Developing Police Patrol Strategies Based on the Urban Street Network

Huanfa Chen

Department of Civil, Environmental and Geomatic Engineering

University College London

Supervisors:

Prof. Tao Cheng

Prof. John Shawe-Taylor

A thesis submitted for the degree of

Doctor of Philosophy in Geographic Information Science

February 2019

Declaration

I, Huanfa Chen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

In urban areas, crime and disorder have been long-lasting problems that spoil the economic and emotional well-being of residents. A significant way to deter crime, and maintain public safety is through police patrolling. So far, the deployment of police forces in patrolling has relied mainly on expert knowledge, and is usually based on two-dimensional spatial units, giving insufficient consideration to the underlying urban structure and collaboration among patrol officers. This approach has led to impractical and inefficient police patrol strategies, as well as a workload imbalance among officers. Therefore, it is of essential importance to devise advanced police patrol strategies that incorporate urban structure, the collaboration of the patrol officers, and a workload balance.

This study aims to develop police patrol strategies that would make intelligent use of the street network layout in urban areas. The street network is a key component in urban structure and is the domain in which crime and policing take place. By explicitly considering street network configurations in their operations, police forces are enabled to provide timely responses to emergency calls and essential coverage to crime hotspots. Although some models have considered street networks in patrolling to some extent, challenges remain. First, most existing methods for the design of police districts use two-dimensional units, such as grid cells, as basic units, but using streets as basic units would lead to districts that are more accessible and usable. Second, the routing problem in police patrolling has several unique characteristics, such as patrollers potentially starting from different stations, but most existing routing strategies have failed to consider these. Third, police patrolling strategies should be validated using real-world scenarios, whilst most existing strategies in the literature have only been tested in small hypothetical instances without realistic settings.

In this thesis, a framework for developing police patrol strategies based on the urban street network is proposed, to effectively cover crime hotspots, as well as the rest of the territory. This framework consists of three strategies, including a districting model, a patrol routing strategy for repeated coverage, and a patrol routing strategy for infrequent coverage. Various relevant factors have been considered in the strategy design, including the underlying structure of the street network and the collaboration among patrollers belonging to different stations. Moreover, these strategies have been validated by the patrolling scenarios in London. The results demonstrate that these strategies outperform the current corresponding benchmark strategies, which indicates that they may have considerable potential in future police operations.

Impact Statement

Policing plays an important role in crime prevention and public safety. Currently, policing relies mainly on human knowledge and experience, but this approach fails to keep pace with the challenges of increasing crime rates and decreasing police funding. Increasingly, police have fewer resources with which to complete their daily activities, which motivates a more intelligent, technology driven approach. In particular, by integrating big data collected by police forces with state-of-the-art computational technologies, intelligent policing offers new opportunities to reduce crime and increase public confidence, while also increasing efficiency and reducing operational costs.

One way to approach this problem is to optimise the way in which the police target crime hotspots while on patrol. This research aims to develop advanced police patrol strategies that make intelligent use of the street network layout in urban areas. Street network is a key component of urban structure and the domain on which crime and policing take place. By explicitly considering the street network configuration in the operations, police forces are capable of providing timely response to emergency calls and essential coverage to crime hotspots. This would lead to a high visibility of police forces in the urban space, hence deterring potential offenders and enhancing the public confidence in policing.

Specifically, this research addresses different aspects of police patrol. First, this research proposes a novel patrol district design based on the street network. This design guarantees short travel distance within districts and immediate response to calls for service. Second, this research puts forward a patrol routing strategy to coordinate a group of patrol officers in an online manner, so as to provide efficient coverage to crime hotspots. Third, this research devises a route design strategy that is capable of generating a set of balanced patrol routes that covers the given locations.

This research has delivered a number of algorithms and tools for patrol planning, in cooperation with the Metropolitan Police Service (MPS) in London. We anticipate that these tools will be adopted in the daily policing practice of MPS. Further, with necessary adjustments, they can be applied to police patrolling in other cities nationally and globally. These algorithms and tools can also solve other relevant problems, such as the design of sales and service districts or route planning in logistics and delivery.

Acknowledgements

This work would not have been possible without the guidance and help of many people around me. First of all, I would like to thank my supervisor, Prof. Tao Cheng. During the past four years, she has been generous in her support and has spent a huge amount of time guiding me through my research. I have always been motivated by the critical comments and warm encouragement from Tao. I would also like to thank my secondary supervisor, Prof. John Shawe-Taylor, who is a great mentor and has made valuable contributions to this research. John is so knowledgeable and insightful that he always provides great inspirations for my research.

I would also like to thank the members of the CPC project (Crime, Policing and Citizenship), especially Dr Sarah Wise, Dr Gabriel Rosser, and Dr Toby Davies. It is a great pleasure to work with you, and your knowledge and experience have been so valuable to me.

I also owe a lot of gratitude to my colleagues at SpaceTimeLab and my friends at UCL. The long list starts with but is not limited to, Dr James Haworth, Jianan Shen, Juntao Lai, Yang Zhang, Monsuru Adepeju, Nilufer Sari Aslam, Dawn Rollocks, and Davide Di Gioia. You have helped make these years some of the most enjoyable moments of my life, and your support and friendliness will be kept in my memory.

