Error consideration in georeferencing police reported crash data

John Bigham
jbigham@berkeley.edu

Safe Transportation Research and Education Center
University of California, Berkeley
www.safetrec.berkeley.edu
About UC Berkeley SafeTREC

Founded in 2000 by with a grant from California Office of Traffic Safety (OTS) to reduce traffic fatalities and injuries through multi-disciplinary collaboration in:

• Education
• Technical assistance
• Information Dissemination/outreach.

Affiliated with the Institute of Transportation Studies and the School of Public Health.

Contact Information
SafeTREC, UC Berkeley
2614 Dwight Way, Berkeley, CA 94720
safetrec@berkeley.edu
Overview

- Police reports are completed for injury involved crashes
- Reports are compiled into a database
- Georeferencing each collision provides a latitude/longitude coordinate for geospatial analysis

However, from the initial police report to the final database there are many obstacles to accurate georeferencing!
Collision Georeferencing Phases

• Police report entry at collision site
• Submission to central authority and database entry (California – CHP)
• State highway reference location (postmile) input
• Georeferencing state highway collisions on a LRS
• Geocoding local road collisions to intersections
• Geocoding fatal collisions for inclusion in FARS database
Collision Georeferencing Flowchart

1. Collision occurs and report is prepared
   - Street location information recorded
   - GPS coordinate recorded (if available)
   - Report added to Local PD Database
   - Copies submitted to CHP

2. CHP inputs into central database
   - State highway collisions sent to DOT to establish postmile
   - SWITRS collision database published
     - Local road collisions geocoded to intersections
     - State highway collisions geocoded via LRS
     - Fatal collisions reviewed and geocoded for FARS database
Police report entry at collision site

1. Collision occurs and report is prepared
   - Street location information recorded
   - GPS coordinate recorded (if available)

2. Report added to Local PD Database
   - Copies submitted to CHP

3. CHP inputs into central database
   - State highway collisions sent to DOT to establish postmile

4. SWITRS collision database published
   - Local road collisions geocoded to intersections
   - State highway collisions geocoded via LRS
   - Fatal collisions reviewed and geocoded for FARS database
Police report entry at collision site

• Invalid initial data entry will affect all subsequent phases
  – Spelling errors
  – Non-existent streets
  – Landmarks instead of streets
  – GPS coordinate incorrectly obtained

• Accurate on-site data entry is the most crucial requirement for valid georeferencing
Submission to central authority and database entry (California – CHP)

1. Collision occurs and report is prepared
   - Street location information recorded
   - GPS coordinate recorded (if available)
2. Report added to Local PD Database
3. Copies submitted to CHP
4. CHP inputs into central database
   - State highway collisions sent to DOT to establish postmile
   - SWITRS collision database published
     - Local road collisions geocoded to intersections
     - State highway collisions geocoded via LRS
     - Fatal collisions reviewed and geocoded for FARS database
Submission to central authority and database entry (California – CHP)

- Local jurisdictions submit paper reports to California Highway Patrol
- CHP manually review and enter collision reports into database
  - Potential for incorrect transcription
  - The second opportunity to invalidate a collision location before georeferencing
State highway reference location (postmile) input

Collision occurs and report is prepared

Street location information recorded
GPS coordinate recorded (if available)

Report added to Local PD Database

Copies submitted to CHP

CHP inputs into central database

State highway collisions sent to DOT to establish postmile

SWITRS collision database published

Local road collisions geocoded to intersections
State highway collisions geocoded via LRS
Fatal collisions reviewed and geocoded for FARS database
State highway reference location (postmile) input

• State highway events are identified by a postmile value

• CHP sends to Caltrans to review collision report and input postmile
  – Incorrect translation can create invalid postmiles
Types of postmile entry errors

• Incorrect postmile within allowed range (not easily detectable)
• Invalid postmile value outside of established ranges for a highway
• Non-existent highway number or direction
• Incorrect county
Georeferencing state highway collisions on a LRS

1. **Collision occurs and report is prepared**
   - Street location information recorded
   - GPS coordinate recorded (if available)
   - Report added to Local PD Database
   - Copies submitted to CHP

2. **CHP inputs into central database**
   - State highway collisions sent to DOT to establish postmile
   - SWITRS collision database published
   - Local road collisions geocoded to intersections
   - State highway collisions geocoded via LRS
   - Fatal collisions reviewed and geocoded for FARS database
Linear Referencing

• Method of storing geographic locations by using relative positions along a linear feature

- Known locations (Postmile Ramp Markers)
- Event Locations (Collisions)
Georeferencing state highway collisions on a LRS

- Causes of LRS calibration errors:
  - Incorrect postmile marker placement (non-sequential order)
  - Incorrect postmile marker placement (not on route)
  - Only one known postmile marker on route
  - Incorrect measures of accumulation
  - No postmile markers near the end of a route
Geocoding local road collisions to intersections

- Collision occurs and report is prepared
  - Street location information recorded
  - GPS coordinate recorded (if available)
    - Report added to Local PD Database
      - Copies submitted to CHP
    - CHP inputs into central database
      - State highway collisions sent to DOT to establish postmile
        - SWITRS collision database published
          - Local road collisions geocoded to intersections
          - State highway collisions geocoded via LRS
          - Fatal collisions reviewed and geocoded for FARS database
Geocoding local road collisions to intersections

• Established software and methods
• However, even with perfect data entry, errors can occur from:
  – Street network out of date or name mis-match
  – Custom offset code for collisions occurring a distance from intersections fails
Geocoding fatal collisions for inclusion in FARS database

Collision occurs and report is prepared

Street location information recorded

Collation added to Local PD Database

Copies submitted to CHP

GPS coordinate recorded (if available)

CHP inputs into central database

State highway collisions sent to DOT to establish postmile

SWITRS collision database published

Local road collisions geocoded to intersections

State highway collisions geocoded via LRS

Fatal collisions reviewed and geocoded for FARS database
Geocoding fatal collisions for inclusion in FARS database

• Collision manually geocoded using custom software – always room for human error
• All collisions must be geocoded for FARS
• Potential errors difficult to identify
An extensive process to georeference collision data
Many different causes of georeferencing errors!
Why is accurate georeferencing important?

• Accurate georeferencing is necessary for accurate geospatial analysis:
  – Collision mapping web tools
  – High concentration collision location identification
  – Safety countermeasure selection
Collision mapping web tools

- Transportation Injury Mapping System (TIMS)
- http://tims.berkeley.edu
High concentration collision location identification

- Highway collision rates
  - Correct georeferencing vs. LRS error
High concentration collision location identification
Safety countermeasure selection

- Benefit/cost analysis of countermeasures
- Apply for Highway Safety Improvement Program (HSIP) funds
Future Directions

• Traffic Records Coordinating Committees (TRCCs)

• Initiatives
  – Standardize statewide data entry forms
  – Automated processes
    • No paper submissions
  – Interactive data entry mapping tools
Funding Support

• Funding for TIMS was provided by a grant from the California Office of Traffic Safety, through the National Highway Traffic Safety Administration.
Questions?

• Thank you!