Data-parallel Implementation of Accident Black Spot Searching Method

Sándor Szénási*, Domokos Jankó**
* J. Selye University, Faculty of Economics, Komárno, Slovakia
** Road Safety Engineering Bureau, Budapest, Hungary
szenasis@ujs.sk, roadsafety@chello.hu

***Abstract*— Identification of accident hot spots of the public roadway (also called accident black spots) is one of the main tasks of road safety experts to avoid further traffic accidents and personal injuries. There are several available methods for this purpose, and one of the most promising of them is based on the GPS (Global Positioning System) coordinates of accidents and uses a well-known data-mining approach called DBSCAN (Density-based Spatial Clustering of Applications with Noise). This method is well parallelizable. Therefore, we can run multiple independent searches from different starting points of the search space. This paper presents a graphics accelerator based implementation of the original sequential algorithm to decrease the processing runtime.**

I. Introduction

The identification of public road accident hot spots (also known as black spots) is a specialty field of road-safety. There are various definitions for "hot spots" which will not be discussed in this paper, but most of the official definitions are based on accident density or accident rate in a given section of the public road network. For example, the officially accepted definition in Hungary (contained in the Governmental Order 176/2011 VIII. 31.) for hot spots 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". Identification and elimination of hot spots need several steps; one of the first steps (and this paper is rather focused on this) is the process in which data is gathered regarding the localization and identification of the accident spots.

To introduce the basic crash data in Hungary, Figure 1 outlines the number of accidents which happened during the past five years on the Hungarian public road system (~160,000 km). Between 2009 and 2013, a total of 80,863 road accidents with personal injuries took place in Hungary. 70% of all these happened in built-up areas. The majority of the accidents occurred in towns, among which the total number of accidents during the five years, happening in Budapest, was 15,800 (that is 28% of all accidents within settlements).

It cannot be exactly stated what percentage of all accidents occurred in black spots, however, according to estimations, this value would be between 5% and 10%. This means that by identifying the actual black spots and introducing appropriate counter-actions, the road safety situation may be significantly improved.

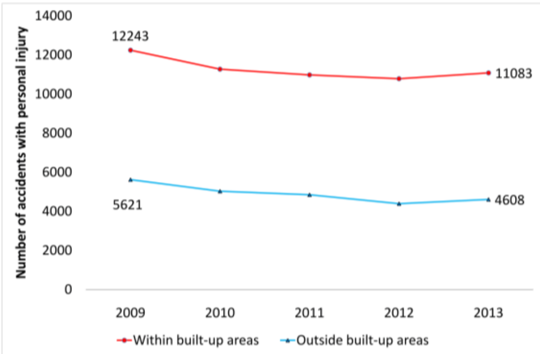

Figure 1. The number of accidents that happened on the Hungarian public road system in the given years

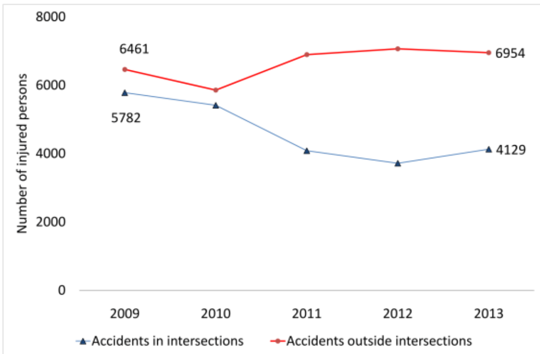

Figure 2. Injured persons within the settlements in the given years

The domestic public roads operating bodies (according to legislation) regularly carry out black spot discovery on their road systems as part of their attempt to improve road safety. The expert working material [1] set up by the working committee of the Hungarian Road Society produces manuals of that activity. However, those mainly deal with the state public roads which do not touch upon roads within settlements under municipality operation. Based on the said instruction, black spots are searched for and discovered following a stipulated criteria, which differs between cases of road sections within and outside built-up areas (numbered road sections) [2], [3]. As per the above: "Within built-up areas: a public road junction or an at least 100 meters long section of the same is considered as a suspected black spot provided that at least four accidents with personal injury took place in the subject section within a three-year period". For the identification of the accumulation points of accidents, the directive recommends the sliding window technique [4], [5] to be used: aggregates of points appearing in windows (of different widths) drawn out over the accident spots on the road section in question will show the accident densification points.

