LEVERAGING CROWD-SOURCED ROAD DEFECT INFORMATION FOR ROAD QUALITY ASSESSMENT

Dimitrios Efthymiou and
Constantinos Antoniou

National Technical
University of Athens

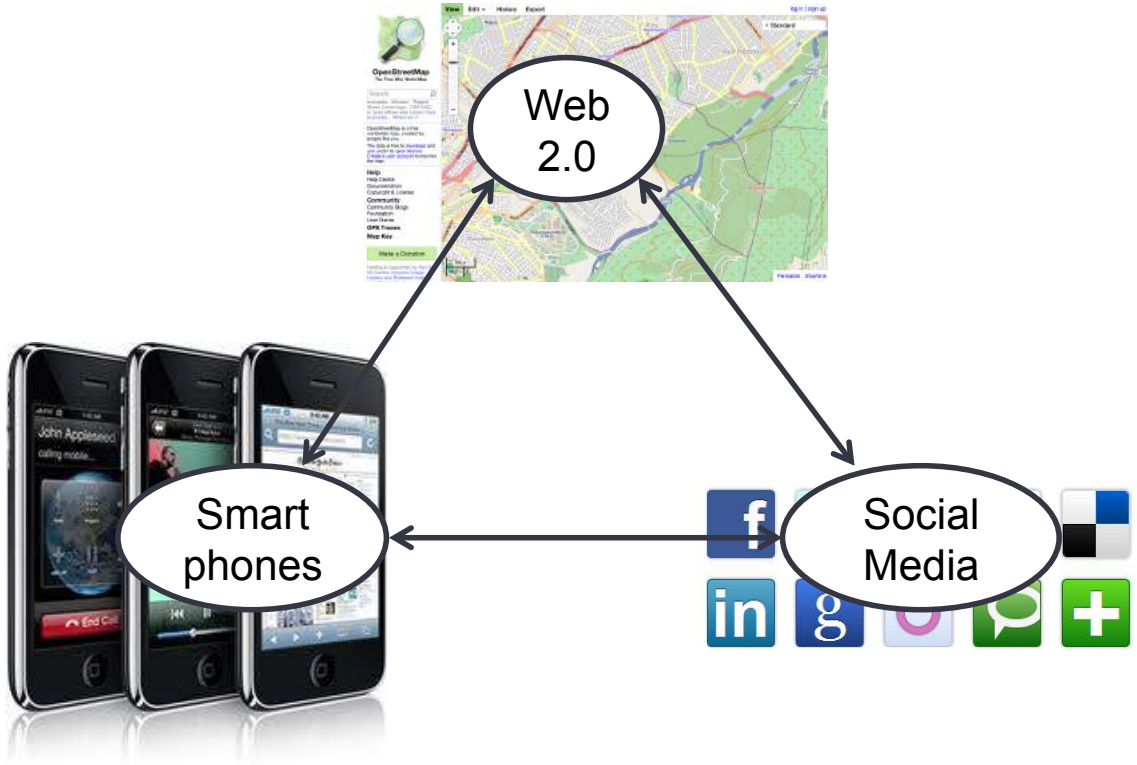
5th National Congress on Road Safety
Volos, 2012

Overview

- Background
- Methodological approaches
- Data sources and collection
- Results and analysis
- Future work

Crowdsourcing

- The general public has the chance to participate in data collection and planning.
- This intentional or not participation of the public in data collection and planning is defined as *crowdsourcing*



Crowdsourcing

- **Volunteers** perform the work that used to be made by professionals
- Well-known examples: *Wikipedia* (wikipedia.org), the free on-line maps *Open Street Maps* (openstreetmaps.org) but also any kind of Open Source software.
- Advantages: 1) **free** for use by the public, professionals and researchers, 2) easily, instantly **accessible**, 3) usually **up-to-date**, and 4) collected by the latest means of **technology**.
- The **quality** of the resulting dataset **varies** among the platforms, and depends on the amount of the participants, the administrators and the time of the data-collection process

Volunteered Geographic Information

- The term *Volunteered Geographic Information (Goodchild and Glennon, 2010)* is used to describe the voluntary public contribution for geographical data and information collection (OpenStreetMaps, OSM)

- *The term Volunteered Geographical Systems* (VGS) *(Savelyev et al., 2011)*, is used for the integration of social networks and VGI.

