
business	0.5442	± 0.009	0.5256	± 0.013	0.5259	± 0.013	0.5252	± 0.011
computers	0.3726	± 0.006	0.3205	± 0.009	0.3230	± 0.011	0.3204	± 0.012
education	0.1486	± 0.006	0.0809	± 0.011	0.0660	± 0.012	0.0709	± 0.014
entertainment	0.3196	± 0.007	0.1935	± 0.019	0.1701	± 0.022	0.1926	± 0.016
ganebase	0.9106	± 0.024	0.5098	± 0.123	0.2576	± 0.092	0.4288	± 0.070
health	0.4473	± 0.010	0.3576	± 0.026	0.3431	± 0.012	0.3425	± 0.019
recreation	0.2773	± 0.005	0.1435	± 0.019	0.0929	± 0.015	0.1204	± 0.019
reference	0.4192	± 0.008	0.3302	± 0.012	0.3472	± 0.011	0.3253	± 0.011
scene	0.4050	± 0.033	0.3840	± 0.027	0.3518	± 0.031	0.3293	± 0.033
science	0.1696	± 0.011	0.0815	± 0.011	0.0689	± 0.018	0.0730	± 0.010
social	0.5309	± 0.009	0.4340	± 0.049	0.4147	± 0.027	0.4398	± 0.020
society	0.3340	± 0.012	0.2074	± 0.049	0.2184	± 0.046	0.1931	± 0.036
yeast	0.1203	± 0.019	0.0988	± 0.016	0.0969	± 0.014	0.1085	± 0.066
rank	1.0		2.43		3.29		3.29	

Experimental Results (Evaluation Method: Hloss & One-error)

Hamming Loss

Dataset	Proposed		BCO_MDP		TCbGA		EMPNGA	
	mean	std	mean	std	mean	std	mean	std
arts	0.0567	± 0.001	0.0631	± 0.002	0.0627	± 0.001	0.0642	± 0.001
business	0.0276	± 0.001	0.0311	± 0.002	0.0290	± 0.001	0.0324	± 0.002
computers	0.0393	± 0.001	0.0426	± 0.001	0.0427	± 0.001	0.0446	± 0.001
education	0.0402	± 0.001	0.0451	± 0.001	0.0441	± 0.001	0.0456	± 0.001
entertainment	0.0529	± 0.001	0.0628	± 0.002	0.0626	± 0.001	0.0636	± 0.002
ganebase	0.0049	± 0.001	0.0277	± 0.006	0.0338	± 0.006	0.0315	± 0.004
health	0.0354	± 0.001	0.0458	± 0.002	0.0465	± 0.002	0.0459	± 0.002
recreation	0.0536	± 0.001	0.0611	± 0.002	0.0615	± 0.001	0.0617	± 0.002
reference	0.0286	± 0.001	0.0366	± 0.002	0.0353	± 0.002	0.0398	± 0.002
scene	0.1459	± 0.008	0.1488	± 0.013	0.1523	± 0.007	0.1567	± 0.009
science	0.0335	± 0.001	0.0392	± 0.002	0.0363	± 0.001	0.0423	± 0.002
social	0.0236	± 0.001	0.0298	± 0.002	0.0301	± 0.001	0.0305	± 0.001
society	0.0527	± 0.001	0.0578	± 0.001	0.0579	± 0.001	0.0576	± 0.001
yeast	0.2251	± 0.007	0.2241	± 0.006	0.2233	± 0.007	0.2253	± 0.005
rank	1.14		2.5		2.64		3.71	

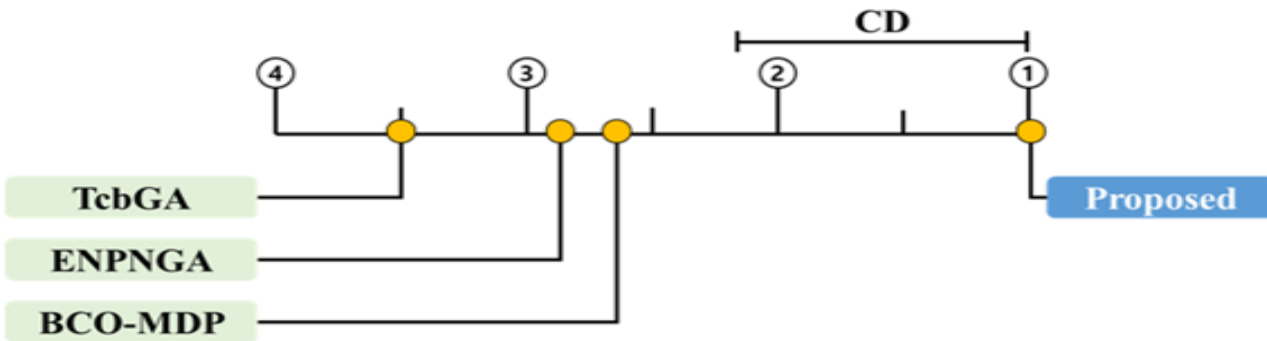
One-error

Dataset	Proposed		BCO_MDP		TCbGA		EMPNGA	
	mean	std	mean	std	mean	std	mean	std
arts	0.6142	± 0.203	0.7076	± 0.155	0.7203	± 0.148	0.7120	± 0.152
business	0.3858	± 0.424	0.3952	± 0.417	0.3933	± 0.418	0.3980	± 0.415
computers	0.3943	± 0.008	0.4428	± 0.018	0.4454	± 0.010	0.4396	± 0.007
education	0.5419	± 0.014	0.6368	± 0.014	0.6440	± 0.020	0.6242	± 0.017
entertainment	0.4520	± 0.007	0.5769	± 0.015	0.6033	± 0.010	0.5782	± 0.021
ganebase	0.5955	± 0.511	0.8045	± 0.303	0.8652	± 0.027	0.8235	± 0.272
health	0.7064	± 0.378	0.7724	± 0.294	0.7741	± 0.029	0.7628	± 0.305
recreation	0.4834	± 0.011	0.6344	± 0.021	0.6867	± 0.015	0.6498	± 0.018
reference	0.6536	± 0.298	0.7140	± 0.246	0.7082	± 0.251	0.7072	± 0.252
scene	0.2919	± 0.022	0.3291	± 0.024	0.3393	± 0.025	0.3264	± 0.031
science	0.5522	± 0.019	0.6881	± 0.020	0.6955	± 0.024	0.6790	± 0.021
social	0.3891	± 0.214	0.4779	± 0.184	0.5045	± 0.175	0.4792	± 0.184
society	0.3936	± 0.010	0.4631	± 0.020	0.4746	± 0.016	0.4652	± 0.013
yeast	0.2617	± 0.023	0.2513	± 0.019	0.2431	± 0.019	0.2652	± 0.020
rank	1.14		2.57		3.57		2.71	

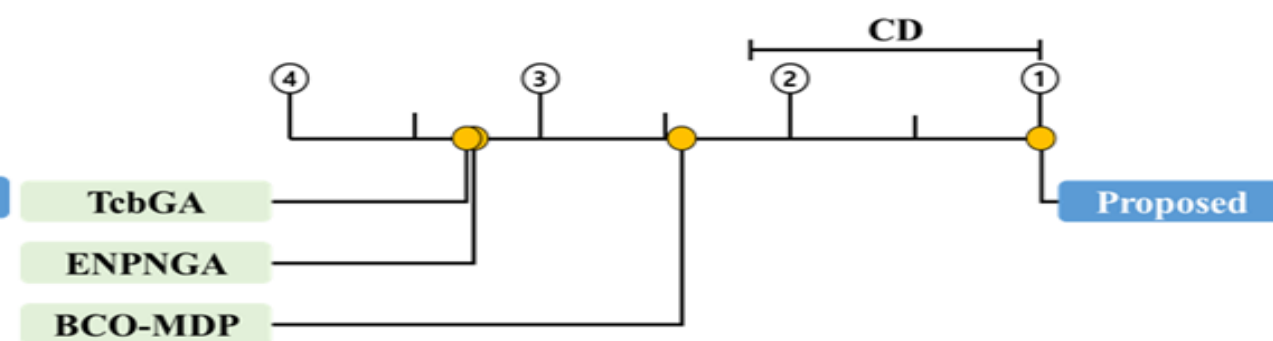
Experimental Results (Evaluation Method: Bonferroni-Dunn test)

- CD indicated in the figure refers to critical distance. If the mean grade is not included in the CD values, it means that there is significant difference in performance.
- Bonferroni-Dunn test is a method of correcting errors that can occur in multiple comparisons. It adopts an alternative hypothesis by rejecting the null hypothesis when the probability of establishing the null hypothesis is low for the observation data in statistical inference.

