RESEARCH ARTICLE

A novel method for multifactorial bio-chemical experiments design based on combinational design theory

Xun Wang[1], Beibei Sun[1], Boyang Liu[2], Yaping Fu[3]*, Pan Zheng[4]*

1 College of Computer and Communication Engineering, China University of Petroleum, Qingdao 266580, Shandong, China, 2 State-owned Asset and Laboratory Management Department, China University of Petroleum, Qingdao 266580, Shandong, China, 3 Institute of Complexity Science, Qingdao University, Qingdao 266071, Shandong, China, 4 Faculty of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus, Kuching 93350, Malaysia

* fuyaping@qdu.edu.cn (YF); pzheng@swiburne.edu.my (PZ)

Abstract

Experimental design focuses on describing or explaining the multifactorial interactions that are hypothesized to reflect the variation. The design introduces conditions that may directly affect the variation, where particular conditions are purposely selected for observation. Combinatorial design theory deals with the existence, construction and properties of systems of finite sets whose arrangements satisfy generalized concepts of balance and/or symmetry. In this work, borrowing the concept of "balance" in combinatorial design theory, a novel method for multifactorial bio-chemical experiments design is proposed, where balanced templates in combinational design are used to select the conditions for observation. Balanced experimental data that covers all the influencing factors of experiments can be obtianed for further processing, such as training set for machine learning models. Finally, a software based on the proposed method is developed for designing experiments with covering influencing factors a certain number of times.

Introduction

The design of experiments, also known as experimental designs, deal with the task of finding relationship among the influencing factors of multifactorial experiments as well as the contribution of each factors to the outcome [1–3]. Experimental design refers to how participants are allocated to the different conditions in an experiment. It involves not only the selection of suitable predictors and outcomes, but planning the delivery of the experiment under statistically optimal conditions given the constraints of available resources [4, 5]. The aim is to design a number of experiments for predicting the outcome by introducing a change of the preconditions. With the experimental data, mathematical models might be built to calculate the outcome from variables or experimental conditions. The three main concerns in experimental design are validity, reliability, and replicability [6, 7]. The mostly

used experimental design methods include Plackett-Burman designs [8], frequentist and Bayesian based approaches [9], response surface methodology [10], central composite design [11] and so on.

In recent years, machine learning methods are used to modelling the bio-chemical experiments with a set of experimental data, see e.g. deep neural network [12], spiking neural P systems [13–17]. Recently, many significant artificial intelligent algorithms and data processing strategies has been applied on data mining, such as a self-adaptive artificial bee colony algorithm based on global best for global optimization [18], the public auditing protocol with novel dynamic structure for cloud data [19], privacy-preserving smart semantic search method for conceptual graphs over encrypted outsourced data [20], a privacy-preserving and copy-deterrence content for image data processing with retrieval scheme in cloud computing [21], and machine learning method have been applied for experimental condition design, see. e.g. a secure and dynamic multi-keyword ranked search scheme over encrypted cloud data [22].

The general idea is to learn from the experimental data, and then achieving a prediction model of the experiment, which matches the known data in a acceptable level. With the model, some optimal conditions for maximizing the outcome or minimizing the cost can be obtained [23–26]. However, the experimental data obtained or collected by classical experimental design methods, such as Plackett-Burman designs, response surface methodology, central composite design, the data is quite unbalanced, are not balanced well, such that we cannot get well fitting models by using machine learning methods.

Combinatorial design theory belongs the field of combinatorial mathematics, which deals with the existence [27–29], construction and properties of systems of finite sets with generalized concepts of balance and/or symmetry [30–32]. It is formulated in [33] that combinatorial designs can provide potential tools in the area of design of experiments, particularly for the design of biological experiments.

In this work, borrowing the concept of "balance" in combinatorial design theory, we propose a novel method for multifactorial bio-chemical experiments design. In the method, balanced templates from the existence and construction in combinational design theory are used to select the experiments which should be done for conditions observation. We can get balanced experimental data from the experiments selected that covers all the influencing factors of experiments for further processing, particularly for machine learning based modelling. Finally, a software with the proposed method is developed, which provides a simulation tool for designing multifactorial experiments, by which the designed experiments can cover influencing factors a pre-designed number of times.

