

# A Random Matrix - based Fraud Prevention Model

Monday Eze

Department of Comp Sc. Babcock University,  
Nigeria

Sam Ogunlere

Department of Comp. Sc, Babcock University,  
Nigeria

## ABSTRACT

Computer-based examination fraud control is an important research area in e-testing. Fraud prevention ensures that the outcome of academic or professional tests approximates the actual capabilities of the candidates in question. Thus, instead of depending wholly on human efforts to monitor a real life examination, requisite computational techniques could be deployed to ensure a more effective invigilation process. Many cases of cheating in examinations involve the collusion of two or more individuals, especially based on the level of familiarity that may have existed before the examination. Thus, an effective control system should strongly incorporate anti-collusion measures. The major contribution of the Matrix-Based Fraud Check technique is the application of randomized algorithms to prevent examination fraud. This research achieves this by first breaking the pre-examination social links that could lead to examination collusions. The strength of this model is that it could analyze the existing seating arrangements and as well suggests the most optimal arrangement that reduces collusion to the barest minimum. The model also generates a watch list of candidates that are most likely to be vulnerable to collusion in a particular examination hall. Such valuable information will no doubt guide the examiners in the invigilation process.

## Keywords

Randomized Algorithms, Matrix-Based Fraud Check, e-Testing, Watch List

## General Terms

Scientific Computing, Computer Applications.

## 1. INTRODUCTION

The Matrix-Based Fraud Prevention Algorithm (MFP) is a new computational technique [1] that could be applied in a class-room setting to prevent examination fraud. In its early developmental stage, the MFP algorithm was partly presented in the IMSCI 2014 Conference [2]. The aim of presenting the earlier work in a peer-reviewed international conference is to capture the critical feedbacks and contributions of the research community [3] with the view to improve its efficiency. This paper therefore encapsulates the most up to date version of the work on matrix-based fraud prevention algorithm. For instance, the original model catered for a maximum examination room size of 30 candidates, made up of six rows by five columns, while the current research could take on a larger examination hall size of up to 81 candidates, equivalent to a data dimension of 9 rows by 9 columns. The Matrix-Based Fraud Prevention (MFP) Algorithm computes a set of integer values known as *Familiarity Index* used to determine the level of familiarity between individual human beings/seats in the examination hall, and also applied in the overall fraud prevention algorithm.

## 2. THE ALGORITHM

The proposed Matrix-Based Fraud Prevention (MFP) Algorithm, also known as random seat arrangement (RSA model) for examination fraud prevention attempts to break the possible effects of established social contacts or familiarity [4] in the examination hall. Previous scientific research points to the fact that widespread examination fraud could easily be perpetrated based on faulty seating arrangement in the examination hall. For instance, a related research [5] reported that 66% of the frauds observed in a particular undergraduate examination came in two categories - 19% were due to collusion between students, while 47% were due to spying of other students' examination scripts. One of the assumptions of the current research is that when two or more persons have pre-examination seating contacts, they could take advantage of their pre-knowledge of each other's academic capabilities to collude with themselves in the examination hall. This is particularly possible if they usually sit close to themselves in the class-room before the exams, and were also allowed to sit close to themselves during the exams. Thus, in order to control examination fraud [6], a computerized algorithm is hereby applied to change the original sitting arrangement. The current algorithm is recommended for internal examinations. These are examinations being conducted by an institution for the students who are currently enrolled. For example, the semester exams, the mid-term tests, among others. The randomized algorithm [7] takes the digitized format of the pre-examination seating arrangement as input, in order to derive the randomized format. The resulting output is the recommended seating arrangement. The details of the MFP algorithm will be presented in this paper.

### 2.1 Algorithmic Details

The MFP algorithm makes use of an original seating arrangement, which is approximated to a two dimensional rectangular or square room. In electronic terms, the 'original seats' are the elements of a matrix, representing a hypothetical class room in Table I. The table element  $C_{xy}$  for instance, represents a person sitting in row  $x$ , column  $y$  in the seating matrix. As earlier mentioned, one of the key aims of the MFP algorithm is to breach the familiarity existing between candidates in the examination hall. This is based on the assumption that, the more the familiarity between two persons, the greater the possibility of collusion. Breaching the familiarity [8] is achieved by introducing a measure of randomness [9]. A computational procedure is also used in this research to estimate the familiarity between any two persons, based on the seating arrangement [10]. The familiarity index between two candidates is defined as the measure of the absolute value (ABS) of the difference between their corresponding indices in the seating matrix. Thus in mathematical terms, the measure of the familiarity between the candidates  $C_{wx}$  and  $C_{yz}$  is given by equation (1),

