CAAP Quarterly Report (Oct-Dec 2018)

Date of Report:	January 15, 2019
Prepared for:	U.S. DOT Pipeline and Hazardous Materials Safety Administration
Contract Number:	Cooperative Agreement #693JK31850011CAAP
Project Title:	Development of a prediction model for pipeline failure probability based on learning from past incidents and pipeline specific data using artificial neural network (ANN)
Prepared by:	Texas A&M Engineering Experiment Station
Contact Information:	James C. Holste j-holste@tamu.edu (979) 845-3384
For quarterly period ending:	December 31, 2018

Business and Activity Section

(a) Contract Activity

The project has been initiated in this quarter.

(b) Status Update of Past Quarter Activities

A team has been formed with Dr. Noor Quddus (postdoctoral research associate) as the team lead for the project. Three graduate students with experiences on data analysis, incident investigation and neural network have been included in the team.

(c) Cost share activity

No cost incurred in this quarter.

(d) Task 1: Development of methodology for creating root-cause analysis reports

Development of methodology for creating root-cause analysis was scheduled for the first quarter.

Detailed discussion and descriptions for the following:

1. Background and Objectives in the 1st Quarter

In order to develop the artificial neural network (ANN) for building a predictive model, data of what contributed to past incidents will have to be gathered. All incidents will need to be investigated to determine what went wrong. Different root-causes analysis usually will identify different causes and investigation reports may express these causes in a variety of ways, although certain management or organization-related

causes may be common to several incidents. As mentioned in the RFP, root-cause failure analysis can have different fields of origin (such as safety-based root-cause failure analysis, production-based root-cause failure analysis *etc.*). Also, what may appear as root cause may actually provoke further questioning to identify deeper hidden causes in another field of origin. Thus, for the selection of inputs, it is essential to determine a reference that defines what actually is termed as root cause. Using taxonomy so that similar terms are used to refer to related root causes can help tackle this issue. At the same time, identified root causes may be present in a way that cannot be used for measurement. This problem is increased when human and organizational factors are identified as a root cause. For example, an investigation may find lack of maintenance as a root cause. This does not provide any measured value that can be used as an input. However, if the percentage of deferred maintenance and the length of deferred maintenance were measured for the system at the time of incident, then it would provide a quantitative assessment of the condition the pipeline system was in at the time of the incident. Thus, investigation findings must be recorded in a manner that will enable information related to the root cause to be expressed in terms of quantifiable indicators. For a similar reason, the output for the investigations must be expressed in terms of severity levels.

Thus, the first challenge is to build a set of cue words or taxonomy so that root-causes analysis conducted for different incidents identify similar causes using similar terms and these causes must be identified in terms of measurable deviations or indicators so that they can later be compared with deviations existing in a system to understand if the system is reaching an unsafe state. If this is done, then there will be a consistency among the reports of root-cause analysis to enable extraction of information from those reports to build a learning model, and then compare them with the current condition of the system to predict failure. If a set of cue words are developed to produce all reports, extraction of information using automated systems based on text mining or data mining can be used.

Task 1 will focus on determining how root-cause analysis can be reported so that the information can be applied for prediction based on current condition.

2. Analysis in the 1st Quarter

No analysis was completed during this quarter.

3. Results and Discussions

No result was produced during this quarter.

4. Future work

Task 1 will be conducted in the next quarter.

References

None.