VARIABLE SIZED PLANAR SLIDING WINDOW TECHNIQUE FOR 
SEARCHING ACCIDENT HOT SPOTS

Sándor Szénási, PhD
1 szenasis@ujs.sk, Universita J. Selye - Slovakia

ABSTRACT

There are several already known methods to find accident hot spots (locations in the national public road network, where the number of accidents is higher than expected based on the average values). One of the most frequently used of these is the sliding window method which is applicable for one-dimensional and two-dimensional spaces too. A particular variant of the one-dimensional method uses a variable sized sliding window, where the length of the investigated interval is not fixed. This paper presents the evaluation of the two-dimensional extension of this method working with planar (GPS) coordinates.

Keywords: GPS, accidents, hot spots, sliding window, planar, variable length

INTRODUCTION

Figure 1 presents the basic annual crash data in Hungary. As is visible from the diagram, the number of accidents happened during the past five years on the Hungarian public road system (the cumulative length of the network is more than 160,000 km). Between 2009 and 2013, a total of 80,863 traffic accidents with personal injuries happened. 70% of all these accidents occurred in built-up areas. The majority of the accidents happened in towns, among which the total number of accidents in the five years happening in Budapest was 15,800.

Figure 1: The number of accidents occurred on the Hungarian public road system in the given years (2009-2013).
The definition of hotspots (also known as black spots) is the following: places in the national public road network, where the accident density is higher than the expected value. The official definition is based on local statutory rules. For example, the locally accepted definition in Hungary (according to a Governmental Order) is: “those sections of the public road network, on which the frequency of accidents with personal injury is over the national average compared to the volume of traffic concerned”. But, the government recommendations vary depending on countries.

Identification and elimination of road accident hotspots are one of the most important road safety procedures. In a broad sense, the black-spot search process contains several steps: data collection, data cleaning, data filtering, localization of suspicious areas, evaluation of these regions, prediction of future conditions, prediction of the effects of possible actions, making a decision that a black spot candidate is a real black spot or not, monitoring the already identified black spots, etc.

According to this, we are trying to find areas of the road network where the accident density is higher than usual. This paper deals only with the localization and evaluation parts especially in the case of accidents in built-up areas. The related work section presents some already existing methods including the “planar sliding method” previously published by the authors; the next section contains the details of a possible improvement named “planar sliding method using variable window size”; the evaluation section shows the test results; and the last section contains the overall experiences and recommendations.

RELATED WORK

There are several well-known procedures [1], [2] to find the accident mentioned above black spot locations, and all of these have several advantages and disadvantages. There also are several theorems [3]–[5] and methods based on these to locate the problematic areas. These techniques can work together, and several articles try to adopt these to particular circumstances [6]–[9]. For example, in Hungary (according to the given official black spot definition) the recommended technique is the traditional “sliding-window method”.

Public roads constitute a large network (which can be represented by a graph data structure) essential elements of which are junctions and road sections (between two junctions). It is also worth differentiating locations within and outside of built-up areas because both the road sections and junctions differ in their traffic technological design and geometric characteristics within these. Within built-up areas, the location identification is usually based on street names and house numbers (although, some large cities are using alternative location identification systems). In the meantime, out of built-up areas, a particular place can be identified by using road numbers and sections (kilometre+metre). Obviously, this duality affects the identification of the locations of traffic accidents.

The traditional sliding window technique is well usable in the case of road number and road section based identification.
The main steps of this method are:

- Filtering the accidents of the given road.
- Sort the accidents using the section number.
- Slide the window over the accidents.
- Collect the locations where the number of accidents under the window is more than a given threshold.

With this method, the sliding window technique calculates the accident density using constant segment lengths. There are several variants of the presented method, for example:

- Using variable window length instead of a regular size.
- Using discrete or continuous window movement.

There are several fine-tuning additional methods in the literature, like future prediction systems (sometimes, it is not enough to analyse the already happened accidents, it is more appropriate giving estimations about the potential future situations and examine these results). It is also important to use some additional filtering options because the product of the sliding window technique is just a set of black spot candidates, these needs further and deeper investigations to decide that the results are real black spots or not. There are many books and papers [10] about this technique, and several years of experience proves that it is well usable in practice.

The technique mentioned above executes the search in a one-dimensional problem space. As a substantial constraint, all accidents of the examination must be on the same road only the road sections (kilometre and metre sections) distinguish the actual positions of the given accidents. But as is visible in Figure 2, a high number of accidents occurs in intersections, and we have to deal with these, too.

![Figure 2: Accidents with personal injuries within settlements in the given years (2009-2013).](image-url)
The basic sliding window technique is not applicable to find black spots located in junctions, because these accidents are usually close to each other, but dispersed to several roads.