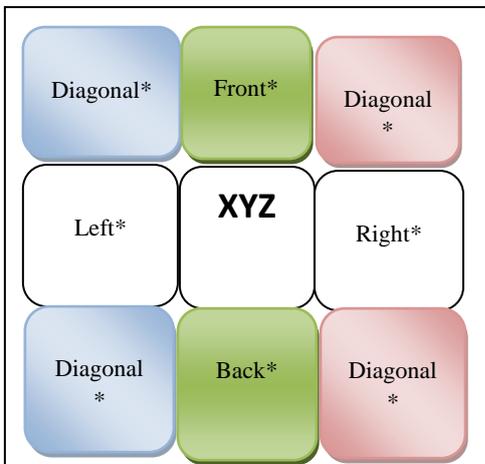
$$FamiliarityIndex (C_{wx}, C_{yz}) = ABS(wx - yz) \quad (1)$$

where, w and x are the rows and columns of the first candidate, y and z are the rows and columns of second candidate, while wx and yz are the indices of the seating matrix [11].

Example, for the seating positions  $C_{wx} = C_{23}$ , and  $C_{yz} = C_{36}$ , the  $FamiliarityIndex (C_{23}, C_{36}) = ABS(23-36)$ , which is 13. Two persons are said to be very close, if their *familiarity index* is in the set {1,9,10,11}. This occurs when a particular candidate ‘XYZ’ in an examination is sitting adjacent to any of the eight positions marked ‘\*’ in Fig. 1. Thus, the seating positions adjacent to the right, left, front, back or diagonal positions of ‘XYZ’, will lead to such *familiarity indices*. Some examples of such positions as can be extracted from Table I are {C22, C23}, {C34, C43}, {C55, C45}, {C52, C63} and so on.

**Table I: A Hypothetical Class Room**

|     |     |     |     |     |
|-----|-----|-----|-----|-----|
| C11 | C12 | C13 | C14 | C15 |
| C21 | C22 | C23 | C24 | C25 |
| C31 | C32 | C33 | C34 | C35 |
| C41 | C42 | C43 | C44 | C45 |
| C51 | C52 | C53 | C54 | C55 |
| C61 | C62 | C63 | C64 | C65 |



**Fig 1: Adjacency Positions**

Thus, the MFP algorithm strives to computationally re-arrange the original seating positions, so as to avoid occurrences of familiarity indices ‘1’, ‘9’, ‘10’ and ‘11’.

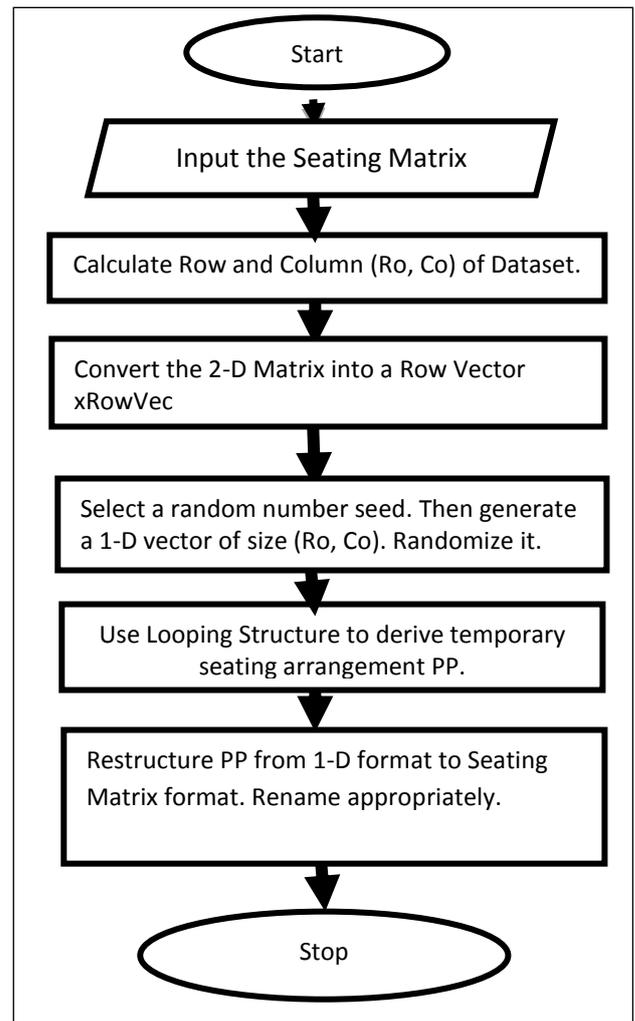
## 2.2 System Implementation and Output

The MFP algorithm was implemented in MATLAB Version 7.8.0.347 (R2009a), and tested using a classroom size of 40 candidates. The input dataset [12] is a two dimensional 5 x 8 array of seating arrangement. In the source code, this original seating matrix is denoted by ‘OrigSeatMatrix’. The resulting output, which is a randomized seating arrangement, is another 5 x 8 array called ‘RandSeatMatrix’. The exact input dataset is shown in Table II.