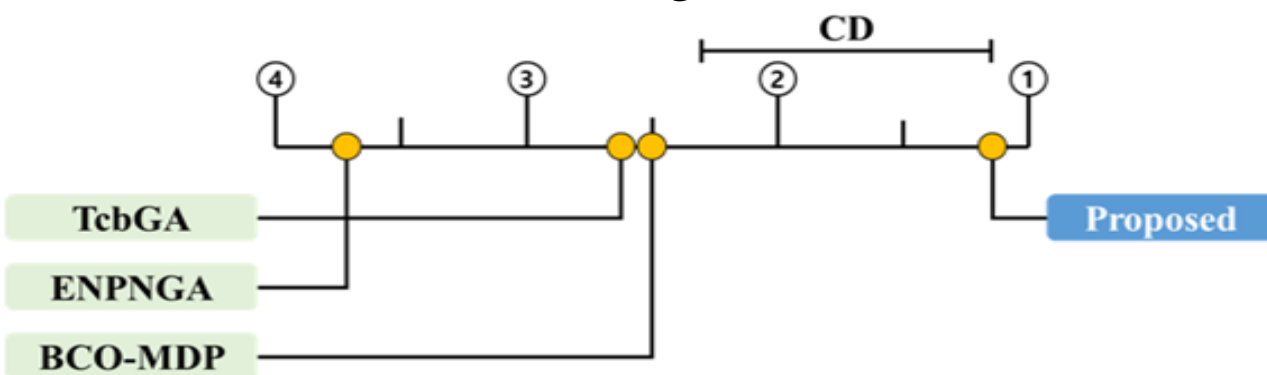
Multi-label accuracy



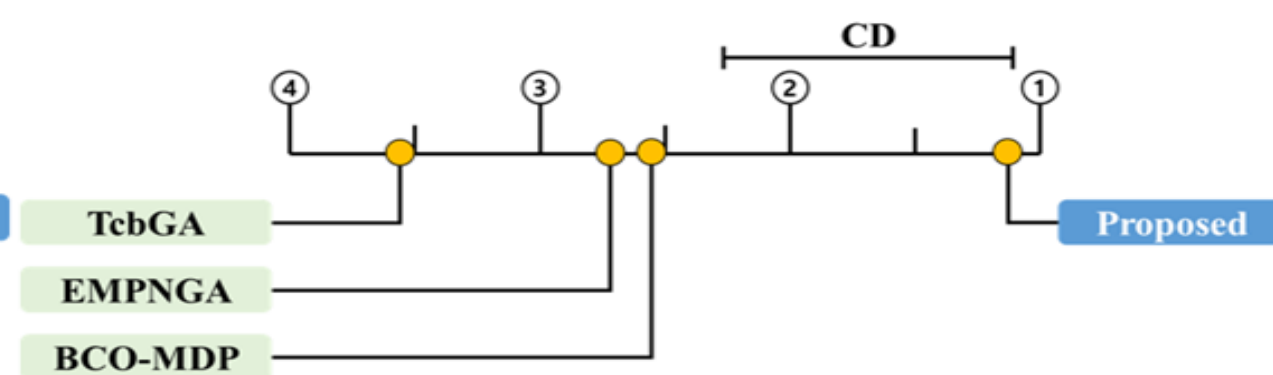
Subset accuracy



Hamming loss

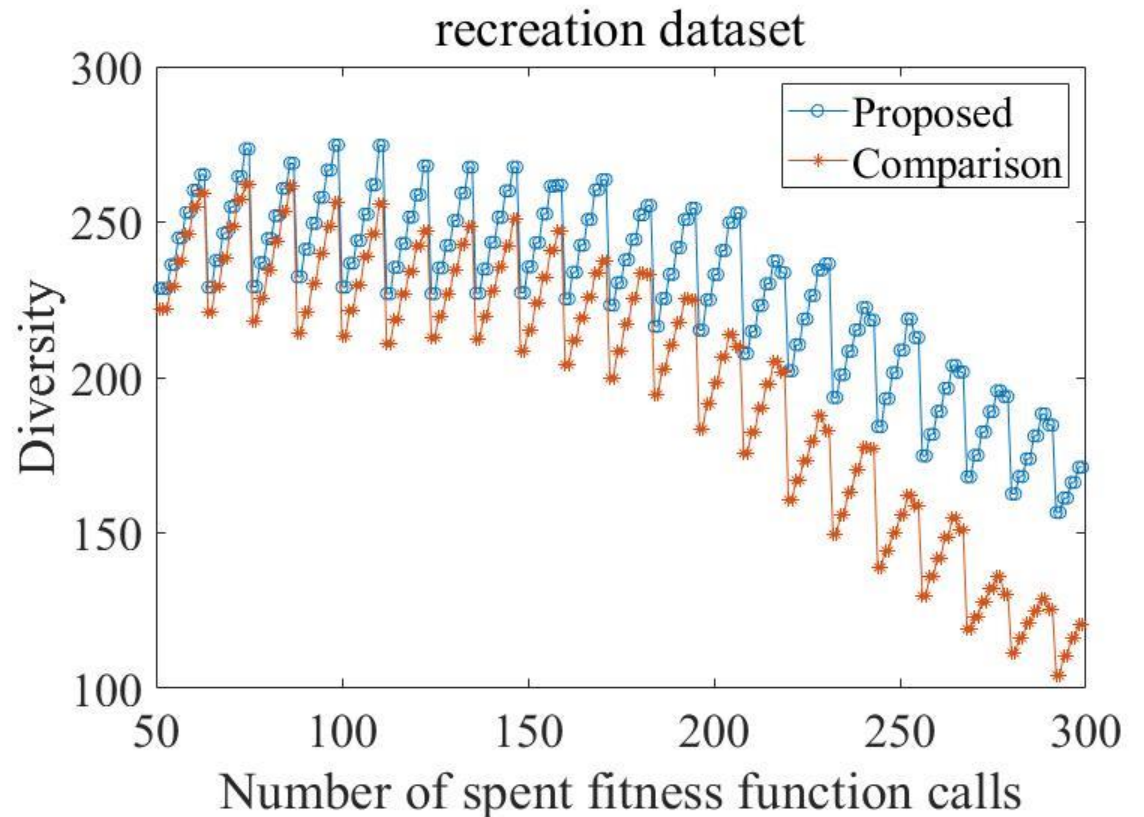
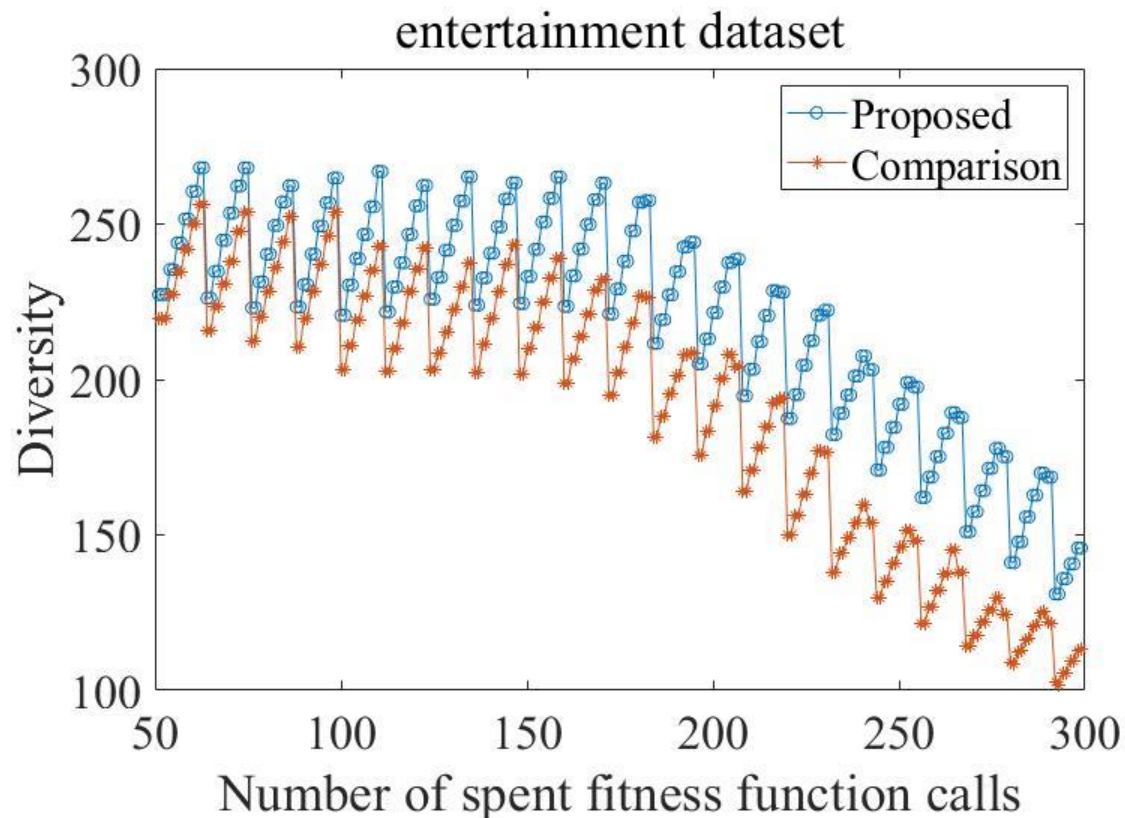


One-error



Analysis

- The diversity of populations appears when the refinement is applied to the migration process or inside the population.
- In the experimental results, the proposed method maintains the diversity of populations, thus increasing the inflow probability of new features and consequently improving the performance.



Conclusion & Contribution

Conclusion

- Refinement is applied to the multi-population genetic algorithm to improve the multi-label classification performance.
- Refinement is a promising method, but the conventional methods do not preserve the diversity of populations, thereby causing a concern of decline in performance.
- To resolve this problem, the refinement for migrated individuals is applied.
- In the experimental results, the proposed method reveals a better performance than the compared algorithms for the 14 multi-label datasets.

Paper published during the degree course:

- Multi-population genetic algorithm for multi-label feature selection based on label complementary communication; **Entropy**, 2020, **Co-author**.

Thank you.

Appendix – Calculation of the Diversity of Populations

Calculation of the diversity of populations

1: Calculate the mean for each feature.

Mean (f_1)	Mean (f_2)	Mean (f_3)	Mean (f_4)	Mean (f_5)	...	Mean ($f_{ F }$)
----------------	----------------	----------------	----------------	----------------	-----	--------------------

2: Calculate the absolute value of difference for each individual with respect to the calculated means.