This PhD is funded by the China Scholarship Council and the UCL BEAMS Dean's Prize. It is part of the CPC project funded by Engineering and Physical Sciences Research Council (EP/J004197/1). I am grateful to the Metropolitan Police Service for the provision of the crime data, and Trevor Adams for valuable discussions on this work. I also thank Holborn police station for accommodating me to observe their patrols in order to gain the first experience of police patrolling.

Finally, I thank my family and friends, especially my parents and grandparents for always supporting me in the pursuit of my research career despite the long distance, and Yiya, for putting up with a PhD student for such a long time. Without your support, I would be nowhere in my PhD.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1 INTRODUCTION

This chapter provides the introduction to this dissertation. In it, the background to, and motivation behind this research are first described. Then, the aim and objectives are defined, followed by a description of the general structure of the thesis, which outlines different phases of this research.

1.1 Background

1.1.1 Police Patrolling

In urban areas, crime and disorder have been chronic societal problems that spoil the economic and emotional well-being of residents. For example, in Greater London, the number of reported offences in the financial year of 2017-2018 is 827,225, which has increased by 6.4% in comparison to the previous year (The Metropolitan Police, 2018). Therefore, predicting, preventing and alleviating crimes is fundamental to the development of the cities.

A crucial approach to deter crime and maintain public safety is through police patrolling. Generally, patrolling is defined as “the act of walking or travelling around an area, at regular intervals, in order to protect or supervise it” (Abate, 1996). In policing, patrolling officers are dispatched to patrol the territory and especially the important locations on a day-to-day basis, in order to increase their visibility to the public and potential offenders. Visible patrols can efficiently deter crime via raising potential offenders’ perceptions of risk in the area (Cook, 1980). This effect that police presence does appear to deter crime is supported by the temporal association of the police movements and the crime incident locations (Cheng *et al.*, 2016).

1.1.2 Police Patrol Strategies

The success of police patrols in deterring crime largely depends on the efficiency of police patrol strategies. Existing police patrol strategies can be roughly divided into two categories: districting and routing strategies.

Districting strategies are concerned with partitioning a territory into smaller sub-areas so that each sub-area can be covered by a patrol team. So far, most of the existing police patrol districting methods use areal units or zones (e.g., beats, sectors, grids, census blocks) as basic units, and fail to consider the street network in the district design.

Routing strategies focus on designing routes for patrol officers that cover important locations in the area efficiently. Routing problems in police patrolling can be categorised into two types – infrequent patrol and repeated patrol routing problem (IPRPs and RPRPs, respectively). The infrequent patrol is involved in covering each target at least once in the time interval, whilst the repeated patrol aims to patrol the targets repeatedly in a given time interval, using a fairly large patrol team.

1.1.3 Limitations of Current Police Patrol Strategies in Urban Areas

Research on police patrol strategies has increased in recent years, but most of the efforts have failed to incorporate the underlying street network structure or consider the peculiarities of police patrols. Specifically, the limitations of recent studies can be summarised around three points:

First, most existing police districting models use areal units or zones (e.g., grids, census blocks) as basic units (Mitchell 1972; Zhang & Brown 2013), which neglects the influence of the underlying street network. The resulting districting plan would be difficult to patrol in operation, as the districts may contain unconnected streets or inaccessible parts. Some recent studies have attempted to incorporate the street network in the design of police patrol areas by using the network distance between historical incidents and patrol teams (Curtin, Hayslett-McCall and Qiu, 2010), but they still use the areal basic units (i.e. patrol sectors) rather than the streets.

Second, the routing problem of police patrol has a number of characteristics: the objective can be either repeated coverage or infrequent coverage of the important locations; patrol officers may have online communication with the control centre; patrol officers can belong to different stations and should start and end their routes in their station; the workload should be evenly distributed among patrol officers. Nevertheless, previous research on the police patrol routing problem fails to account for these characteristics. For example, the existing studies on route design of a group of patrol officers assume that all officers start from the same station (Ahr and Reinelt, 2006; Benavent *et al.*, 2009; Willemse and Joubert, 2012). Therefore, new models should be proposed to fit the actual needs of police patrol.

Finally, all proposed police patrol strategies should be validated using real-world scenarios or scenarios of similar size. This suggests that such strategies should be scalable. Existing models, however, have only been tested in small-sized or hypothetical problem instances. Previous districting models have been tested in regions with up to several hundred units (Camacho-Collados, Liberatore and Angulo, 2015; Liberatore and Camacho-Collados, 2016), but it is unknown whether the models and algorithms are applicable to larger instances, such as a London borough with several thousands of street units. Similarly, the routing strategy for repeated coverage has been tested on hypothetical environments with fewer than 100 locations to visit (Portugal and Rocha, 2013a), and application of this routing strategy in real-world patrolling scenarios has not been attempted.

1.2 Research Aim and Objectives

As described in the previous sections, this research evolved from the necessity to explicitly incorporate the underlying street network into the design of patrol strategies. Therefore, the main aim of this work is to propose a framework for developing police patrol strategies based on the street network, in order to effectively cover crime hotspots, as well as the entire street network.

