Detection of Suspicious Patterns in Online MCQ Exams

Anindya Sundar Dey

Abstract*: The purpose of this paper is to find ways to identify or detect suspicious patterns in multiple choice questions and highlight them allowing for greater scrutiny and helping curb malpractices in examination halls. For this purpose an algorithm has been developed to detect suspicious answer patterns in online MCQ exams which can detect if multiple students have been taking outside help using applications like Team Viewer. To obtain data for this project we obtained results of an MCQ test where two groups of students in who completed a MCQ test of moderate difficulty. While both groups were kept under scrutiny. Unlike group B group A had a select number of "special examinees" who were being helped by outside sources(teachers). We later ran the test results of both the groups under the algorithm. The Algorithm was able to detect 17 Students with Suspicious Patterns in Group A. While no student was detected in Group B.

Index Terms*: Suspicion Threshold, Online Mcq, Suspicion Factor, Answer Key, Array of Converted Answers, Comparison Point.

I. INTRODUCTION

Malpractices in exams is nothing new. It is an evil that persists even with the strictest of interventions.

As the methods to prevent malpractices has evolved so have the methods to commit malpractices.

Around 75% of college under graduates in America admit that they have cheated on tests. It was found that obtaining information from another student about a test was the most common form of cheating . Many students accept cheating without concern. Forty percent of the college students surveyed did not disapprove of cheating on tests, 29% did not feel guilty about cheating, and only 1% said that they would report cheating if they observed it. (Baird) [21].

It was also found that the cheating was not uniform certain groups are more likely to cheat: men, students with low grades, underclassmen, business majors and students involved in few extracurricular activities (Baird) [21].

II. MALPRACTICE BY PLAGIARISM ON ONLINE MCQ EXAMS

Our paper focuses specifically on one subset of them. Multiple Choice Question (MCQ) Exams. To be even more specific, Online MCQ Exam

One of the biggest examples would be the Staff Selection Committee Exams that are held in India.

It is easy to understand why malpractices during government exams are happening. These are esteemed government job posts which only a select few can get into after a rigorous series of examinations.

However what most people don't know is HOW they are cheating. From corrupt invigilators to gangs who are dedicated to giving their best for their customers. The depth of this rabbit hole is unfathomable.

Using a software called Team Viewer they can access the examinee's computer screen giving answers in behalf of the examinees. Each gang containing 4-5 people could answer the questions of up to 200 candidates. However India is not the only country to experience such examination scandals.

A very famous US education scandal known as the Atlanta Cheating Scandal was one that shook the nation to its roots where the teachers themselves helped the students cheat or changed their answers post exams to correct ones. You may notice a lot of similarities between the SSC Scandals and the Atlanta Cheating Scandal. However, you may notice that Erasure Analysis was done in the examinations involved in the Atlanta Cheating Scandal and SSC Exams.

Unfortunately due to an incredible number of them being online, Erasure Analysis is no longer possible in multiple cases.

This brings us to our Question:

Is it possible to detect malpractice in online MCQ exams using Algorithms ??

This is what we have done for our paper. We have developed a methodology or an algorithm that should be able to detect certain patterns in Exam answer sheets of different students as long as the answer keys and the student's answers are entered correctly.

With the ability to detect suspicious patterns in this algorithm it could virtually eliminate all external malpractices occurring during Online MCQ exams.

III. METHODS OF PREVENTION OR DETECTION OF MALPRACTICE IN MCQ EXAMS

The earliest attempts to prevent these were of inspection of observed distributions of the number of identical wrong for pair of examinees (Bird)[20].

Many derived ratios for identical errors between two examinees. One used binomial standard deviation (Anikeef)[17], while others used percentages, like the Identical Error Percentage (Dickenson)[18].

Using similarity of errors has been used by multiple people time and time again (Belleza)[1].

However, many have also tended to use a more statistical approach using statistical models and indices. Including the Nominal Response Model (Wollack)[10][12][15].

Revised Manuscript Received on May 09, 2019.

Anindya Sundar Dey, School of Information Technology and Engineering, Vellore Institute of Technology, Vellore (Tamil Nadu), India.

Sotaridona et.al used statistical indices to detect answer copying by using Poisson distribution of matching incorrect answers and incorporate matching correct answers to the matching incorrect answers (Sotaridona)[8].
In terms of Prevention, having Graded Multiple Questions and rotating the question numbers (Denyer G)[3]. Rotating the answer responses alongside rotating the question numbers (Houston) are among the few that have been suggested.
However, most of these detection or prevention methods were developed before the advent of Online MCQs and the ability to manipulate answers of examinees with the help of external sources.
The algorithms work on the premise that two students near each other maybe possible cheaters. These algorithms also cannot determine who was cheating and who was it being cheated from.
Due to the very nature of external help being provided to the examinees, there is no need to determine a cheater or a cheated in this situation.
The only research remotely similar to these is that of Levitt S.D who developed his own algorithm for detecting teachers cheating with student papers (Levitt)[9].