2-1: Repeat as many times as the number of individuals

(| Mean (f_1) – number of individuals (f_1) |)

2-2: Repeat the above operation as many times as the number of features

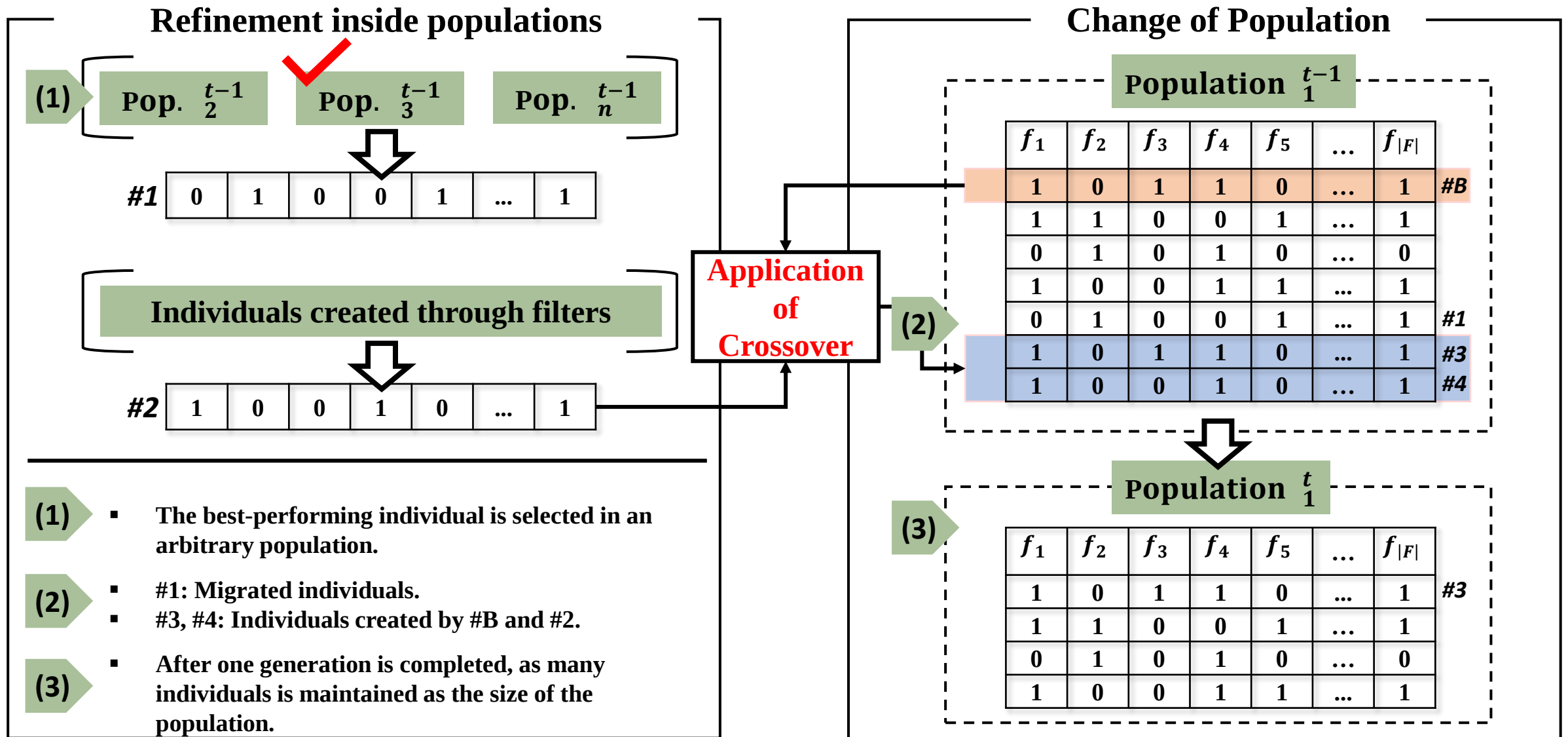
(| Mean ($f_{1 \sim |F|}$) – number of features ($f_{1 \sim |F|}$) |)

2-3: Add all the values derived through 2-1 / 2-2 ➡ The diversity of each population

3: Repeat the above process for every population and calculate the average by using the number of populations.

Diversity of each population / Number of populations

Appendix – refinement Process inside Populations



Appendix – Experimental Results (Evaluation Method: One-Error)

Dataset	Proposed			BCO_MDP			TCbGA			EMPNGA		
	mean	std		mean	std		mean	std		mean	std	
arts	0.6142	±	0.203	0.7076	±	0.155	0.7203	±	0.148	0.7120	±	0.152
business	0.3858	±	0.424	0.3952	±	0.417	0.3933	±	0.418	0.3980	±	0.415
computers	0.3943	±	0.008	0.4428	±	0.018	0.4454	±	0.010	0.4396	±	0.007
education	0.5419	±	0.014	0.6368	±	0.014	0.6440	±	0.020	0.6242	±	0.017
entertainment	0.4520	±	0.007	0.5769	±	0.015	0.6033	±	0.010	0.5782	±	0.021
ganebase	0.5955	±	0.511	0.8045	±	0.303	0.8652	±	0.027	0.8235	±	0.272
health	0.7064	±	0.378	0.7724	±	0.294	0.7741	±	0.029	0.7628	±	0.305
recreation	0.4834	±	0.011	0.6344	±	0.021	0.6867	±	0.015	0.6498	±	0.018
reference	0.6536	±	0.298	0.7140	±	0.246	0.7082	±	0.251	0.7072	±	0.252
scene	0.2919	±	0.022	0.3291	±	0.024	0.3393	±	0.025	0.3264	±	0.031
science	0.5522	±	0.019	0.6881	±	0.020	0.6955	±	0.024	0.6790	±	0.021
social	0.3891	±	0.214	0.4779	±	0.184	0.5045	±	0.175	0.4792	±	0.184
society	0.3936	±	0.010	0.4631	±	0.020	0.4746	±	0.016	0.4652	±	0.013
yeast	0.2617	±	0.023	0.2513	±	0.019	0.2431	±	0.019	0.2652	±	0.020
rank	1.14			2.57			3.57			2.71		

Appendix – Experimental Results (Evaluation Method: Multi-label Accuracy)

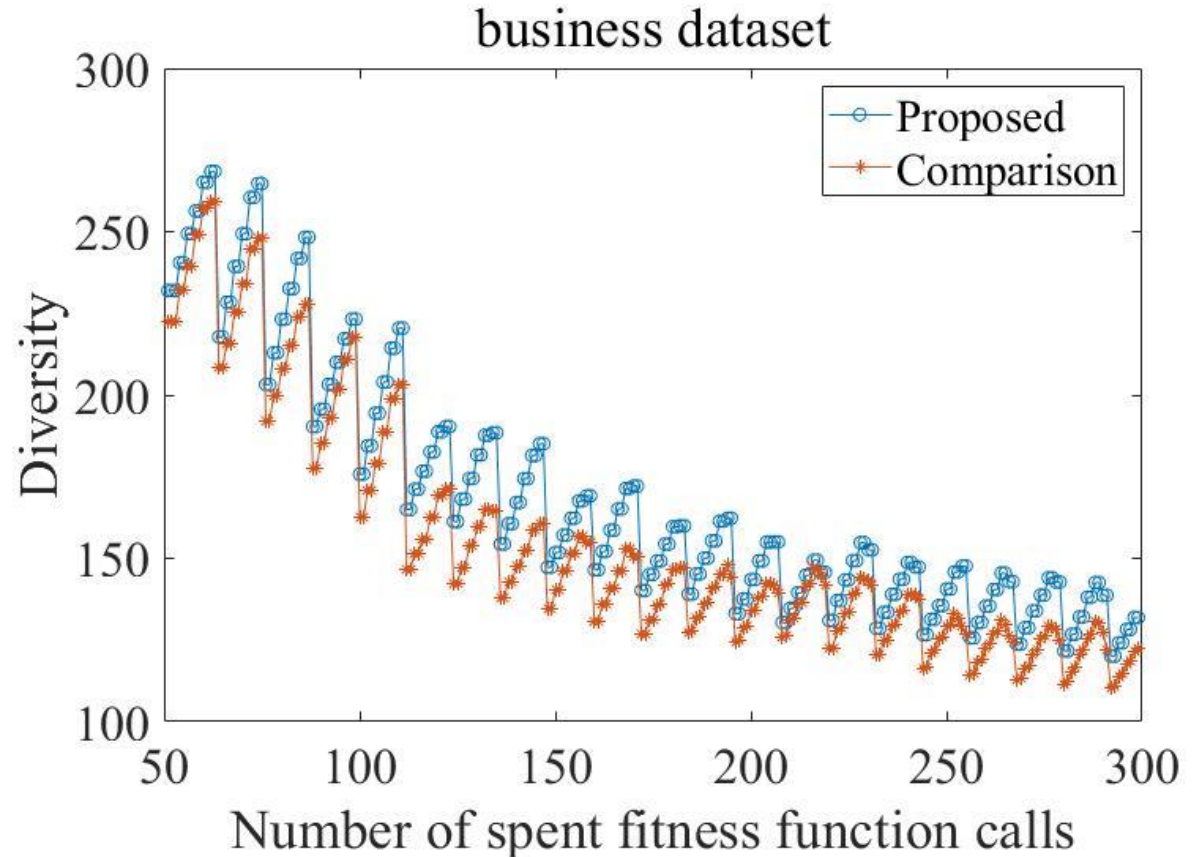
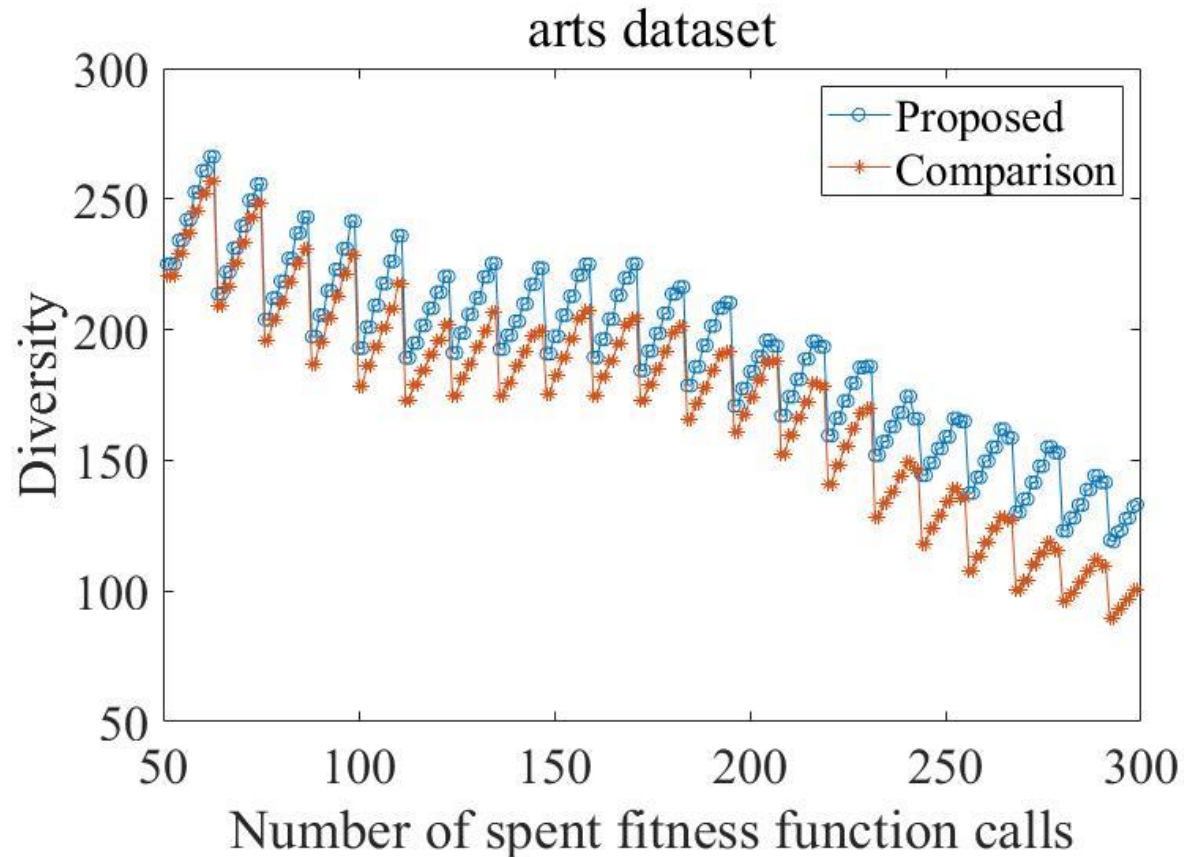
Dataset	Proposed			BCO_MDP			TCbGA			EMPNGA		
	mean	std		mean	std		mean	std		mean	std	
arts	0.1951	±	0.004	0.1007	±	0.018	0.0853	±	0.014	0.0939	±	0.020
business	0.5442	±	0.009	0.5256	±	0.013	0.5259	±	0.013	0.5252	±	0.011
computers	0.3726	±	0.006	0.3205	±	0.009	0.3230	±	0.011	0.3204	±	0.012
education	0.1486	±	0.006	0.0809	±	0.011	0.0660	±	0.012	0.0709	±	0.014
entertainment	0.3196	±	0.007	0.1935	±	0.019	0.1701	±	0.022	0.1926	±	0.016
ganebase	0.9106	±	0.024	0.5098	±	0.123	0.2576	±	0.092	0.4288	±	0.070
health	0.4473	±	0.010	0.3576	±	0.026	0.3431	±	0.012	0.3425	±	0.019
recreation	0.2773	±	0.005	0.1435	±	0.019	0.0929	±	0.015	0.1204	±	0.019
reference	0.4192	±	0.008	0.3302	±	0.012	0.3472	±	0.011	0.3253	±	0.011
scene	0.4050	±	0.033	0.3840	±	0.027	0.3518	±	0.031	0.3293	±	0.033
science	0.1696	±	0.011	0.0815	±	0.011	0.0689	±	0.018	0.0730	±	0.010
social	0.5309	±	0.009	0.4340	±	0.049	0.4147	±	0.027	0.4398	±	0.020
society	0.3340	±	0.012	0.2074	±	0.049	0.2184	±	0.046	0.1931	±	0.036
yeast	0.1203	±	0.019	0.0988	±	0.016	0.0969	±	0.014	0.1085	±	0.066
rank	1.0			2.43			3.29			3.29		

