Triggering Solar-Powered Vehicle Activated Signs using Self Organising Maps with K-means

Diala Jomaa, Siril Yella, Mark Dougherty Department of Computer Engineering, Dalarna University 78170 Borlänge, Sweden dio@du.se, sve@du.se, mdo@du.se

Abstract— Solar-powered vehicle activated signs (VAS) are speed warning signs powered by batteries that are recharged by solar panels. These signs are more desirable than other active warning signs due to the low cost of installation and the minimal maintenance requirements. However, one problem that can affect a solar-powered VAS is the limited power capacity available to keep the sign operational. In order to be able to operate the sign more efficiently, it is proposed that the sign be appropriately triggered by taking into account the prevalent conditions. Triggering the sign depends on many factors such as the prevailing speed limit, road geometry, traffic behaviour, the weather and the number of hours of daylight. The main goal of this paper is therefore to develop an intelligent algorithm that would help optimize the trigger point to achieve the best compromise between speed reduction and power consumption. Data have been systematically collected whereby vehicle speed data were gathered whilst varying the value of the trigger speed threshold. A two stage algorithm is then utilized to extract the trigger speed value. Initially the algorithm employs a Self-Organising Map (SOM), to effectively visualize and explore the properties of the data that is then clustered in the second stage using K-means clustering method. Preliminary results achieved in the study indicate that using a SOM in conjunction with K-means method is found to perform well as opposed to direct clustering of the data by Kmeans alone. Using a SOM in the current case helped the algorithm determine the number of clusters in the data set, which is a frequent problem in data clustering.

Keywords: Solar-powered vehicle activated signs; Self Organising Maps; K-means clustering; Trigger speed

I. INTRODUCTION

An excessive or inappropriate speed is often a reason for traffic fatalities. Therefore an important consideration for traffic municipalities is to reduce speeding by either modifying the roadway infrastructure or introducing additional and more effective signage. Modifying road infrastructure is more costly than deploying additional signs. Therefore, a range of road safety signs has been developed and deployed to encourage drivers to adapt to the speed limit or to warn drivers when they are approaching a hazard. Solar vehicle activated signs (VAS) are one type of signs that are widely used on roadways. Typically, solar VAS are speed warning sign powered by batteries that are recharged by solar panels. These signs are more desirable than other simple battery driven VAS due to their low cost of installation. A main source of the reduction in cost is due to

the fact that no external power supply is needed. These signs usually consist of radar that is mounted inside the sign in order to detect vehicles and measure their speed. The sign displays a message when vehicle speed exceeds a pre-set threshold, which is called the trigger speed. The trigger speed is usually set to a constant value, which is often equal, or relative, to the speed limit on a particular segment of road. Earlier studies reviewing the effectiveness of variable message signs or vehicle activated signs have been reported by [1]. Such studies have reviewed relevant work published between 2000 and 2005 and have mainly investigated the influence of VMS on human behaviour. This study is an update to the earlier studies in that this study includes data from 2006-2009 [2]. The prior studies showed that these signs have a significant impact on driver behaviour, traffic safety and traffic efficiency. In most cases, the signs have vielded reductions in the mean speed and in speed variation as well as in longer headway. However, most of the experiments were performed with the signs set to a certain static configuration under specific conditions. Since some of the aforementioned factors are dynamic in nature, it is felt that the earlier researchers did not consider the aspects of sign configuration carefully enough. The previous studies lack a clear statement describing the relationship between the trigger value and its consequences under different conditions [3]. Efficiently setting up the radar speed threshold helps prevent the battery from running out of power. Previous authors have reported different strategies for calculating the appropriate trigger speed. In one reported experiment the trigger speed was set at 10% over the speed limit plus an additional 2mph, i.e., in a 30mph speed limit the trigger speed would be set at 35mph) [4]. The trigger speed was set at the 50^{th} percentile of the speed that was detected prior to the installation of the VAS. This was intended to target half of the drivers. In other previous studies, the trigger speeds were set at between the 75th and 81th percentile speeds [5]. In Mattox et al., the predetermined trigger speed was set to the posted speed limit with a 3- mph buffer [6]. The method that is used predominantly in the United Kingdom is one in which the trigger speed is set to the 85th percentile of the average speed that is measured before installation.