In various countries, the official definition of hotspots varies [5]–[10], therefore, we do not strictly adhere to following this rule. With our method, we are trying to find areas where the accident density is higher than a given threshold value. In a broad sense, the black-spot search procedure 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.

There are several techniques working together and several articles trying to adopt these to particular circumstances [11]–[15]. In this paper, we are dealing only with the localization and evaluation parts (and marginally with the data collection) especially in the case of accidents in the built-up areas. The rest of the paper is structured as follows: Section 2 contains the related work about black spot searching methods, then Section 3 presents the idea of the DBSCAN data-mining algorithm. Section 4 contains the adaptation of the DBSCAN method to find accident black spots identified by GPS coordinates. Section 5 shows the parallelization method and the next section contains some real-world examples from Budapest. Finally, the last section contains the conclusions and the further development plans.

II. RELATED WORK

Public roads constitute a network, the core elements of which are road sections and junctions. Both the road sections and junctions differ in their geometric characteristics and traffic technological design within and outside of built-up areas. The latter means the road network within built-up areas, which consists of named streets (although, numbered public roads can also be found there, which however are named ones again [16]). Out of the built-up areas lead the numbered roads, where "kilometre + metre" sections are allocated. Therefore, identifying the place of an accident will also be different for the above.

The place of an accident on road sections within a built-up area will be identified by the "street name + house number", while in junctions by "street name + street name". In principle, (together with the denomination of the district concerned in case of larger settlements) this is the allocation which identifies the spot of an accident. Unfortunately, however, several problems occur in practice making the spot identification process uncertain.

Traditional black spot searching techniques within built-up areas (not in numbered public roads but the network of named streets) cannot be used efficiently for the following reasons:

- While searching for black spots, the distance of accident spots from one another has particular importance, which is rather more complicated to calculate within a built-up area. Traditionally, in the case of two accidents on the same road, the difference between the two road sections is considered as the distance. Street numbers, however (especially when it comes to different streets) do not give such an evident and fast result.
- A significant amount of accidents within built-up areas can be linked to junctions. In such cases, depending on the exact spot of the crash, it differs to which street the accident will be allocated by the police officer. There may be up to four or five different street names appearing at the same junction (even considering the case where the junction concerned is also registered as a named square in the system). Thereby, the accidents will appear in the database as spread among all these. Traditional black spot searching methods examine one road only, providing no choice to handle these accidents as having had occurred in the one location.
- Another disadvantage of built-up areas is the lack of reliable and complete databases. For example in Hungary, numbered roads are maintained by a central organization, but streets of towns are by the local government of the settlement. In the latter case, there are no comprehensive street-level traffic databases, GIS systems containing road geometry [17] etc. (except some larger cities, but the format of these is not unified. Therefore, we cannot use these for country-wide analysis).

Unfortunately, today's best black spot search methods are based on multiple databases, but the lack of information made these unusable in these cases. However, there are thousands of accidents in these settlements; therefore, we should focus on there.

The aim of our project was to create a methodology for black spot searching in this heterogeneous environment. Fortunately, the accident database is reliable and complete because it is maintained by the police, using a country-wide, uniform format. The usage of GPS coordinates eliminates the issues caused by the different place identification methods. Based on these, we have developed an algorithm, which can search black spots in built-up areas, based on these GPS coordinates and some major accident data.

Road safety projects involves several steps:

- collection of data
- cleaning of data
- localization of suspicious areas
- evaluation of these locations
- prediction of future accident count
- prediction of the effects of optional actions

Computer programs can help to find black spot candidates, but nowadays, these methods also need some human assistance. It is quite hard to decide that if a given blackspot candidate a real black-spot or not. This requires a deeper analysis of the details of the accidents. Experts can acquire this necessary data with a targeted data collection process, in the deeper investigation phase of the black spot evaluation process. This is one of the reasons why we emphasize that the results of our method are only black spot candidates, experts need some additional manual work to give a clear decision about the eventual and actual safety level.

