Disentangle crime hot spots and displacements in space and time: an analysis for Chicago from 2001 to 2016
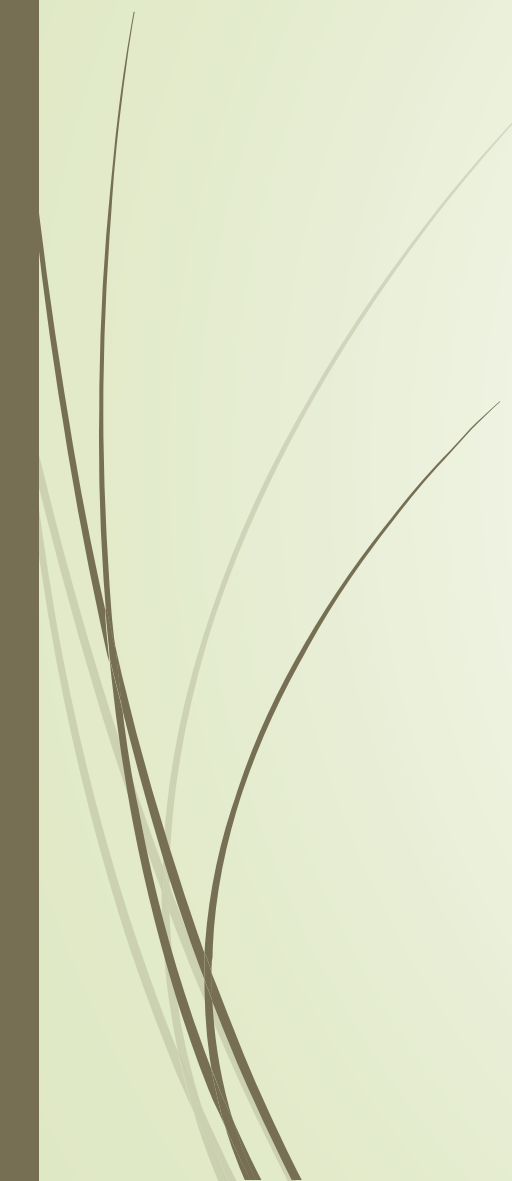
Kai Wang, Xiaolu Zhou, Lixin Li

