AN ENHANCEMENT OF PLANNED RANDOM ALGORITHM APPLIED ON AN ONLINE QUESTIONNAIRE

Jemuel Fontila[a], Jiro Mark Garcia[a], Mark Jimwell Lagarta[a], Mark Christopher Blanco[a], Dan Michael Cortez[a]
[a]**Pamantasan ng Lungsod ng Maynila, Manila, *Philippines***

Abstract

Like many other innovations, randomization has been a key factor in various fields such as games, science, art, statistics, and other areas. But although treated as common knowledge, randomization is often misunderstood by users when they start to perceive relationships or connections between randomly distributed data, this tendency is known as apophenia. It is observed mostly when a randomized dataset shows the same elements in a row or in multiple repetitions, an occurrence that is plausible given the nature of randomization. Nevertheless, this may appear as biased or non-random to humans, thinking that since an element had appeared once or a couple of times then it should not appear next to or within a set amount of time, another user misunderstanding known as gambler's fallacy. The Planned Random algorithm (PR algorithm) was created as a solution to this problem, as it introduces a non- random process to lessen the repetitions of data in the returned dataset and making the result less random to appear more random. This study aims to enhance the Planned Random algorithm by increasing its apparent randomness by further minimizing the repetition of data in a resulting dataset. 5 final test simulations were done for both the existing PR algorithm and enhanced PR algorithm wherein the enhanced algorithm showed a promising 22.22% decrease in both total and average repetitions compared to the existing algorithm. In conclusion, the enhanced algorithm had seen a significant increase in apparent randomness when compared to the existing PR algorithm.

Introduction

Randomness is the unpredictability or lack of pattern in a set of events. Randomly generated numbers are often used in various applications such as statistics, encryption, science, gaming, etc. The numbers are generated by various standard algorithms, which exist and are predefined in libraries of almost all programming languages. Though the generated numbers are unpredictable, they may seem repetitive or appear to be following a pattern.

Computer-generated random numbers are meant to be unpredictable, but they have often shown pattern or repetition. This is because computers are systemized, and the numbers generated rely on the time or the CPU clock. A computer can be programmed to return a random number array from a data set of numbers, with the returned data set having the possibility of containing repeated values. Because of this, the returned data set may appear to be biased or not random to us humans. This can be attributed to apophenia, which is the tendency to perceive meaningful connections between unrelated things, and in this case, it is the tendency to discern patterns out of total randomness from the numbers in the data set. Because of this, the data sets that are generated from the various standard algorithms are difficult to appear to be random.

Considering all of this, an algorithm was proposed which was devised to increase the apparent randomness and distribute values that will prevent the tendency to see a pattern from these values. Because of this, an element of planning within the proposed algorithm with the way it returns values is used to achieve this goal.

The Planned Random Algorithm works by using time-based randomization that is aimed at increasing the apparent randomness of the shuffled list. The algorithm is comparable to the Modern Fisher–Yates Algorithm in terms of efficiency due to its time complexity being the same at O(n) [1]. This will be ideal since according to a study, the advantage of using Modern Fisher–Yates Algorithm is the

effectiveness of the randomization method due to it being unbiased and the optimal complexity of the algorithm [2].The use of the Planned Random Algorithm at its current state is general and only applies to a specific list of numbers, which compares the primitive values.

```
Original List: 1 2 2 5 5 6 7 8 10 10
Shuffled List: 6 5 2 8 10 5 7 1 10 2
Process returned 0 (0x0)    execution time : 0.058 s
Press any key to continue.
```

Fig. 1. List Comparison of the Input and Output. Source: [1]

The algorithm will certainly randomize the occurrence of the number/object in a brief period of the loop albeit it would not appear random if applied to a list of games, movies, songs, and questions. For instance, you have 5 questions in a list, but the distance between those questions can be near in the shuffling of items as the algorithm does not consider other factors such as the categories of the questions such as Math, Science, English, Filipino, Abstract, the subcategories of the question such as Geometry, Biology, Grammar, and the difficulty of the question such as Easy, Average, Hard. We want to inhibit the questions with the same properties to appear in a brief period and decrease the likelihood of the chance to be with the synonymous value after every question in a list.

Below is the statement of the researchers that can be improved in the future to increase the algorithm which means it must be scaled to practical application to truly achieve the randomness.

7 Future Work

The algorithm can be extended to all practical applications and can even be made to learn by using machine learning. The algorithm can grow every day with its use.

For example: In the application of Songs and Shuffling, it can monitor the number of times a song is played. As a human brain will tend to perceive the occurrence of the most played song as highly repetitive, the algorithm will prevent playing of that song in the beginning.