Again, there are several black spot definitions, but this paper does not strictly adhere to any of these. We are trying to find areas of the public road network where the density of the accidents (number of accidents / area) is higher than a given threshold value.

III. GPS LOCATION BASED DATA MINING

GPS-based location identification has been introduced for several years, and this offers many benefits compared to the traditional "road number and kilometre section" based location identification system. The first obvious benefit of these is the increase in accuracy and reliability, given that accident investigators get data accurate to the metre within seconds, not including the errors of the measurements that existed before. This is critical in the case of black spot seeking, given that measurement errors of a few hundred metres and other uncertainties in location identification can lead to an entirely different result.

Another benefit is that an accident that takes place at the same geographic location but on different road sections (for example, different directions of intersections and divided lanes, and rebuilding of roads) will be close to each other based on the coordinates specified by the global positioning system. Therefore, there will be no need for resource-intensive corrections for seeking accident hot spots. Obviously, this aspect has to be taken into account in the analyses, too.

Traditional algorithms using fixed and mobile sectioning (for example, the well-known sliding window method) are hard to use for the data produced this way. Using GPS positions, the exact location of the accident is specified using a pair of numbers instead of a section number. The existing algorithms can be adapted to this, for example, the section numbers can be replaced by a simple two-dimensional grid, but the results produced by this are less valuable, and the implementation of the grid structure using variable window size is a very resource-intensive task. Moreover, it is capable of recognising only rectangular black spots.

It is worth taking a look at some clustering techniques already used in the field of data mining. The main principle of density-based search methods is quite simple: the density of elements within a cluster is much higher than between groups, that is how clusters and outliers can be identified. These methods are general; to calculate density, a distance concept needs to be defined. Fortunately, in the accident black spots, this is very simple.

One of the most basic and efficient density-based methods is the DBSCAN (Density-Based Clustering of Application with Noise) algorithm [18], [19]. This method gradually increases sufficiently dense areas into clusters and defines them as a domain of points with dense connections.

A huge benefit of the DBSCAN algorithm is that it is capable of recognising clusters of any shape in geographic databases, and can be used effenciently in cases of significant noise (in the field of road accidents, an element is called noise if it does not belong to any cluster, which very much applies to accident data, as accidents data includes all accidents that are not found in one black spot).

The proposed DBSCAN algorithm uses two input parameters [20]:

- ε, which is a radius-type value
- *MinPts*, which is the limit for element numbers

The space within a radius of ε of an element is called the *ε-environment* of this element. If the ε-environment contains at least *MinPts* items, this is known as an *internal element*. For a given domain of elements, one element is *directly densely accessible* from other internal elements if it is the first element's ε-environment. The definition of *dense reachability* is similar, only here, it is permitted for one element to be accessible from another only through a chain of directly densely accessible elements. Two elements are *densely connected* if there is an element from which both are densely reachable with the given parameters. The *density-based cluster* is a domain of densely connected items that shows maximum accessibility of density.

Based on these definitions, our goal is to find domains of accidents in the public road network in which all elements are densely connected based on the specified parameters and no further expansion is possible.

IV. DBSCAN ALGORITHM FOR BLACK SPOT SEEKING

As usual, the general data-mining algorithm needs to be adapted to the specific task, in this case, to find accident black spots. First, the concept of "distance" needs to be defined. This can be done in two ways:

- The traditional approach is that if the accidents are on the same road, this can be considered to be the difference between kilometre section numbers. However, in practice, this raises several problems: a) how to calculate the difference of two accidents on a road that has been rebuilt during the time between the two accidents and sectioning changed b) this measurement is not usable in the case of two accidents on different roads.
- If we use the GPS-based location identification, the distance is offered naturally. We can use the simple Euclidean distance between the two geographic coordinates.