However, solar VAS are often challenged by a limitation is the capacity of the power supply that is required for the sign to remain operational. In order to be able to operate the sign more efficiently, it is proposed that the sign trigger should take into account the prevalent conditions. Triggering the sign depends on many factors such as the prevailing speed limit, road geometry, traffic behaviour and number of hours of daylight (more daylight implies more solar power available, but on the other hand the sign needs to shine brighter during daylight, which implies higher power consumption). As a consequence of the lack of power capacity, the sign may be triggered with a high value but it should preserve the impact of the sign on vehicle speed reduction. To determine the optimal trigger speed in order to minimize power consumption while simultaneously maximizing vehicle speed reduction is a nontrivial problem. Thus, the optimal trigger speed will be accomplished by first collecting traffic data using various trigger speeds, preprocess the data and extract the trigger speeds that reveal the information and relationship hidden in these data for example, the relationship between the time of day and the traffic conditions. The objective of this paper is to therefore develop an intelligent algorithm that searches for the optimal trigger speed, which yields the best compromise between reducing both vehicle speeds and the power consumption of the sign. The algorithm is mainly done by combining two clustering techniques, Self-Organising Map (SOM) and Kmeans. In this algorithm, traffic data will be first clustered by a SOM in order to effectively visualize, explore the properties of the data and determine the preliminary number of clusters as well as determine the centre of each cluster. After using the SOM, the clusters are refined by using Kmeans. The remainder of this paper is organized as follows. Section 2 details the experimental design and discusses data acquisition. Section 3 presents issues relevant to solar power consumption for the sign. The developed algorithm and its trigger speed are described extensively in sections 4 and 5. The results and discussion are reported in section 6. Lastly concluding remarks will be presented.

II. SOLAR POWER CONSUMPTION

Solar powered signs are allow for signs to be easily installed in locations that are far away from the power grid. They are designed to run all year round, particularly running more in the winter when there is less light. The radar is always on and this draws somewhere in the region of 40-50mA. However, the only other power draw is the controller and 3 LEDs on the rear of the sign, which require about 20mA in total. The total current from the 12V battery is then around 70mA. When the sign faced is active and at full brightness, the current consumption is 1.8A. This reduces in steps as the ambient light reduces right down to under 0.5A. The modem that is fitted to this unit is also drawing current all of the time this is in the region of 50mA and has an impact on battery life. The current taken by the unit depends on the quality of the network signal that is available at the site and the amount of data that is being moved, so it is not an easy sum to calculate. Therefore there is a need to consume the energy efficiently in order to ensure that the VAS has an adequate amount of power. The established sign runs on 40W solar panels with 35 Ah batteries. In this study, the energy consumption is calculated as follows. If the vehicle speeds exceed the proposed threshold speed then:

- Calculate the length of flash f for each vehicle speed
- Sum all flashes
- Calculate energy consumption at full brightness in Ah

III. EXPERIMENT DESIGN AND DATA ACQUISITION

Solar Traffic agencies have a specific policy for the use of interactive signs such as VAS. These policies stipulate where the sign can be placed and the possible roadways for which they are suitable. Such signs are mostly placed at or near speed limit changes or at sites where a high collision rate exists. Sites near junctions and pedestrian crossings are normally avoided, as vehicle speeds are generally reduced in these areas [7]. In this study, two test sites were selected in Borlänge, Sweden. The first site is referred to as the Korsgård test site. The Korsgård test site is located between the Tuna and Hugo Hedström roadways. The second site is referred to as the Mjälga test site. Note that traffic flows in both directions and the posted speed limit at the test sites is 40 km/hr. Furthermore both sites are located in rural area and are notorious for speeding. At both sites, two radars (radar 1 and radar 2) were used in order to study the reduction of vehicle speed before and after triggering the sign. Radar 1 was positioned 100 meters before the VAS and radar 2 was positioned in line with the VAS. The VAS is triggered at 100 m in distance before the location of the sign. The sign used in the current study is a typical solar powered vehicle activated sign (VAS). The sign displays two warning messages in succession. The first is a reminder of the posted speed limit, which is 40km/hr, which is followed by a "SÄNK FARTEN" (reduce speed) message. Typically the messages are displayed only when the vehicle speed exceeds the trigger speed. The sign is equipped with radar and a data logger to detect and record vehicle speed. The sign is also equipped with a general packet radio service (GPRS) modem to facilitate communication to the radar and for authorized users to download and upload data. Note that it has been possible to alter the radar settings remotely. Such a setup has facilitated alteration of the trigger speed, thereby permitting the study to investigate the effect of different trigger speeds on various driving speeds. Upon request the data stored in the collection module is uploaded to a web server for the user to download.

