Problems With Bounding Space for the Analysis of Crime

- George F. Rengert
- Department of Criminal Justice
- Temple University
- Philadelphia, PA 12122
- grengert@temple.edu
- (215) 204-7994

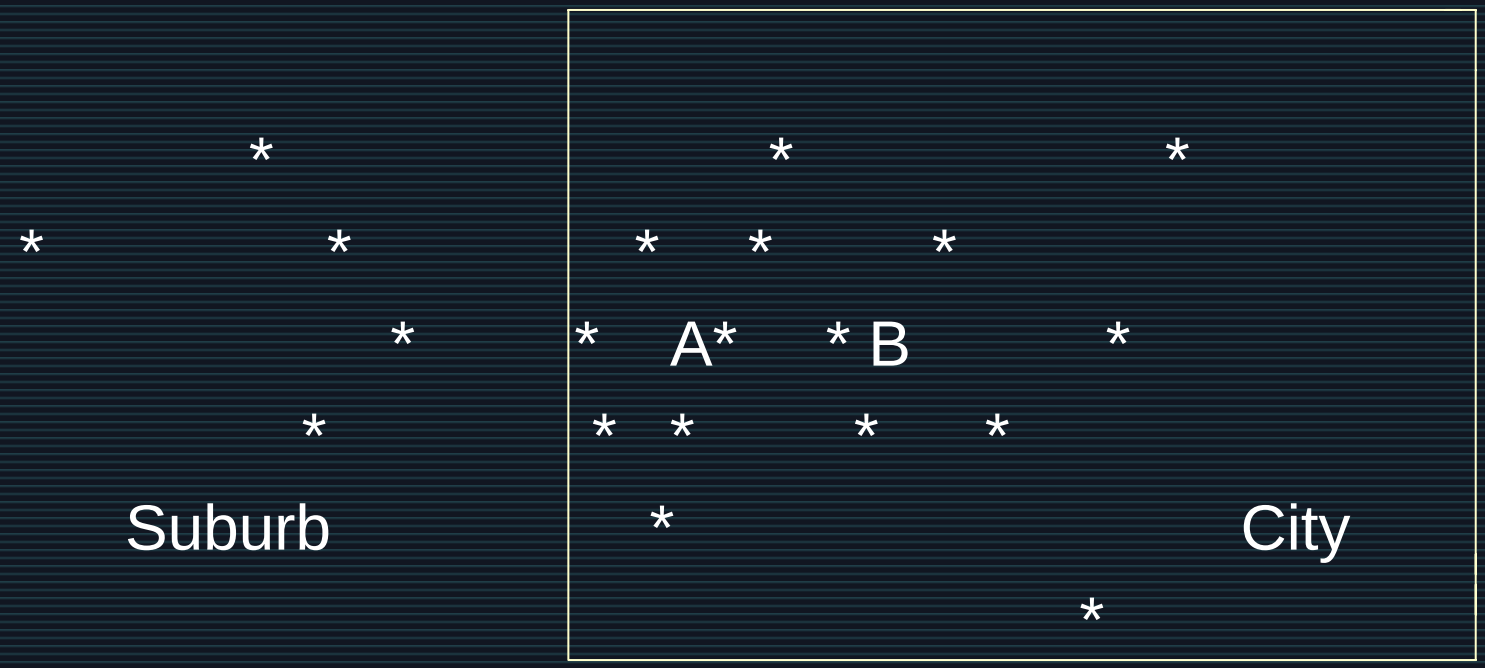
- Purpose of Regionalization in science.
  - Maximize between region variance and minimize within region variance.
    - Political administrative boundaries seldom do this.
- Problems of Political Boundaries.
  - Static East Coast Boundaries.
    - In some of these cities, political boundaries have not changed for over a hundred years.
      - 1854 for Philadelphia.
  - Data collected is not always shared with neighbors.
    - Just as important, may resemble their neighbors.
      - More variance within city than between city and neighboring suburbs.

        This is counter to what we desire in statistics.

- Census bureau decided to do something about this problem.
  - Metropolitan Areas
    - Urban population of over 50,000.
    - Commuting patterns indicate strong economic interaction.
  - Micropolitan Areas
    - Central county with min. 1,000 people/sq. mile.
    - Urban cluster population of between 10,000 and 50,000.
    - Commuting patterns indicate strong economic interaction.
  - This made metropolitan areas more alike internally and different from the surrounding area.
    - What we want when we regionalize.

- But criminals may not recognize (or pay attention to) these boundaries in choosing crime sites.

- ‘B’ is the predicted anchor point of the serial offender if information is available only for the bounded area.