Methods

In this section, we introduce the method proposed for experimental design based on combinational design theory.

Before introducing the method, we clarify the meaning of some involved symbols. Let m be the number of influencing factors for an experiment. These factors are denoted by $f_1, f_2, \ldots, f_m$. By n_j with $j = 1, 2, \ldots, m$, we denote the possible values of factor f_j. Parameter s is set to be covering times of all the factors. Among the n_j values of any factor f_j, it is not necessary to select all the values, but t values is sufficient to cover all the cases of factors. The number of experiments needed to cover all the cases of factors and all the values is denoted by r. In other word, we need design r experiments to cover all the factors s times with each time considering t values of any factor.

The mathematical model of experimental design

Let P be a $m \times n$ matrix recording the m influencing factors, where $n = \max\{n_1, n_2, \ldots, n_m\}$. Mathematically, matrix $P_{m \times n} = (p_{ij})_{m \times n}$ is denoted as follows

$$
P = \begin{bmatrix}
p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\
p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_{m1} & p_{m2} & p_{m3} & \cdots & p_{mn}
\end{bmatrix}.
$$

By p_{ij}, we denote the the jth possible value of the ith influencing factor to the experiment. The problem of experimental design is to find suitable values of r, s and t, with which we can obtain an experiment design, i.e., a group of experiments, $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_g\}$. Each experiment $\gamma_k \in \gamma$ is of the form $\gamma_k = \{p_{1k_1}, p_{2k_2}, \ldots, p_{mk_m}\}$. The experimental design to be found should satisfy the following items:

- $\bigcup_{k=1}^{g} \gamma_k = \{p_{ij} \mid i = 1, 2, \ldots, m, j = 1, 2, \ldots, n\}$;

- for any $\gamma_k = \{p_{1k_1}, p_{2k_2}, \ldots, p_{mk_m}\}$, its elements are from matrix P, where p_{jk_j} is elected from the jth row of matrix P with $j = 1, 2, \ldots, m$.

- in total $g \times m$, that is, r experiments in the design;

- given t elemental elements (crucial to the experiment) from matrix P, and these elements present in γ for s times.

When given the t elemental elements, the object is to finding certain experiment design γ, which has minimal value of r and maximal value of s.

Theoretical support from combinational design theory

With the notations defined above, the experimental design can be transferred as finding a suitable values and combination of parameters t, s and λ. In combinational design theory, the concept of difference set have some common features with parameters t, s and λ. This provides a way to design experiments by the way of finding difference set developed in combinational design theory. Let's briefly recall some basic concepts of difference set.

Let G be an Abelian group in modern algebra theory. A (v, k, λ) difference set is a subset D of a group G such that the order of G is v, the size of D is k, and every nonidentity element of G can be expressed as a product $d_1 d_2^{-1}$ of elements of D in exactly λ ways (when G can be written with a multiplicative operation) [34]. For any element g in G, if subset D is a difference set, then it holds $g \cdot D = \{g \cdot d: d \in D\}$ is also a difference set, which is named as a translate of D. It is known that the set of all translates of a difference set D can achieve a symmetric block design. In such a design there are v elements and v blocks. In each block of the design, it has k points, and each point is contained in k blocks [30].

In combinational design theory, there are some theorems on designing difference sets with distinguished values of parameters v, k and λ). With the concepts of difference set in combinational design theory, it is not hard to find that finding suitable and reasonable values of s, t and r is similar of designing a different set. The strategy used in constructing different sets can be used to determine values of r, s, and t, thus achieving a way for experimental design.

The algorithm for experimental design

In this subsection, we propose an algorithm for finding suitable values of parameters r, s, and t, like finding different sets in combinational design theory. The flowchart of the algorithm is shown in Fig 1.

The general process of the algorithm is as follows.