In the case of built-up areas, the situation is even worst, because we cannot use road numbers, the accident identification is based on street names and house numbers. It is common in the case of large junctions that accidents of five or more roads belong to the same black spot cluster.

**METHODOLOGY**

Spreading of GPS (Global Positioning System) technology in the road accident scene investigator work leads to the extensive use of WGS84 (World Geodetic System) coordinates. Nowadays, it is common to use both the road number + section and GPS coordinate based positioning systems to localize an accident. Consequently, road safety analysts can use either data according to the chosen processing algorithm. In the case of the traditional sliding window, the regular road number + section pair is more appropriate. But if they would like to analyse accidents near junctions, the usage of GPS coordinates can be more efficient.

We have already developed a novel planar sliding window method [11]. This has the advantage that it can use a two-dimensional window instead of the traditional one-dimensional one. The current paper presents an improved variant of this method, using variable window length.

Based on the above, the input of this new algorithm is the following:

- accidents (locations are identified by GPS coordinates)
- window size constraints
  - minimum/maximum width (metre)
  - minimum/maximum height (metre)
  - step lengths in both directions
- minimum density (minimal number of accidents/area of the window)

There are major differences between the parameters of the planar and the original method. Firstly, accident locations are identified by GPS coordinates instead of road number and section pairs. The window itself is similar, but in this case, it is two-dimensional. Therefore the algorithm needs two parameters: window width and window height. In the particular case of variable sized window approach, it is necessary to specify the minimal and maximal values for window height and width.
The meaning of the minimum density limit is also different in this case. The dimension of the accident density was accident/metre for the traditional sliding window method. The planar sliding window method uses two-dimensional windows; according to this, the dimension of the crash density is accident/metre$^2$. This new approach needs some further investigations to find the best parameter values [12].

The pseudo code of the new algorithm is shown in Figure 3. The algorithm is self-describing and very similar to the already published non-variable sized window planar sliding window method.

**EVALUATION OF THE NEW METHOD**

It is hard to make a decision about the usability of a new method, because there is no adequate definition of hot spots. Therefore, it is difficult to check that the result (the collection of black spot candidates) is valid or not. However, there are some ways to verify the results of a new method. A widely accepted procedure is the calculation of the following relative values (presented by Chang, Washington, and Montella [6]).
Chang and Montella suggest to check the followings to test the consistency of a method:

- Site consistency: if a site is considered as a hot spot in time period A, it should also contain a high number of accidents in a further time period B.
  
  \[ SC = \frac{\sum \text{Accidents in } i^{th} \text{ window of period } A \text{ during period } B}{\sum \sqrt{\text{window height} \times \text{window width}}} \]

- Method consistency: if a site is considered as a hot spot in time period A, it should also be found as a hot spot in time period B.
  
  \[ MC = \frac{\text{Number of black spots identified both in } A \text{ and } B \text{ period}}{\text{Number of black spots identified only in } A \text{ or } B \text{ period}} \]

- Rank difference: if we run the black spot searching method in two different time periods (A and B), the results must be similar. We can also assume that the order of the black spots candidates will be similar too.
  
  \[ RD = \frac{\sum \text{Rank difference between hot spots in } A \text{ and } B \text{ period}}{\text{Number of hot spots}} \]

We ran several tests using the real-world accident database of Budapest. The chosen input parameters were:

- Window size: 0-150m x 0-150m (step: 10m)
- Minimum accident density: 0.0003 accidents/metre²

The values of the consistency variables mentioned above are (due to length limitations, we are not able to present the details of the tests):

- Site consistency: 0.0068 (accident/m)
- Method consistency: 0.8684
- Rank difference: 275.67

The following list contains the same values in the case of traditional fixed sized window length sliding window method:

- Site consistency: 0.0029 (accident/m)
- Method consistency: 0.8223
- Rank difference: 180.68

As it is visible, the rank difference is bigger in the original case, but the other two consistency values are significantly better in the case of the new method.

**CONCLUSIONS**

The primary purpose of the current project was to develop and implement a new accident black spot searching method using GPS coordinates. We have already developed a novel method using planar sliding window technique. This paper presents an extension of this, the planar sliding window method using variable window size.
The evaluation of the process shows that it performs well and applicable for practical use. The consistency values used for evaluation are very promising compared to other methods. The main strength of the algorithm is the potential of using it inside built-up areas.

ACKNOWLEDGMENT

This publication is the partial result of the Research & Development Operational Programme for the project "Modernisation and Improvement of Technical Infrastructure for Research and Development of J. Selye University in the Fields of Nanotechnology and Intelligent Space", ITMS 26210120042, co-funded by the European Regional Development Fund.

REFERENCES


[10] A. Bogárdi-Mészöly, A. Rövid, H. Ishikawa, S. Yokoyama, and Z. Vámossy,