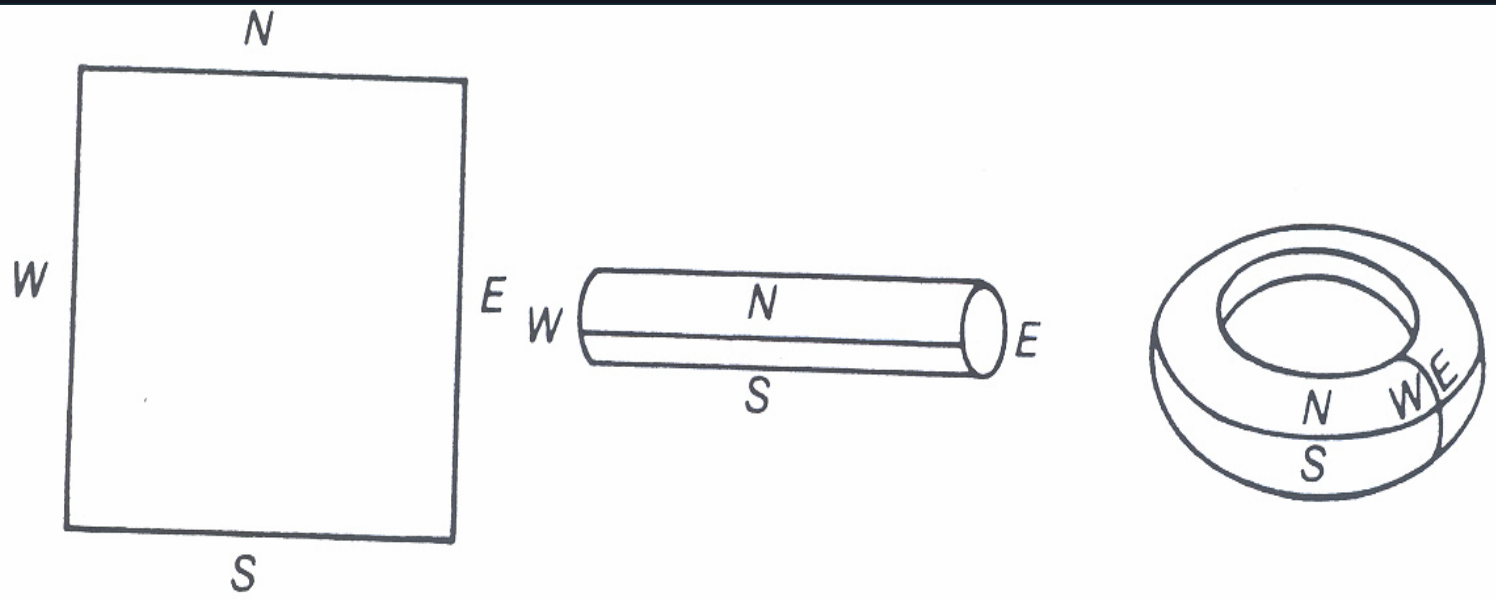
‘A’ is the predicted anchor point of the serial offender if information is available for the entire area.

‘B’ is the predicted anchor point of the serial offender if information is available only for the bounded area.

- Problems Effects ‘Geographic Profiling.’
  - Truncates data.
  - Shifts predicted high risk area toward center of city.
    - Actual high risk area from existing data should be shifted toward suburb.