This aim – the desired outcome – could only be achieved by breaking the work down into several smaller objectives, which can be summarised as follows:

1. To optimally design police districts, based on the street network, by:
    a. Defining and formulating the problem of street network-based police districting;
    b. Designing efficient algorithms (exact, approximation or heuristic) from which to derive high-quality solutions in an acceptable time.
2. To design balanced patrol routes for officers from different police stations, and to cover each crime hotspot at least once, by:
    a. Defining and formulating the routing problem;
    b. Designing the algorithms to generate (near-)optimally balanced patrol routes quickly.
3. To design an online patrol routing strategy in order to cover crime hotspots repeatedly, this strategy being useful in situations where patrol officers have online communication with a control centre, by:
    a. Defining the metrics for evaluating the performance of online patrol routing strategies;
    b. Designing an online patrol routing strategy.
4. To evaluate case studies to test the proposed police patrol strategies, based on hypothetical and real-world data.

1.3 Thesis Structure

This thesis is organised into nine chapters. A roadmap of the thesis structure is presented in Figure 1.1. The contents of the chapters are as follows:

In Chapter 2, entitled *Literature Review: Police Patrol Strategies Based on Street Networks*, the existing research on police patrol strategies is reviewed. The chapter focuses on three facets of the strategies and models described in the literature: 1) the types of model that have been used, broadly separated into districting models and routing models; 2) the ways in which street networks have been incorporated into the models; and 3) the algorithms used to solve the models. The findings from this chapter motivated the direction of this study, as reported in the following chapters.

In Chapter 3, entitled *Methodology*, a framework for police patrol strategy is presented. This framework includes a districting strategy and two routing strategies. Further, these can be coupled to form a more comprehensive strategy.

In Chapter 4, entitled *Information about the Data*, two case study areas and the datasets used are introduced. The case study was conducted in the London Borough of Camden and the South Side of Chicago. The datasets used include the street network data, the police station locations, crime records data, the predictive crime risk map, and the land cover dataset.

In Chapter 5, entitled *A Police Districting Model Based on Street Networks*, a model for designing patrol districts based on a street network is proposed. The aim of this model is to partition a region into a given number of districts such that each district is contiguous, compact and incorporates a relatively equal workload. The generated districts can be further policed using the patrol routing strategies described in Chapter 6 and 7, or other alternatives.

In Chapter 6, entitled *Infrequent Patrol Routing: A balanced route design*, a routing strategy that was developed so that each crime hotspot is covered at least once while the route length of different patrols is balanced, is described. This strategy was based on the Min–Max Multiple-Depot Rural Postman Problem (MMMDRPP), which is formulated and solved in this chapter.

In Chapter 7, entitled *Repeated Patrol Routing: An online approach*, a routing strategy is proposed that provides repeated coverage of crime hotspots in a street network. A set of measures was developed to evaluate the performance of this routing strategy.

In Chapter 8, entitled *Coupling Different Strategies in Police Patrol*, the coupling of distinct police patrol strategies is discussed. Strategy coupling is very useful, bringing about several potential advantages. The focus is placed on coupling the districting and the online repeated patrol routing strategies. A case study was carried out to validate the coupled strategy.

Chapter 9 contains the conclusions of this thesis. The major outcomes of the research are stated, as well as the expected contributions to the scientific literature on this topic. This is followed by a discussion of the limitations and future research directions of the study.