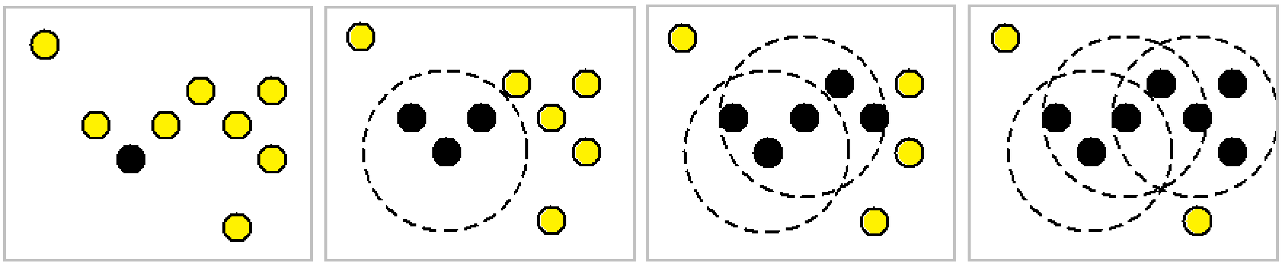

Figure 3. The main steps of the DBSCAN method a) select first item b) investigate ε environment c) extend the black spot candidate if possible d) final phase of the algorithm

Both distance definitions can lead to an acceptable solution; it depends on the application that is the most usable in the given particular case.

Using the presented DBSCAN algorithm, we can calculate the *transitive closed domain* of directly dense accessibilities (the maximum domain of densely accessible accidents from the starting location) for all points of the search space. This can be done by selecting one of the points of the space as the starting point and investigating its ε-radius environment. If there are not at least MinPts pieces of accidents in this area, we can step to the next starting accident.

If there are at least one MinPts pieces of accidents in this area, the point is considered as an internal point, so that a new cluster is created using this as a starting point. In this case, the directly densely accessible accidents from this location are collected and included in the cluster, after which the search towards the new elements continues recursively (Figure 3).

The ε-radius environment of these new accidents is also investigated, and if there are accidents that meet our conditions, these are added to the cluster, and the search continues. If there are not any new elements, the recursion is finished, and if possible, we start it again from a new starting accident. We can launch the search from any locations of the search space; however, it is reasonable to select only the position of each accident of the road network as a starting point. Although this does not necessarily provide the best solution it gives an acceptable approximation.

In contrast with the original DBSCAN algorithm, in this case, the goal is not to obtain extensive clusters; we need the group of only those accidents that "really" belong to the same hot spot. Therefore, we have extended the search method with a *minimum density limit* (MinDns).

To do that, we have to define the area of a given cluster (which can be done based on the points that make up the cluster), and then an additional minimum limit can be introduced. According to this, only those new accidents are included in the black spot that do not reduce the density of the cluster to below the given threshold. It is worthwhile to take into account the outcome of the accident as well as if it is based on the possibility to define the weight of a black spot candidate (the sum of the weights of the accidents in the black spots [21], [22]) and the black spot density (the ratio of the weight and area).

This extension helps avoid accident hot spots, that are close to each other, from melting into a single large cluster. Figure 4 shows an example in which we can see two different black spots; however, if we strictly follow the steps of the original DBSCAN algorithm using the ε-distances, the algorithm would merge these because they are densely connected.

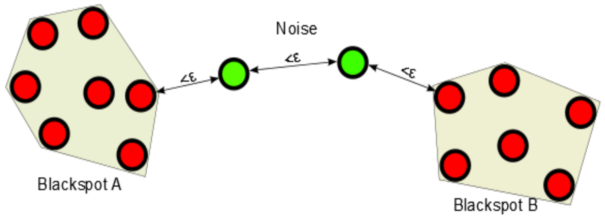

Figure 4. Two different black spots connected by densely accessible accidents

For similar reasons, in contrast to the original DBSCAN method, we do not make combinations when expanding black spots. We need black spots that are close to but different from each other not to merge into a big cluster. Executing the presented algorithm from each accident of the road network will produce overlapping black spot candidates (for example, starting the algorithm from the accidents belonging to the same black spot, all searches will return the same black spot candidate).

