

Development of System for Automated & Secure Generation of Content (ASCGS)

संदीप सिंह यादव (Sandeep Singh Yadav)

राष्ट्रीय विद्यालय तंत्रिका महाविद्यालय, बेंगलूरु, ५६००१०, भारत
(Rāshtrīya Vidyālaya Tāntrika Mahāvīdyālaya/ISE, Bengaluru, 560010, Bharat)
E-mail: sandeepsy90@gmail.com

मंदीप सिंह यादव (Mandeep Singh Yadav)

एच एम आर तंत्रिका एवं प्रबंध संस्थान, नई दिल्ली, ११००३६, भारत
(HMR Tāntrika & Prabandha Sansthan/CSE, Nayi Dilli, 110036, Bharat)
E-mail: mandeepsy95@gmail.com

Abstract—Automation of manual work and systems is a vast growing trend as it brings in efficiency and quality in work. Online and offline aptitude exams are conducted by organizations and institutions for accessing the skills of students. An aptitude test is an effective method for testing the students however the creation of database for conducting aptitude tests is an uphill task, because the organization has to maintain to a very large database consisting of thousands of questions and the content has to be changed periodically to avoid repeating of questions. An Automated & Secure Content Generation System (ASCGS) is a framework that provides automated creation of aptitude questions by converting a single of type of question into multiple questions by altering its variables. Also the system automates the process of calculating the answer for every question by using its formula. Security in the system is provided by the means of encryption. The present system lags in many perspectives like every question in has to be created manually and also the answer for the question has to be computed manually. Since the entire work is being done manually so there is a high risk that some questions may contain error due to human fault and also the cost and effort required to create the content is large. The proposed system overcomes these shortcomings of the existing system as only one format is required to be created for one type of question thus saving time and human effort. Also it is no longer required to do mathematical calculation manually as it is done by the system, the user has to insert only the formula. The present system requires the organization very long time ranging from few weeks to few months for generating ten thousand questions. With the new system the same work can be done within few days time and with minimal cost. With the development of the system organizations will be relieved from the tedious work of content creation and also management of content will become easier and efficient. The software will enable the institutions to create aptitude questions of different levels thus enabling the institute to conduct aptitude tests for all the students

of different classes based on their levels.

Index Terms—Automation, Secure, Randomization, Aptitude Test, Manual.

I. INTRODUCTION

Computer technology has brought a revolutionary change in the world. In every field of life we are moving from manual to automated systems. With response to the growing number of students entering universities and many students coming for recruitment, many universities and organizations prefer to conduct online or computer based examinations which are suitable for mass evaluation to measure student's aptitude skill. The computer based method [1] [2] is a quick and efficient method to test the student's aptitude skills and measure their performance. This method could evaluate student's achievement, and helps to give fair scores. The main objective of Content Generation System is to efficiently create the content for the aptitude test thoroughly through a fully automated system that not only saves lots of time but also is very efficient. Online aptitude [3] test is basically made for reducing paper work of any traditional aptitude test. Also it is available all the time, anywhere so the students need not go to examination center. Students can easily get registered and take the test. After the test is over, their result will be displayed within seconds, this process reduces time of submitting and checking the test papers. ASCGS is a java application that generates the content for the aptitude test to take place. Since the data for such systems is very large consisting of thousands and lakhs of questions. Creating such big data base of aptitude questions manually is a very tiresome, here ASCGS comes to use for creating the content easily and efficiently. Rather than constructing different questions of same format multiple times, only single format is constructed and multiple questions are derived out of it by altering the value.

II. RELATED WORK

Quite a lot researches have been performed on implementing automated content generation in various fields. Some of the developments with respect to automated content generation are listed below. A random test paper generation algorithm has been designed and implemented for examination system. The paper [4] discusses how to use optimized genetic algorithm to generate test papers from the question bank. Also fishnet algorithm has been introduced to generate test papers automatically for getting a better, more fair and more objective test paper. An ant colony algorithm [5] of automatic sets method has been implemented for generating automated content. First, a mathematical model of constraint is built according to the requirements of papers, and by using the ant colony algorithm, the optimal solution of grouping is obtained. Simulation results show that, compared with the traditional sets algorithm, the ant colony algorithm improves the efficiency of grouping, increase rate of grouping, and satisfies the current network on-line examination system of real-time higher requirements of grouping. Research has been done for dealing with automatic classification of Tamil documents and also the information extraction from those documents to construct the knowledge base. The system [6] makes use of Naive Bayes (NB) [7] [8] algorithm for classifying Tamil documents to one of the predefined categories. Integration of an algorithm [9] and multi-agent system for automatic generation of test paper with higher efficiency has been proposed. Algorithm first generates a quantization model of test paper then the test items are selected from the database according to the quantization model.