Fig. 2. Future Work. Source: [1]

The use of additional factors came from the implemented solution by Apple for their playlist shuffling where they introduced a feature called "Smart Shuffle," which controls how the shuffling works by taking into consideration the song artist and song album. This was implemented because their users usually heard the same artist twice in a row, and they came to a judgement that the shuffling feature could not be random [3].

The proposed algorithm will be applied to an Online Questionnaire. For instance, the values will be questions and their attributes will be the category, subcategory, and difficulty. Then, these attributes will be checked to prevent questions with similar attributes from appearing in quick succession. The enhancements will serve as a cheating deterrent by changing the question sequence for each output test. In fact, in an article entitled "7 Assessment Challenges of Moving Your Course Online (and a Dozen+ Solutions)" It was stated that assessments or quizzes that were effective in the face-to-face environment may require alterations or replacement to be effective in the current online setting due to problems of academic dishonesty that arises in online learning environments. Test question shuffling was introduced as one of the solutions to deter examinees from cheating as two students that are seated adjacent to each other have a reduced likelihood of being able to take the same test together due to different question orders [4,5,6].
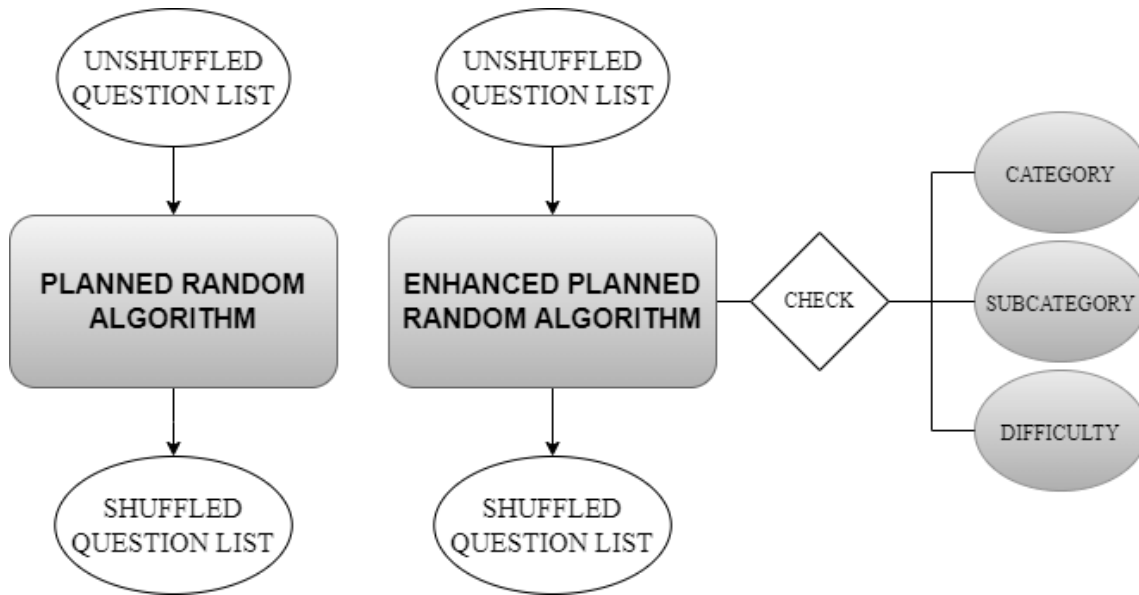
Fig. 3. Visualization

Methods

The methodology that the researchers used in this application is a quantitative research method specifically, an experimental research design to determine and analyze the effects and improvements of the enhancements that will be applied to the existing Planned Random Algorithm. Experimental research design divides the simulation into two groups. The two simulation groups consist of the control simulation and the experimental simulation. The control simulation will be the existing PR Algorithm that will serve as the baseline for this experiment. The experimental simulation will be the enhanced algorithm with the proposed changes wherein the algorithm will take into consideration additional properties when distributing the questions. The simulations will be using a quiz data set to produce results and compare the existing algorithm and the proposed enhanced algorithm to see the difference between the two, this will determine if the changes applied will make a difference and improve the randomness of the shuffled list.

For both tests, an array of 50 elements with the same category, but various subcategory and difficulty rating is used for all test cases, taken with repetition of subcategory and difficulty rating, is shuffled using for both simulation groups. The array will be shuffled according to their respective algorithms. Afterwards, a new shuffled list would be constructed again in the next iteration. This simulation will transpire for 1000 iterations and be carried out five times. The sole objective of this test is to determine how random the algorithm appears to be. Considering this, at the end of each iteration, the current shuffled list and the previous shuffled list will be compared. The two lists are utilized to check for a match. The first item up until the last item of each list, there will be a match checking that will be conducted to compare their corresponding subcategory and difficulty rating,if they are equal in both values, a variable counter is incremented by one to signify repetition of pattern. Consequently, at the end of 1000 iterations, the counter is presented on the screen. The results can then be compared to determine the difference between the respective randomness of the two algorithms.