Therefore, an additional processing step is required to eliminate this redundancy. The principle of this method is that it sorts all black spot candidates by weight, after which black spots with decreasing weights are investigated gradually starting from the largest one, and only those that have no overlap with another black spot with a greater weight will be deemed to be valuable in the final result.

V. Data-parallel DBSCAN

As it is visible, we have to launch more than one cluster search from different starting accidents, and these should not interfere with each other. Because the order of these investigations does not affect the final result, it is possible to run these searches in a parallel fashion. In the case of multi-core central processing units (CPU), we can achieve linear speed-up if we use more than one processor core [23], [24]. There is not any communication demand between these threads; therefore, we can fully utilize the processing resources of all processor cores without unnecessary synchronizations.

It is worthwhile to implement the whole algorithm in a graphics processing unit (GPU) too. All processing threads have to execute the same algorithm using different data sources; this is the main attribute of the problem class, called "embarrassingly parallelizable" problems. In the case of graphics accelerators, there are thousands of processing units integrated into the card; therefore, we need a high number of parallel thread executions to utilize the available resources fully. This process can fulfil this requirement because we can start the DBSCAN algorithm from each accident, and we usually have thousands of accidents in the search space.

GPU programming has some special requirements [11], [25], according to this, the GPU implementation is based on the following steps:

1. Load the necessary accident data (location, outcome, etc.) from the database to the main memory.
2. Move this dataset to the memory of the graphics accelerator.
3. Execute the DBSCAN kernel on the GPU using as many threads as the number of accidents. If the number of accidents is larger than the number of physical cores in the GPU, the graphics card will automatically schedule the execution.
4. Copy the final result (accident black spot candidates) back to the host memory.
5. Display the results (using a table, map, etc.).

The 3rd step uses the presented modified DBSCAN algorithm implemented in CUDA C language to make it capable of running on the GPU. In our first implementation, each thread runs separate DBSCAN iterations starting from a given accident of the public road network.

Figure 5 shows the results of the runtime tests. As is visible, the GPU was faster in all cases. The speed-up was greater in the case of a large number of threads. In the case of 1000 accidents (1000 parallel threads), the GPU was about 15 times faster.

VI. Real World Examples

This paper contains two real-world examples to show the basic functionality and the results of the introduced algorithm.

We have launched a search using the accidents in the 15th district of Budapest. The search covered the period from 2002 to 2014. It is worth noting that professionals recommend using a 5-year extended period for black spot searching, but in this case we chose a larger dataset to demonstrate the capabilities of the algorithm.

Input parameters had selected as the followings:

- ε= 50 metre
- MinPTS = 5 accident
- MinAREA = 100 metre^2
- MinSCR= 0.001 weighted accident/metre^2
- Weights for accidents:
    - Fatal accidents = 10
    - Serious accidents = 3
    - Light accidents = 1
    - Accidents with property damage: 0
- Weight factors for injured persons:
    - Killed = 0
    - Seriously injured = 0
    - Lightly injured = 0

The result of the algorithm is a list of black spot candidates. This paper presents two of them (the primary criterion for the selection was to demonstrate the capabilities of the algorithm).

A. *Example 1 - Szentmihályi Street, Budapest*

The first case study shows accidents located on the same street. Table I indicates that all the accidents of the black spot are on Szentmihályi Street. House numbers identify the correct location; however it is very hard to calculate distance based on these (it makes it more difficult that the accidents are located on both sides of the street). As is visible in Figure 6, the usage of GPS coordinates solved this problem. The black spot is visible and well-defined.