At both sites, the data were collected 24-hours a day. At the Korsgård test site, data collection were done from 1 September 2012 through 31 December 2012 and at the Mjälga test site, data collection were established between from 1 May 2013 and 1 August 2013. The data collected consists of records about vehicle speed, the direction the vehicle is traveling, the length of vehicle, a time stamp and a time gap and the date. Only the vehicle speed, direction and the time of day were used in this study.

IV. A TWO STAGE CLUSTERING APPROACH USING SELF-ORGANISING MAP WITH K-MEANS CLUSTERING

is an unsupervised classification Clustering of observations, data items, or objects into different groups or clusters so that the data in each cluster share the same properties [8]. Grouping is usually done on the basis of similarities or distances without any assumptions concerning the number of groups that should be pertained to the data. There are various methods that are used for cluster analysis. Self-Organising Maps (SOM) and K-means are two methods that are commonly applied in several fields such as marketing, pattern recognition and traffic analysis. However, the performance of these methods can differ depending of the characteristics of the data, for example the size of the data, the number of clusters, and the type of data. Nevertheless, none of these methods outperforms the others in all data conditions. For instance, SOM is sensitive to the size of the data set; its speed of convergence is slower than K-means [9]. K-means is a simple and fast algorithm but it is very sensitive to the selection of the initial number of clusters [10]. In this paper, a Self-Organizing Map (SOM) was initially used to visualize, and understand data properties such as the number of prevalent clusters. Preliminary identification of the number of clusters is deemed useful because it enables the researcher to decide the number of clusters (K) while implementing the K-means algorithm [11].

A. Khohonen self-organising maps

The Self-organising map (SOM) is an artificial neural network that learns the properties of data via an unsupervised learning algorithm [12]. It consists of two layers of artificial neurons; an input layer and an output layer. Every input neuron is connected to every output neuron by a weighting value. The Euclidian distance is calculated between the input vector and the incoming weighted vector for each output. The output neuron with the smallest distance is declared as the winner and its weights modified to be closer to the input vector. In fact SOM is an iterative process where the connections' weights are modified according to the following equations (1) and (2) [13, 14].

$$w(t+1) = w(t) + h(t)(x(t) - w(t))$$
(1)

$$h(t) = \alpha \cdot e^{-\frac{d^2}{2\sigma^{2(t)}}}$$
(2)

Where w(t) is the connection weight at time t,

 $\mathbf{x}(t)$ is the input vector

h(t) is the neighbourhood function,

 α is the learning rate,

d is the Euclidian distance between the winning unit and the current unit,

 σ is the neighbourhood width parameter

B. K-means clustering

The k-means clustering is an optimization clustering algorithm where clusters are formed by optimizing some measure of cluster goodness. The centre of the cluster k is the mean of the data items within the cluster. The K-means algorithm proceeds by first randomly selecting k of the items where each selection is done by partitioning data items into k initial clusters. Each item is assigned to the cluster to which it is the most similar, based on the distance between the item and the cluster mean. It then computes the new mean for each cluster and assigns the new mean as the new cluster centre. This process iterates until a stopping condition is reached.

In this paper a two stage clustering algorithm is applied as follows:

- a) Data pre-processing: The data is initially preprocessed by simple filtering. In this case the filtering consists based on the direction the vehicles are traveling. Vehicles traveling in the same direction are filtered into the same group. Later, the data is grouped into two classes: cars and trucks. Motorcycles and long trucks are excluded from this study by assuming that their speed is low.
- b) Feature extraction: The main features considered in this study are: time mean speed μ , traffic flow q and the standard deviation of speed σ . Assuming that all of the vehicles are moving with v km/hr. The number of vehicles counted at a certain point in one hour is the traffic flow q. Time mean speed is the average of spot speed V_i or simply the average of n vehicles passing a point during a certain period of time [15]. Time mean speed μ is given by equation (3):

$$\mu = 1/n \sum_{i=1}^{n} v_i \tag{3}$$

Speed standard deviation is approximately the square root of the sum of the squares for the difference between each vehicle speed v_i and the mean speed μ [16].