III. EXISTING SYSTEM

Aptitude test is one of the competitive exams conducted to test a student's skill. Fig. 3.1 shows the existing system for creation of content.

The steps for the creation of content in existing system are described as follows:

Manual Creation of Question: The first step includes the formation of questions. Here only the language of question is created based on its type. The question can be of any type.

Assume Values of Variables: Here the values for all the variables are assumed. The values are not repeated while assuming.

Manual Computation of Answer: The values are then put in the formula, and the answer is computed. After computation, based on the results there are two possible outcomes

- Case I Answer value is as per requirement: In this case the question is considered valid and it is stored in the server
- Case II Answer value is not as per requirement: In this case, backtracking is done, and again the

admin needs to assume different values for variables and keep doing it until answer value come as per requirement.

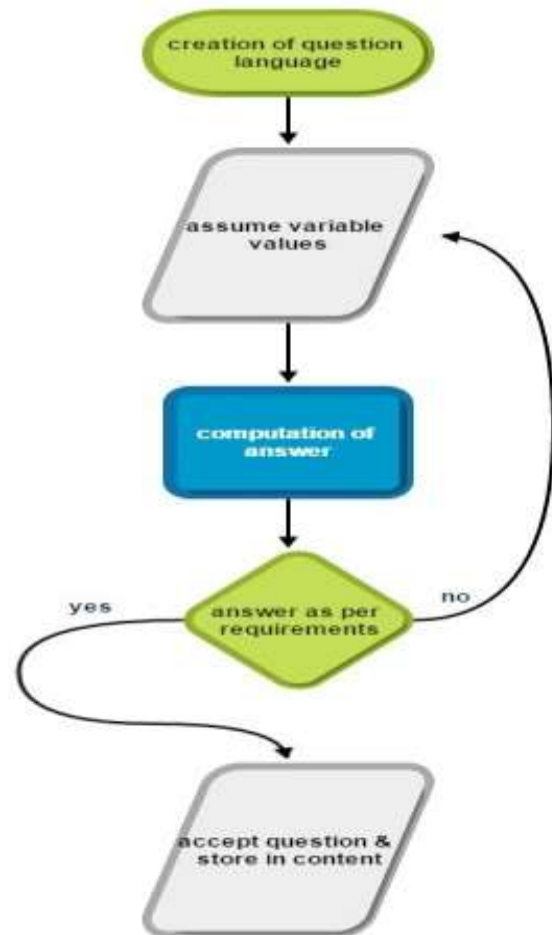


Fig.3.1. Existing Content Generator System

A. Review of Experimental Studies on Existing System

On reviewing the existing system it is found that the present system faces two major issues, the first is that it is very time consuming and the second is it is unreliable. The system is called time consuming as it is very hard for humans to find out the combinations of variables suiting the required output answer and also it is very tiring to do repeated calculations involving typical mathematical calculations constantly. Secondly the system is called unreliable as humans are prone to create error and since there is a very large content that needs to be created so it becomes highly probable that some questions will carry incorrectness. In short the entire process in the system is being done manually and it requires more time and manpower for developing the content and the main risk in the present system is the probability of having incorrect questions.

IV. PROPOSED MODEL

Fig. 4.1 shows the system architecture of the system. The main components are;