Step 1. Initialize parameters m, n and set $s = n$, $\gamma = \emptyset$;

Step 2. Generating matrix $P_{m \times n}$ with the n values of each of the m influencing factors.

Step 3. Selecting the t elemental elements from matrix P.

Step 4. Generating γ_i by randomly select one element from each row of matrix P.

Step 5. Updating $\gamma = \gamma \cup \gamma_i$.

Step 6. If there exits some other γ_i that can be generated, then go to **Step 4**.; otherwise go to **Step 7**.

Step 7. Check if the generated γ covers all the t pro-defined elemental elements of matrix P. If so, go to **Step 8**.; otherwise, updating $s = n - 1$ and go to **Step 1**. to repeat the process.

Step 8. Check if the generated γ matches the request of similarly by comparing any two groups of experiments. If so, go to **Step 9**.; otherwise, go to **Step 4**. to re-design the experiments.

Step 9. Check if the generated γ matches the request of balance, which is calculated by the rate between the minimal times and maximal times of the pairs of two influencing factors in the designed experiments. If so, halt the algorithm and output the designed experiments γ; otherwise, go to **Step 4**. to re-design the experiments.

Simulation tools

A software based on Visual Studio 2010 is developed for the the simulation of the proposed algorithm. The simulation tool produces a group of experiments with generalized concepts of balance and/or symmetry in combinational design theory. We set a similarly comparison mechanism to avoiding similar groups of experiments are designed. The starting page of the software is shown in Fig 2.

The meaning of the parameters in Fig 2 is as follows.

- The number of influencing factors is represented by m.

- The maximal number of possible values of all the influencing factor is n.

- Input 11 is the value of the minimal similarity between any couple of γ_i and γ_j in the designed γ. If there exists any two groups of designed experiments having similarity less the value of Input 11, then it repeats the algorithm to generate a new group of experiments.

- Input 12 is the number of total experiments that need be designed.

- Input 13 is the balance measure of the designed experiments, which is calculated by the rate between the minimal times and maximal times of the pairs of two influencing factors in the designed experiments.

For example, we can design experiments with $m = 6$, $n = 3$, Input 11 be 8, Input 12 be 10 and Input 13 be 0.001. The salutation page is shown in Fig 3.
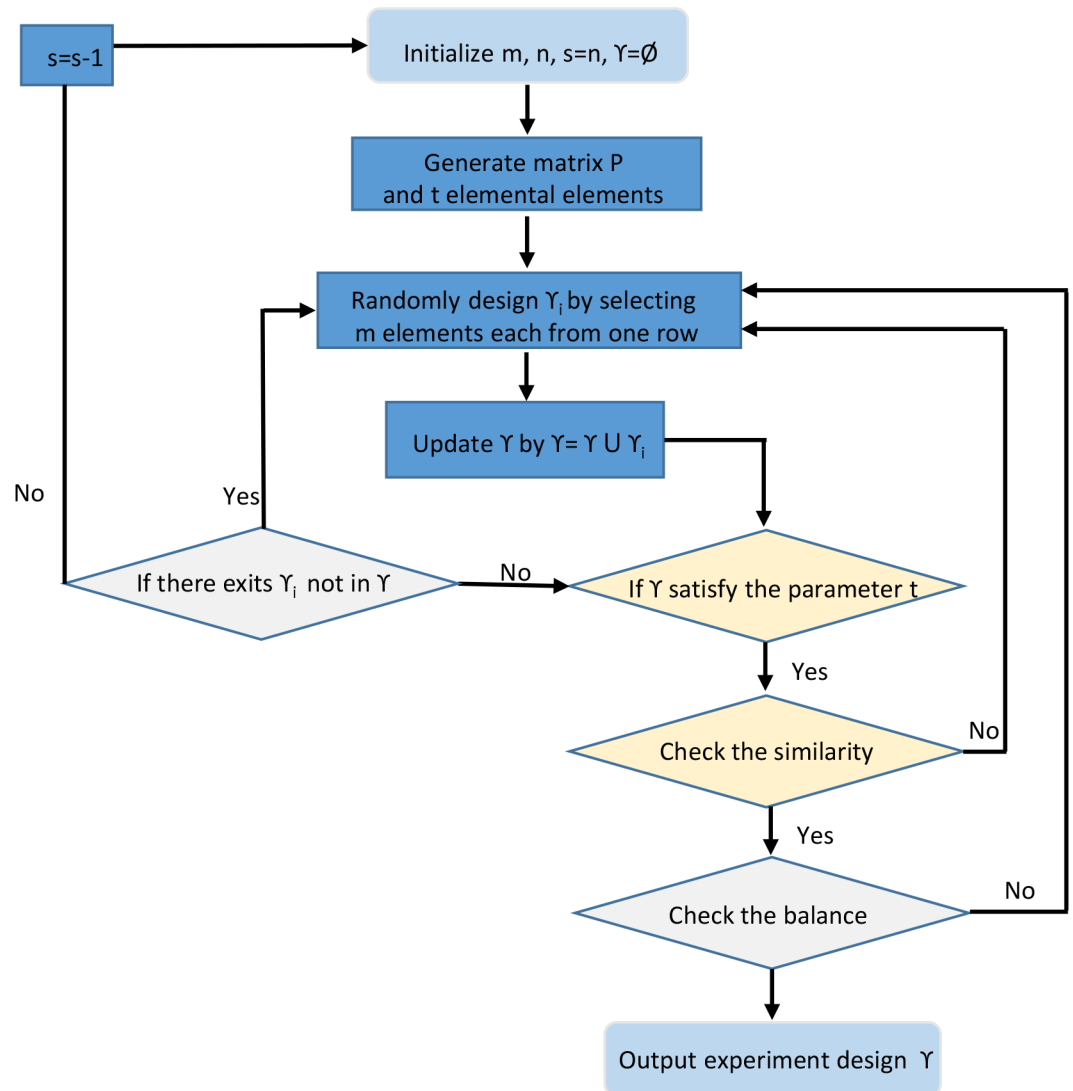
Fig 1. The flowchart of the algorithm.