Speed standard deviation is given by equation (4): $\sigma_{=} \sqrt{1/n^2 \sum_{i=1}^{n} (v_i - \mu)}$ (4)

- c) The SOM clustering stage: Obtain the initial number of clusters by SOM. The input vectors assigned to the SOM input neurons consist of 3 dimensions; time mean speed, traffic flow and standard deviation. In this experiment, a 24 by 3 layer of neurons is used to classify the input into 3 clusters. This means a layer of 72 neurons spread out in a 24 by 3 grid. After training the network with 100 iterations, the map is well distributed with regard to the input space.
- d) The K-means clustering stage: Refine the cluster centroid by K-means. The number of clusters and the cluster centres obtained from SOM can be used as the initial input of the K-means algorithm.
- e) Trigger speed setting: According to the clustering results obtained from the K-means, the trigger speed is considered as the median speed of the cluster.

V. RESULTS AND ANALYSIS

The two-stage clustering algorithm was applied to the experimental data obtained from both test sites; Korsgård and Mjälga. A choice of K=3 is based on the initial clustering done by SOM. Figure 1 (a) and (b) show how the K-means partitioned the traffic flow into 3 clusters at both the Korsgård and Mjälga test sites, respectively. This means grouping the time of day into three clusters that correspond to the number of vehicles passing the sign. At the Korsgård test site, the traffic flow is low in cluster 1, average in cluster 3, high in cluster 2, however at the Mjälga test site the traffic flow is low in cluster 2, and high in cluster 1. In fact, the partition of the time of day for the traffic flow is not the same in both sites.



Figure 1. Clustering traffic flow respective to time; (a) at the Korsgård test site; (b) at the Mjälga test site

The main idea in using K-means clustering is to retrieve the median speed centroid, which is the expected trigger speed threshold. Fig. 2 (a) and (b) shows a box plot of the median speed for the three clusters at the Korsgård test site and at the Mjälga test site, respectively. The central mark in this box plot is the median, the edges of the box are the 25th and 75th percentiles and outliers are plotted individually. The whiskers in Cluster 2 at the Korsgård site and in cluster 3 at the Mjälga site are more extended than other clusters due to a larger deviation between median speeds. In fact when the traffic flow is high at the Korsgård test site, the median speed is high (cluster 2) but when the traffic is low at the Mjälga site, the median speed is high (cluster 3). Bear in mind, that both test sites have a speed limit of 40 km/hr.



Figure 2. Box plot of median speed for the three clusters; (a) at the Korsgård test site; (b) at the Mjälga test site

For the purpose for this study, the 24 hours of the day were grouped into three time period clusters. The time periods obtained from the clustering algorithm ensured that the trigger speed is applied to the entire clusters, rather to the individual hours. The trigger speeds are considered as the centroid median speeds for the clusters obtained from the experiment. Tables 1 and 2 present the centroid median speed and the corresponding standard deviation of all of the clusters from the mean speed at the Korsgård and Mjälga test sites, respectively.

TABLE 1. THE TIME PERIOD RANGE, THE CENTROID'S MEDIAN SPEED AND THE CENTROID STANDARD DEVIATION FOR EACH CLUSTER AT THE KORSGÅRD TEST SITE

Clusters number (K)	Cluster's time period range	Centroid median speed	Centroid Standard deviation
1	{07,09,10,11,12,19,20,21}	49	09
2	{00,01,02,03,04,05,06,22,23}	51	13
3	{08,13,14,15,16,17,18}	49	08

Clusters number (K)	Cluster's time period range	Centroid median speed	Centroid standard deviation
1	{00,01,02,03,04,05,06,22,23}	51	13
2	{07,09,10,20,21}	49	10
3	{08,11,12,13,14,15,16,17,18,19}	48	09

TABLE 2. THE TIME PERIOD RANGE, THE CENTROID'S MEDIAN SPEED AND THE CENTROID STANDARD DEVIATION FOR EACH CLUSTER AT THE MJÄLGA TEST SITE

Both sites provide nearly identical results for the centroid median speed and the centroid standard deviation. However, the partitioning of the day is not the same in all clusters. The partitions are the same at night-time but it differs during the daytime.