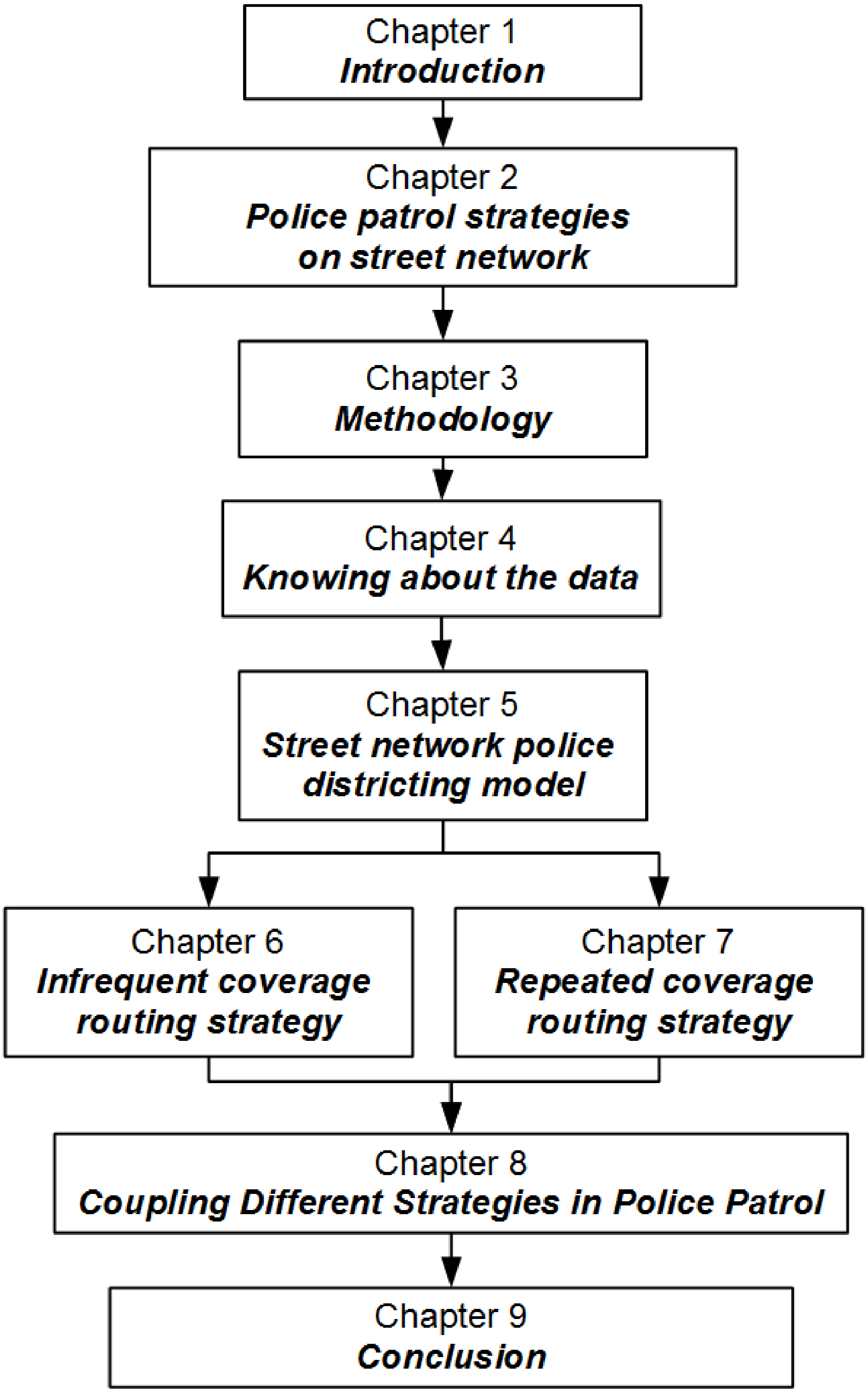


Figure 1.1 Thesis structure

Chapter 2

LITERATURE REVIEW: POLICE PATROL STRATEGIES BASED ON STREET NETWORKS

2 LITERATURE REVIEW: POLICE PATROL STRATEGIES BASED ON STREET NETWORKS [1] [2]

2.1 Introduction

In policing operations, police districts have been adopted for many years. Police districting problems (PDPs) arise from the need to meet response time thresholds and to balance workloads across districts (Bruce, 2009). Meanwhile, routing problems in police patrol are motivated by the necessity to guarantee highly responsive patrols or to reduce travel distance. This chapter presents a walkthrough of the districting and routing problems associated with police patrolling.

In Section 2.2, it is discussed how previous research has addressed PDPs, and what attributes have been considered in the models. Existing methodologies and approaches for solving these problems and generating patrol districting plans are also covered.

In Section 2.3, the existing research on police patrol routing problems and the methods and algorithms used to obtain feasible and efficient routes for patrol officers, are outlined. The advantages and limitations of different models are discussed.

Section 2.4 highlights the attempt to incorporate street networks into police patrolling. It starts by explaining why street networks are essential in devising police patrol strategies, followed by a discussion of the relevant research.

2.2 Police Districting Problems (PDPs)

This section presents the generic problem of defining the districts, which is then contextualised in the framework of police patrolling. It also provides a thorough review of the attributes considered and the methodologies adopted, as well as a discussion on the limitations of the existing research.

---

[1] Part of this chapter was presented in the following publication: Chen, H., Cheng, T. & Wise, S., 2017. Developing an online cooperative police patrol routing strategy. Computers, Environment and Urban Systems, 62, pp.19–29.

[2] Part of this chapter was presented in the following publication: Chen, H., Cheng, T. & Shawe-Taylor, J., 2018. A Balanced Route Design for Min-Max Multiple-Depot Rural Postman Problem (MMMDRPP): a police patrolling case. International Journal of Geographical Information Science, 32(1), pp.169–190.

2.2.1 The Districting Problem

Districting is the problem of grouping small geographic areas (called basic units), into larger geographic clusters (called districts), such that the districts are balanced, contiguous, and compact, according to a number of relevant attributes or criteria. Depending on the practical context, districting is also called territory design, territory alignment, sector design, or zone design (Kalcsics, 2015). Districting has been applied to a number of fields, including:

- Political districting and the definition of electoral areas (Mehrotra, Johnson and Nemhauser, 1998; Cirincione, Darling and O'Rourke, 2000; Bozkaya, Erkut and Laporte, 2003);
- Sales districting (Fleischmann and Paraschis, 1988; Zoltners and Sinha, 2005; López-Pérez and Ríos-Mercado, 2013);
- Service districting (Blais, Lapierre and Laporte, 2003; Galvão *et al.*, 2006);
- Electric power districting (Bergey, Ragsdale and Hoskote, 2003a, 2003b);
- Emergency service districting (Larson, 1974; Iannoni, Morabito and Saydam, 2009);
- Internet networking (Park *et al.*, 2000);
- Health information systems (Braa and Hedberg, 2002);
- Public transportation network districting (Tavares-Pereira *et al.*, 2007; Tavares Pereira *et al.*, 2009);
- School districting (Schoepfle and Church, 1991; Caro *et al.*, 2004);
- Social facilities districting (Minciardi, Puliafito and Zoppoli, 1981); and
- Winter service districting (Muyldermans, 2003; Muyldermans, Cattrysse and Van Oudheusden, 2003)

These applications share certain criteria with police districting, such as contiguity and the compactness of the districts. Some methods used in these applications can be used to address PDPs. What distinguishes police districting requirements from those of other applications are an additional considerations for patrol officer workloads and response times to calls for service (D'Amico *et al.*, 2002). These considerations are discussed in this chapter.