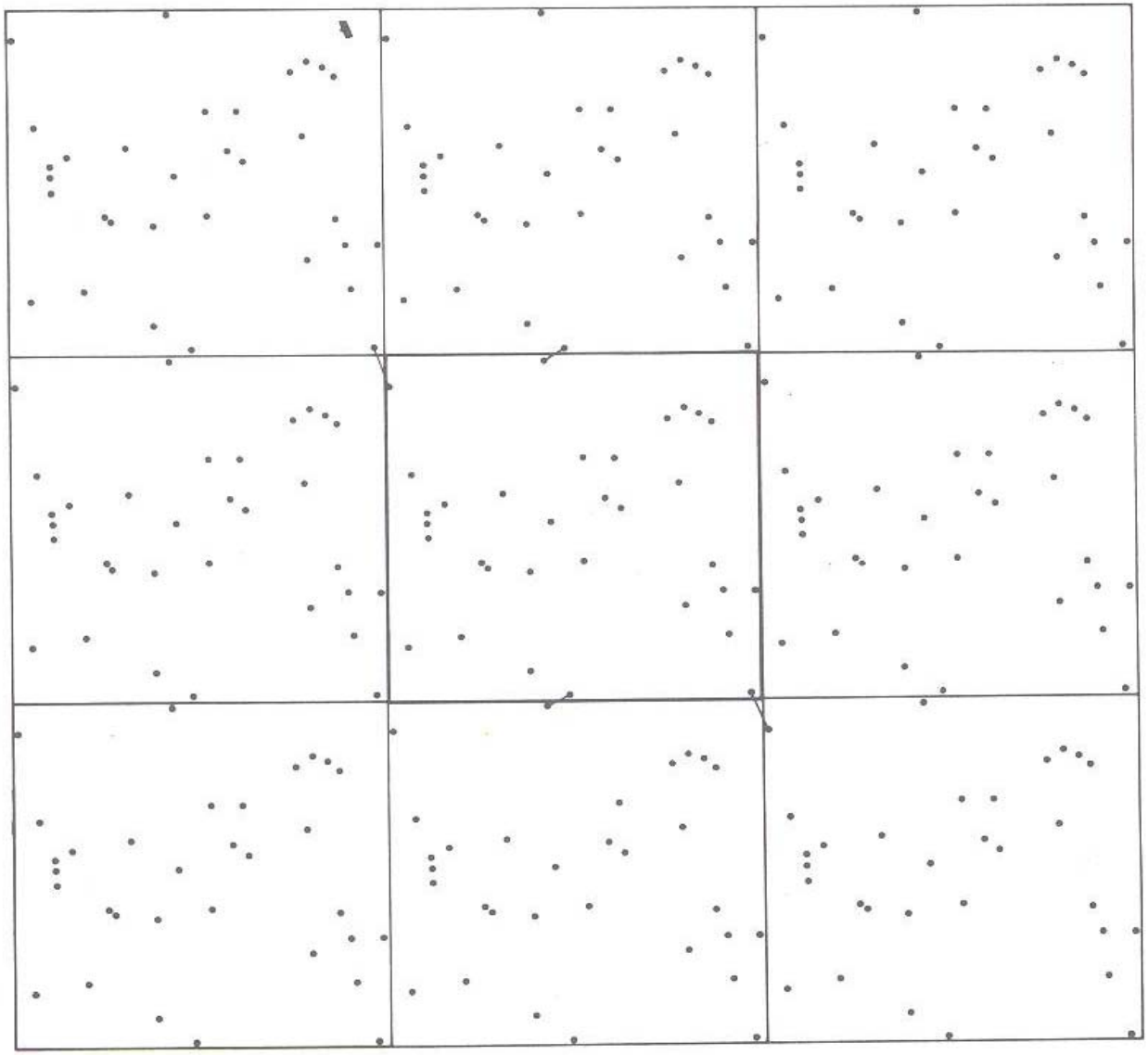
- Boundary Problems Often Referred to as ‘Edge Effects.’
  - Effects measures of spatial arrangement.
    - Clustered, uniform, or random pattern.
    - Different processes assumed:
      - Clustered = something attracting phenomena to a point.
      - Uniform = phenomena tend to be as far away from each other as possible.
      - Random = there is nothing effecting the spatial arrangement of the phenomena.
  - Nearest Neighbor analysis.
    - The nearest neighbor of a point near boundary may be outside area.
      - How can we correct this?

- Creation of a ‘Torus’
- that has no edges.



- Torus assumes the same process effecting the arrangement of points is operating outside boundary as inside it.
  - Then, Torus works.
    - In practice, create 8 maps identical to study area and surround the study area keeping orientation the same.

A



- A torus only works when we have a square or a rectangular shaped map.
  - When we have an irregularly shaped map as in the case of Philadelphia, we must use other methods of correcting for edge effects.
    - One method is to ignore all points that are closer to the boundary than to another point in the distribution.
      - This has the disadvantage of losing information that may be critical when there are few points in the pattern.
    - A similar method is to use Geographic Information Systems to buffer in from the boundary the distance of those points that are closer to the boundary than to another point.
      - In this case, the points in the buffer can be used to measure nearest neighbors of points outside the buffer.

        Information is lost in this case also.

- Finally, Donnelly has shown that when N is larger than 7, a correction factor can be added to the nearest neighbor formula of Complete Spatial Randomness.
  - This formula works if the study area is not highly irregular.
    - $E(d_i) = 0.5 \sqrt{(A/N)} + (0.0514 + 0.041 / \sqrt{N}) B/N$.