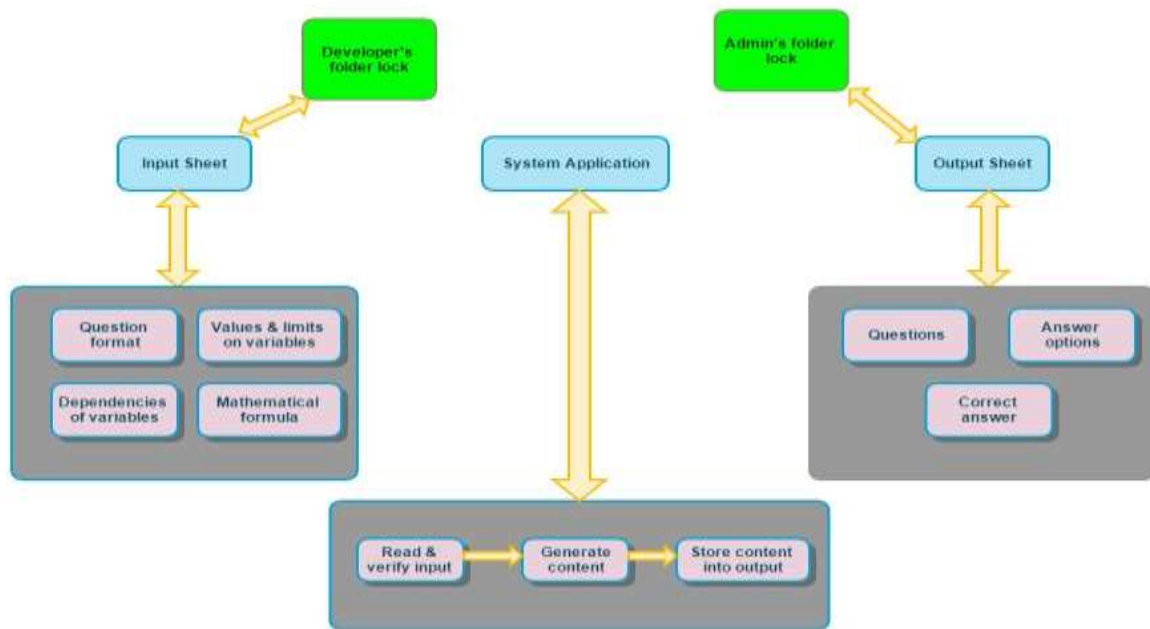


Fig.4.1. System Architecture

A. Input Sheet

The input sheet is managed by the developer. Developer is responsible for the development, modification, updating or any other operation involved on the sheet. The sheet contains question format, the formula for computing answer and the limits present on the input and output variables. The sheet is protected by folder lock and hence the developer needs to provide authentication before being able to modify it.

B. System Application

This component forms the core area of the system as it is responsible for the generation of the content. The application is developed in Java as it is a portable and very efficient language. The interaction between application and user is made through GUI (Graphical User Interface).

C. Output Sheet

The resultant data of the system is stored in output excel sheet. The output sheet is generated by the system after processing of the input sheet. The system is responsible for the generation of the automated content from the input sheet. The system firstly process the input sheet and verifies the content type, if it is matching according to the format specified then the system generate the content as pe the demand of developer and stores the same into the output sheet which is retrieved by the developer. The admin and developer share a key between them. This key is used by developer for encryption of output sheet before being sent to the admin. The admin uses the key for decrypting the output sheet and then stores it on the server where it is locked and not visible to other users.

V. ASCGS

In this section we describe the system in detail.

A. Structure Chart

The system represents a tree where the leaves are the subtasks and the subtasks are very small and they can be coded with a very few programming statements. Fig. 5.1 shows the structure chart of the system. It depicts the structure and calling relationships in the main module.

The Structure is defined as follows:

Developer creates the data of question formats into an excel sheet. Then developer logs into the system to generate the output sheet. After retrieving the output sheet the developer logs into folder lock software to secure the generated file. In the next step, the file is encrypted with shared key and forwarded to admin. Now the admin retrieves the file and decrypts it with shared key and then verifies it for correct content. For securing the files system admin makes use of folder lock, firstly the files are locked and then uploaded over the server.

B. User Interface

The purpose of this component is to act as an interface for the user to build interaction with the system. The user gives commands to the system using this module. The tool being developed should devise a way for a user to interact with the system in order to work on it. This interface is needed to provide security as well as robust usage to the user. The interface brings GUI into the system hence the usage becomes very easy.

Fig. 5.2 shows the flow chart for the User Interface Portal. The sequence of execution steps is as follows.

The sequence of execution steps is as follows.

Developer first logs in the details. Then developer specifies the path of input and output files, unauthorized developers are denied for system access for wrong login

details. Before retrieving the output file, developer enters the number of questions required.