Multiple students being helped by a small group of people in a short amount of time. This gives rise to something unseen during the times these previous detection methods were created, patterns.

IV. ANALYSIS OF PATTERNS IN MCQ EXAMS

The way the proposed algorithm works is very simple. The answer key alongside the answers and IDs of the students giving the exam are given to the program(read from a file).
NOTE: The omitted answers are written as 0.
STEP 1: The File Name containing the student's answers and the Answer Key itself is given as input.
STEP 2: It is then checked whether the File exists. If it does not exist go to STEP 20.
STEP 3: The File is then Read and the ANSWER KEY is stored as the first element of the array of a designated Class Object with an ID called ANSWER KEY.
STEP 4: The algorithm breaks the ANSWER KEY into groups of 4 and 5 and stores them.
STEP 5: The File is Read Again and the ID and Answer Key of the Student are Fetched and Stored
STEP 6: The answers of the student are then broken into two groups of 4 and 5. The broken up answers are stored in designated arrays present in the class as data members.
STEP 7: The stored grouped answers are compared with the grouped answers of the ANSWER KEY. If an Answer does not Match go to STEP 9.
STEP 8: Matched answer is converted to 0 and stored.
STEP 9: Looks for Next Group of Answer in the object of the Student. If present Go to STEP 7.
STEP 10: Check for more students. If available. Go to STEP 5.
STEP 11: Select Student's Array of Converted Answers (ACA) as the comparison point (CP.)
STEP 12: The CP is then compared with another ACA omitting those elements whose value is 0 (all correct answers). If it does not match go to STEP 14.
STEP 13: The suspicion factor of the CP student increases by 1. Each student has two suspicion factors one for groups of 4 and one for groups of 5.
STEP 14: Find another ACA for comparison with the CP. If present go to STEP 12.
STEP 15: Check student for CP. If present go to STEP 11.
STEP 16: Compare Suspicion Factor of Student with Suspicion Threshold. If does not exceed go to STEP 18.
STEP 17: Display Student ID in High Suspicion List.
STEP 18: Check whether more students available. If Yes. Go To STEP 16.
STEP 19: Close FILE.
STEP 20: END
When either exceeds the suspicion threshold. The ID(name) of the student is displayed under the suspicion list.

The suspicion threshold is determined by the number of students attending the exam say n. The nearest number to n which is divisible by 5 becomes the suspicion threshold of grouped by 4, while that number when multiplied by 0.8 becomes the suspicion threshold of grouped by 5.

V. RESULTS AND DISCUSSION

A. Can the algorithm detect ?

In short, the answer is yes. We use the same samples used by Levitt S.D et.al in
PREVALENCE AND PREDICTORS (Levitt)[9]
There are two classes of the same grade attempting the same paper. Let us say, class A and class B. Class A has a strength of 22 students while Class B has a strength of 18 students. Accordingly, the suspicion threshold for both of the classes would be 20 for grouped by 4 and 16 for grouped by 5.

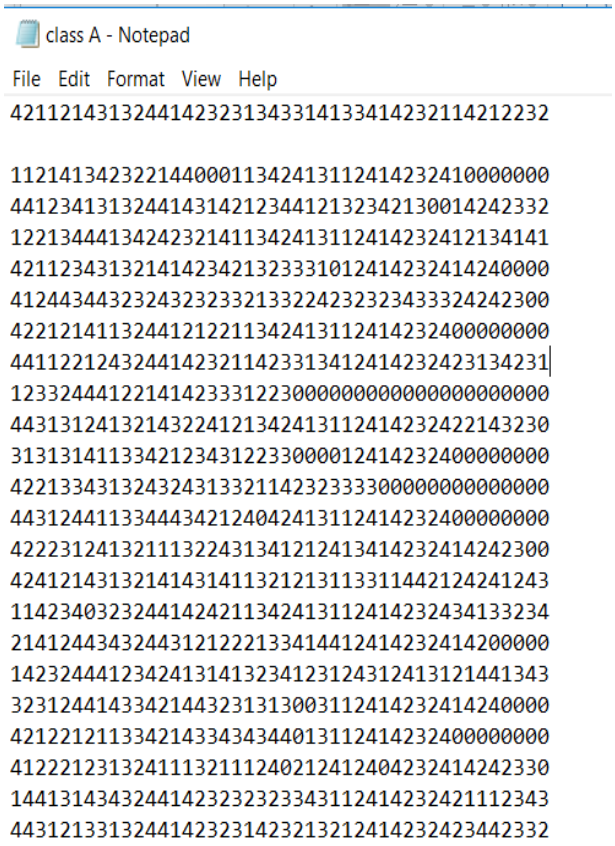

Fig.-1 Class A (the topmost part is the answer key)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