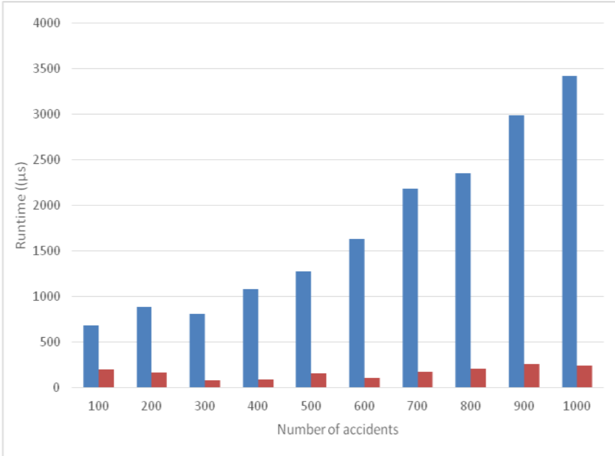

Figure 5. CPU (blue) and GPU (red) runtimes

Table I. Accidents on Szentmihályi Street, 15th district of Budapest (2002-2013). The table shows the time and the location of accidents (GPS position and house number), the severity of the accidents (F=fatal accident, S=with serious injuries, L=with slight injuries), and the number of killed/seriously injured/slightly injured persons.

#	Acc. time	Street num	A.S.	Killed	Se.I.	Sl.I.	GPS LAT	GPS LON
1	5/29/2011	12	L	0	0	1	47°33'32.99	19°07'19.39
2	12/29/2013	7	S	0	1	0	47°33'34.06	19°07'19.96
3	12/8/2011	2	L	0	0	1	47°33'34.75	19°07'18.77
4	10/30/2013	8	L	0	0	1	47°33'33.59	19°07'18.38
5	12/15/2005	6	S	0	1	3	47°33'33.85	19°07'17.92
6	1/24/2005	12	L	0	0	1	47°33'32.99	19°07'19.39
7	1/31/2002	6	F	1	0	0	47°33'33.85	19°07'17.92

This black spot consists of 7 accidents. As is visible in Figure 6, these are almost positioned in line and close to each other. The black spot is near to a high traffic intersection, where two one-way streets join into a two-way street. The density of the black spot is quite high, 0.0195 weighted accident/metre^2, which is significantly greater than the given threshold. After the detailed investigation of the accidents, we accepted this result as a real black spot.

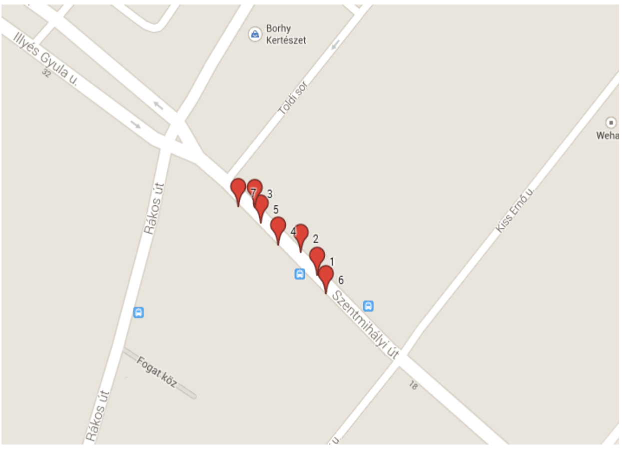

Figure 6. Detected black spot on Szentmihályi Street, Budapest (viewed using Google Maps, 2014)

This example shows that the algorithm can find black spots in built-up areas on a single street. The results are quite similar as in the case of traditional black spot search methods. The main difference is the location identification basis: instead of road number and sections, we have a street name and house numbers (and the corresponding GPS coordinates).

B. *Example 2 - Hubay Jenő Square, Budapest*

Figure 7 shows another example. It is a black spot located at the intersection of five streets.

In the next example, we have used the same searching parameters (it is just another example from the result set). As it is visible, the algorithm can find black spots near road intersections. Table II shows the detailed data of the accidents.

Table II. Accidents of a black spot in the 15th district of Budapest (2002-2013). The table shows the time and location of the accidents (GPS position and street name), the severity of the accidents (S=with serious injuries, L=with slight injuries), and the number of killed/seriously injured/slightly injured persons.