VI. VALIDATION AND ENERGY CONSUMPTION

Validating the data is not a straightforward process. There are several cluster validation techniques. One of these techniques which were initially proposed by Hauser and Schere, is to break down the data into subsets and then try to reach the same clusters from the original data [17]. In this study, the validation is done by comparing the effect of the sign that was triggered with the clustering model to the effect of the sign that was triggered with other static trigger speeds. The static trigger speeds are obtained from the 15th, 50th and 85th percentiles of the vehicle speeds. The 15th percentile speed is the speed up to which 15% of vehicles travel. For example if the 15th percentile speed was 100 km/hr then 15% of the traffic would travel at 100 km/hr or lower. The same definition applies to the 50th and 85th percentile speeds respectively. Bear in mind that 15th, 50th and 85th percentiles are reflecting a vehicle's speed at the specific test sites (as measured before the experiment began). To limit the effect to only weekdays, the evaluations were based only on the data that were collected on three Mondays. Table 3 shows the various trigger speeds applied at both test sites.

TABLE 3. THE VARIOUS TRIGGER SPEED (KM/H) AT THE TEST SITES

Korsgård test site		Mjälga test site	
Trigger speed basis	Trigger speed (TS)	Trigger speed basis	Trigger speed (TS)
15th percentile	42	15th percentile	46
50th percentile	47	50th percentile	49
85th percentile	52	85th percentile	50
Clustering	52-47	Clustering	52-49
model		model	

The effect of the sign is done by calculating:

- The variation in the mean speed of a vehicle travelling before triggering the sign as well as after triggering the sign.
- Energy consumption

Fig. 3 and Fig. 4 show the energy consumed on the clustering model is actually near the 50^{th} percentile and larger than the 85^{th} percentile at both test sites.



Figure 3. Energy consumption in Ah for the clustering model compared to the models based on the 15th, 50th and 85th percentiles at Korsgård test site



Figure 4. Energy consumption in Ah for the clustering model compared to the models based on the 15th , 50th and 85th percentiles at Mjälga test site

Fig. 5 and fig. 6 shows the effect of the clustering model on speed reduction compared to 15th, 50th and 85th percentiles. The speed reduction was based on the speed variation 100 meters before and 100 meters after triggering the sign. As seen in the figure, over the cause of one day, data were entered into three clusters according to the clustering algorithm.

The clustering model performed better at the Mjälga test site than at the Korsgård test site. At Korsgård, the greatest reduction in speed was in cluster one (nighttime) that had a trigger speed in the 15th percentile. At Mjälga, the greatest reduction that occurred at night was with a trigger speed that was in the 85th percentile.



Figure 5. Speed reduction for the clustering model compared to the models based on 15th, 50th and 85th percentiles at korsgård test site



Figure 6. Speed reduction for the clustering model compared to the models based on 15th, 50th and 85th percentiles at Mjälga test site

In fact, the results obtained from both test sites were not similar where both sites have the same speed limit (40 km/hr). Note that. VAS should be individually configured and adapted to the location and its traffic conditions in order to result in optimal effectiveness.

VII. CONCLUSION AND FUTURE RESEARCH

This paper presents an intelligent algorithm that would help optimise the trigger point at which the sign should be operative. A self-organizing map helps the algorithm find the number of cluster and extract the right input dataset that is required to perform K-means clustering. K-means clustering divides the 24 hours of the day into 3 clusters. Kmeans clustering provides the algorithm to determine the trigger speed threshold simply by extracting the centroid median of the clusters. The comparison in energy consumption and speed reduction between the algorithm and other models shows that the algorithm is able maintain a reasonable level of energy consumption while also positively affecting driver behaviour. The algorithm was applied to the data set using traffic flow, median speed and standard deviation as the main features to categorize the time of day. To better assess the usefulness of this algorithm, the data set can be extended with more features. The data were collected by varying the trigger speed threshold value at each site. This data can be also extended by altering both the trigger speed threshold and the trigger distance threshold. Thus, it is beneficial to configure the sign in a way that reflects the actual traffic situation and provides optimal energy consumption.