class B - Notepad

File Edit Format View Help

```
42112143132441423231343314133414232114212232

42314314222413124413344221113413323424312434
41111111132243423131223232423311422314224331
44211242142143222131231213411400000000000000
32213434433242121123343000000000000000000000
34112141232443443131112314233400000000000000
43313431122242444221342314232223240000000000
23113242112424311413423134212423333413232000
43223444112442322421223211113412212113310000
43112233432441422333223212222232322443224230
34131143133314422143421342332342242234331334
41313141222141221124243113112100000000000000
41213141113423433221213314124421312440000000
31411333432414214131424314343122413223214240
44213143132441423421323214123413412113224321
42113443232241422434123114123313412120000000
41341431132441433432134124122434113211344212
12141121112241422431223114123432413200000000
42113342122441423231113212133411222120000000
```

Fig.-2 Class B(the topmost part is the answer key)

After running the algorithm on the inputs, we get the outputs:

```
Enter Student Name: S

Enter answer of student:41222123132411132111240212412404232414242330

Enter Student Name: T

Enter answer of student:14413143432441423232323343112414232421112343

Enter Student Name: U

Enter answer of student:21412443432443121222133414412414232414200000

Enter Student Name: V

Enter answer of student:44312133132441423231423213212414232423442332

Suspicious Patterns Detected in Papers of Students:

A
C
D
F
G
I
J
L
M
N
O
Q
R
S
T
U
V
```

Fig.-3 Class A list of suspicious students.

```
Enter answer of student:41313141222141221124243113112100000000000000

Enter Student Name: L

Enter answer of student:41213141113423433221213314124421312440000000

Enter Student Name: M

Enter answer of student:31411333432414214131424314343122413223214240

Enter Student Name: N

Enter answer of student:44213143132441423421323214123413412113224321

Enter Student Name: O

Enter answer of student:42113443232241422434123114123313412120000000

Enter Student Name: P

Enter answer of student:41341431132441433432134124122434113211344212

Enter Student Name: Q

Enter answer of student:12141121111224142243122311412343241320000000

Enter Student Name: R

Enter answer of student:42113342122441423231113212133411222120000000

Suspicious Patterns Detected in Papers of Students:
```

Fig.-4 Class B list of suspicious students

We can clearly see that there are 17 students who exceed the suspicion factor in class A, while class B has no students who have exceeded the suspicion factor.

B. Comparison among existing algorithms

Let us compare, what Levitt S.D et.al found with his algorithm using the same inputs.