Goodchild M. and J. A. Glennon (2010). Crowdsourcing geographical information for disaster response: A research frontier. *International Journal of Digital Earth,* Vol. 3, No. 3, pp. 231–241

Savelyev A., Janowicz K., Thatcher J., Xu S., Mulligann C. and Luo W (2011). Volunteered geographic services: Developing a linked data driven location-based service.

The Role of Social Media

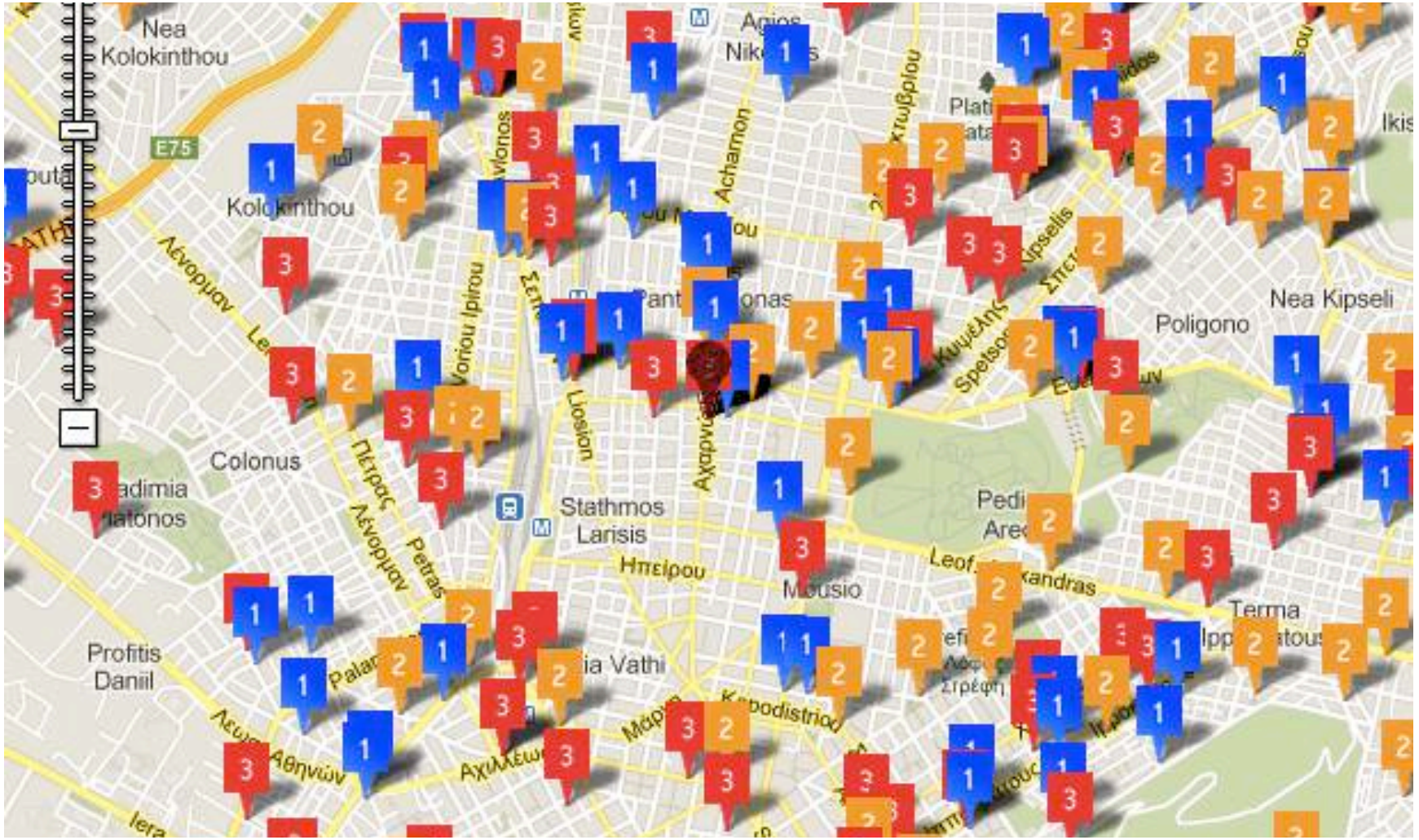
- Social Media are web 2.0 applications that allow their users to share their thoughts on-line.
- They are addressed to people with specific aims (e.g. www.academia.org for exchange of scientific knowledge, www.linkedin.org for job-finding or recruiting), or to the general public (e.g. plus.google.com, www.twitter.com, www.facebook.com and www.flickr.com).
- They have been used for different purposes: 1) political campaigns, 2) marketing, 3) news updates, 4) recruitment or even for 5) protests revolutions
- The multi-variation of the demographic characteristics of their users (different ages, nationality, education level etc.) renders these platforms, tools suitable for data collection. Social networks have been used for transportation-related applications