**Table II: Dataset for Seating Arrangement Runs**

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 |
| 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 |
| 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 |

The flowchart [13] used for the system implementation is shown in Fig. 2, while the corresponding source code is shown in the appendix A. The output of the MFP algorithm is shown in Fig. 3.



**Fig 2: Flowchart for MFP Algorithm**

## 3. SYSTEM VALIDATION

A validation program [14] was developed as part of the post MFP model runs. The aim of this is to scan through the result of the randomization operation, and detect isolated cases where the familiarity indices of 1,9,10 and 11 remained unchanged. Such isolated cases form what is termed the *watch list* [15] of candidates, for closer attention by the invigilators. The source code for the watch list is called watchProgram.m, and is listed in Appendix A of this research paper. Comparing the final output in Fig 3 with the watch list evolution

procedure as explained using Fig 1, it follows that the resulting watch list consists of the five pairs of candidates: {32, 31}, {17,16}, {53,52}, {26,27} and {23,12}. Thus, these isolated positions are either changed manually, or left as a watch list for the invigilators.

#### 4. CONCLUSION

The computational objectives of the MFP algorithm are assumed to be fully realized with the generation of a new randomized seating arrangement, with precise and minimal number of items in the watch list. Future research will explore other algorithms that randomize the examination contents. Future focus will also be extended to integration of artificial intelligence into the process of e-testing, as a means of controlling inherent frauds.

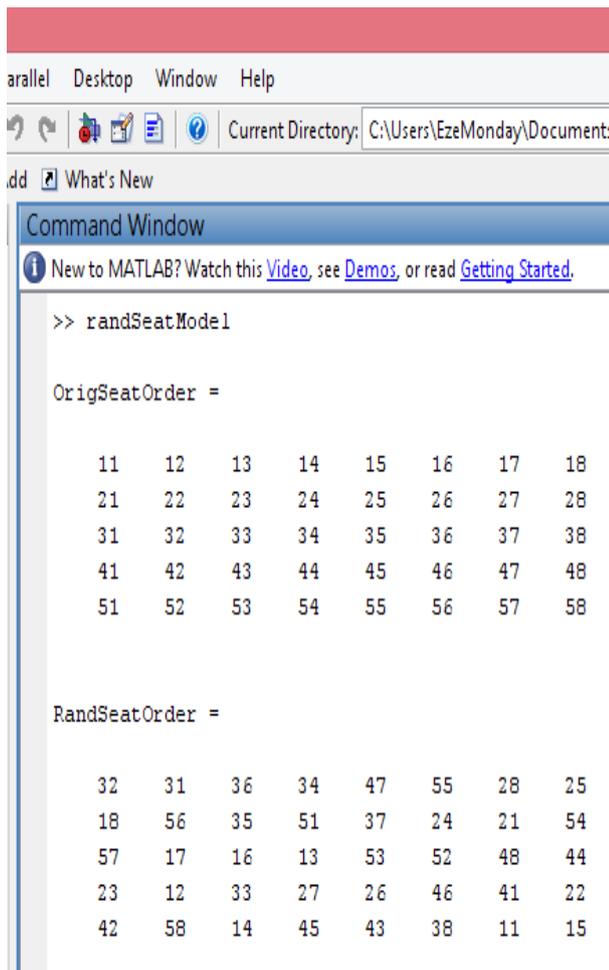


Fig 3: Output of MFP Model Runs

#### APPENDIX A

```
%*****SOURCE CODE*****
% Aim of Program: Matrix-Based Fraud Prevention Model
% Prog. Lang: MATLAB
% Module Name: RSeatArr.m
%DataSet: The Matrix of Original Seating
% Arrangement
%*****
% Bring in the DataSet
OrigSeatMatrix= [11 12 13 14 15 16 17 18;
```

```
21 22 23 24 25 26 27 28;
31 32 33 34 35 36 37 38;
41 42 43 44 45 46 47 48;
51 52 53 54 55 56 57 58]
```

```
%Take note of the rows and columns .
[numrow,numcol]=size(OrigSeatMatrix);

%The 2-dimentional array is first converted to a row vector.
xrovec= reshape(OrigSeatMatrix',numrow*numcol,1);

%Randomize the Row Vector:
rand(1000);
rpam= randperm(numrow*numcol);

%Generate the new Matrix gg in Randomized Format.
% The rpam positions will be used to randomize gg thru the
index
for k=1:numrow*numcol
    gg (rpam(k)) =xrovec (k);
end

%Reshape back into a 3 by 4 matrix
RandSeatMatrix=reshape(pp,numrow,numcol)

%*****SOURCE CODE*****
% Aim of Program: Used to generate an isolated list
% of Candidates in the Watch List for
% invigilators attention.
% Prog. Lang: MATLAB
% Module Name: watchProgram.m
%DataSet: Output of the model run as
% described in the paper.
%*****

% Assign the variable name 'a' to the RandSeatMatrix.
a=RandSeatMatrix
for m=1:5
    for wk=1:8
        if (wk < 8)
            x =abs(a(m,wk) - a(m,wk+1));
            if (x==1)
                m;
                wk;
                a (m,wk)
                a(m,wk+1)
            end
        end
    end
end
end
for m=1:5
    for wk=1:8
        if (wk < 8)
            x =abs(a(m,w) - a(m,wk+1));
            if (x==9)
                m;
                wk;
                a (m,wk)
                a(m,wk+1)
            end
        end
    end
end
end
for m=1:5
```

```
for wk=1:8
  if (wk < 8)
    x =abs(a(m,wk) - a(m,wk+1));
    if (x==10)
      m;
      wk;
      a (m,wk)
      a(m,wk+1)
    end
  end
end
end
for m=1:5
  for wk=1:8
    if (w < 8)
      x =abs(a(m,wk) - a(m,wk+1));
      if (x==11)
        m;
        wk;
        a (m,wk)
        a(m,wk+1)
      end
    end
  end
end
end
```

