

Responsive Cities – Open Data Challenge

28 October 2016.

How can we use open data to reduce cable theft in Tshwane?

Background

Copper cable theft is a serious problem all around South Africa. While there are existing solutions in place and in the market to address this concern, none of them are preventative. Most of the solutions involve placing some sort of identification on the cables to deter the trade of this copper with metal scrap dealers.

Companies such as Eskom, Telkom and Transnet are incurring a great deal of financial loss at the hands of copper thieves who are targeting their infrastructure.

What is needed is a way to stop the theft before any major disruptions and loss is caused.

Problem Statement

The challenge requires the use of open data to develop a cost-effective solution to reduce copper theft in one or more areas in Tshwane:

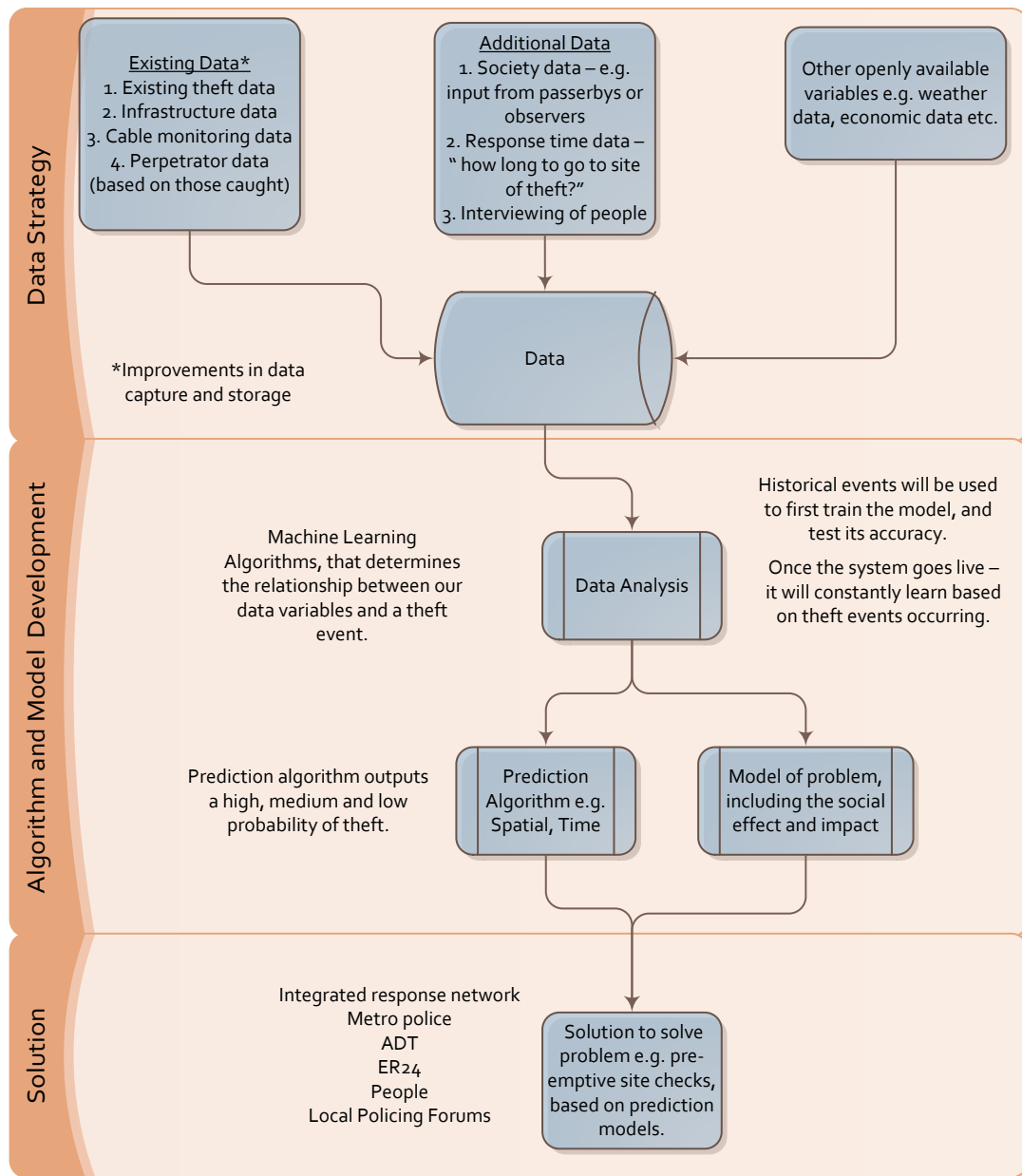
- Mountainous terrain of the Tshwane area.
- In townships where thieves target house connections.
- In the city centre and surrounding metropolis.

Proposed Solution

The proposed strategy to define a solution has been broken up as follows:

- Data Strategy
- Algorithm and Model Development
- Solution Development

The figure on the next page presents an overview of this approach.



• **Figure 1: Proposed Solution for Cable Theft in Tshwane**

1. Data Strategy

The aim of this phase is to collate data from the different sources. The objective is to ensure that we have considered all factors or “features”, that could have a direct impact or have a secondary impact on the prediction algorithm and hence the solution. It is for this reason that the following data will be looked at initially:

- Existing theft data
- Infrastructure data
- Cable monitoring data
- Data about the people that have stolen cables and that have been caught
- Situational data (time of day of theft, weather, stock price, etc.)

- Possible sensor data

Other useful data may be identified during the course of this project and incorporated into our strategy. This data will be stored in a central database.

2. Algorithm and Model Development

The data analysis phase would do the following:

- Formatting and cleaning the raw data into a form suitable for data analysis.
- Determine relationships as well as patterns between data and the “features”.

This information would then be used to model the problem, as well as determine a prediction algorithm so that the time and location of a potential cable theft could be predicted.

During the data analysis phase the model will be trained and its accuracy tested. As new data becomes available, it will constantly update to account for new information provided to it. Thus the model will evolve with time and become more accurate.

3. Solution Development

Both the model and predictive algorithm would also indicate if additional data is required and thus sensors can be used to get this additional data. The model will also help to determine if an integrated solution is required, as the problem of cable theft is a multifaceted one.

The final solution would have to include the following:

- Prediction of the next possible cable theft in real time.
- Identification of a real theft situation or a decoy.
- Would need to ensure that response to stop the crime in progress would be within a certain timeframe. This will be different for each situation.
- The evidence collected must be feasible and allowed in a court of law.

Conclusion

Proper data collection, storage and analysis are essential to predicting the criminal’s next move. In this solution we pit the cable thieves against an artificial intelligence system that will try to figure out the next move based on the relevant data. The challenge also lies in responding quickly enough based on this prediction and this is where an integrated response network is of value.