IDEA GROUP PUBLISHING **ITB10841**

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

This chapter appears in the book, *Geographic Information Systems and Crime Analysis*, edited by Fahui Wang. © 2005, Idea Group Inc.

Chapter III

Garbage In, Garbage Out: Geocoding Accuracy and Spatial Analysis of Crime

Tess McCarthy, Charles Sturt University, Australia

Jerry Ratcliffe, Temple University, USA

Abstract

Advances in computing technology and analytical techniques have given crime analysts increasingly powerful toolboxes with which to unlock the spatial patterns and processes of crime. However, the utility of such tools is still bounded by the "garbage in, garbage out," maxim, whereby analytical output is only as reliable as the analytical input. Therefore, this chapter reviews some of the sources of spatial data inaccuracy that must be considered when analyzing crime. Given the prevalence of street addresses as a spatial location identifier for crime events, particular attention is given to the accuracy and optimum parameters for geographically referencing address data. Example data drawn from burglary records in the city of Wollongong, Australia, illustrate the significance of the issues and the impact that poor address management can have on the analysis of crime.

The chapter emphasizes the practical, by outlining address correction options and summarizing recent research that identifies optimum settings for geocoding software tools.

Introduction

Modern crime fighting techniques, such as SARA (Eck & Spelman, 1987), problem-oriented policing (Mazerolle & Terrill, 1997) and CompStat (Walsh, 2001), are increasingly dependent on spatial analyses of crime to effectively and efficiently allocate crime reduction resources. The accuracy of these analyses is tied in part to the accuracy of the spatial crime data on which they are based. Not to be confused with precision, accuracy is the degree to which a measurement matches the accepted truth, while precision refers to the level of measurement. The following discussion will focus on three types of spatial data inaccuracy: conceptual, positional and attribute, using the example of burglaries recorded in the city of Wollongong, on the southeast coast of Australia, approximately 40 miles south of Sydney.

Conceptual Accuracy

Conceptual accuracy refers to the legitimacy of conceptual models used to simplify and represent complex, real-world features or events. In spatial analysis, this may relate to the choice of spatial model used to summarize and describe the geography of crime events – the choice of a point, line or area as the basic geographic or topological unit. Points have no spatial size – that is, no length or area – and have come to represent the vast majority of crime events. Lines represent objects that exist in one dimension only, objects with length. There are few examples of crime events being mapped as lines, though LeBeau used road segments to suggest areas of risk for patrolling officers based on the number of incidents occurring along a road segment over a given period of time (LeBeau, 2000). Regions, such as police beats, exist in two dimensions, possessing both length and area, and may be represented by a polygon (in a vector format) or a collection of pixels (in a raster system). Though in reality most crime events exist in three spatial dimensions, finite resources and the scale at which the data will ultimately be visualized and analyzed usually demands that the mapping unit be collapsed and the event represented by a simple point or line. Indeed, of the major crime categories recorded by the New South Wales Police