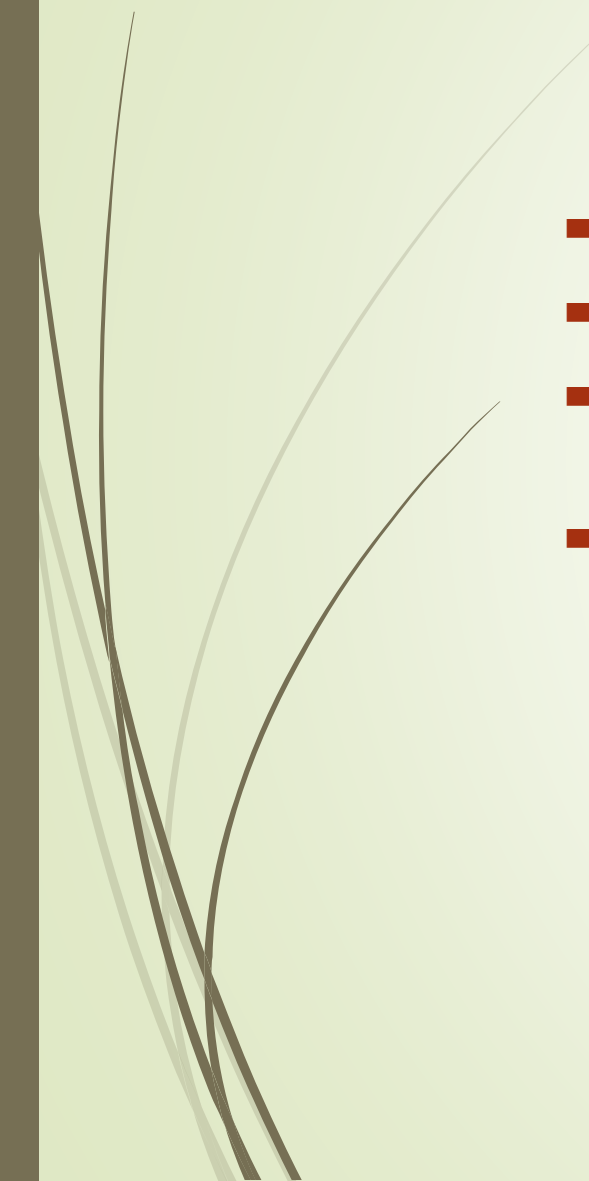Georgia Southern University

11-07-2017

Outline

- Motivation
- The data source
- Methods and algorithms
- Results
- Conclusions

Crime data analysis for humanities

- Digital age has made crime data increasingly available to the public
- Patterns in crime data can be analyzed to benefit society and humanities
- Police can use the analytical results from crime data to optimize the dispatch of resources
- Residents and visitors can use the patterns from crime data to guide daily travel and social meeting

Chosen data for analysis

- Chicago crime data from 2001~2016 with 6,086,589 unique records downloaded from Chicago Data Portal

- Each data record contains the following fields: id, case_number, date, block, iucr, primary_type, description, location_description, arrest, domestic, beat, district, ward, community_area, fbi_code, x_coordinate, y_coordinate, year, updated_on, latitude, longitude, location

- Analysis tools: PostgreSQL database + Custom Java/C++ application

- Visualization tools: ArcGIS; R with plotly

Methods and algorithms

- Space-Time Kernel Density Estimation (STKDE) [Delmelle et al]:

$$f(x, y, t) = \frac{1}{nh_s^2 h_t} \sum_i I(d_i < h_s, t_i < h_t) k_s\left(\frac{x - x_i}{h_s}, \frac{y - y_i}{h_s}\right) k_t\left(\frac{t - t_i}{h_t}\right),$$

- Spatial coordinates: grid based coordinates calculated based on conversion formulas from longitude, latitude and zoom level:

$$x = \frac{2^z (lon\_deg + 180)}{360}$$

$$y = \frac{2^z (1 - \log(\tan(lat\_rad) + \sec(lat\_rad))/\pi)}{2}$$

- Time coordinates: varied granularities (hour, day, week, month, year)

Efficient STKDE calculation

Algorithm 1: STKDE calculation for all 3D points (x, y, t).

Input: A table of crime records including time and location for each record, space bandwidth h_s and time bandwidth h_t, kernel functions k_s and k_t

Output: A list of STKDE values that associates a density d with each point (x, y, t)

1 Initialize m as a hashmap that maps each point (x, y, t) to a density d

2 **for** *each crime record in the input table with coordinates (x, y, t)* **do**

3   for each point with coordinates (x', y', t') that is within space bandwidth h_s and time bandwidth h_t from the point (x, y, t), add the contribution of density of the point (x, y, t) into the entry in the hashmap m for the point (x', y', t') based on Formula (2)

4 **return** m

Displacement graph creation and analysis

- Divide a long time period T (say 10 years) into short time intervals t (say a month)

- For each short time interval, select top k (say k=5) spatial hot spots based on STKDE calculation results

- Create a directed graph with kT/t nodes. Arcs are added between nodes representing hot spots of adjacent spatial grids.

- In digraphs so created, a long path of length N represents a hot spot which remained a crime hot spot for N months.

Displacement analysis using longest path graph algorithm

Algorithm 2: longest path in Directed Acyclic Graph (DAG).