2.2.2 The Police Districting Problem (PDP)

Police districts have been adopted in policing operations for many years. For most of the 20th century, police districts were created by police officers on a map with a marker-pen, by following the major streets, but without making much efforts to achieve a geographical or workload balance (Bruce, 2009). This is termed the Bud-Shell (Bruce, 2009) or map-and-marker (Camacho-Collados, Liberatore and Angulo, 2015) method. Since the seminal study by Mitchell (1972), a number of mathematical optimisation models have been proposed for the PDP. With advances in Geographic Information Systems (GISs), which provides affordable software and reasonable computation times, automated and advanced methodologies for creating police districts have gained popularity among practitioners (Wang, 2012).

In many countries, police departments adopt a hierarchical structure in the partitioning of territory. As a typical example, in the city of Buffalo, New York State (USA), this structure consists of the following layers (Sarac *et al.*, 1999): command districts (or precincts); patrol sectors (or beats); and reporting districts (or r-districts). Operationally, Buffalo is divided into five command districts, and patrolling operations among the command districts are independent of each other. Furthermore, each district is divided into several police sectors. The patrol operations among sectors of the same district are not independent. In fact, cars may pass through the borders of sectors in the same district. Each sector is further subdivided into about 400 reporting districts, the smallest geographical unit for which statistics are kept (Sarac *et al.*, 1999). In Greater London (UK), the area policed by the Metropolitan Police Service (MPS) is called the Metropolitan Police District (MPD), and the MPD is divided into 32 Borough Operational Command Units (BOCUs), which directly align with the 32 London boroughs. The patrol cars or officers assigned to a borough are expected to attend to all calls for service in the borough, and rarely pass through the borough borders.

The PDP concerns the optimal grouping of small units (e.g., r-district or census blocks) into a number of patrol districts, subject to planning criteria. Different criteria have been employed in the PDP, as introduced in the next section.

2.2.3 Criteria

This section provides an overview of the common criteria considered in the PDP and the various ways used to measure them. Most PDP applications adopt criteria ranging from workload balance, the geometry of districts, to response time and others. In the formulation of PDPs, some criteria are optimised in the objective function, while others are expressed as constraints.

2.2.3.1 Workload Balance

Defining the workload of a district is difficult, due to the complexity of policing operations and the variability of the tasks an officer may undertake. Nevertheless, an accurate definition of the workload is desirable, as it ensures homogeneity regarding the burden on police officers, and equalises the amount of preventive patrols and the quality of service in the districts (Bodily, 1978).

The first definition of the workload was given by Mitchell (1972), in which the workload was computed as the sum of the expected service and the travel time. Curtin et al. (2010) used the number of calls occurring in each district as a proxy for the workload. This definition has been criticised as being too simplistic, as different levels of service demand different service times. The workload has also been defined as the fraction of working time that an officer spends attending to calls (Bodily, 1978; Benveniste, 1985; D'Amico *et al.*, 2002). Recently, the workload has been computed as a combination of different characteristics. In Kistler (2009), a convex combination of total hours worked, the number of calls, and the population was adopted. Zhang and Brown (2014) considered both the average travel time and the response time. These definitions require specific data regarding the location, time, and service time of historical calls, and this requirement makes such definitions inapplicable when the data is unavailable. To solve this problem, Camacho-Collados et al. (2015) defined the workload as the crime risk in each basic unit. The crime risk can be informed by experienced practitioners, or computed by advanced crime mapping and predictive policing models. As the accurate estimate of the crime risk is an indicator of the potential need for a police service, it is a better measurement of workload in comparison to the number of historical calls.

The workload balance among districts is of critical importance. First, a balanced workload ensures the service quality of each district and avoids potential work overload and dissatisfaction of officers. Second, if the patrol car or officer is always busy responding to a call when another incident occurs, a car from a neighbouring district would have to respond. This leads to a domino effect (Mayer, 2009), meaning that cars pulled from their assigned districts would leave the district unattended, and therefore more vulnerable to criminal incidents.

2.2.3.2 Geometry of the Districts

In the context of the PDP, the geometric attributes of districts are relevant in terms of both efficiency and administrative reasons. A district with an appropriate shape would be easy to patrol and would reduce travel distances. The typical geometric criteria considered in PDPs include district size, contiguity, compactness, and convexity.

District size (Kistler, 2009; Camacho-Collados, Liberatore and Angulo, 2015). This criterion is quantified by the measurement of the area size (if polygon units are used) or the sum of the unit length (if line units are used).

Contiguity (Sarac *et al.*, 1999; D'Amico *et al.*, 2002; Camacho-Collados, Liberatore and Angulo, 2015). This criterion requires that each district remains connected so that the patrol team can traverse the district without having to going outside of it. If the basic units are lines or polygons, it is easy to define the neighbourhood relationship between the units, and a district is contiguous if the basic units of the district induce a connected subgraph in the overall neighbourhood graph (Kalcsics, 2015).