The simulation result can be used to design experiment of fungal fermentation experiments. There are 6 influencing factors: (1) inoculum concentration, (2) the volume of liquid, (3) temperature, (4) PH value, (5) yeast concentration and (6) amylaceum concentration. Each influencing factor has 3 possible values, which is shown in Table 1.

Using the simulation result of γ, we can obtain the designed experiments as follows.

- $\gamma_1 = \{\{5g/L, 75ml, 32.5°C, 7, 4g/L, 70g/L\}, \{6g/L, 50ml, 30°C, 6, 8g/L, 45g/L\}, \{7g/L, 100ml, 27.5°C, 8, 6g/L, 85g/L\}\}$.

- $\gamma_2 = \{\{5g/L, 100ml, 30°C, 8, 6g/L, 45g/L\}, \{6g/L, 75ml, 27.5°C, 7, 4g/L, 85g/L\}, \{7g/L, 50ml, 32.5°C, 6, 8g/L, 70g/L\}\}$.

- $\gamma_3 = \{\{5g/L, 50ml, 27.5°C, 6, 8g/L, 85g/L\}, \{6g/L, 100ml, 32.5°C, 8, 6g/L, 70g/L\}, \{7g/L, 75ml, 30°C, 7, 4g/L, 45g/L\}\}$.