The Pin-Project

- www.msfree.gr/pin/
- Launched in 2007 by the Road Safety Institute Panos Mylonas (IOAS)
- Motivate the users of the Greek road network to report voluntarily any constructional defects
- Create a spatial database with the Greek streets defects.
- Push the policy makers to take action

- 650.000 visitors until May 2009
- 12.421 points had been "pinned" and
- According to the users, 7.468 could be fixed immediately.
- 1.081 of these points were characterized as dangerous for serious traffic accidents, while the majority of the pins (69%) were about the existence of dangerous potholes on the road surface (Danelli-Mylonas V, 2009. *Road Safety: Actions and perspectives*. In Swedish Trade Council Conference, R.S.I. "Panos Mylonas")

Pin-Project's Points on Map



The Illegal-Signs Project

- www.illegalsigns.gov.gr
- Volunteers report the existence of illegal marketing signs across the Greek street network.

- *Signs that attract out attention and are placed on roundabouts, side-walks ad points where could cause accidents, perpendicular to the traffic flow, are illegal, dangerous and should be dismantled (*Hellenic Highway Directives).

- *Death caused by an accident on these signs, is a murder by negligence* (The *Prosecutor of Athens*).

- At least 8 persons per year die in car accidents where an illegal sign was involved, while the one tenth of the car accidents is because of the crash on illegal signs

- The signs are first reported by volunteers (who indicate the exact location, attach a photo and make comments), and then are removed by the local authorities.

Data Collection

- Pin-Project (4128 data points)
  - Geo-location (X, Y coordinates)
  - Type of defect
  - Short description (text)