Compactness (Sarac *et al.*, 1999; D'Amico *et al.*, 2002; Kistler, 2009; Camacho-Collados, Liberatore and Angulo, 2015). This criterion requires that each district is geographically compact and undistorted. It can be measured by the diameter (i.e., the maximum distance between any pair of points) of a district (Camacho-Collados, Liberatore and Angulo, 2015) or the ratio of the longest Euclidean path and the square root of the area (D'Amico *et al.*, 2002).

Convexity (D'Amico *et al.*, 2002; Camacho-Collados, Liberatore and Angulo, 2015). A district is said to be convex if, for all pairs of units, the shortest path distance in the district is equal to the shortest path distance in the whole territory.

In PDPs, the criteria of contiguity, compactness, and convexity share a similar goal, which is to reduce the travel distance within the district.

2.2.3.3 Response Time

In policing, response time is the time between the arrival of a call for service and the arrival of a police unit at the location. A reduction in overall or average response time leads to multiple positive effects (Bodily, 1978), such as an increased likelihood of stopping a crime in progress, having a deterrent effect on criminals, and increasing confidence in the police.

Response time is the sum of two parts: the queueing time of the calls (i.e., the time from receiving the call to dispatching an officer), and the travel time to the location of the call. Most researchers only consider the travel times or the travel distances. The only study that has considered the queuing time is that of D'Amico et al. (2002), in which the Patrol Car Allocation Model was utilised.

2.2.3.4 Other criteria

Some PDP criteria have been introduced that do not fall into any of the categories discussed above. These criteria attempt to capture the complexity of real-world district planning. The Buffalo Police Department redesigned the r-districts such that the existing boundaries of the districts would be obeyed, and their use by other agents would be easy (Sarac *et al.*, 1999). For the Tuscon police in Arizona (USA), some additional considerations in redrawing their division boundaries included the boundaries of gang territories, city council wards, neighbourhood associations and the Davis–Monthan Air Force Base (Kistler, 2009). Moreover, Curtin et al. (2010) introduced backup coverage into patrol area design, which is achieved when more than one patrol officer can cover an incident within the service distance.

Apart from the criteria above, Wang (2012) discussed new criteria that should be considered in police districting, including minimising total costs (or response time) and the number of districts (or dispatch centres), and maximising equal accessibility.

2.2.4 Methods for PDP

Many different methods and approaches have been proposed to solve PDPs, that can roughly be divided into set partitioning methods, location-allocation methods, and meta-heuristics. Note that these are not mutually exclusive, and two or more methods can be combined to solve a PDP. A brief description of these methods is presented in this section. Detailed reviews on methods for generic districting problems can be found in Kalcsics et al. (2005) and Ricca et al. (2013).

2.2.4.1 Set Partitioning Methods

As districting is essentially a partitioning problem, classical set partitioning approaches can be used to solve a districting problem (Kalcsics, 2015). This usually follows two steps: 1) candidate districting plans that meet the given constraints are generated; 2) a districting plan is selected from the set of candidates that optimises the overall workload balance or other objectives. Mitchell (1972) proposed a set partitioning model to group geographic units into patrol beats. The model can be solved by analytical or heuristic methods. Due to the computational complexity, only small instances can be solved optimally.

2.2.4.2 Location-Allocation Methods

Location-allocation analysis is aimed at determining the optimal location for one or more facilities that serve a given set of spatially distributed demands. Each demand can be assigned to one or more facilities according to factors such as shortest distance or minimum transport cost, among others (Lei, Church and Lei, 2015). This approach can be used to solve the districting problem. Basic units correspond to the demands, and a district consists of the basic units that are assigned to a facility.

Curtin et al. (2010) applied the maximal covering model to define efficient police patrol areas. Their objective was to maximise the number of incidents served, while meeting an acceptable response time. In the case study, with 267 incident locations and five beats, this model was solved optimally using the CPLEX optimisation software for integer programming applications.

2.2.4.3 Meta-heuristics

A meta-heuristic is a high-level procedure that is designed to find a heuristic that can provide a sufficiently good solution to an optimisation problem, especially with incomplete or imperfect information or limited computational capacity. Meta-heuristics are well-suited to tackling the districting problems (Kalcsics, 2015), as such problems are computationally demanding, especially when they are medium or large-sized. For this reason, a wide range of meta-heuristics have been applied to districting problems, including simulated annealing, tabu search and genetic algorithms, among others.

D'Amico et al. (2002) proposed a simulated annealing algorithm to solve the PDP, subject to the constraints of response time, contiguity, compactness, convexity, and size. Liberatore and Camacho-Collados (2016) proposed and compared three meta-heuristic algorithms to tackle a multi-criteria PDP. The algorithms used include simple hill climbing, steepest descent hill climbing, and tabu search. Their case study, in the Central District of Madrid (Spain), demonstrated that the solutions identified by the tabu search outperform the other algorithms.

A major advantage of meta-heuristics in solving districting problems is their ability to generate sufficiently good solutions in an acceptable time. Another advantage is their flexibility to include almost any practical criterion and measure for the design of the districts (Kalcsics, 2015).

2.2.5 Summary of PDP

In summary, the existing research on PDP has involved a range of relevant attributes, and different methods have been utilised to solve the PDP. This lays the foundation for defining the PDP based on the street network, which is to be discussed in Section 2.4 and Chapter 5.