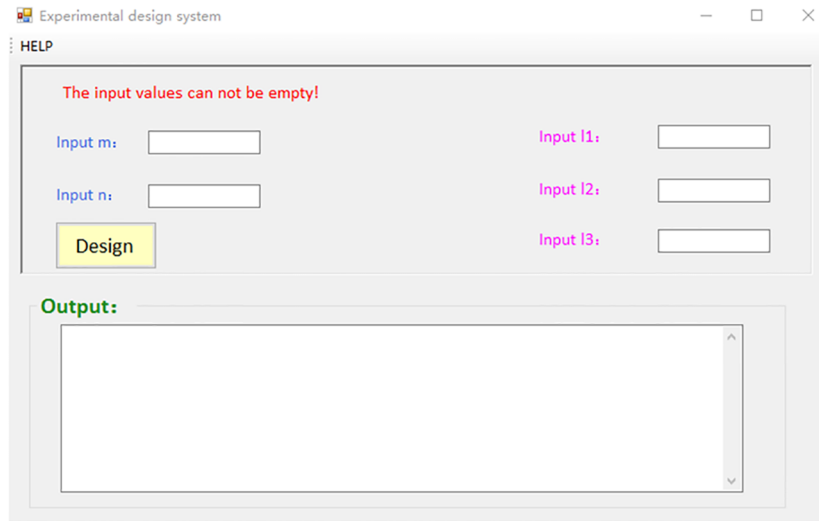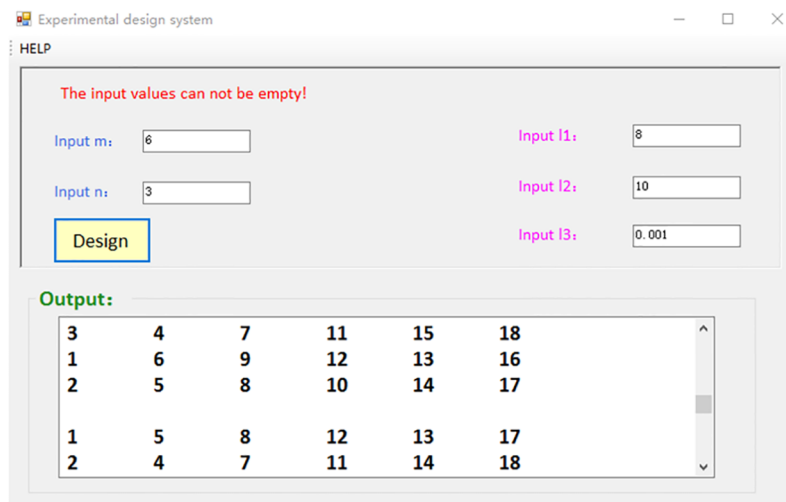**Fig 2. The starting page of the simulation software.**

https://doi.org/10.1371/journal.pone.0186853.g002

Fig 3. An example with inputs $m = 6$, $n = 3$, Input l1 be 8, Input l2 be 10 and Input l3 be 0.001.

https://doi.org/10.1371/journal.pone.0186853.g003

Table 1. Possible values of the influencing factors.

inoculum concentration	the volume of liquid	temperature	PH value	yeast	amylaceum
5g/L	50ml	27.5˚C	6	4g/L	45g/L
6g/L	75ml	30˚C	7	6g/L	70g/L
7g/L	100ml	32.5˚C	8	8g/L	85g/L

https://doi.org/10.1371/journal.pone.0186853.t001

The similarity of each pair of γ_i and γ_j with $i \neq j$ and $i, j \in \{1, 2, 3\}$ is less than 0.001. The balance measure of the designed experiments is 1/3.

Conclusion

In this work, we propose a novel method for multifactorial experiments design, where the generated concept of "balance" in combinatorial design theory is considered. In our method, balance and similarity are both used to select the experiments which should be done for conditions observation, with which we can get balanced experimental data from the experiments selected that covers all the influencing factors of experiments for further processing.

Since there is theoretical support in combinational design theory to ensure the exists of certain combination of parameters, our algorithm just starts from a random point and proceeds to find the possible designed experiments. By checking the similarity and balance designed experiments, we can choose to repeat the algorithm or output the desired experimental design. In complexity theory, starting from a random point would bring some extra computation consumption. It is of interest to design intelligent algorithms, such as GA with fitness function relating the similarity and balance, to improve the performance of our method.

For further research, some newly developed evolution computing models and algorithms, see e.g. [35–43], can be used to improve the balanced data sampling method. As well learning and training requests on the data for neural-like computing models and spiking neural networks [14, 44] should be considered in design balanced templates. In DNA computing, it needs to design DNA probes and DNA sequences [45–48], in which balanced template might provide useful tools. As well, some recently developed data processing and mining methods, such as the speculative approach to spatial-temporal efficiency for multi-objective optimization in cloud data and computing [49], privacy-preserving smart similarity search methods in simhash over encrypted data in cloud computing [49], k-degree anonymity with vertex and edge modification algorithm [50], kernel quaternion principal component analysis for object recognition [51], might be used for optimizing experiment design with intelligent methods.