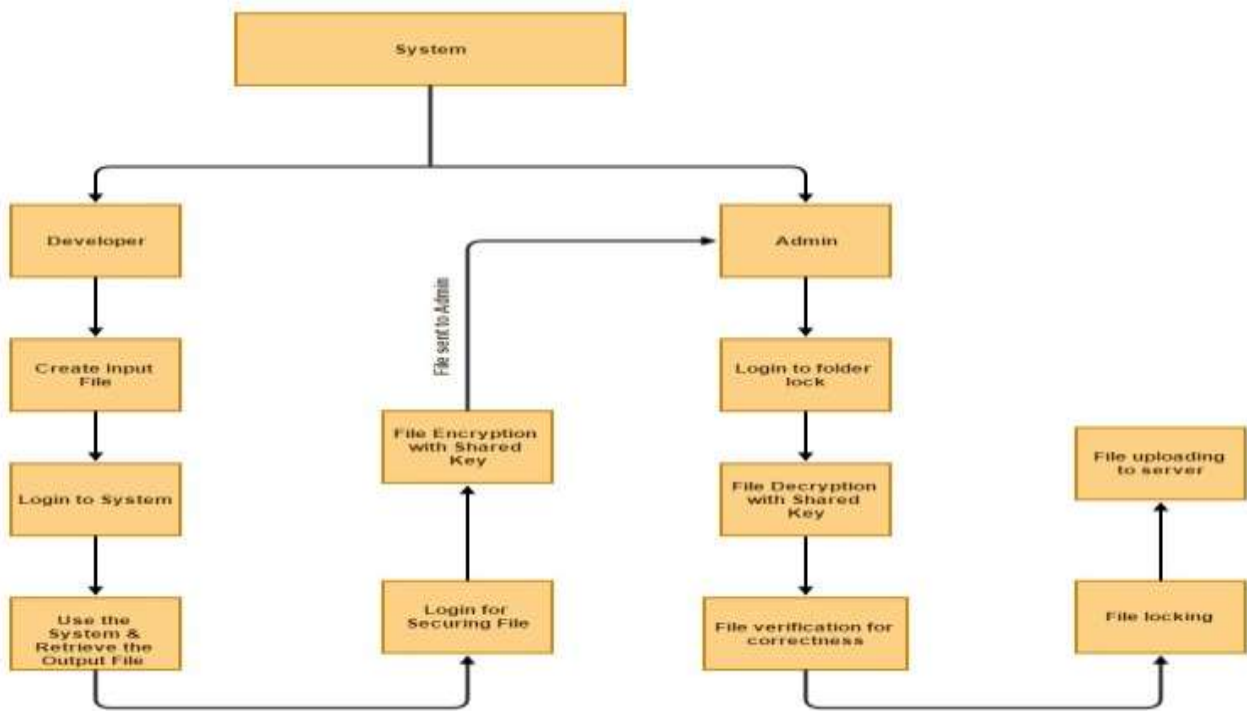


Fig.5.1. Structure Chart of the System

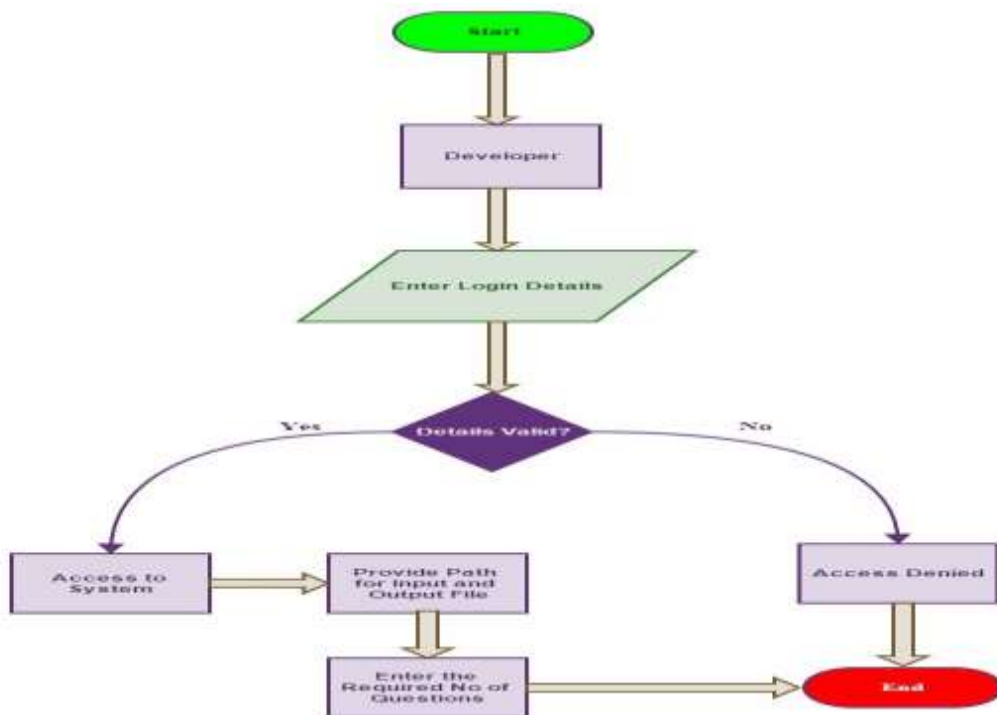


Fig.5.2. Flow chart of user interface

C. Input Sheet

The purpose of the input excel datasheet is to act as the data source point of the input content. It is used by the system for developing the content. It acts as the store

point for the input which the system utilizes for generating the required content. The input excel sheet is developed by the developer in JSON format.

Fig. 5.3 shows the content stored in input sheet:



Fig.5.3. Content Stored in Input Excel Datasheet

Question Format: This column contains the language of the aptitude question. The variables present in the question format are replaced by random values and remaining question remains the same in the question format. These variables are altered to have multiple values for generating multiple questions from single format of question.

Maximum Variation: This column contains the limit up to which the incorrect answers can vary from the correct answer.

Minimum Answer: The column is used to specify the minimum value that all the four answers should have.