2.3 Police Routing Models

This section introduces the routing problems associated with police patrolling, which is based on the concept of crime hotspots and hotspot policing. After that, the routing problem is categorised into RPRPs and IPRPs, depending on the scenarios and objectives. A review of these two problems is provided below.

2.3.1 Police Patrol Routing Problem

In recent decades, empirical research has demonstrated that crime is concentrated in a range of spatial scales, from neighbourhoods and census blocks to street segments and street corners to individual streets (Rosser *et al.*, 2017). For instance, Sherman et al. (1998) reported that, in Minneapolis (USA), 50% of calls for service came from 3.3% of the city’s addresses, whilst Bowers (2004) found that 80% of thefts in bars in London (UK) happened in only 20% of the bars. There are varying definitions of crime hotspots, but with one common definition being an area with “a greater than average number of criminal or disorder events” (Eck *et al.*, 2005).

Hotspot policing involves the deployment of police resources to crime hotspots (Weisburd, 2005), and interventions include increased police patrols, problem-oriented policing, or offender-based strategies. Among them, hotspot patrolling is one of the most important strategies. The effectiveness of hotspot patrolling in reducing crime and disorder has been proved by a range of experiments, such as the Minneapolis Hot Spots Patrol Experiment (Sherman & Weisburd 1995), and another study conducted in Philadelphia (USA) (Ratcliffe et al. 2011). In patrolling, when there is more than one hotspot to cover, police officers will typically rotate randomly between the hotspots, as, for example, in the field trial in Sacramento, California (USA) (Telep et al. 2012). This randomised strategy cannot be applied to situations where police resources are limited and there are many hotspots areas, however. Rather, the successful operation of patrolling to cover the hotspots requires a detailed patrol routing strategy.

Informally, depending on the specific objectives and the ratio of patrolled targets to patrollers, patrol routing problems can be categorised into two types: infrequent and repeated patrol routing problem. The infrequent patrolling concerns covering each target (hotspot) at least once in the time interval, in cases where available patrol forces are limited and the targets are many. On the contrary, the repeated patrolling aims to patrol the targets (hotspots) repeatedly in a given time interval, using a fairly large patrol team. These two problems are different in their application situations, measures, and objectives, and will be reviewed separately.

2.3.2 Infrequent Patrol Routing Problem

Here, an overview of the infrequent patrol routing problem (IPRP) is provided. IPRPs are concerned with covering each target once or twice in the given time interval, as the available patrol forces are limited and the targets are many (Wolfler Calvo and Cordone, 2003; Willemse and Joubert, 2012). As this study focuses on the patrol routing problem on the street network, in which the patrol targets are certain street segments, the corresponding IPRP can be modelled as an arc routing problem (ARP).

This part begins with an overview of the ARP, including its definition and applications. Then, an introduction to the Rural Postman Problem (RPP) is provided, as this is an important variant of the ARP and is relevant to the IPRP. This is followed by a review on the approaches for and solution evaluation of RPP and its variants. After that, an overview of the existing research on the IPRP is presented.

2.3.2.1 Arc Routing Problem (ARP)

Informally, routing problems include any problem that involves of generating a tour, or a set of tours, on a network or a subset of a network, given a set of constraints and the requirement to optimise one or several fixed objectives (Jozefowiez, Semet and Talbi, 2008). Generally, research related to routing problems falls into three categories: node routing, arc routing, and general routing problem. One essential difference between them is the coverage objective. In the node routing problem (NRP), the objective is to find a route which covers a subset of nodes. In ARPs, the objective is to find a route that covers a subset of required arcs. In general routing problems, a subset of nodes and a subset of arcs have to be covered on the objective route. ARPs are more relevant to this research, as this study was focused on traversing the hotspot segments.

According to Dror (2000), ARPs have many real-world situations. The routing of security guards (Wolfler Calvo and Cordone, 2003; Willemse and Joubert, 2012; Shafahi and Haghani, 2015), and school buses (Delgado and Pacheco, 2001), the delivery of newspapers to customers (Applegate *et al.*, 2002), waste collection (Lacomme, Prins and Sevaux, 2006), and snowplow routing (Rao, Mitra and Zollweg, 2011) are existing applications of ARPs.

2.3.2.2 Rural Postman Problem (RPP)

Three important ARPs have been derived from the general routing problem: the Chinese Postman Problem (CPP), RPP, and Undirected Capacitated Arc Routing Problem (UCARP). The CPP was introduced by Kwan (1962). Its aim is to detect the shortest tour that traverses a complete arc set, *E*, of a given graph, *G*. The RPP was introduced by Orloff (1974); in this, not all of the arcs need to be traversed, and only a given subset, *R*, of arcs are required to be covered by the tour. Note that the CPP can be seen as a special example of the RPP, as the RPP is transformed into the CPP if $R = E$. The UCARP was introduced by Golden and Wong (1981), in which each arc has a predefined traversal cost, and only certain arcs are required to be traversed. There is a single depot that contains multiple postmen, each with limited capacity. The objective of the UCARP is to determine a set of minimum total costs, where every tour starts and ends at the depot, the total demand covered by each tour does not exceed the vehicle's capacity, and each required edge is serviced. It is worth noting that the RPP is a special case of the UCARP, in which the vehicles are not capacitated, and the demand on each arc in *R* is the same (or, for example, all the arcs in *R* have the same weights or priorities). Among these problems, the RPP is the most relevant to this research and is the topic of the following discussion.