Student Answer Strings (each row represents one student's answers)	Student Test Scores: Year t-1	Year t	Year t+1
Suspected Cheating Classroom			
112A4A342CB214D0003ACD24A3A12DADBCB4A0000000	1.9	5.3	4.4
1B2A34DA4C42D23B141ACD24A3A12DADBCBA2134141	4.3	5.6	4.3
DB2ABAD1ACBDDA212B1ACD24A3A12DADBCB4000000000	3.0	6.5	5.1
1142340C2CBDDADB4B1ACD24A3A12DADBCB43D133BC4	3.6	6.3	4.9
D43A3A24ACB1D32B412ACD24A3A12DADBCB422143BC0	5.2	5.9	4.9
D43AB4D1ACDD43422240D24A3A12DADBCB400000000	4.8	5.3	3.6
DBA2BA21AC2D2AD3C4C4CD4DA3A12DADBCB400000000	1.9	6.1	3.6
DBAA4ADC4CBD24DBCB2A1110A3A12DADBCB400000000	3.3	6.3	6.2
144A3ADC4CBDDADBCBC2C2CC43A12DADBCB4211AB343	3.0	6.8	4.9
D43ABA3CACBDDADBCBCA42C2A32D12DADBCB42344B3CB	4.8	7.1	6.6
214AB4DC4CBDD31B1B2213C4AD412DADBCB4ADB00000	3.6	6.1	4.3
313AJAD1AC3D2A23431223C0000D12DADBCB400000000	3.8	4.7	5.1
D4AAB2124CBDDADBCB1A42CCA3412DADBCB423134BC1	5.5	6.6	7.7
3B3AB4D14C3D2AD4CBCAC1C003A12DADBCB4ADB40000	3.0	6.5	6.6
DBAAB3DCACB1DADBC42AC2CC3104D12DADBCB4ADB40000	3.8	7.1	5.6
DB223A24ACB11A3B24CACD12A241CDADBCB4ADB4B300	4.9	6.5	5.8
D122BA2CACBD1A13211A2D02A2412D0DBCB4ADB4B3C0	3.6	6.1	6.2
1423B4D4A23D24131413234123A243A2413A21441343	4.9	2.5	5.6
DB4ABADCACB1DAD3141AC212A3A1C3A144BA2DB41B43	5.9	6.5	7.7
DB2A33DCACBD32D313C21142323CC3100000000000000	3.8	4.4	5.6
1B33B4D4A2B1DADBC3CA22C0000000000000000000000	5.0	4.4	7.2
D12443D43232D32323C213C22D2C23234C332DB4B300	3.3	3.8	3.6
D4A2341CACBDDAD3142A2344A2AC23421C00ADB4B3CB	6.4	5.9	6.2
	4.1	5.8	5.5
	Average Test Scores		
Typical Classroom			
34ABAD12CBDD3D4C1CA112CAD2CCD000000000000	3.8	5.6	6.4
D33A3431A2B2D2D44B2ACD2CAD2C2223B400000000	4.6	4.9	5.8
DB3A431422BD131B4413CD4221A1CDA332342D3AB4C4	4.0	5.1	5.1
D1AA1A11ACB2D3DBC1CA22C23242C3A142B3ADB243C1	4.6	5.9	5.3
D42A12D2A4B1D32B21CA2312A3411D000000000000	4.5	3.8	6.4
3B2A34344C32D21B1123CDC0000000000000000000	3.3	2.8	5.1
23AA32D2A1BD2431341342C13D212D233C34A3B3B000	3.3	4.4	4.9
D32234D4A1BDD23B242A22C2A1A1CDA2B1BAA33A0000	5.1	5.6	5.9
D3AAB23C4CBDDADB23C322C2A22222232B443B24BC3	4.7	5.6	7.0
D13A14313C31D42B14C421C42332CD2242B3433A3343	2.2	3.8	4.9
D13A3AD122B1DA2B11242DC1A3A12100000000000000	4.5	4.1	5.9
D12A3AD1A13D23D3CB2A21CCADA24D2131B4400000000	3.6	5.3	5.9
314A133C4CBD142141CA424CAD34C12413223BA4B40	3.3	4.7	4.4
D42A3ADCACBDDADBC42AC2C2ADA2CDA343BAA3B24321	5.6	6.9	8.5
DBAA34DC2CB2DADB24C412C1ADA2C3A341BA20000000	5.0	5.9	7.0
D1341431ACBDDAD3C4C213412DA22D3D1132A1344B1B	3.8	5.3	5.3
1BA41A21A1B2DADB24CA22C1ADA2CD12413200000000	4.3	5.3	6.8
DBAA33D2A2BDDADBCBCA11C2A2ACCDA1B2BA20000000	4.5	6.8	7.9
	4.2	5.1	6.0

FIGURE I
Sample Answer Strings and Test Scores from Two Classrooms

Fig.-5 Class A and Class B by Levitt S.D et.al

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

NOTE: Levitt S.D. et.al wrote all correct answers as A, B, C and D while writing the incorrect ones as 1, 2, 3 and 4. Omitted ones are written as 0 (Levitt)[9].

As observed, Levitt's algorithm was able to find 15 students who had similar patterns in class A while not a single student in class B was found to have this kind of pattern.

The reason being the teacher themselves altered answers of many students to make them get higher marks (Levitt)[9].

However, while Levitt's algorithm was able to find only 15. If one closely scrutinizes the results one can easily see that the two answers directly below also share the pattern just not exactly.

Instead of 12DADBCB4, the one directly below has 1CDADBCB4. While the one below that shares 12D0DBCB4. They both share the pattern of DBCB4 alongside the marked ones however they were not included in the suspicious patterns.

The proposed algorithm was able to detect and display the other two as well. This can be considered as an improvement over the algorithm used by Levitt et.al in their work Rotten Apples: An Investigation of the Prevalence and Predictors of Cheating (Levitt)[9].