Maximum Answer: The column is used to specify the maximum value that all the four answers can have.

Output Type: The column specifies the type of answer which is allowed for the question. It can have the value like int or float.

Formula: This column contains the formula that the system uses to calculate the answer for the question.

Variables: This column contains the variables. The variables also have the constraints. The constraints that are specified on variables are:

Type: Here the type of value is specified that the input variables can have like int, float or string.

Minimum: This is the minimum limit for the input variable

Maximum: It is the maximum limit for the input variable

Dependency: The variable is required when one variable is dependent on some other variable, for example it may be required that variable v3 should always be greater than variable v1 and less than variable. When variables are interdependent for their values, dependency columns comes into use.

D. Output Excel Datasheets

The purpose of the output excel sheet is to act as the source point of the output/content. This is the file that contains the resultant aptitude questions. It acts as the data store point for the output data once the system generates the required content. The output excel sheet is

generated by the system on processing the input sheet.

Fig. 5.4 shows the content present in the output sheet. Now we describe the contents here:

Question: This column contains the question. Now the variables have been replaced with numerical values.

Four Answer Options: Next four answer options are present out of which one is correct.

Correct Answer Option: Here the correct answer option is provided. The correct option is always random, it always random, it can be anyone out of the four options.

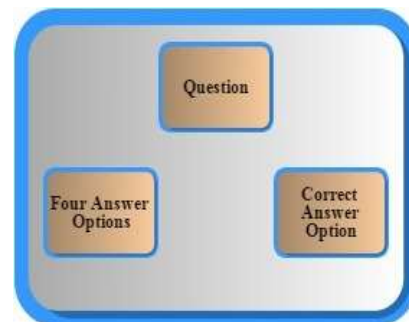


Fig.5.4. Content Getting Stored in Output Excel Datasheet

Fig. 5.5 shows the flow chart for content generator. The system reads the question format, identifies the variables present in it and allocates memory space for them in the system. Then the system reads in the formula to compute the answer. Next the system reads the limits present on the input variables Random values within given limits are inserted into the formula and answer is computed. If the answer lies within the max and min answer limit, system moves forward and read the variation limits present for the incorrect answers else if the answer does not lie within the limit system move back to step number 4. Now 3 incorrect answers within the variation limits are found. Comparison of incorrect answers is done with the Max and Min answers, answers satisfy the limits system stores the data into excel sheet otherwise it moves one step backwards at step no 6.

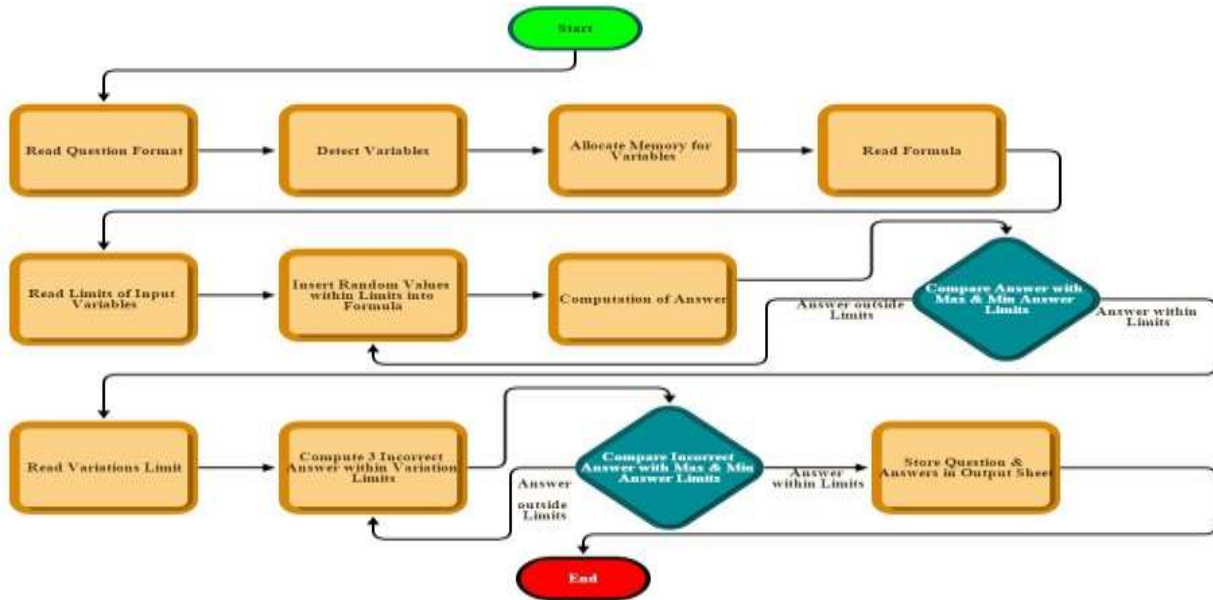


Fig.5.5. Flow Chart of Content Generator

VI. BENEFITS

Time saving: As the system is involved in creating multiple questions out of single format and also system is involved in doing in the mathematical calculations, therefore a lot of time is saved.