Appendix – Experimental Results (Evaluation Method: Subset Accuracy)

Dataset	Proposed			BCO_MDP			TCbGA			EMPNGA		
	mean	std		mean	std		mean	std		mean	std	
arts	0.1951	±	0.004	0.1007	±	0.018	0.0853	±	0.014	0.0939	±	0.020
business	0.5442	±	0.009	0.5256	±	0.013	0.5259	±	0.013	0.5252	±	0.011
computers	0.3726	±	0.006	0.3205	±	0.009	0.3230	±	0.011	0.3204	±	0.012
education	0.1486	±	0.006	0.0809	±	0.011	0.0660	±	0.012	0.0709	±	0.014
entertainment	0.3196	±	0.007	0.1935	±	0.019	0.1701	±	0.022	0.1926	±	0.016
ganebase	0.9106	±	0.024	0.5098	±	0.123	0.2576	±	0.092	0.4288	±	0.070
health	0.4473	±	0.010	0.3576	±	0.026	0.3431	±	0.012	0.3425	±	0.019
recreation	0.2773	±	0.005	0.1435	±	0.019	0.0929	±	0.015	0.1204	±	0.019
reference	0.4192	±	0.008	0.3302	±	0.012	0.3472	±	0.011	0.3253	±	0.011
scene	0.4050	±	0.033	0.3840	±	0.027	0.3518	±	0.031	0.3293	±	0.033
science	0.1696	±	0.011	0.0815	±	0.011	0.0689	±	0.018	0.0730	±	0.010
social	0.5309	±	0.009	0.4340	±	0.049	0.4147	±	0.027	0.4398	±	0.020
society	0.3340	±	0.012	0.2074	±	0.049	0.2184	±	0.046	0.1931	±	0.036
yeast	0.1203	±	0.019	0.0988	±	0.016	0.0969	±	0.014	0.1085	±	0.066
rank	1.0			2.43			3.29			3.29		

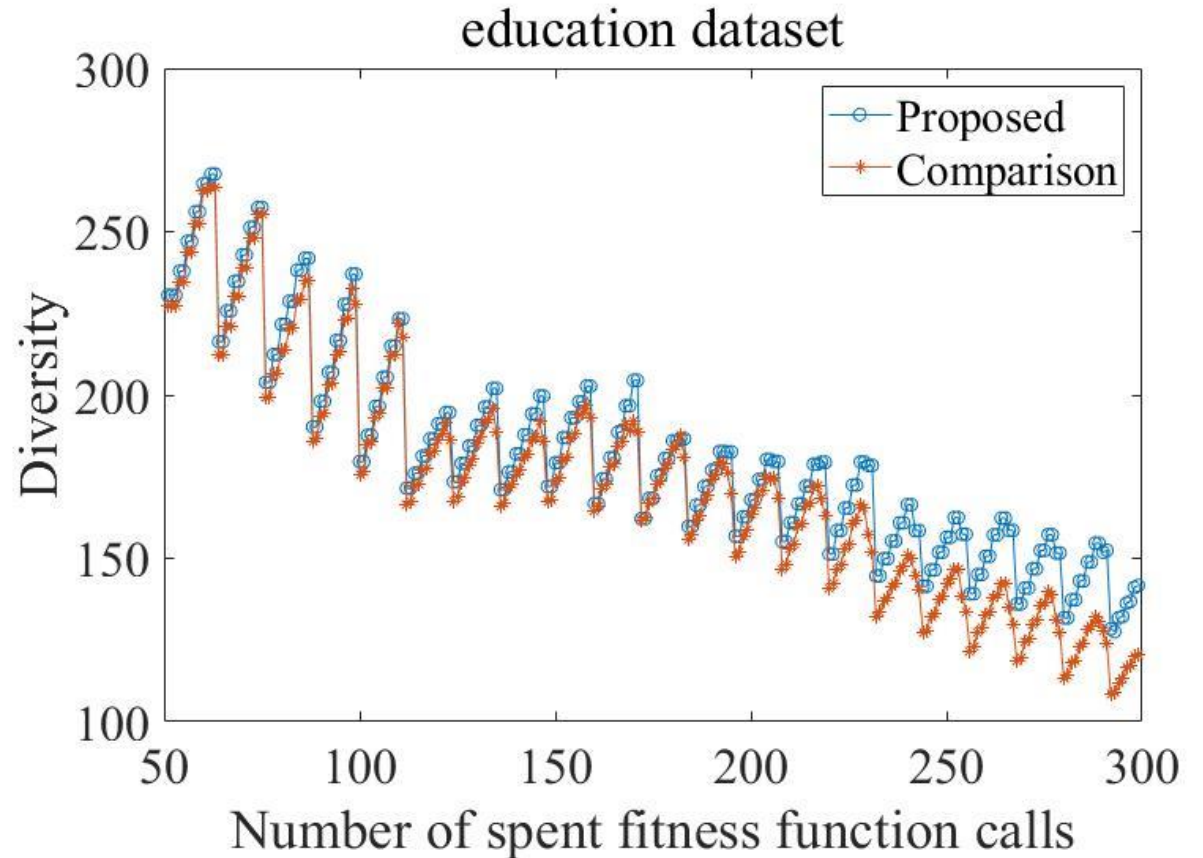
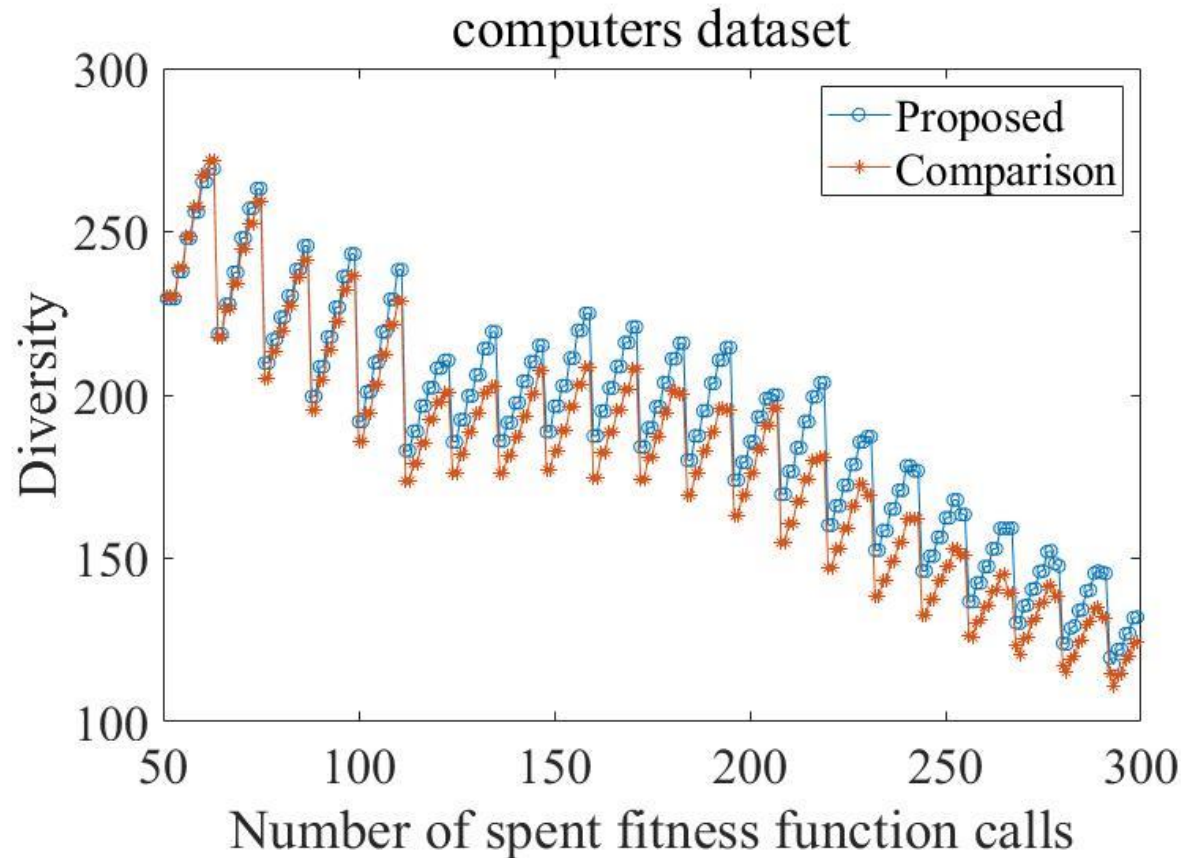
Appendix – Experimental Results of Population Diversity

- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.



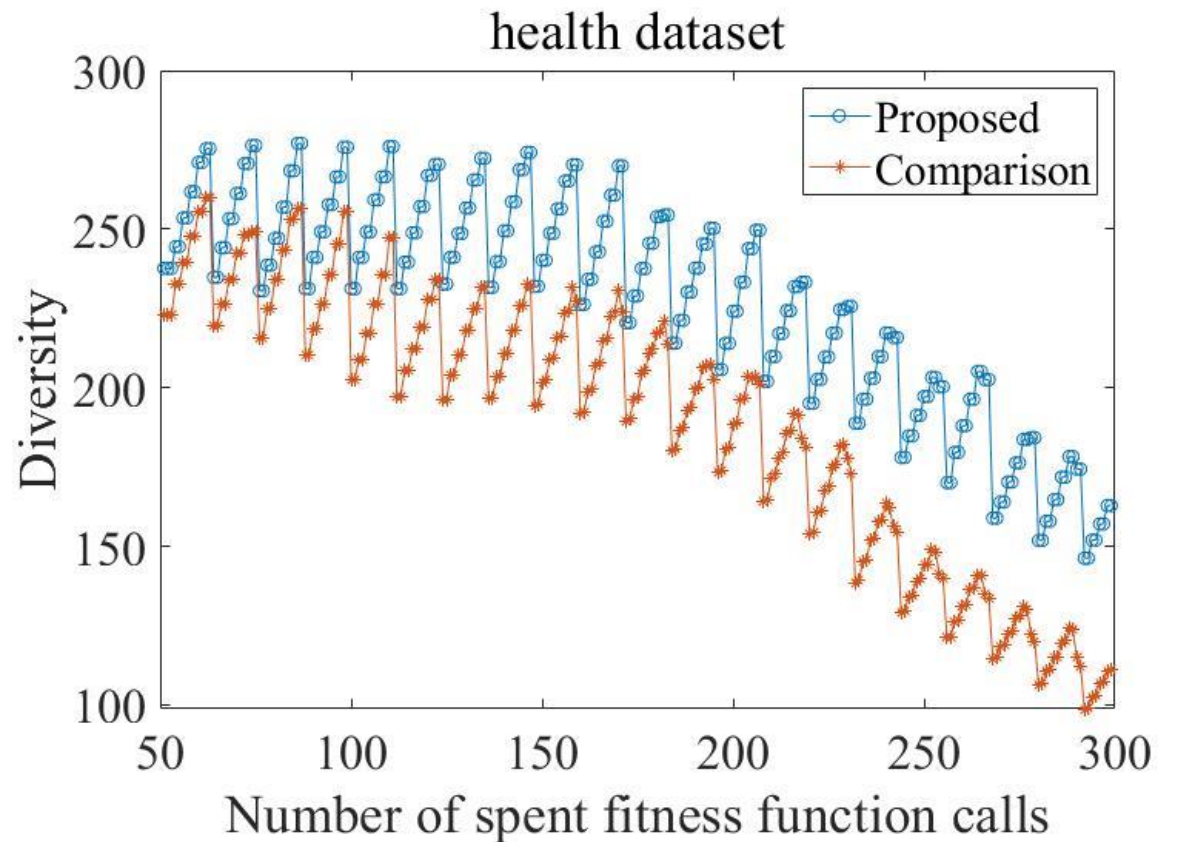
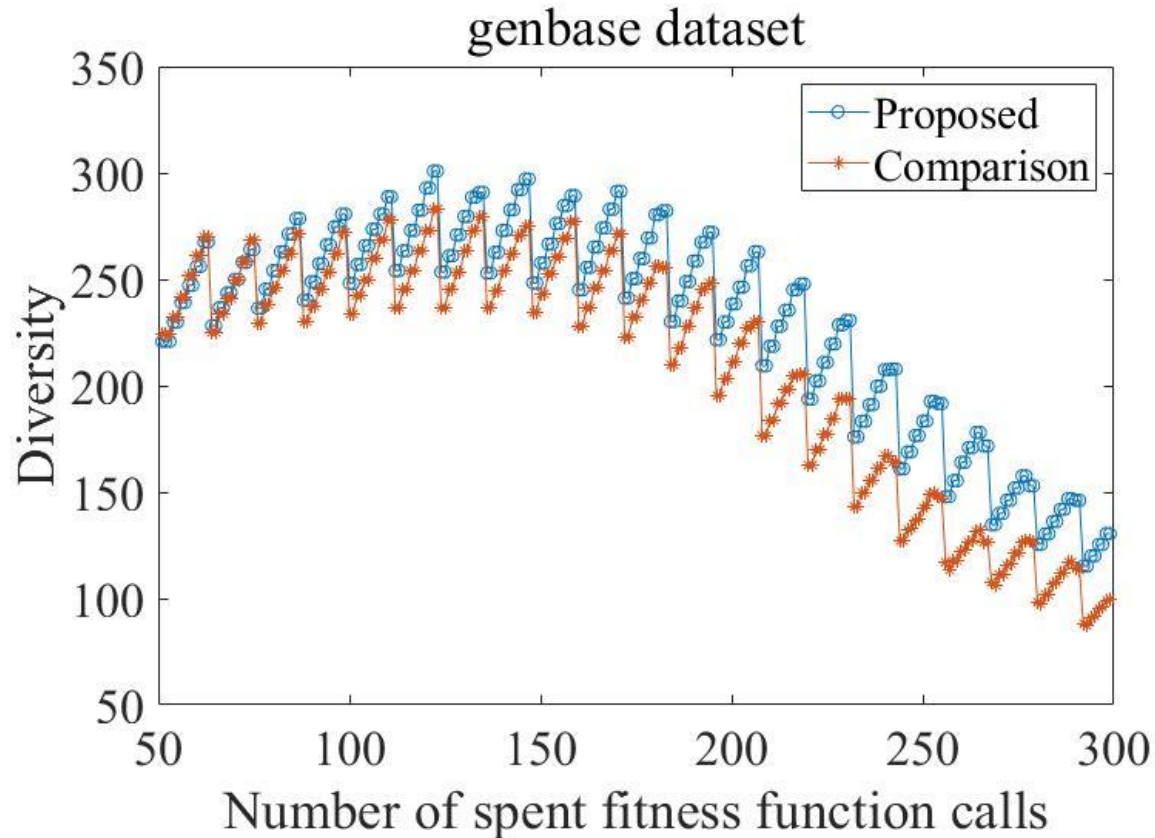
Appendix – Experimental Results of Population Diversity

- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.