C. Limitations of proposed algorithm

The proposed algorithm is solely created for situations when multiple students are helped by a same source during an MCQ test, preferably online. That is this algorithm is only useful for out of class plagiarism and helps to detect whether the invigilators themselves are involved or not.

VI. CONCLUSION

The proposed algorithm is to detect when external plagiarism occurs during any online MCQ exam scenarios. The data obtained from Atlanta Public School Scandal has been used as test material for the algorithm even though the tests were not online. The external help of invigilators is similar to the requirements of our proposed algorithm.

Hence, it was able to detect seventeen students from a class with a corrupt teacher while no students were detected in the class without a corrupt teacher.

REFERENCES

1. Bellezza, F.S. and Bellezza, S.F. (1989) Detection of Cheating on Multiple-Choice Tests by Using Error-Similarity Analysis. *Teaching of Psychology*, **16**(3): 151–155.
2. Bresnock, A.E. (1989) Multiple-Choice Testing: Question and Response Position. *Journal of Economic Education*, **20**(3), 239–245.
3. Denyer, G. and Hancock, D. (2002) Graded Multiple Choice Questions: Rewarding Understanding and Preventing Plagiarism. *Journal of Chemical Education*, **79**(8), 961–964.
4. Frary, R.B. (1993) Statistical Detection of Multiple-Choice Answer Copying: Review and Commentary. *Applied Measurement in Education*, **6**(2), 153–165.
5. Houston, J.P. (1983) Alternate Test Forms as a Means of Reducing Multiple-Choice Answer Copying in the Classroom. *Journal of Educational Psychology*, **75**(4), 572–575.
6. Houston, J.P. (1986) Classroom Answer Copying: Roles of Acquaintanceship and Free versus Assigned Seating. *Journal of Educational Psychology*, **78**(3), 230–232.
7. Moeck, P.G. (2002) Academic Dishonesty: Cheating among Community College Students. *Community College Journal of Research and Practice*, **26**(6), 479–491.
8. Sotaridona, L.S. and Meijer, R.R. (2001) *Two New Statistics To Detect Answer Copying*. Research Report, Twente University, Enschede (Netherlands). Faculty of Educational Science and Technology. [BBB31588].
9. Levitt S.D. & Jacob B.A. (2003) Rotten Apples: An Investigation of the Prevalence and Predictors of Cheating, *The Quarterly Journal of Economics*, **118**(3), pg 843-877
10. Wollack J.A. (1997) A Nominal Response Model Approach for Detecting Answer Copying, *Applied Psychological Movement*, **21**(4),pg 307-320
11. Van der Linden, W. J., & Sotaridona, L. (2006). Detecting Answer Copying When the Regular Response Process Follows a Known Response Model. *Journal of Educational and Behavioral Statistics*, **31**(3), 283–304.
12. Wollack, J. A. (2003). Comparison of Answer Copying Indices with Real Data. *Journal of Educational Measurement*, **40**(3), 189–205.
13. Angoff, W. H. (1974). The Development of Statistical Indices for Detecting Cheaters. *Journal of the American Statistical Association*, **69**(345), 44–49.
14. Saupe, J. L. (1960). An Empirical Model for the Corroboration of Suspected Cheating on Multiple-Choice Tests. *Educational and Psychological Measurement*, **20**(3), 475–489.
15. Wollack, J. A. & Cohen, A. ***S.*** (1998). Detection of answer copying with unknown item and trait parameters. *Applied Psychological Measurement*, **22,** 144-152.
16. Scheers, N. J., & Dayton, C. M. (1987). Improved estimation of academic cheating behavior using the randomized response technique. *Research in Higher Education*, **26,** 61-69.
17. Anikeef, A. M. **(**1954**).** Index of collaboration for test administrators. *Journal of Applied*
18. *Psychology*, **38,** 174-177**.**
19. Dickenson, H. F. (1945). Identical Errors and Deception. *The Journal of Educational Research*, **38**(7), 534–542.
20. Cody, R. P. **(**1985**).** Statistical analysis of examinations to detect cheating. *Journal of Medical*
21. *Education*, **60**, 136-137**.**
22. Bird, C. **(**1929**).** An improved method of detecting cheating in objective examinations. *Journal*
23. *of Educational Research*, **19,** 341-348**.**
24. Baird, J. S., Jr. (1980). Current trends in college cheating. *Psychology in the Schools*, **17**, 515-522

AUTHORS PROFILE

Anindya Sundar Dey is pursuing his Master of Computer Applications from Vellore Institute of Technology Vellore. He graduated his Bachelors degree in stream of Computer Science from Calcutta University and has great inclination towards Data Sciences and Manipulation.