Input: a directed acyclic graph G

**Output:** longest path length in G

1 Perform a topological sort on G

2 $L \leftarrow 0$

3 **for** *each node v in G in reverse order of the sort* **do**

4   Find the longest path length L_v from v based on the topological sort of G

5   **if** $L_v > L$ **then**

6       $L \leftarrow L_v$

7 **return** L

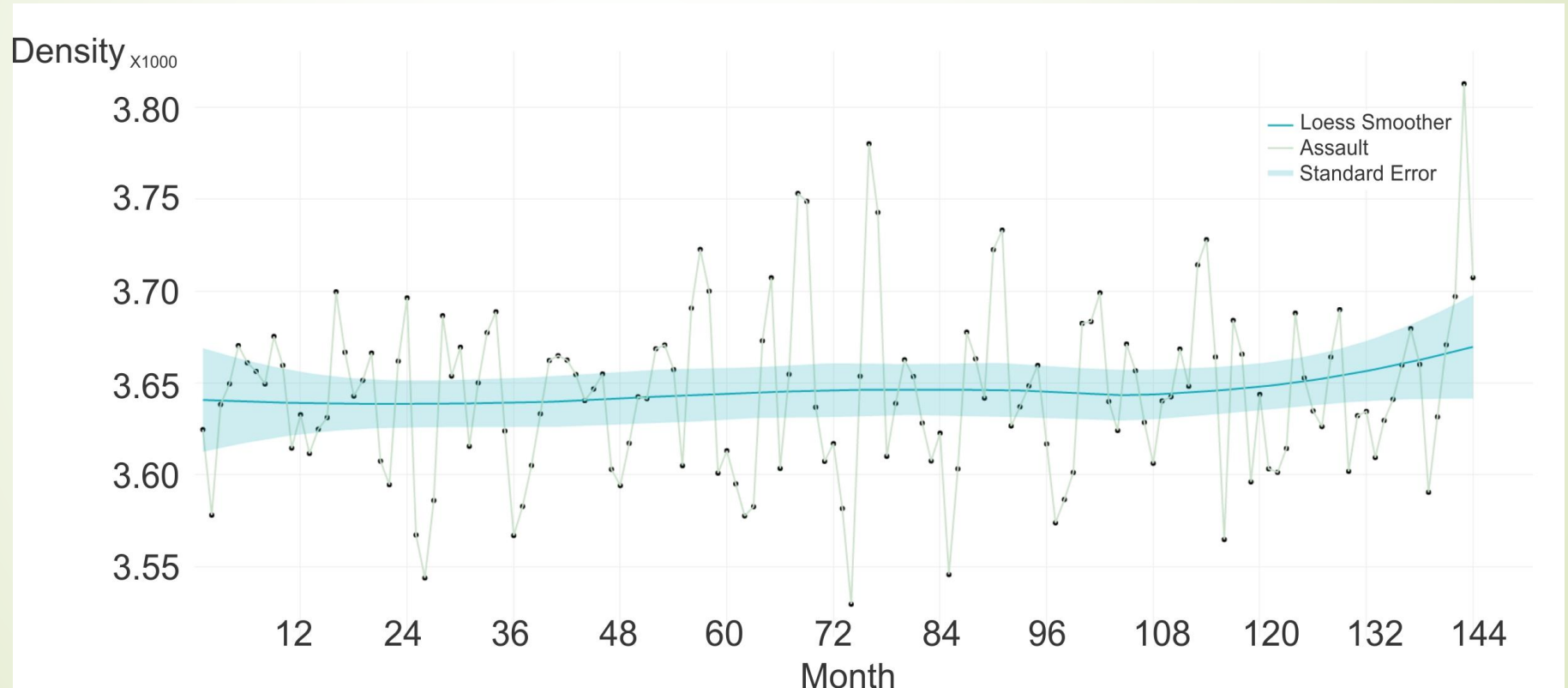# Visualization of analysis results

- STKDE visualization using ArcGIS

- Visualization of displacement graphs using R with the plotly package