Appendix – Experimental Results of Population Diversity

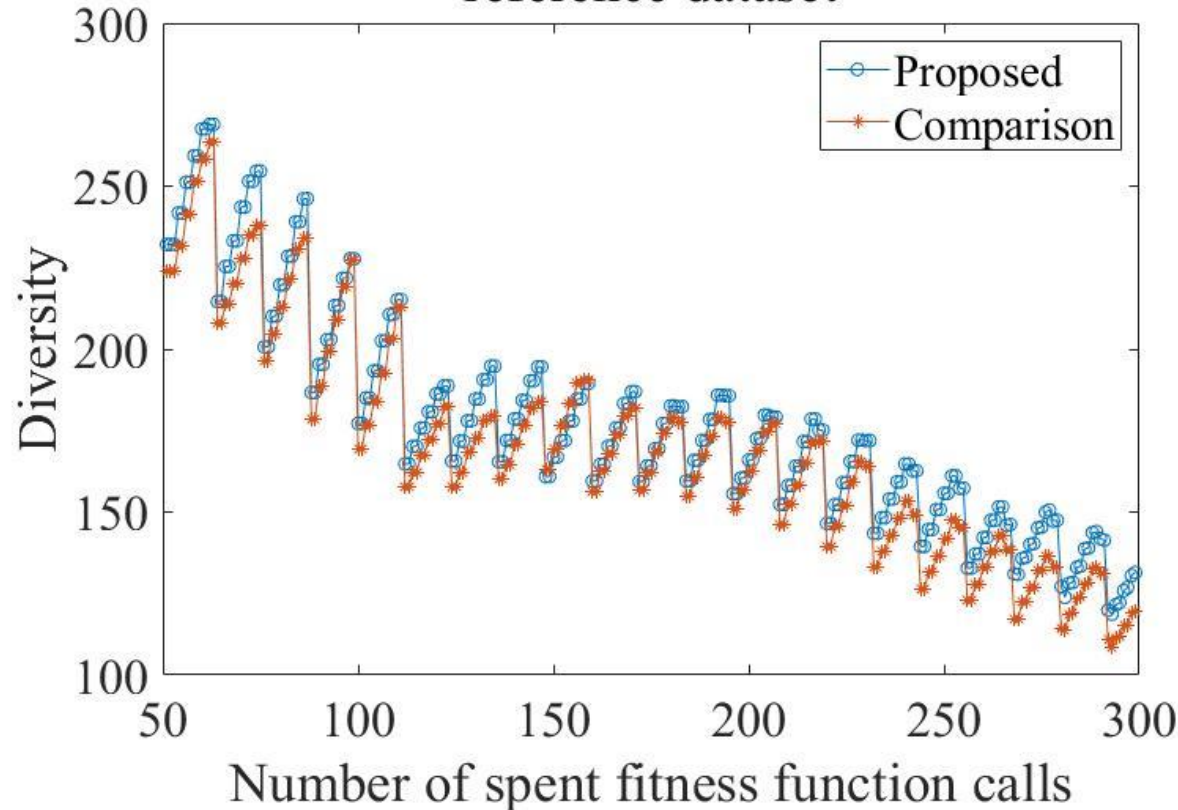
- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.



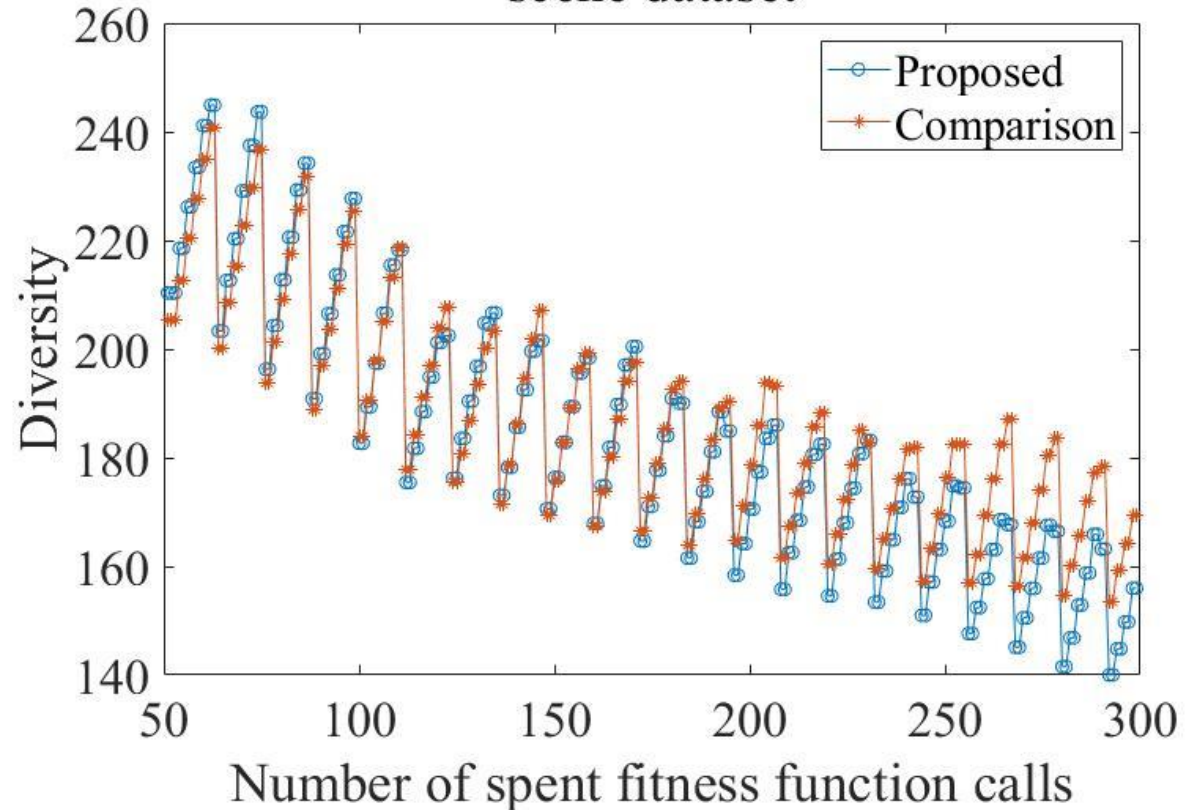
Appendix – Experimental Results of Population Diversity

- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.