Probability of error is reduced: If we humans are involved in doing the mathematical calculations, then certainly the probability of an error is higher as compared to when the computer is performing the same question.

Cost saving: Organizations have to hire large number of well qualified people to create dataset of aptitude questions, but after using this system a lot of cost savings can be done in this area.

Ease of modification: If there is a modification required in a certain type of question, then it can be easily performed as only one question format needs to be altered rather than altering hundreds of questions of that type.

Less human effort: Effort required by humans is reduced to a big extent as we only create one single format for that type of question and also we are no longer required to do the mathematical questions to get the correct answer of the questions.

VII. CONCLUSION

Automated Content Generation is a technique of improving the quality along with quantity and efficiency of content by making use of software for the creation of content. The system is of very useful as manual management of content is a very tough task. With this system the entire concept of content creation is brought in an automated and efficient method. As this application involves huge amounts of quality data generated by the system in an efficient and easy manner, it is an ideal candidate for illustrating automated content generation. Automated & Secure Content Generation System fulfils the main objective of creation of a platform to create content in an automated way. The system ensures the security of the content as the content is stored in an encrypted form. Also the system is beneficial for the students taking the aptitude test as questions with errors are removed, this is because system does calculation of answers in an automated way. Another benefit the students get is that questions can be set as per the standard of particular class, it is done by keeping limits on the type and range of answer. Since the system is very much beneficial in terms of cost and time saving and is working efficiently so the existing manual system for content generation can be replaced by the new system.