## 5. REFERENCES

- [1] Claire J. T., Ian M., Alexandre M. B., and Meeko O. 2003. Computational Techniques for the Verification of Hybrid Systems. Proceedings of the IEEE, Vol. 91, No. 7, p1-29
- [2] Eze M.O. 2014. Computational Models for Examination Fraud Control Using Randomized Algorithms. The Proc. of Multi-Conf. on Society, Cybernetics & Infor. Jul. 15 - 18, Orlando, Florida.
- [3] Luisa S., Catarina C., Carlos A., Rui C. 2011. Computational Techniques and Validation of Blood Flow Simulation. WSEAS Trans. on Bio. & Biomedicine, Issue 4, Vol. 8, p145-155
- [4] Ahmed S.M, Beck B, Maurana C.A, Newton G. 2004. Overcoming Barriers to Effective Community-Based Participatory Research in US Medical Schools. Edu. for Health Vol. 17, No. 2, p141-151
- [5] Michael I. N., Jeana H. F., Dan A. 2007. Less Is More: The Lure of Ambiguity, or Why Familiarity Breeds Contempt, J. of Personality and Social Psy. Vol. 92, No. 1, p97-105
- [6] Oduwaiye, R.O. 2014. Students Perception of Factors and Solutions to Examination Malpractices in Nigerian Universities: A Case study of the University of Ilorin. <http://www.unilorin.edu.ng>, Accessed Feb 1, 2014.
- [7] Lilian C. and Ndudzo, D. 2014. Students and Staff Perceptions on Examination Malpractice and Fraud in Higher Education in Zimbabwe. Asian J. of Humanities and Social Sc. Vol. 2, Issue 2, p78-90
- [8] Kirsten E. M. 2011. TMI (Too Much Information): The Role of Friction and Familiarity in Disclosing Information, Business and Professional Ethics. Journals, Vol. 30, No.1-2, p4-32
- [9] Nafis Alam. Challenges Faced by Sudanese Banks in Implementing Online Banking: Bankers' Perception. Journal of Internet Banking and Commerce, Aug. 2010, Vol. 15, No.2, p1-9
- [10] Pedro M.A, Antonio G. and Ruiz, M. 2013. A New Randomness Test for Bit Sequences, Informatica, Vol. 24, No. 3, 339-356.
- [11] Simmons, K., Carpenter, L., Crenshaw, S. and Hinton, V..M. 2015. Exploration of Classroom Seating Arrangement and Student Behavior in a Second Grade Classroom. Georgia Educational Researcher: Vol. 12: Issue. 1, Article 3, p50-68.
- [12] Hammang, A.J. 2012. The Effect of Seating Assignment on Student Achievement in Biology Classroom. An MSc thesis submitted to Science Education Department, Montana State Univ.
- [13] Wickham, H. 2014. Tidy Data. Journal of Statistical Software, Vol. 59, Issue 10, p1-23
- [14] Stewart, M., Brown, M.E and Weatherstone, A. 2009. Interactive Scenario Design: The Value of Flowcharts and Schemas in Developing Scenario-based Lessons for Online and Flexible Learning Contexts, J. of Dist. Learning, Vol.13, No. 1, p71-90.
- [15] Mala.V.P. and Nageswara, A.M.Y. 2011. Importance of Data Collection for Systematic Software Development Process. Internat. Journal of Comp. Sc and IT. Vol 3, No 2, p260-278.
- [16] Jamil S. 2009. The Growing Accountability Agenda in Tertiary Education: Progress or Mixed Blessing?. World Bank Education Working Paper Series, No. 16, Washington, DC, p1-30.