The PIN Project
ΣΕ ΣΥΝΕΡΓΑΣΙΑ ΜΕ ΤΟ ΙΟΑΣ

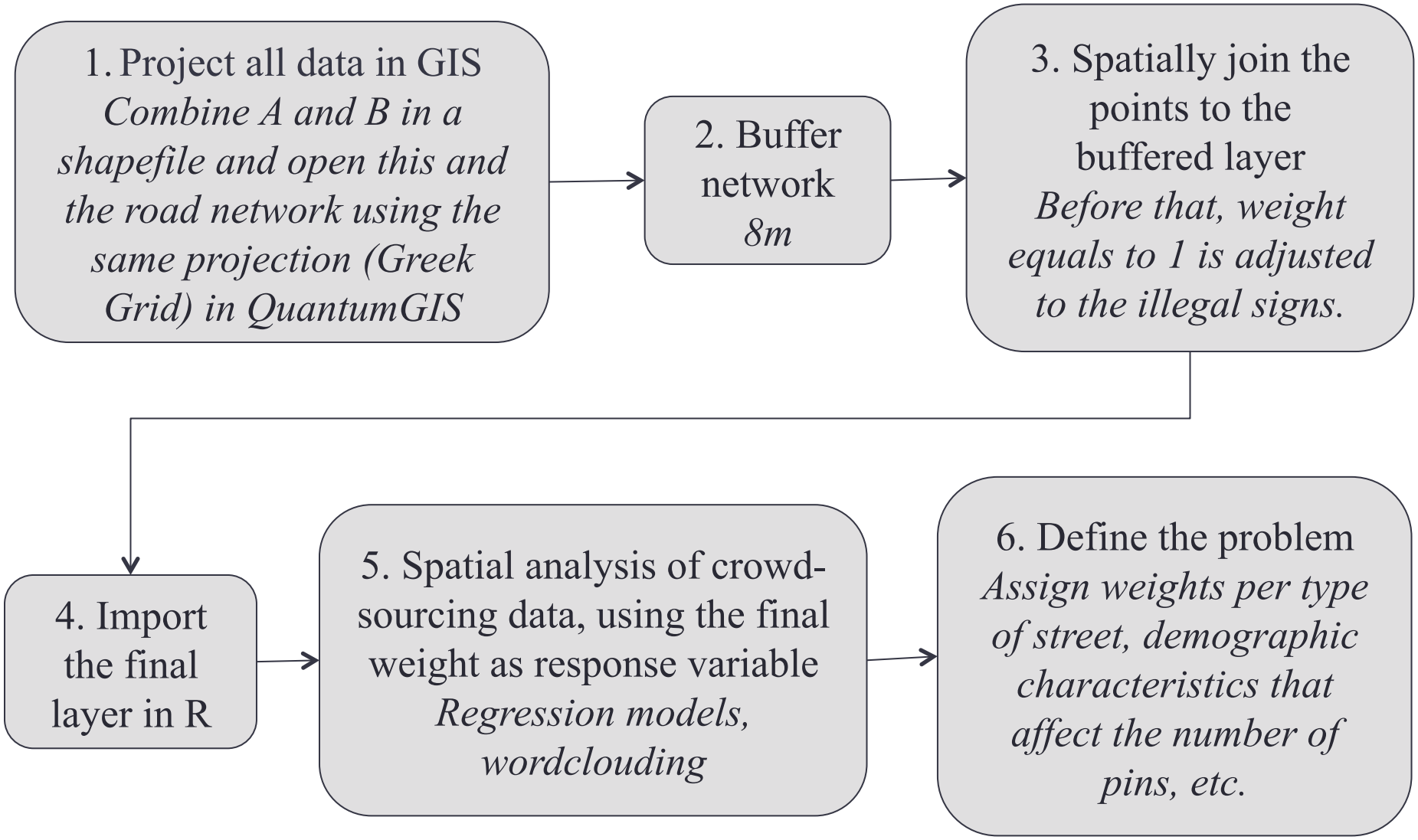
Χωρίς παράνομες διαφημιστικές πινακίδες
Αναδεικνύουμε το δημόσιο χώρο με τη βοήθειά σου

- Illegal signs project (762 data points)
- Open Street maps (street network)

Type of road	Length (m)	Links	Total score
Residential	174641	1957	4072
Motorway	35889	115	196
Primary	113407	721	1466
Secondary	100556	809	1612
Tertiary	75458	740	1613
Trunk	12282	60	202