Results and discussion

The figure below shows the comparison of the results regarding the total repetitions and average repetitions of the simulation runs when applying the existing Planned Random algorithm and the proposed enhanced version of the algorithm.

Table 1. Simulation Results

Planned Random Algorithm		Enhanced Proposed Algorithm	
Simulation	Values	Simulation	Values

Jemuel Fontila et.al(2022)

1	4,291	**1**	3,403
2	4,360	**2**	3,286
3	4,293	**3**	3,385
4	4,277	**4**	3,223
5	4,190	**5**	3,356
Total Repetitions: 21,411		**Total Repetitions:** 16,653	
Avg. Repetitions: 4,282.2		**Avg. Repetitions:** 3,330.6	

The results in the table show that the existing Planned Random Algorithm has 21,411 total repetitions in the 5 test runs, averaging 4,282.2 repetitions. While the results in the table for the Enhanced Proposed Algorithm show that it has 16,653 total repetitions in the 5 test runs, averaging 3,330.6 repetitions.

The data collected from these simulations exemplifies that it can be concluded that the Proposed Enhanced Algorithm increased the apparent randomness of the shuffle significantly, as it led to a 22.22% decrease in both the average repetitions and total repetitions per case from the PR algorithm

Impact

This manuscript can provide technological impact to various areas of discipline related to the usage of randomization. Such applications include science, statistics, games, music, arts, and many others where the enhanced algorithm could be potentially applied to.

Conclusions

In the current setting of remote learning, this study about the enhancement of the Planned Random Algorithm is useful and beneficial for schools and universities that utilize online questionnaires to conduct their examinations. This is because the algorithm will properly distribute and balance the questions throughout the questionnaire or exam. Therefore, it will discourage students from memorizing answers and performing academic dishonesty. This in turn will encourage the students to study diligently and improve the accuracy of the results of the exams.

Furthermore, with the results of the simulation runs done, it was clearly depicted that the enhancement for the Planned Random Algorithm showed that there was a 22.22% decrease in both average and total repetitions per simulated case. Therefore, this enhancement research study boasts a successful improvement upon the original algorithm.

The researchers recommend applying the Planned Random Algorithm in different game genres such as Role- Playing Games (RPG) where monster difficulty, density, and distribution plays a crucial role in user experience and overall game mechanics, or in Adventure type games wherein treasure distribution across the map is a fundamental aspect in making the player roam the entirety of the map or level.

Furthermore, another recommendation is to apply the algorithm to a different type of application, such as song shuffling wherein attributes like the song artist, genre, and rating of the song are considered to further even the distribution of songs throughout the playlist.

Lastly, the researchers recommend improving the scalability of the current algorithm with regards to the number of attributes that the algorithm can consider (currently only 3). This is due to the growing quantity of data present in applications that may prove useful in improving the algorithm further or making a more even distribution of elements within the application where it is utilized. It can also be combined with time restraints and other functions in an online exam application that can reduce the chances of cheating as the researchers primarily focused on enhancing the algorithm and not the overall utilization of the algorithm regarding exams.

Conflict of interest

There is no conflict of interest

References

1. A. Pant, K. Agrawal and B. Tripathy, "Planned Random Algorithm", *Advances in Intelligent Systems and Computing*, pp. 119-127,(2018). https://doi.org/10.1007/978-981-10-7563-6_13
2. F.Panca Juniawan, H.Arie Pradana, Laurentinus and D.YunySylfania,"Performance Comparison of Linear Congruent Method and Fisher-Yates Shuffle for Data Randomization", *Journal of Physics: Conference Series*, vol. 1196, p. 012035, (2019). https://doi.org/10.1088/1742-6596/1196/1/012035
3. B. Cohen,"Spotifymadeitsshuffle featurelessrandom sothatitwould actually feelmorerandom to listeners here's why", *Business Insider*, 2020. [Online]. https://www.businessinsider.com/spotify-made-shuffle-feature-less-random-to-actually-feel-random-2020-3?international=true&r=US&IR=T.
4. Moore,E.,2013. 7Assessment Challenges of MovingYourCourse Online(andaDozen+Solutions).Faculty Focus. https://www.facultyfocus.com/articles/online-education/7-assessment-challenges-of-moving-your- course- online-solutions/
5. Ibrahim, S. Jafar Ali, and M. Thangamani. "Enhanced singular value decomposition for prediction of drugs and diseases with hepatocellular carcinoma based on multi-source bat algorithm based random walk." *Measurement* 141 (2019): 176-183.
6. Ibrahim, S. Jafar Ali, and M. Thangamani. "Prediction of Novel Drugs and Diseases for Hepatocellular Carcinoma Based on Multi-Source Simulated Annealing Based Random Walk." *Journal of medical systems* 42, no. 10 (2018): 188.