Future research should also include a comparison between the effectiveness of the VAS during the day and at night as well as a comparison of the effectiveness of the VAS during the week and on the weekend. Meanwhile, the traffic data shall further be integrated with weather data sources. The results differ between the test sites. This means that the location of the sign can have an impact on the study. Obviously in order to optimise the effect of the trigger speed on a driver's speed, the VAS should be further adapted to a greater diversity of traffic and road conditions. In the long term the goal is to develop an "intelligent" VAS which selfadapts to traffic conditions on site in order to automatically operate at the optimum trigger speed.

REFERENCES

- S. Nygårdhs and G. Helmers, 'VMS Variable Message Signs. A Literature Review', Swedish National Road and Transport Research Institue VTI rapport 570A, Linköping, 2007.
- [2] S. Nygårdhs, 'Literature Review on Variable Message Sign 2006- 2009', Swedish National Road and Transport Research Institue VTI rapport 15A, Linköping, 2011.
- [3] D. Jomaa, S. Yella, and M. Dougherty, 'Review of the effectiveness of vehicle activated signs', Journal of Transportation Technologies, Vol.3 No.2, 2013.
- [4] Department for transport, 'Vehicle-Activated Signs', Traffic advisory unit, London, 2003.
- [5] M. A. Winnett and A. H. Wheeler, 'Vehicle Activated Signs a Large Scale Evaluation' TRL Report TRL 548, TRL Limited, UK, 2002.
- [6] J. Mattox, W. Sarasua, J. Ogle, R. Eckenrode, and A. Dunning, 'Development and Evaluation of a Speed-Activated Sign to Reduce Speeds in Work Zones', Transportation Research, Record of the National Academies, Washington, 2007, pp. 3-11.
- [7] L. K. Walter and J. Knowles, 'Effectiveness of Speed Indicator Devices on reducing vehicle speeds in London', Transport Research Laboratory TRL 314, Transport for London, London Road Safety Unit, 2008.
- [8] U. A. Kumar and Y. Dhamija, 'Comparative Analysis of SOM Neural Network with K-means Clustering Algorithm', 2010 IEEE International Conference on Management of Innovation and Technology (ICMIT), Singapore, June 2010, pp .55-59.
- [9] W. Huai-bin, Y. Hong-liang, X. Zhi-jian, and Y. Zheng, 'A Clustering Algorithm Use SOM and K-Means in Intrusion

Detection', 2010 International Conference on E-Business and E-Government (ICEE), Guangzhou, May 2010, pp. 1281 - 1284.

- [10] X. Wang, W. Cottrell, and S. Mu, 'Using K-Means Clustering to Identify Time-of-Day Break Points for Traffic Signal Timing Plans', Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems, Vienna, Austria, September 2005, pp. 13-16.
- [11] J. Vesanto, J. and E. Alhoniemi, E., 'Clustering of the Self-Organizing Map', IEEE Transactions on Neural Networks, vol. 11, no. 3, May 2000, pp. 586-600.
- [12] Y. Chen, J. Hu, Y. Zhang, and X. Li, 'Traffic Data Analysis Using Kernel PCA and Self-Organizing Map', Proceedings of the 2006 IEEE on Intelligent Vehicles Symposium, Tokyo, Japan, June. 2006, pp. 472 – 477.
- [13] M. J. Watts and S. P. Worner, 'Estimating the risk of insect species invasion: Kohonen self-organising maps versus k-

means clustering', Ecological Modelling, Vol. 220, Issue 6, 2009, pp. 821-829.

- [14] S. K. Shukla, S. Rungta, S., and L. K. Sharma, 'Self-Organizing Map based Clustering Approach for Trajectory Data', International Journal of Computer Trends and Technology, Vol. 3, No.3, 2012, pp. 321-324.
- [15] T. V. Mathew, 'Fundamental Relations of Traffic Flow. Traffic Engineering and Management', IIT Bombay, 2012.
- [16] G. J. Kerns, 'Introduction to Probability and Statistics Using R', First Edition, 2011.
- [17] Hauser, T.A. & Scherer, W.T., 'Data Mining Tools for Real-Time Traffic Signal Decision Support and Maintenance', Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Tucson, AZ, Vol. 3, 2001, pp. 1471-1477.