APPENDIX A SCREENSHOTS OF INPUT & OUTPUT SHEETS

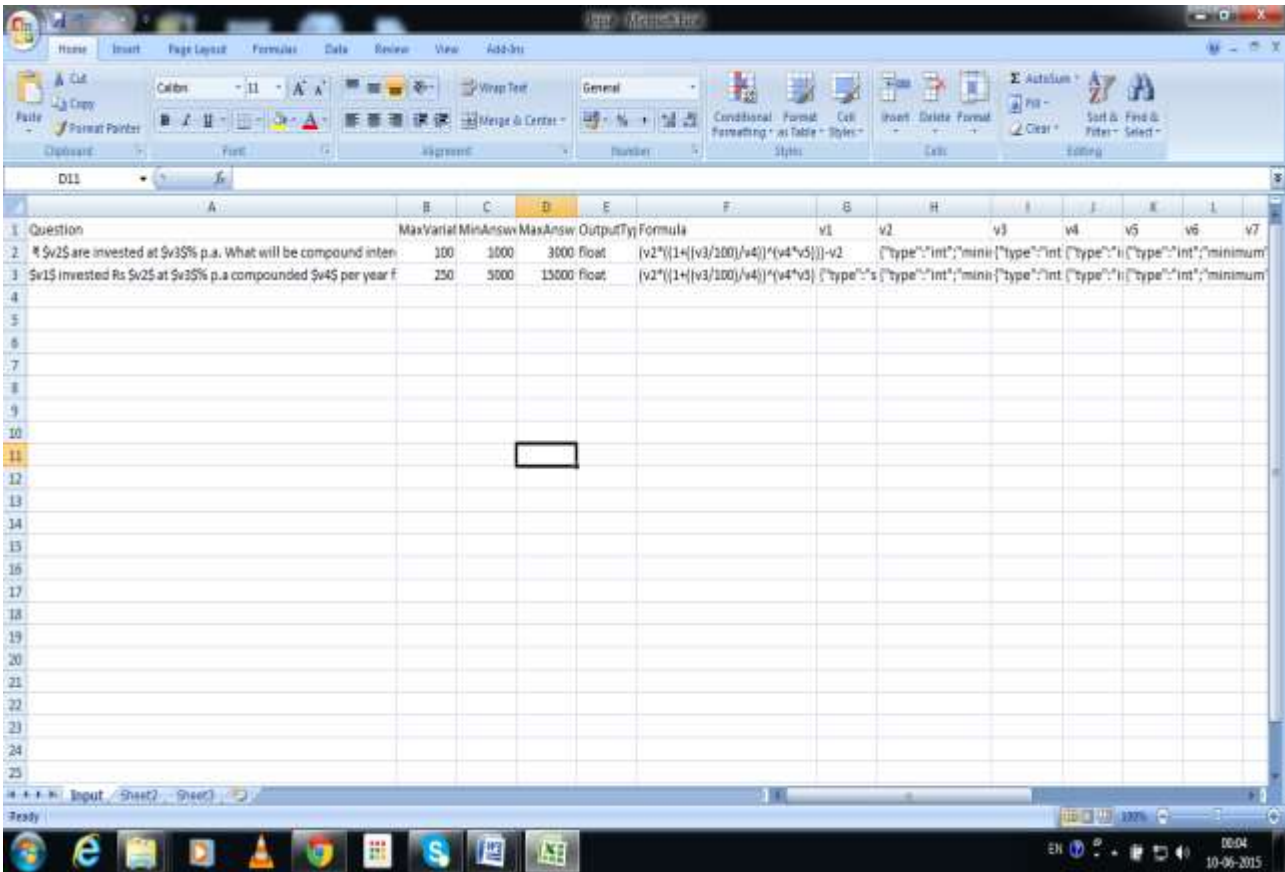


Fig.A.1. Screenshot of Input Sheet

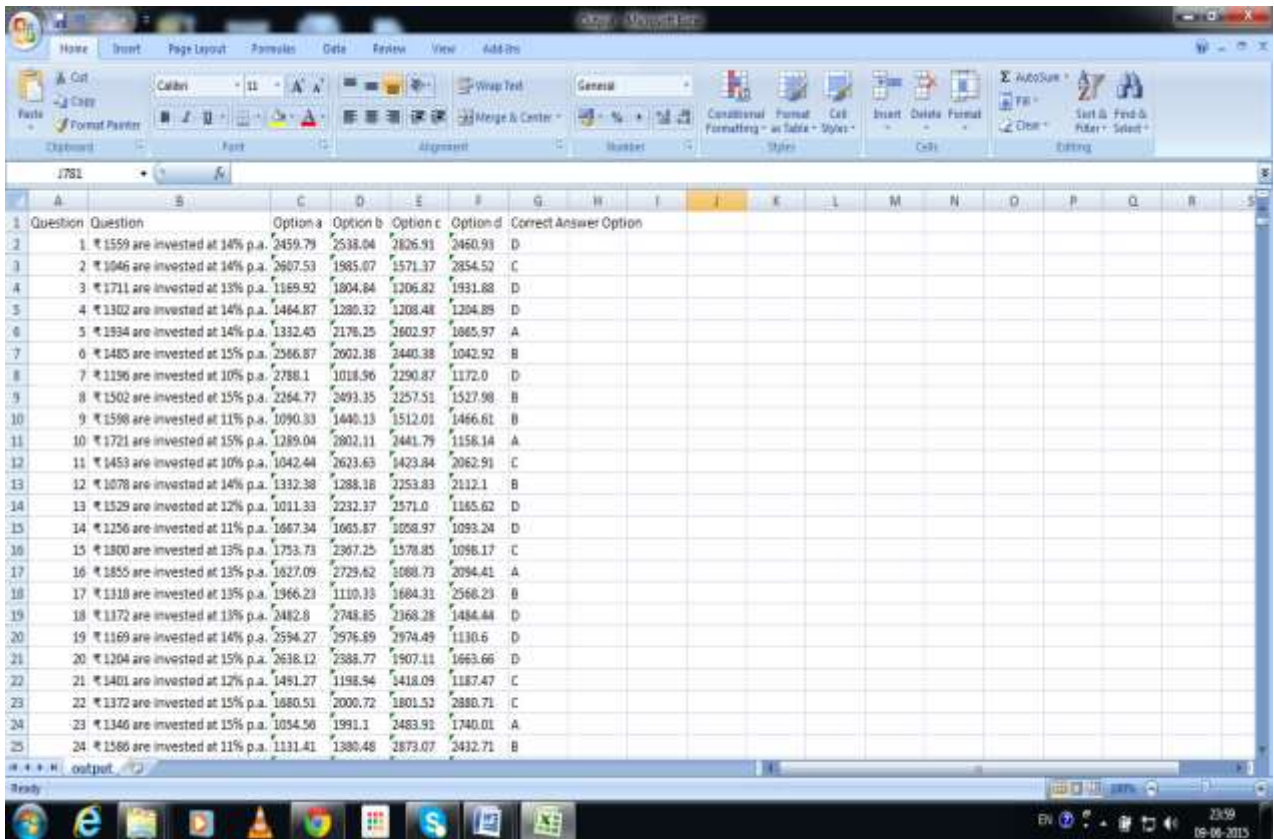


Fig.A.2. Screenshot of Output Sheet

ACKNOWLEDGEMENT

Thanks to Rāshtrīya Vidyālaya Tāntrika Mahāvīdyālaya, Bengaluru and HMR Tāntrika evam Prabandha Sansthan, Nayi Dilli for immense support in the project. Special thanks to our beloved parents for their unconditional love & support.