Acknowledgments

Author Contributions

Investigation: Xun Wang, Pan Zheng.

Methodology: Yaping Fu.

Software: Beibei Sun.

Validation: Boyang Liu.

References

1. Kazdin AE. Single-case experimental designs: Strategies for studying behavior change. New York: Oxford University Press; 1982.

2. Cox GM, Cochran WG. Experimental designs. JSTOR; 1957.

3. Kleijnen JP. Regression and Kriging metamodels with their experimental designs in simulation: a review. European Journal of Operational Research. 2017; 256(1):1–16. https://doi.org/10.1016/j.ejor.2016.06.041

4. Campbell DT, Stanley JC. Experimental and quasi-experimental designs for research. Ravenio Books; 2015.

5. Atkinson A, Donev A, Tobias R. Optimum experimental designs, with SAS. vol. 34. Oxford University Press; 2007.

6. Krishnan SS, Sitaraman RK. Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs. IEEE/ACM Transactions on Networking. 2013; 21(6):2001–2014. https://doi.org/10.1109/TNET.2013.2281542

7. Montgomery DC. Design and analysis of experiments. John Wiley & Sons; 2017.

8. Tyssedal J. Plackett–burman designs. Encyclopedia of statistics in quality and reliability. 2008;. https://doi.org/10.1002/9780470061572.eqr020

9. Martinez-Cantin R. Bayesopt: A bayesian optimization library for nonlinear optimization, experimental design and bandits. The Journal of Machine Learning Research. 2014; 15(1):3735–3739.

10. Khuri AI, Mukhopadhyay S. Response surface methodology. Wiley Interdisciplinary Reviews: Computational Statistics. 2010; 2(2):128–149. https://doi.org/10.1002/wics.73

11. Ahmadi M, Vahabzadeh F, Bonakdarpour B, Mofarrah E, Mehranian M. Application of the central composite design and response surface methodology to the advanced treatment of olive oil processing wastewater using Fenton's peroxidation. Journal of Hazardous Materials. 2005; 123(1):187–195. https://doi.org/10.1016/j.jhazmat.2005.03.042 PMID: 15993298

12. Yang Y, Heffernan R, Paliwal K, Lyons J, Dehzangi A, Sharma A, et al. SPIDER2: A Package to Predict Secondary Structure, Accessible Surface Area, and Main-Chain Torsional Angles by Deep Neural Networks. Prediction of Protein Secondary Structure. 2017; p. 55–63. https://doi.org/10.1007/978-1-4939-6406-2_6

13. Wang X, Song T, Zheng P, Hao S, Ma T. Spiking Neural P Systems with Anti-Spikes and without Annihilating Priority. Rammian Journal of Science and Technology. 2017; 20(1):32–41.

14. Song T, Gong F, Liu X, Zhao Y, Zhang X. Spiking neural P systems with white hole neurons. IEEE transactions on nanobioscience. 2016; 15(7):666–673. https://doi.org/10.1109/TNB.2016.2598879 PMID: 28029614

15. Wang X, Song T, Gong F, Zheng P. On the computational power of spiking neural P systems with self-organization. Scientific Reports. 2016; 6:27624. https://doi.org/10.1038/srep27624 PMID: 27283843

16. Song T, Wang X, Zhang Z, Chen Z. Homogenous spiking neural P systems with anti-spikes. Neural Computing and Applications. 2014; 24(7-8):1833–1841. https://doi.org/10.1007/s00521-013-1397-8

17. Song T, Wang X. Homogenous spiking neural P systems with inhibitory synapses. Neural Processing Letters. 2015; 42(1):199–214. https://doi.org/10.1007/s11063-014-9352-y

18. Xue Y, Jiang J, Zhao B, Ma T. A self-adaptive artificial bee colony algorithm based on global best for global optimization. Soft Computing. 2017;(8):1–18.