3D Visualization of Yearly STKDE for Assault Crimes in Chicago (2011-2016)
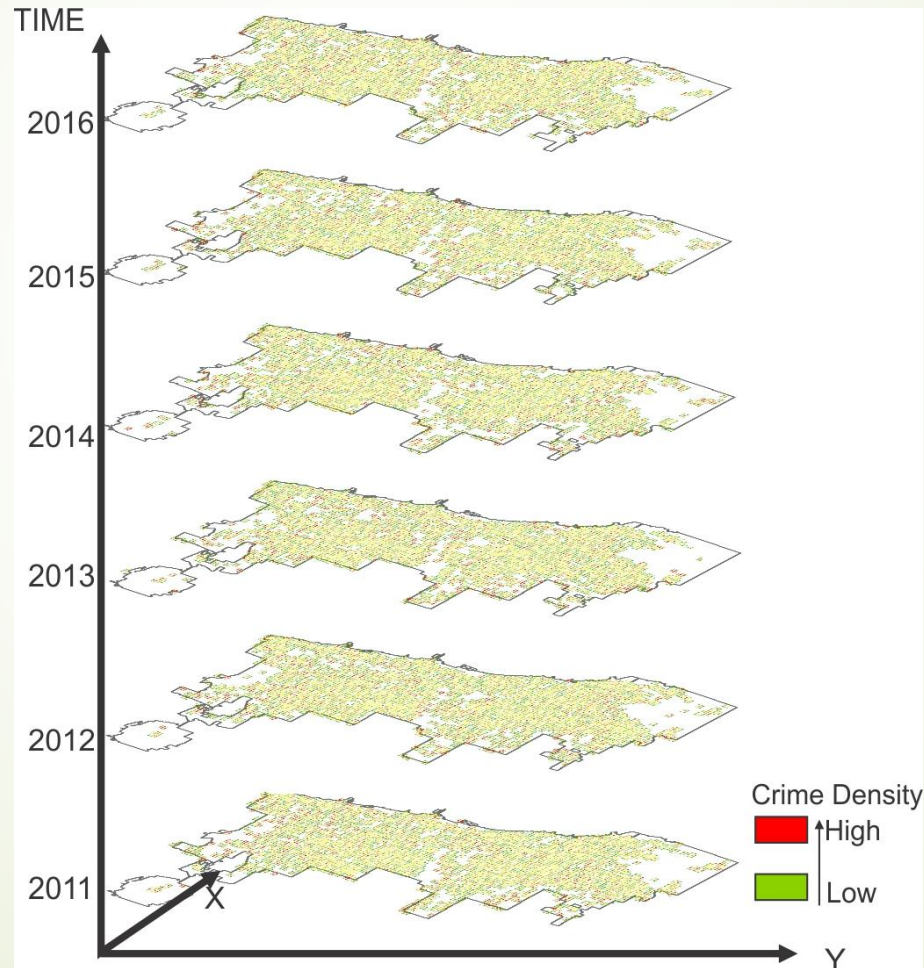
# Average Monthly Density for Assault Crimes in Chicago (2005-2016)

Displacement graphs

Digraph	Crime Type	Space Threshold	#Nodes	#Arcs
G1	Assault	1	897	704
G2	Assault	2	1126	1049
G3	Weapons	1	1036	800
G4	Weapons	2	1343	1255

Digraph	#Components	Largest Component	Longest path length	Ave path length	Std of path lengths
G1	193	34	18	2.65	2.81
G2	77	177	52	6.01	9.47
G3	236	45	23	2.36	2.75
G4	88	441	102	5.36	14.29