REFERENCES

- [1] Guzman, E.; Conejo, R., "Self-Assessment in a Feasible, Adaptive Web-based Testing System", *IEEE Trans. Education*, vol.48, no.4, 2005, pp. 688-695.
- [2] Guang Cena; Yuxiao Dong; Wanlin Gao; Lina Yu; Simon See; Qing Wang; YingYang; Hongbiao Jiang, "Implementation of an Automatic Examination Paper Generation System", *Mathematical and Computer Modelling*, vol. 51, 2010, pp. 1339–1342.
- [3] Christopher J. Sangwin, "Implementing Computer Algebra Enabled Questions for the Assessment & Learning of Mathematics", *International Journal of Technology in Mathematics Education*, vol. 15 no. 1, 2008, pp. 3-18.
- [4] Chen Yuan-Bin; Dai Jie, "Design on Algorithm of Automatic Test Papers Generation for Examination System of Electric Energy Measurement", *Proceedings of the IEEE Int'l Conference on Computer Science & Service System*, Nanjing, Jiangsu, China, 2012, pp. 1397-1400.
- [5] Dan Liu; Jianmin Wang; Lijuan Zheng, "Automatic Test Paper Generation Based on Ant Colony Algorithm", *Journal of Software*, vol. 8, no. 10, 2013, pp. 2600-2606.
- [6] Kohilavani S; Mala T; Geetha T V, "Automatic Tamil Content Generation", *Proceedings of the IEEE Int'l Conference on Intelligent Agent & Multi-Agent Systems*, Chennai, Tamil Nadu, Bharat, 2009, pp. 1-6.
- [7] Yuguang Huang; Lei Li, "Naive Bayes Classification Algorithm Based on Small Sample Set", *Proceedings of the IEEE Int'l Conference on Cloud Computing and Intelligence Systems (CCIS)*, Beijing, China, 2011, pp. 34-39.
- [8] Martinez-Arroyo M.; Sucar L.E., "Learning an Optimal Naive Bayes Classifier," *Proceedings of the 18th International Conference on Pattern Recognition (ICPR)*, Hong Kong, China, 2006, vol. 3, pp. 1236-1239.
- [9] Wang Hairui; Wang Hua, "Research and Implementation of Multi-agent Based Test Paper Generation Algorithm", *Proceedings of the IEEE Int'l Conference on Computer Science & Software Engineering*, Wuhan, Hubei, China, 2008, pp. 493-496.
- [10] Anthony J. DolRocco; Sally A. Chambless, "Automated Content Generation", *US Patent 8,798,989 B2*, Aug. 5, 2014.
- [11] Qian Huang; Zhu Liu; Rosenberg A.; Gibbon D.; Shahraray B., "Automated Generation of News Content Hierarchy by Integrating Audio, Video, and Text Information", *Proceedings of the IEEE Int'l Conference on Acoustics, Speech, and Signal Processing*, Phoenix, Arizona, USA, vol.6, 1999, pp. 3025-3028.
- [12] Ge Guo; Jin Jin; Xijian Ping; Tao Zhang, "Automatic Video Text Localization and Recognition", *Proceedings of the Fourth International Conference on Image and Graphics*, Sichuan, Chengdu, Tibet, 2007, pp. 484-489.
- [13] Do Van Nhon, "Constructing Intelligent Systems For Computation – Research A now Development of Knowledge Representation Models To Design Systems

for Automated Solving Problems", PhD. thesis, National University of Ho Chi Minh City, Vietnam, 2001-2002.

- [14] Yang Yu, Hongyan Wang, "Adaptive Online Exam Questions Based on Systematic Analysis and Design", *Wuhan University of Technology*, vol. 4, 2008, pp. 30-35.
- [15] Wang Yu-Ying; HOU Shuang; GUO Mao-zu, "Algorithm for Automatic Test Paper Generation", *Journal of Harbin Institute of Technology*. 2003, pp. 342-346.

Author's Profiles



महाविद्यालय, बेंगलूरु, भारत (Rāshtrīya Vidyālaya Tāntrika Mahāvīdyālaya, Bengaluru, Bharat).

He did his internship from SkillWiz, Bengaluru and worked as SOFTWARE DEVELOPER. His field of interest is Avionics.



भारत (HMR Tāntrika & Prabandha Sansthan, Nayi Dilli, Bharat). His field of interest is Software Engineer.

How to cite this paper: Sandeep Singh Yadav, Mandeep Singh Yadav, "Development of System for Automated & Secure Generation of Content (ASCGS)", *International Journal of Information Technology and Computer Science(IJITCS)*, vol.7, no.11, pp.81-88, 2015. DOI: 10.5815/ijitcs.2015.11.10