19. Shen J, Shen J, Chen X, Huang X, Susilo W. An Efficient Public Auditing Protocol With Novel Dynamic Structure for Cloud Data. IEEE Transactions on Information Forensics & Security. 2017; 12(10):2402–2415. https://doi.org/10.1109/TIFS.2017.2705620

20. Fu Z, Huang F, Ren K, Weng J, Wang C. Privacy-Preserving Smart Semantic Search Based on Conceptual Graphs Over Encrypted Outsourced Data. IEEE Transactions on Information Forensics & Security. 2017; 12(8):1874–1884. https://doi.org/10.1109/TIFS.2017.2692728

21. Xia Z, Wang X, Zhang L, Qin Z, Sun X, Ren K. A Privacy-Preserving and Copy-Deterrence Content-Based Image Retrieval Scheme in Cloud Computing. IEEE Transactions on Information Forensics & Security. 2016; 11(11):2594–2608. https://doi.org/10.1109/TIFS.2016.2590944

22. Xia Z, Wang X, Sun X, Wang Q. A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data. IEEE Transactions on Parallel & Distributed Systems. 2016; 27(2):340–352.

23. Li Z, Xin Y, Wang X, Sun B, Xia S, Li H, et al. Optimization to the Culture Conditions for Phellinus Production with Regression Analysis and Gene-Set Based Genetic Algorithm. BioMed research international. 2016; 2016.

24. Li Z, Sun B, Xin Y, Wang X, Zhu H. A computational method for optimizing experimental environments for Phellinus igniarius via genetic algorithm and BP neural network. BioMed research international. 2016; 2016.

25. Banga JR, Balsa-Canto E. Parameter estimation and optimal experimental design. Essays in biochemistry. 2008; 45:195–210. https://doi.org/10.1042/bse0450195 PMID: 18793133

26. Balsa-Canto E, Alonso AA, Banga JR. Computational procedures for optimal experimental design in biological systems. IET systems biology. 2008; 2(4):163–172. https://doi.org/10.1049/iet-syb:20070069 PMID: 18681746

27. Wang X, Wu D. The existence of almost difference families. Journal of Statistical Planning and Inference. 2009; 139(12):4200–4205. https://doi.org/10.1016/j.jspi.2009.06.004

28. Wang X, Wu D, Cheng M. The existence of (q, k, λ, t)-ADFs for $k = 4, 5, 6$. Journal of Statistical Planning and Inference. 2010; 140(11):3243–3251. https://doi.org/10.1016/j.jspi.2010.04.022

29. Tso R, Miao Y. A Survey of Secret Sharing Schemes Based on Latin Squares. In: International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Springer; 2017. p. 267–272.

30. Beth T, Jungnickel D, Lenz H. Design theory. vol. 69. Cambridge University Press; 1999.

31. Hall M. Combinatorial theory. vol. 71. John Wiley & Sons; 1998.

32. Wang X, Miao Y, Cheng M. Finding motifs in DNA sequences using low-dispersion sequences. Journal of Computational Biology. 2014; 21(4):320–329. https://doi.org/10.1089/cmb.2013.0054 PMID: 24597706

33. Assmus EF, Key JD. Designs and their Codes. vol. 103. Cambridge University Press; 1994.

34. Moore EH, Pollatsek HSK. Difference Sets: Connecting Algebra, Combinatorics, and Geometry. vol. 67. American Mathematical Soc.; 2013.