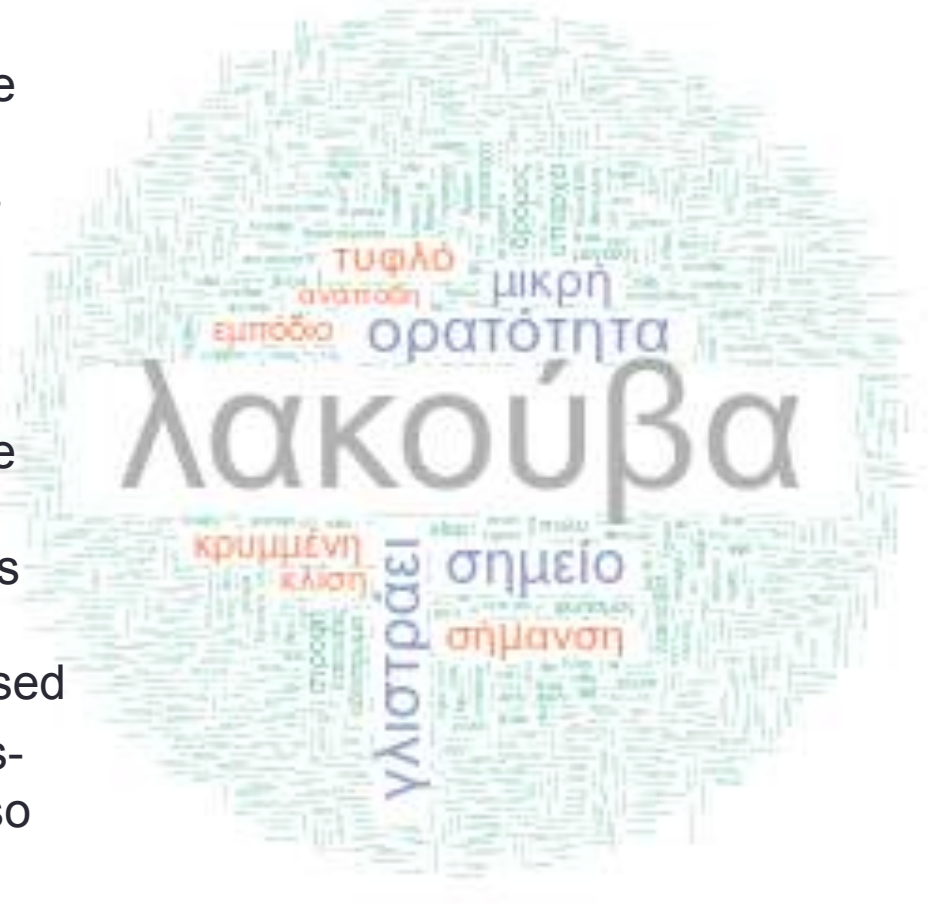
OpenStreetMap
The Free Wiki World Map

Methodology



Word-cloud

- A word-cloud was created using the descriptions of the road defects added by the volunteers in the Pin-Project
- Words of length less than 3 letters and numbers were not included
- The figure reveals quantitatively the weight of the words used by the volunteers to describe the problems in each case. "Pothole" (in Greek: "λακούβα") is the dominant word used
- Additional information from the less-frequently appearing words can also be used



Modeling the Quality

$$Y = \beta_{rs}X_{rs} + \beta_{pr}X_{pr} + \beta_{sc}X_{sc} + \beta_{tr}X_{tr} + \beta_{mt}X_{mt} + \ \ldots$$

Variable	Description	β	Standard error	t-value
βrs	Type: Residential	1.21	0.16	7.65
βmt	Type: Motorway	0.91	0.23	4.02
βpr	Type: Primary	1.18	0.17	7.00
βsc	Type: Secondary	1.18	0.16	7.37
βtr	Type: Tertiary	1.34	0.16	8.28
βtn	Type: Trunk	2.44	0.28	8.75
βedu	High education	1.79	0.35	5.17
βden	Low density (<500 persons/*km2)*	1.00E-05	4.00E-06	3.57
βlen	Logarithm of link length	0.10	0.029	3.42

Suggestions for Improvement

- **Integrate** Social Media (e.g. flickr) to Pin-project and Illegal signs projects to facilitate the uploading of photos of the defects
- Increase the **interaction** between the users
- Ask for **more data from the users**, such as demographics (e.g. age, sex, education level) and travel patterns (e.g. how often they use the car, which mode they use to go to work)
- **Integrate** all platforms to one
- Greater **convergence** with the National iRAP projects.

LEVERAGING CROWD-SOURCED ROAD DEFECT INFORMATION FOR ROAD QUALITY ASSESSMENT

Dimitrios Efthymiou and
Constantinos Antoniou

National Technical
University of Athens

5th National Congress on Road Safety
Volos, 2012