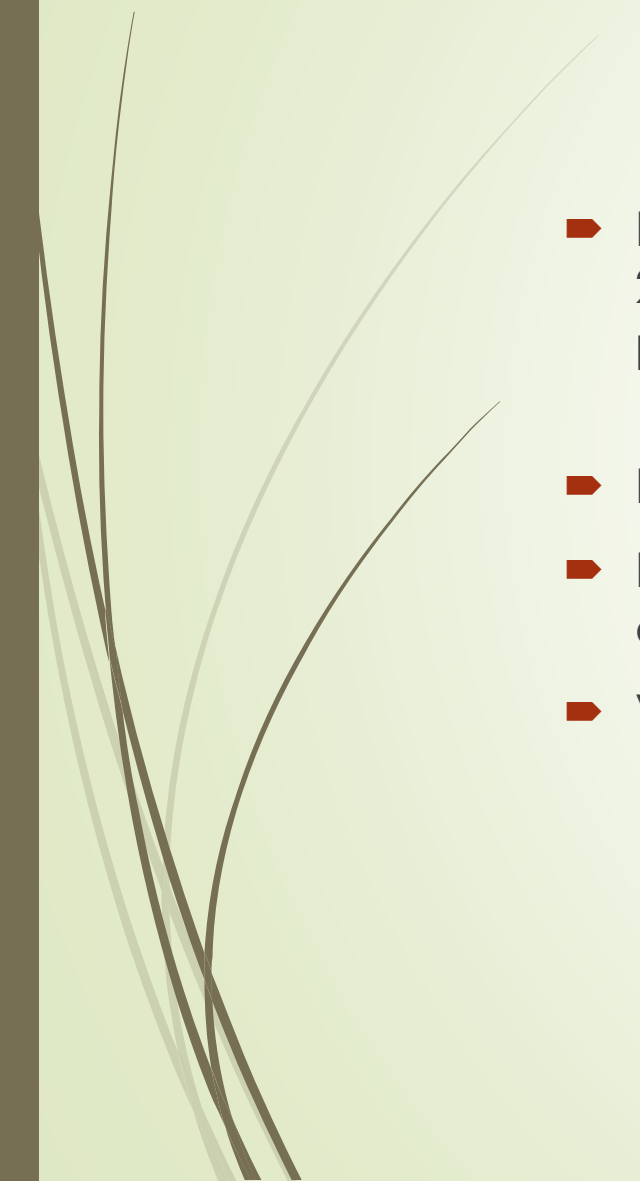# 3D Visualization of Monthly Assault Crime Hot Spot Displacement Digraph G2
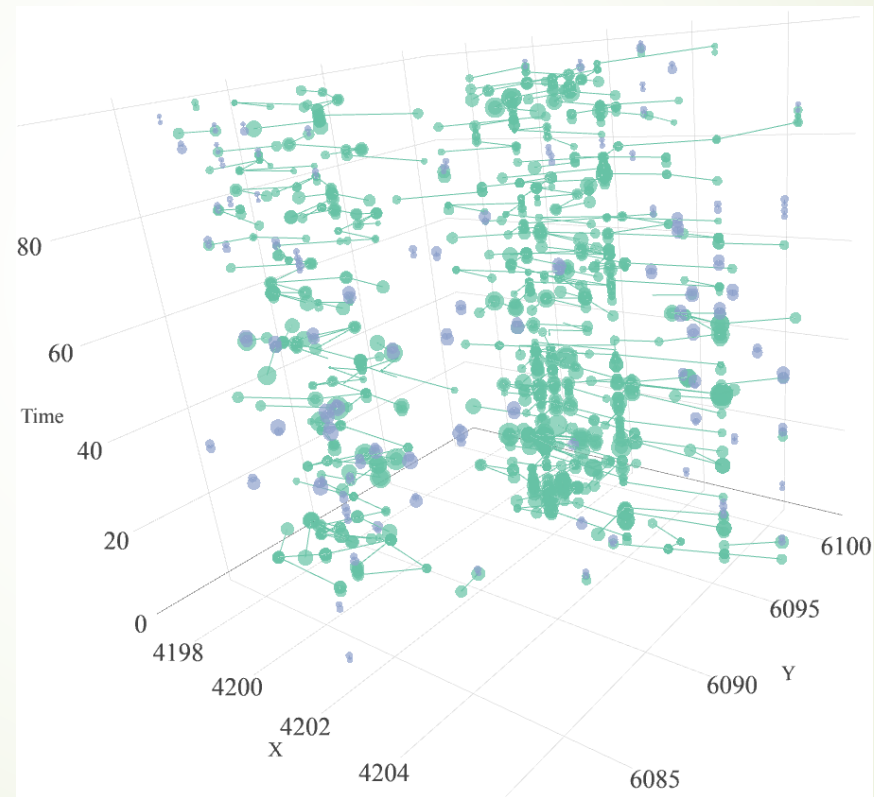
3D Visualization of Monthly Weapon Crime Hot Spot Displacement Digraph G3
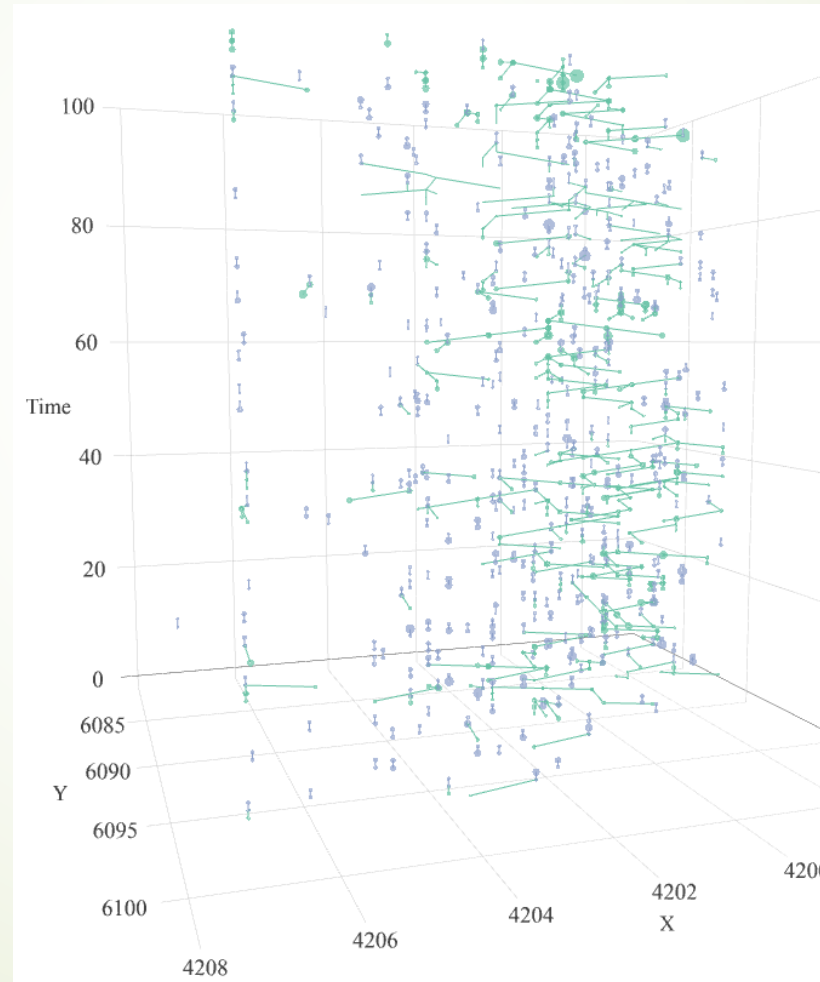
References & Conclusions

- Eric Delmelle, Coline Dony, Irene Casas, Meijuan Jia, and Wenwu Tang. 2014. Visualizing the impact of space-time uncertainties on dengue fever patterns. *International Journal of Geographical Information Science* 28, 5 (2014), 1107–1127.

- Presented a fast algorithm for calculating STKDE

- Presented new graph analytical methods to find the patterns in hot spot displacement

- Visualization of the results

Thank you

- Questions?