#	Acc. time	Street name	A.S.	Se.I.	Sl.I.	GPS LAT	GPS LON
1	10/3/2011	Deák	L	0	1	47°33'47.28	19°06'56.90
2	1/19/2012	Illyés Gy.	L	0	1	47°33'43.34	19°06'58.87
3	1/6/2006	Hubay J.	S	1	2	47°33'45.89	19°06'55.90
4	10/16/2008	Eötvös	S	1	0	47°33'44.02	19°06'55.70
5	5/23/2009	Eötvös	L	0	1	47°33'43.93	19°06'57.01
6	4/21/2004	Eötvös	L	0	1	47°33'44.02	19°06'55.70
7	12/28/2008	Bácska	L	0	1	47°33'44.76	19°06'54.34
8	3/17/2007	Bácska	L	0	1	47°33'44.76	19°06'54.34
9	11/6/2005	Eötvös	L	0	1	47°33'44.02	19°06'55.70
10	7/23/2006	Deák	L	0	1	47°33'46.47	19°06'57.18

The black spot consists of accidents located on the roads near Hubay Jenő Square. These streets are: “Árpád Street”, “Illyés Gyula Street”, “Deák Street”. In the case of road safety investigations, accidents close enough (usually 50-100 metres) to intersections are considered as belonging to the intersection. Table II shows that police officers have saved these accidents to the database using different street names. This makes the identification of the black spot very hard using the traditional methods. The new method based on the GPS coordinates clarifies the situation.

The area of the black spot is 5,120 metre^2. It is the area of the polygon spanned by accidents 1, 2, 4, and 8. The accumulated weight of accidents is 14. The ratio of these is the accident density: 0.0027 weighted accident/metre^2.

This is acceptable as a black spot candidate too. The accident density is higher than expected. The map in Figure 7 shows the location of the accidents, and the table shows the detailed information about them.

Figure 7. Detected black spot near Hubay Jenő square, Budapest (viewed using Google Maps, 2014).

VII. Conclusions

The main objective of this project was to adopt new data analysis methods in addition to the related developments over the last decades: spreading of GPS technology, the appearance of new computer architectures (GPGPU - General-purpose computing on graphics processing units), data-mining and the existence of the critical amount of accident data.

We have developed various methods, regulations and computer applications [26], [27] to integrate these advancements into the daily work of road safety experts. The necessary tools and applications have been implemented and tested.

The final step of this project is to speed-up the processing methods. We have adapted the proposed DBSCAN method to run on graphics processing units. We have implemented the algorithm using the CUDA C programming language.

As the real-world examples show, the method is effectively-usable in practice, it can identify black spot candidates based on the GPS coordinates of accidents. The GPU implementation of the algorithm is significantly faster the original CPU one while the result of this is the same as the original one. Therefore, we can use this implementation in the future without any drawbacks.

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

[1] E. Hóz, “Csomópontok és útvonalak balesetveszélyességi értékelési módszertanának kidolgozása,” 2005.

[2] A. S. and A. M.V.L.R., “Safety analysis of urban signalized intersections under mixed traffic,” *J. Safety Res.*, vol. 52, no. 0, pp. 9–14, 2015.

[3] Y. Zhang, Y. Xie, and L. Li, “Crash frequency analysis of different types of urban roadway segments using generalized additive model,” *J. Safety Res.*, vol. 43, no. 2, pp. 107–114, 2012.

[4] S. Lee and Y. Lee, "Calculation Method for Sliding-window Length: A Traffic Accident Frequency Case Study," in *Easter Asia Society for Trasportation Studies*, 2013, vol. 9, pp. 1–13.

[5] R. Elvik, "A survey of operational definitions of hazardous road locations in some European countries," *Accid. Anal. Prev.*, vol. 40, no. 6, pp. 1830–1835, 2008.

[6] R. Delorme and S. Lassarre, "A new theory of complexity for safety research. The case of the long-lasting gap in road safety outcomes between France and Great Britain," *Saf. Sci.*, vol. 70, no. 0, pp. 488–503, 2014.