reference dataset



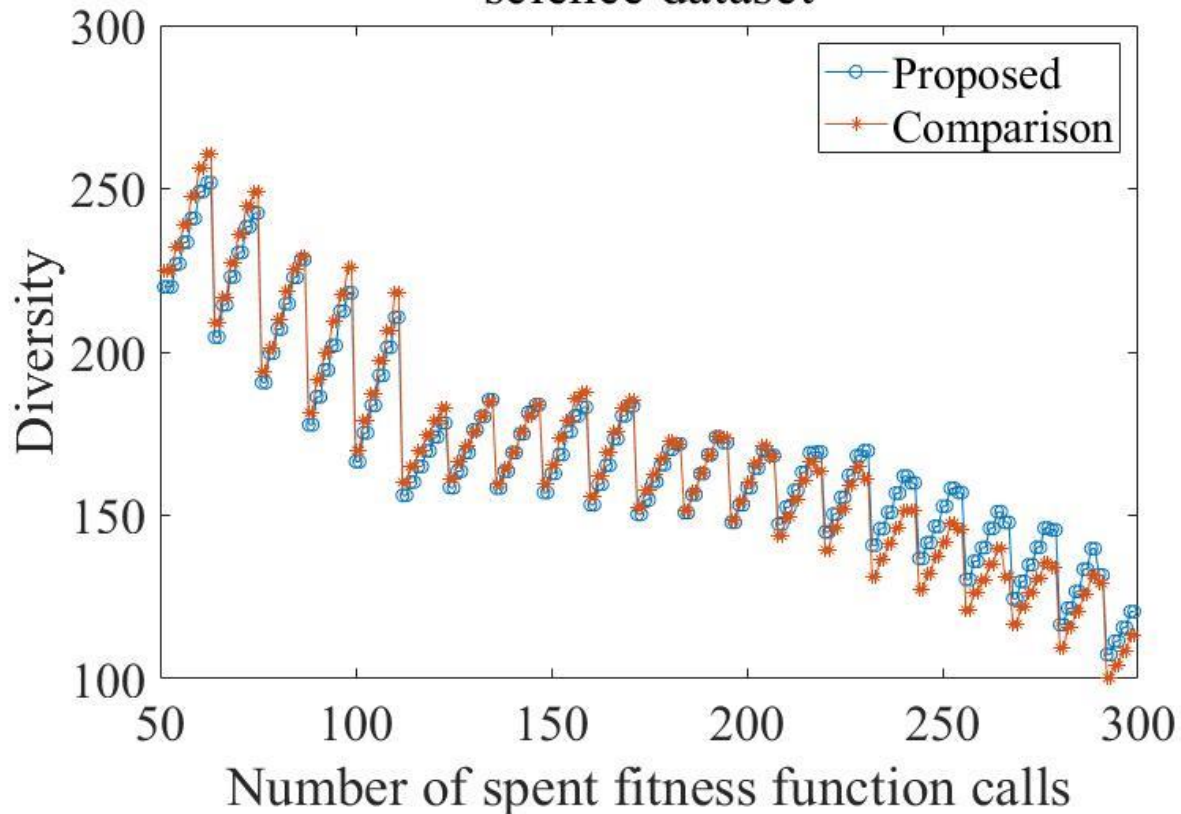
scene dataset



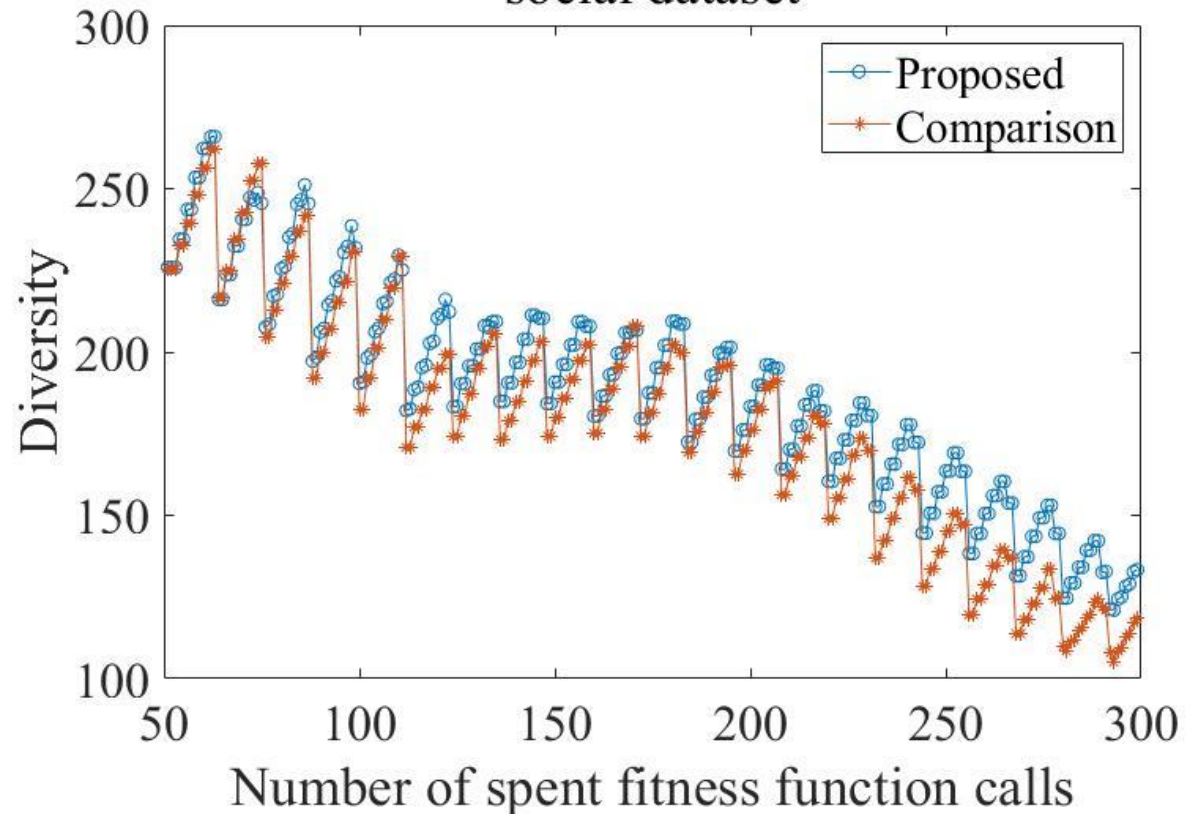
Appendix – Experimental Results of Population Diversity

- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.

science dataset



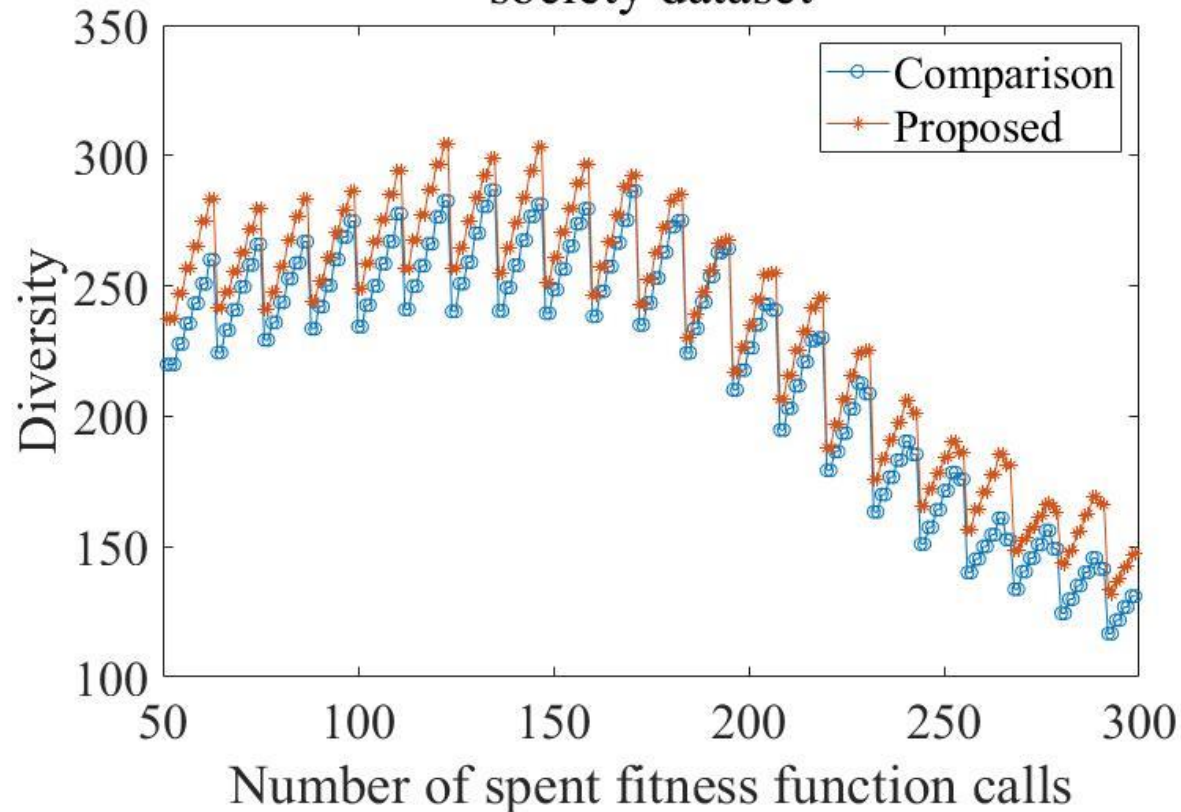
social dataset



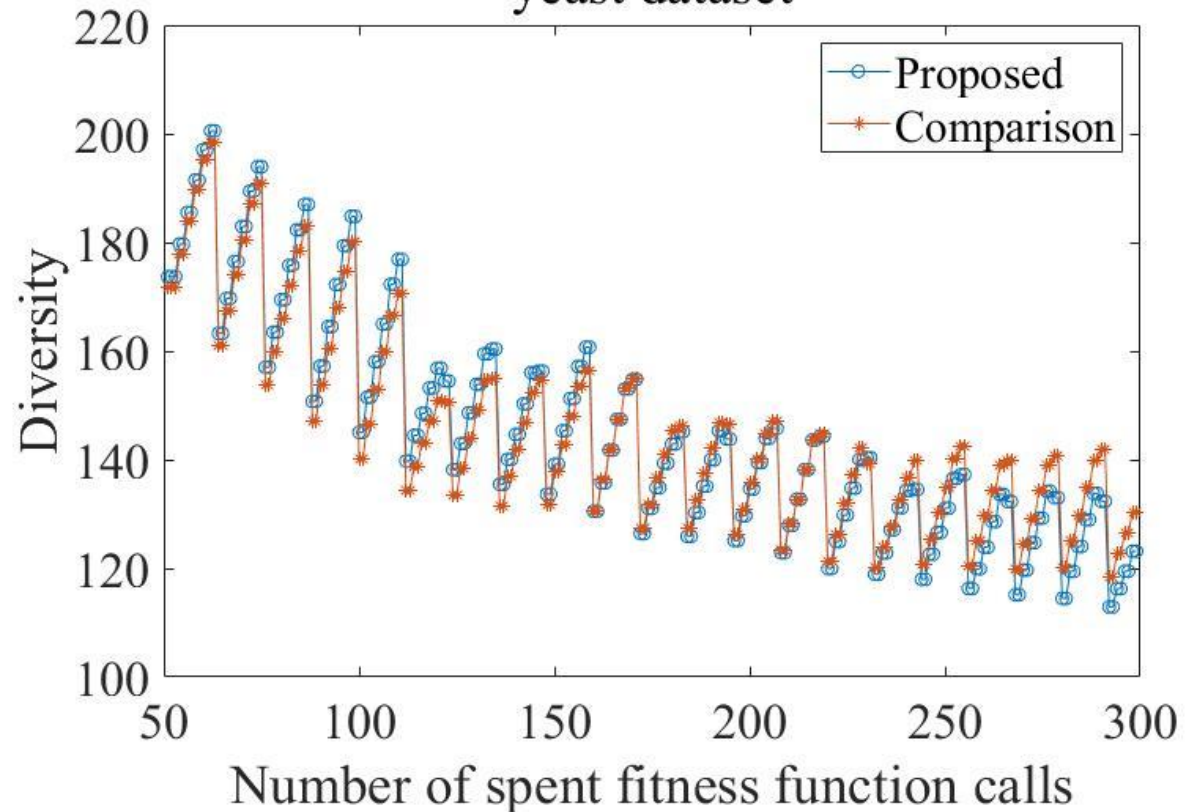
Appendix – Experimental Results of Population Diversity

- An experiment that compares the diversity of population between the proposed algorithm and the comparison algorithm.
- The comparison algorithm applies the refinement to the inside of the population.

society dataset



yeast dataset



Appendix

- Performance comparison of the refinement application methods

Multi-label accuracy

Dataset	Proposed			Comparison		
	mean	std		mean	std	
arts	0.2497	±	0.005	0.2475	±	0.007
business	0.6948	±	0.008	0.6962	±	0.007
computers	0.4494	±	0.007	0.4467	±	0.006
education	0.1818	±	0.007	0.1816	±	0.007
entertainment	0.3765	±	0.008	0.3760	±	0.008
ganebase	0.9513	±	0.014	0.9494	±	0.011
health	0.5777	±	0.006	0.5772	±	0.008
recreation	0.3329	±	0.007	0.3325	±	0.007
reference	0.4572	±	0.010	0.4583	±	0.014
scene	0.5537	±	0.018	0.5514	±	0.016
science	0.2068	±	0.008	0.2061	±	0.009
social	0.5843	±	0.007	0.5851	±	0.007
society	0.4165	±	0.039	0.4296	±	0.010
yeast	0.4651	±	0.017	0.4675	±	0.014
rank	1.285			1.714		

