Development of Corruption Detection App Using K-means Clustering

Md. Fazle Rasul, Nahin Kumar Dey and Dr. MMA Hashem
Khulna University of Engineering and Technology, Khulna, Bangladesh.
Email: fazlerasul3@gmail.com, nahin333@gmail.com and mma.hashem@outlook.com

Background

Corruption means the abuse of entrusted power for private gain. Corruption has been a regular issue for Bangladesh. The main problem lies in the fact that most people are not likely to directly protest against it.

Corruption can be classified as grand, petty and political based on the amount lost and sector where it occurs[1]. Corruption can be in many forms like: Bribery, Embezzlement and Fraud, Extortion, Conflict of interests, Favoritism, Nepotism, Cronyism, Political Corruption etc.[2] affecting different sectors[3]. So an intelligent system (app) has been proposed to detect corruption based on anonymous feedback from service receivers by using k-means clustering and modified k-means clustering algorithms. This system (app) will sort respective persons according to their corruption level. Service receivers can also complain against corrupted persons with evidence in the form of audio, video, pdf or image. Suggestion and complaint portal [4], ratings and reviews systems in e-commerce [5] or online services [6], teaching evaluation report system [7] can use this app to detect corruption in respective fields.

Motivation and Problem Formulation

System data is generated from opinions given by service receivers and internal employees through the app. App users will register using NID and email. There is an optional field for members who will take actions against these complaints. Complainer's identity will be hidden from them but to ensure that the complaints are not coming out of spite, complainer's information will be available to specialized persons only. Service receivers will give opinions through answering five questions. They will choose among five options (Strongly agree, agree, neutral, disagree, strongly disagree).Each option will bear point from 1 to 5. Similar opinion will be taken from service providers internally. All user information and evidence submitted by app users will be stored in the cloud and will be visible to the concerned authority only.

Through this app we can identify corrupted persons and more importantly information about opinion holders will be hidden. The main limitation is it is an online based idea and hence workable over internet only. In Bangladesh it is difficult to access over the internet for each and every one. It is also difficult to detect internal corruption. The app works best in cases where service receivers directly interact with the service provider. The main task is to organize opinions to apply different algorithms on them and cluster service providers in honest, average and corrupted group.

Proposed Methodology

Proposed model is shown in fig 1. and algorithms used in this app are shown in fig 2. There will be different interfaces in the app to have data from users.

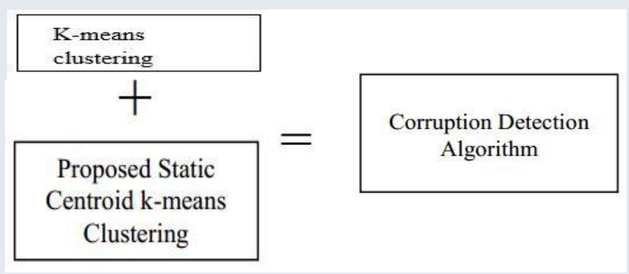

Fig 2. Proposed algorithm for corruption

Static Centroid K-means clustering algorithm :

1: Select static K cluster centroids U1,U2,…Uk € R
2: Repeat {
3: for i = 1 to m
4: Ci = index (from 1 to K) of cluster centroid
5: closest to Xi
6: for K = 1 to K
7: Uk= average (mean) of points assigned to cluster k }

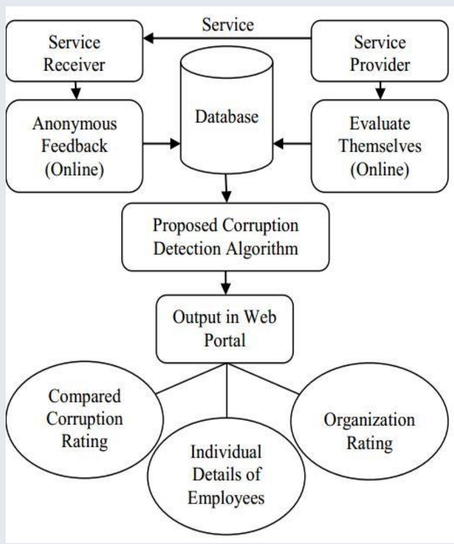

Fig 1. Proposed model to detect corruption.

Proposed static centroid K-means clustering algorithm is on the left side. It is to applied on the result obtained by applying K-means clustering.

Experimentation

Users will submit evidence and other information as shown in fig 3.

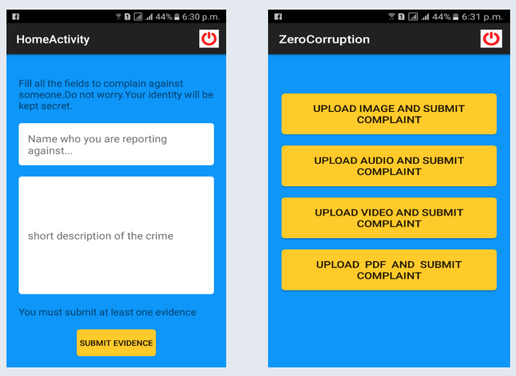

Fig 3: Submitting complaints

User opinions will be taken for all employees like fig 4 and K-means clustering will be applied on these data. Centroids will be fixed from now on and based on distances from centroids, employees will be distributed to the nearest centroid of that employee constructing a total number of three clusters.