[7] W. Murray, J. White, and S. Ison, "Work-related road safety: A case study of Roche Australia," *Saf. Sci.*, vol. 50, no. 1, pp. 129–137, 2012.

[8] E. Hóz, C. Koren, and T. Mocsári, "A strategic approach for safety: putting knowledge into practice," in *XXIVth World Road Congress*, 2011, pp. 1–13.

[9] C. Koren, P. Lányi, and T. Mocsári, "Implementation of the EU Directive on Road Infrastructure Safety Management in Hungary," in *Proc. 11th International Scientific Conference (MOBILITA11)*, 2011, pp. 164–170.

[10] A. Montella, D. Andreassen, A. P. Tarko, S. Turner, F. Mauriello, L. L. Imbriani, and M. A. Romero, "Crash Databases in Australasia, the European Union, and the United States," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2386, no. 1, pp. 128–136, 2013.

[11] "European Commission DG Energy and Transport High Level Group Road Safety Road Infrastructure Safety Management Report of the Working Group on Infrastructure Safety Status : Public Version of 19 December 2003," 2003.

[12] SweRoad, "Black Spot Manual." SweRoad, 2001.

[13] M. Sørensen and S. K. Pedersen, "Injury severity based black spot identification," in *Annual Transport Conference*, 2007, pp. 1–10.

[14] M. Sørensen, "Best Practice Guidelines on Black Spot Management and Safety Analysis of Road Networks," 2007.

[15] J. Han and M. Kamber, *Data Mining. Concepts and Techniques.* Elsevier, 2001.

[16] X. Qu and Q. Meng, "A note on hotspot identification for urban expressways," *Saf. Sci.*, vol. 66, no. 0, pp. 87–91, 2014.

[17] I. B. Gundogdu, "Applying linear analysis methods to GIS-supported procedures for preventing traffic accidents: Case study of Konya," *Saf. Sci.*, vol. 48, no. 6, pp. 763–769, 2010.

[18] B. Borah and D. K. Bhattacharyya, "An improved sampling-based DBSCAN for large spatial databases," in *Intelligent Sensing and Information Processing, 2004. Proceedings of International Conference on*, 2004, pp. 92–96.

[19] S. Kisilevich, F. Mansmann, and D. Keim, "P-DBSCAN," in *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application - COM.Geo '10*, 2010, p. 1.

[20] S. Sergyán and L. Csink, "Automatic Parametrization of Region Finding Algorithms in Gray Images," in *4th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, 2007, pp. 199–202.

[21] S. Da Costa, X. Qu, and P. M. Parajuli, "A Crash Severity-Based Black Spot Identification Model," *J. Transp. Saf. Secur.*, vol. 7, no. 3, pp. 268–277, 2015.

[22] K. Geurts, G. Wets, T. Brijs, K. Vanhoof, and D. Karlis, "Ranking and selecting dangerous crash locations: Correcting for the number of passengers and Bayesian ranking plots," *J. Safety Res.*, vol. 37, no. 1, pp. 83–91, 2006.

[23] D. Arlia and M. Coppola, "Experiments in Parallel Clustering with DBSCAN," Springer Berlin Heidelberg, 2001, pp. 326–331.

[24] X. Xu, J. Jäger, and H.-P. Kriegel, "A Fast Parallel Clustering Algorithm for Large Spatial Databases," in *High Performance Data Mining*, Boston: Kluwer Academic Publishers, 1999, pp. 263–290.

[25] NVIDIA, "Whitepaper. NVIDIA's Next Generation CUDATM Compute Architecture: Kepler TM GK110." 2012.

[26] G. Kertesz and E. Hajnal, "Special issues in the development of a large user based web application," in *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2014, pp. 141–145.

[27] A. Bogárdi-Mészöly, A. Rövid, H. Ishikawa, S. Yokoyama, and Z. Vámossy, "Tag and Topic Recommendation Systems," *Acta Polytech. Hungarica*, vol. 10, no. 6, pp. 171–191, 2013.