Subset accuracy

Dataset	Proposed			Comparison		
	mean	std		mean	std	
arts	0.1951	±	0.004	0.1944	±	0.005
business	0.5442	±	0.009	0.5447	±	0.012
computers	0.3726	±	0.006	0.3745	±	0.006
education	0.1486	±	0.006	0.1481	±	0.007
entertainment	0.3196	±	0.007	0.3195	±	0.008
ganebase	0.9106	±	0.024	0.9008	±	0.033
health	0.4473	±	0.010	0.4424	±	0.012
recreation	0.2773	±	0.005	0.2754	±	0.005
reference	0.4192	±	0.008	0.4146	±	0.011
scene	0.4050	±	0.033	0.4037	±	0.021
science	0.1696	±	0.011	0.1727	±	0.007
social	0.5309	±	0.009	0.5310	±	0.008
society	0.3340	±	0.012	0.3339	±	0.012
yeast	0.1203	±	0.019	0.1219	±	0.009
rank	1.357			1.642		

Appendix

- Performance comparison of the refinement application methods

Hamming loss

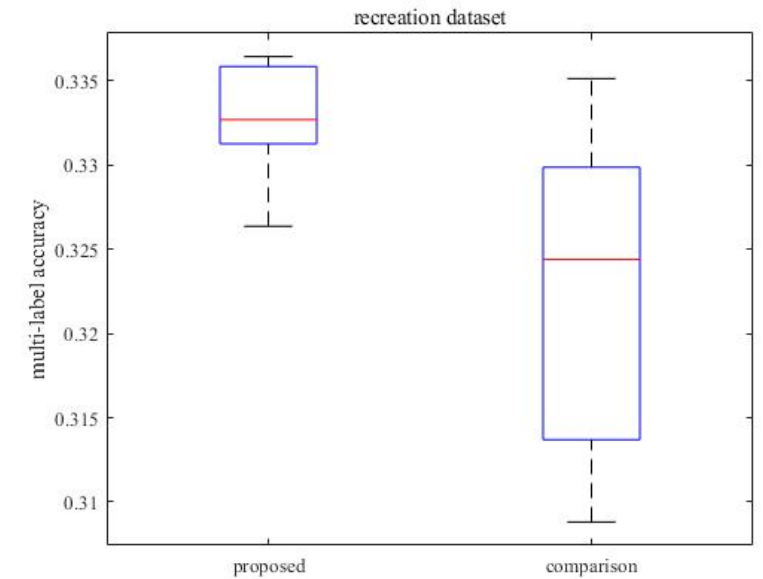
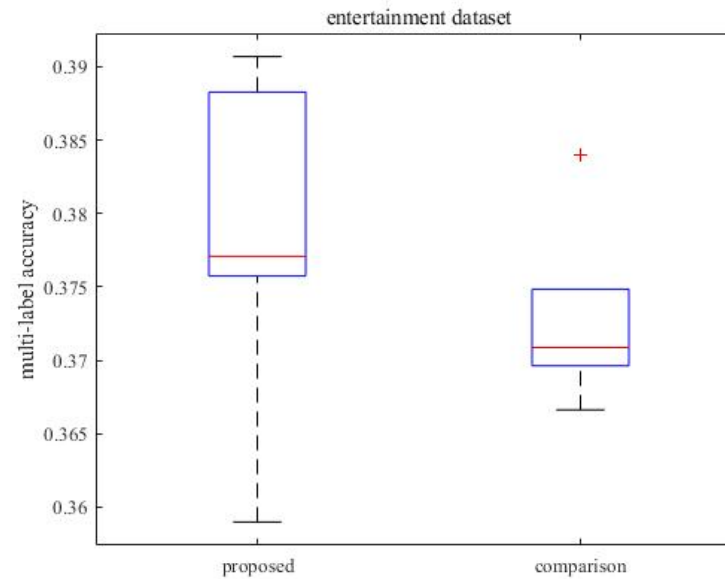
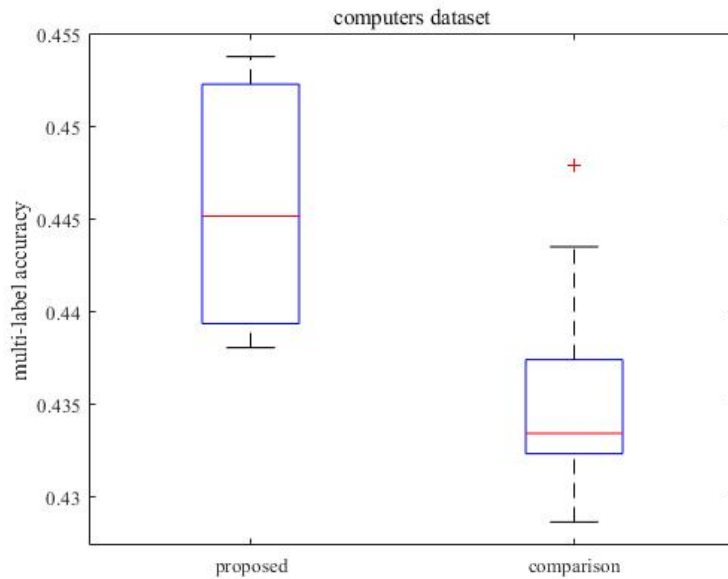
Dataset	Proposed			Comparison		
	mean	std		mean	std	
arts	0.0567	±	0.001	0.0567	±	0.001
business	0.0276	±	0.001	0.0274	±	0.001
computers	0.0393	±	0.001	0.0394	±	0.001
education	0.0402	±	0.001	0.0402	±	0.001
entertainment	0.0529	±	0.001	0.0527	±	0.001
ganebase	0.0049	±	0.001	0.0059	±	0.002
health	0.0354	±	0.001	0.0355	±	0.001
recreation	0.0536	±	0.001	0.0537	±	0.001
reference	0.0286	±	0.001	0.0287	±	0.001
scene	0.1459	±	0.008	0.1445	±	0.007
science	0.0335	±	0.001	0.0333	±	0.001
social	0.0236	±	0.001	0.0238	±	0.001
society	0.0527	±	0.001	0.0527	±	0.001
yeast	0.2251	±	0.007	0.2254	±	0.007
rank	1.285			1.714		

One-error

Dataset	Proposed			Comparison		
	mean	std		mean	std	
arts	0.6142	±	0.203	0.6200	±	0.200
business	0.3858	±	0.424	0.3862	±	0.423
computers	0.3943	±	0.008	0.3951	±	0.007
education	0.5419	±	0.014	0.5448	±	0.014
entertainment	0.4520	±	0.007	0.4532	±	0.008
ganebase	0.5955	±	0.511	0.5962	±	0.510
health	0.7064	±	0.378	0.7087	±	0.375
recreation	0.4834	±	0.011	0.4846	±	0.007
reference	0.6536	±	0.298	0.6553	±	0.296
scene	0.2919	±	0.022	0.2925	±	0.020
science	0.5522	±	0.019	0.5521	±	0.018
social	0.3891	±	0.214	0.3897	±	0.214
society	0.3936	±	0.010	0.3965	±	0.012
yeast	0.2617	±	0.023	0.2582	±	0.024
rank	1.142			1.857		

Appendix

- In the figures, a boxplot for each refinement application is depicted to compare their performances.
- Since the input values of the boxplot are a matrix, the mark at the center of each box represents the median, and the lower and upper edges of the box represent the 25th percentile and the 75th percentile, respectively.
- The whisker extends to the maximum or minimum data points, which is not regarded as an outlier. (The outlier floats individually utilizing the + symbol).



Reference

- [1] Lee, J., Kim, D.W. (2015). Mutual information-based multi-label feature selection using interaction information. *Expert Syst. Appl.*, 42:2013-2025.
- [2] Lee, J., Kim, D.W. (2015). Fast multi-label feature selection based on information-theoretic feature ranking. *Pattern Recognit.*, 48:2761-2771.
- [3] Rehman, A., Javed, K., Babri, H.A. (2017). Feature selection based on a normalized difference measure for text classification. *Inf. Process. Manag.*, 53:473-489.
- [4] Zhang, Y., Gong, D.W., Sun, X.Y., Guo, Y.N. (2017). A PSO-based multi-objective multi-label feature selection method in classification. *Sci. Rep.*, 7:376.
- [5] Lu, Y., Liang, M., Ye, Z., Cao, L. (2015). Improved particle swarm optimization algorithm and its application in text feature selection. *Appl. Soft Comput.*, 35:629-636.
- [6] Lin, K.C., Zhang, K.Y., Huang, Y.H., Hung, J.C., Yen, N. (2016). Feature selection based on an improved cat swarm optimization algorithm for big data classification. *J. Supercomput.*, 72:3210-3221.
- [7] Shi, X., Long, W., Deng, Y.W. (2020). Research on the performance of multi-population genetic algorithms with different complex network structures. *Soft Comput.*, 1-19.
- [8] Lee, J., Kim, D.W. (2017). Multi-label feature selection based on scalable criterion for large label set. *Pattern Recognit.*, 66:342-352.
- [9] Seo, W., Kim, D.W. (2019). Generalized Information-theoretic Criterion for Multi-label Feature Selection. *Ieee Access.*, 7:122854-122863

Reference

- [10] Ma, B., Xia, Y. (2017). A tribe competition-based genetic algorithm for feature selection in pattern classification. *Appl. Soft Comput.*, 58:328-338.
- [11] Zhang, W., He, H., Zhang. (2019). A novel multi-stage hybrid model with enhanced multi-population niche genetic algorithm: An application in credit scoring. *Expert Syst. Appl.*, 121:221-232.
- [12] Wang, H., Tan, L., Niu, B. (2019). Feature selection for classification of microarray gene expression cancers using bacterial colony optimization with multi-dimensional population. *Swarm Evol. Comput.*, 48:172-181.
- [13] Zhang, M.L., Peña, J.M., Robles, V. (2009). Feature selection for multi-label naive Bayes classification. *Inf. Sci.*, 179:3218-3229.