All the data are saved in the cloud in real-time using firebase. Every time an app user wants to see corruption results, necessary data will be fetched from cloud and will be manipulated according to the algorithms as stated in the previous sections and it will be calculated in the app internally giving us three clusters denoting honest , less honest and corrupted group of people. Only last fifty opinion for each employee will be stored and older opinions will be overwritten with new ones. Thus corruption result changes with opinion change.

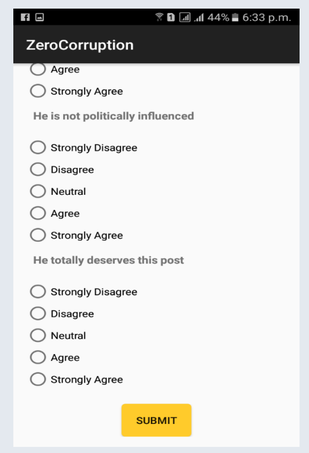

Fig 4. Taking user opinions.

Findings

The app shows tremendous outcome though it depends on unbiased opinions from both internal employees and service receivers. Biased opinion from opinion holders can judge an employee wrongly. But as the result depends on large number of opinions, it can be assumed that most of the opinions will be true. And as the result changes with more number of opinions it can be regarded to be a perfect one to judge honesty of employees. It accurately clusters employees into honest, less honest and corrupted ones and reflects the public opinion regarding any employee or an organization.

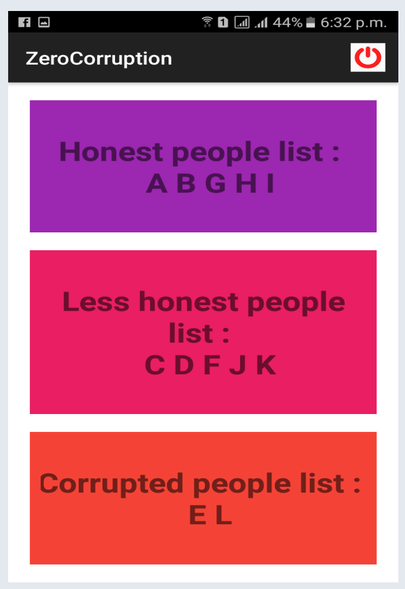

Fig 5: clustered employees

Conclusion and Future Work

There are a lot of future scopes of this app. Auto mail sending option to concerned organization after specific time can be developed. It is possible to generate history graph of improvement of employees. The model will be capable to data synchronization for transfer issue and synchronized with an SMS based feedback system. This system can be used as a digital complain and suggestion box.

In summary, this model will be effective in society if the corrupted people are being faced punishment. It is also important to make awareness against corruption and its bad effect. Education from family can be more effective to the growth of morality. Everyone must be made aware of the consequences of corruption.

References

1. e.V., T. (2018). Transparency International - What is Corruption?[online] Transparency.org. Available at: https://www.transparency.org/what-is-corruption [Accessed 12 May 2018].
2. Assignment Point. (2018). Corruption in Bangladesh - Assignment Point. [online] Available at: http://www.assignmentpoint.com/arts/social-science/corruption-bangladesh.html [Accessed 5 Apr. 2018]
3. Bliss, B. (2018). Bangladesh Corruption Report. [online] Business Anti-Corruption Portal. Available at: https://www.business-anti-corruption.com/country-profiles/bangladesh/ [Accessed 12 May 2018].
4. Just.edu.jo. (2018). Complaints and suggestion system. [online] Available at: http://www.just.edu.jo/Centers/HealthCenter/Pages/Complaints-and-suggestion-system.aspx[Accessed 5 Apr. 2018].
5.Online Shipping Blog | Endicia. (2018). Customer Feedback and Why Your E-commerce Store Needs It - Online Shipping Blog | Endicia. [online] Available at: https://online-shipping-blog.endicia.com/customer-feedback-and-why-your-e-commerce-store-needs-it/ [Accessed 5 Apr. 2018].
6. Team, L., Directors, B., News, P., Us, C., Programs, S. and Videos, A. (2018). How You Can Use Customer Feedback to Improve Your Business - Insightly. [online] Insightly. Available at: https://www.insightly.com/blog/how-you-can-use-customer-feedback-to-improve-your-business/ [Accessed 5 Apr. 2018].
7. P. Tan, M. Steinbach, A. Karpatne and V. Kumar, Introduction to data mining. New York, NY: Pearson Education, 2018.