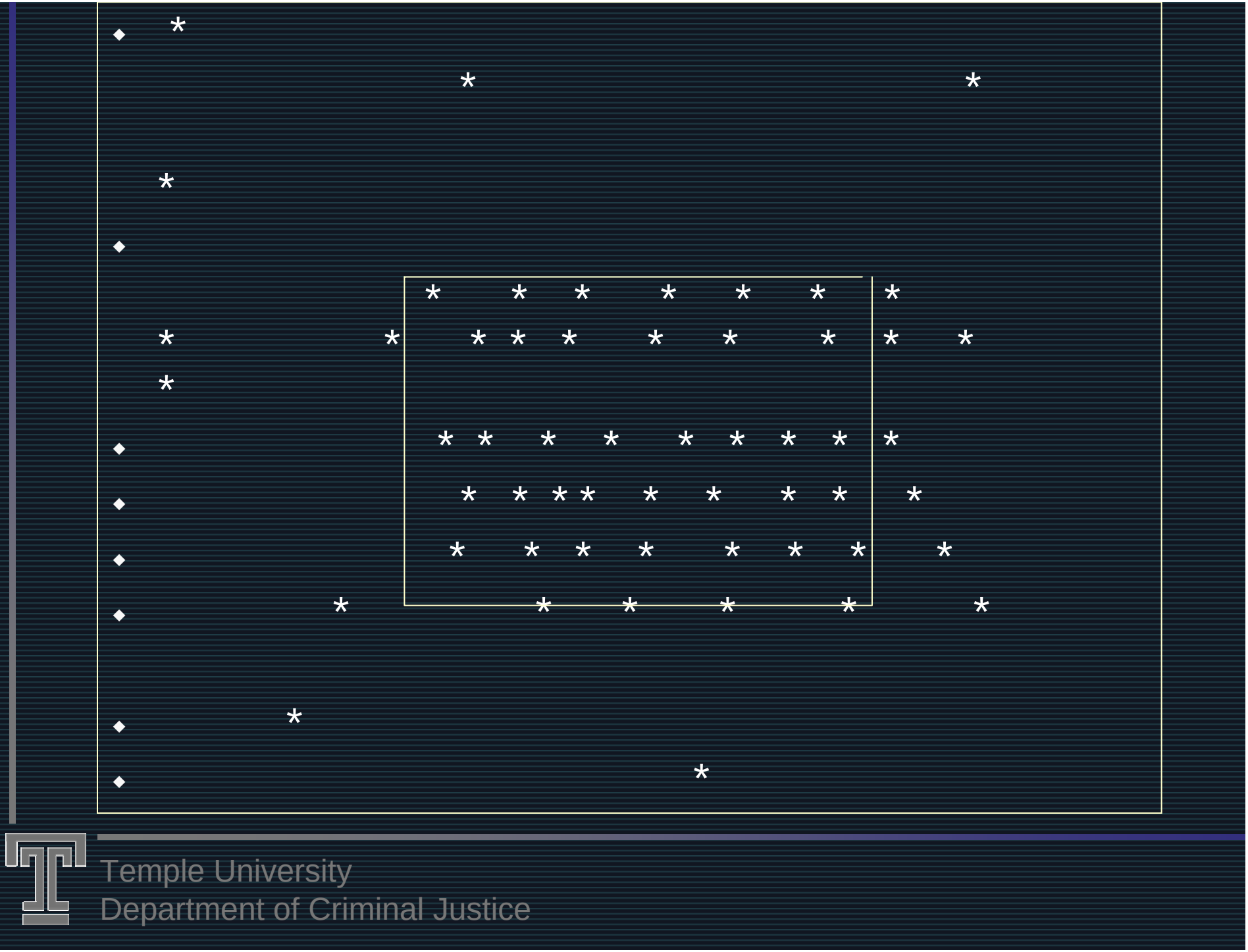
      - Where:

        A = the size of the area.

        N = the number of points in the area.

        B = the length of the boundary.

- Boundaries can change the predicted spatial arrangement of points.
  - Referred to as the Modifiable Area Unit Problem
  - Following example:
    - From a clustered to a random or uniform pattern.



- Finally, boundaries tend to mask variation in spatial data.
  - Spatial variation is obscured.
    - When boundaries are drawn around spatial data, values are averaged for the new area.
      - Also, variance measures are reduced in most cases.

- Example of how this happens:

Problems of Aggregation.

A

1	7	3	5	7	5
3	7	3	3	3	8
4	9	5	8	4	2
6	6	7	0	0	8
9	3	4	5	2	7
5	1	3	8	6	4

B

1 7 4.5 3 7	3 6 3.75 3 2	7 5 5.75 3 8
4 9 6.25 6 6	5 8 5.0 7 0	4 2 3.5 0 8
9 3 4.5 5 1	4 5 5.0 3 8	2 7 4.75 6 4

Adopted from Yeates (1974).

The numbers in A indicate the number of drug sales arrests per street face. The center number in B is the average number of drug sales arrests per block when four streets are aggregated together. Note that the average masks a considerable amount of variation.

- Notice that the ‘range’ in variables in A are between zero and nine.
  - The ‘range’ in B is between 3.75 and 6.25.
    - The range is considerably reduced when we further aggregate the data.
      - We lose information.

- From this discussion, it should be clear that researchers must be careful in drawing boundaries.
  - Not to obscure patterns.
  - Not to lose too much information.
  - Not to cause misleading interpretations.
    - Geographic profiling example.

- Again, the purpose of boundaries (regionalization) is to bound like areas so they can be contrasted with unalike neighboring areas.
  - Minimize within area variance, maximize between area variance.

Temple University
Department of Criminal Justice