35. Zhang L, Pan H, Su Y, Zhang X, Niu Y. A Mixed Representation-Based Multiobjective Evolutionary Algorithm for Overlapping Community Detection. IEEE Transactions on Cybernetics. 2017; 47 (9):2703–2716. https://doi.org/10.1109/TCYB.2017.2711038 PMID: 28622681

36. Bin Gu, Sheng Victor S., Tay Keng Yeow, Romano Walter, and Li Shuo, Incremental Support Vector Learning for Ordinal Regression, IEEE Transactions on Neural Networks and Learning Systems, 2015; 26(7):1403–1416. https://doi.org/10.1109/TNNLS.2014.2342533

37. Tian Y, Cheng R, Zhang X, Cheng F, Jin Y, An indicator based multi-objective evolutionary algorithm with reference point adaptation for better versatility, IEEE Transactions on Evolutionary Computation, 2017, in press. https://doi.org/10.1109/TEVC.2017.2749619

38. Zhang X, Tian Y, Cheng R, Jin Y. A decision variable clustering based evolutionary algorithm for large-scale many-objective optimization. IEEE Transactions on Evolutionary Computation, 2016, in press. https://doi.org/10.1109/TEVC.2016.2600642

39. Zhang X, Duan F, Zhang L, Cheng F, Jin Y, Tang K. Pattern recommendation in task-oriented applications: a multi-objective perspective. IEEE Computational Intelligence Magazine, 2017, 12(3):43–53. https://doi.org/10.1109/MCI.2017.2708578

40. Chen X, Wang C, Tang S, Yu C, Zou Q. CMSA: a heterogeneous CPU/GPU computing system for multiple similar RNA/DNA sequence alignment. BMC bioinformatics. 2017; 18(1):315. https://doi.org/10.1186/s12859-017-1725-6 PMID: 28646874

41. Zeng X, Lin W, Guo M, Zou Q. A comprehensive overview and evaluation of circular RNA detection tools. PLoS computational biology. 2017; 13(6):e1005420. https://doi.org/10.1371/journal.pcbi.1005420 PMID: 28594838

42. Zou Q, Wan S, Ju Y, Tang J, Zeng X. Pretata: predicting TATA binding proteins with novel features and dimensionality reduction strategy. BMC systems biology. 2016; 10(4):114. https://doi.org/10.1186/s12918-016-0353-5 PMID: 28155714

43. Liu Y, Zeng X, He Z, Zou Q. Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. IEEE/ACM transactions on computational biology and bioinformatics. 2016;.

44. Song T, Zou Q, Liu X, Zeng X. Asynchronous spiking neural P systems with rules on synapses. Neurocomputing. 2015; 151:1439–1445. https://doi.org/10.1016/j.neucom.2014.10.044

45. Shi X, Wang Z, Deng C, Song T, Pan L, Chen Z. A novel bio-sensor based on DNA strand displacement. PLoS One. 2014; 9(10):e108856. https://doi.org/10.1371/journal.pone.0108856 PMID: 25303242

46. Shi X, Wu X, Song T, Li X. Construction of DNA nanotubes with controllable diameters and patterns using hierarchical DNA sub-tiles. Nanoscale. 2016; 8(31):14785–14792. https://doi.org/10.1039/C6NR02695H PMID: 27444699

47. Xu J. Probe machine. IEEE transactions on neural networks and learning systems. 2016; 27(7):1405–1416. https://doi.org/10.1109/TNNLS.2016.2555845

48. Shi X, Chen C, Li X, Song T, Chen Z, Zhang Z, et al. Size-controllable DNA nanoribbons assembled from three types of reusable brick single-strand DNA tiles. Soft matter. 2015; 11(43):8484–8492. https://doi.org/10.1039/C5SM00796H PMID: 26367111

49. Liu Q, Cai W, Shen J, Fu Z, Liu X, Linge N. A speculative approach to spatial-temporal efficiency with multi-objective optimization in a heterogeneous cloud environment. Security & Communication Networks. 2016; 9(17):4002–4012. https://doi.org/10.1002/sec.1582

50. Ma T, Zhang Y, Cao J, Shen J, Tang M, Tian Y, et al. KDVEM: a (k)-degree anonymity with vertex and edge modification algorithm. Computing. 2015; 97(12):1165–1184. https://doi.org/10.1007/s00607-015-0453-x

51. Chen B, Yang J, Jeon B, Zhang X. Kernel quaternion principal component analysis and its application in RGB-D object recognition. Neurocomputing. 2017.