Before further discussion, however, it is worth noting that the terms 'postman' and 'postmen' represent the persons or vehicles that travel to carry out the given tasks, and the 'depot' represents the location(s) where the postmen start and/or end their tour. In the context of police patrolling, 'depot' refers to the police station(s) where the officers start and end their patrols.

Studies of the RPP have been expanded to address different aspects. First, it can be further broadened to include multiple postmen, known as k-RPP. The objective there is to minimise either the total distance travelled, or the total time taken, by the postmen. Second, the objective of the k-RPP can be adjusted to minimise the length of the longest route, this new problem being known as the min-max k-RPP (MM k-RPP, or MMRPP). As Ahr and Reinelt (2006) stated, the min-max objective function is preferable when the required edges have to be traversed as early as possible, and when more balanced routes or workloads are required. Other relevant contributions include those of Arkin et al. (2006), Benavent et al. (2009) and Willemse & Joubert (2012).

Another important extension of the RPP is that in which postmen belong to different depots, called a multi-depot RPP (MDRPP); however, the MDRPP has received minimal attention in the literature. The only contribution to this problem in the literature is that of Pereira & Fernández (Pereira and Fernández, 2015). On the other hand, the MDRPP can be understood as a special case of the multi-depot UCARP, in which the depots are incapacitated, such that the existing solution for the multi-depot UCARP can be adapted to solve the MDRPP. The problem, however, is that the adapted algorithm is not efficient for the MDRPP. A similar question arises with the Multi-Depot k-CPP (MDCPP) (Platz and Hamers, 2013; Akyurt, Keskinturk and Kalkanci, 2015; Shafahi and Haghani, 2015), but this problem requires that all arcs in the graph be traversed, and so is different from the MDRPP.

2.3.2.3 Existing Research on IPRPs

There are several studies that have applied ARP models to solve the IPRP. Wolfler Calvo and Cordone (2003) introduced the overnight security service problem, which designs routes for the overnight patrolling of streets and the inspection of buildings. The model obtained includes a number of objectives, including cost reduction, a high level of service, an effective response to alerts, fair task assignment, and practicability. This is the first paper to have formulated and solved an IPRP. Recently, Willemse and Joubert (2012) applied the Min-Max k-Postmen Problems to design routes for security guards. The objective was to minimise the length of the longest route, so as to impose a balanced workload. This problem was solved by a two-step heuristic approach, with tabu search being the second step used to improve the solutions. Both approaches, however, assumed that the patrollers started from the same depot, failing to consider a multiple depot problem.

2.3.2.4 Summary of the IPRP

In summary, existing approaches for the arc routing problems are not suitable to solve the infrequent patrol routing problems (IPRP), as there is lack of research regarding the multi-depot routing problems. In police patrolling, there are multiple officers that start from different police stations (i.e. depots), and they cooperate to cover the crime hotspots. Therefore, more research is needed to solve the multiple-depot problem that aims to balance the route lengths, and the new model and algorithm can be used to solve the IPRP, which is the focus of the Chapter 6 in this thesis.

2.3.3 Repeated Patrol Routing Problem (RPRP)

This subsection provides an introduction to RPRPs. Note that the police RPRP is similar to the multi-agent patrolling problem (Almeida *et al.*, 2004), or multi-robot patrolling problem (Portugal and Rocha, 2011), which focuses on surveillance tasks using multiple mobile robots that visit important locations in the environment frequently; both are reviewed here.

2.3.3.1 Definition

Informally, RPRP are used in planning routes for patrollers to repeatedly and effectively cover predefined targets. There are several essential issues in their definition. First, such problems are not restricted to police patrols. They include applications in police patrolling to cover crime hotspots, in multi-robot patrol to monitor important sites, and in the patrolling of estates, as these applications bear many similarities. Second, one basic assumption for this problem is that the targets have been identified and provided. In terms of crime and policing, the targets are crime hotspots, which are usually identified by crime mapping techniques (Ratcliffe, 2010). Third, this problem should be distinguished from the coverage problem that determines the efficient spatial distributions of police patrol areas to provide maximal and multiple coverage of incidents (Curtin, Hayslett-McCall and Qiu, 2010). The latter problem is focused on the location of the centres of patrol areas, and do not consider a detailed routing strategy for the patrol teams.

As the PRPR focuses on repeated patrolling of targets, the measure that is usually used to evaluate the strategy performance is the average time lag between consecutive visits to the targets, which is called idleness in the literature (Machado *et al.*, 2002). The basic objective of the routing is therefore to minimise the average or maximum idleness among the targets.

2.3.3.2 Routing Strategies for Repeated Patrolling

Varying solutions have been presented for planning repeated patrol routes. Depending on the basic ideas, they can be categorised into greedy strategies (Machado *et al.*, 2002; Almeida *et al.*, 2004; Portugal and Rocha, 2013a), operations research strategies (Chevaleyre, 2004; Elmaliach, Agmon and Kaminka, 2009; Portugal and Rocha, 2010), interaction strategies (Reis *et al.*, 2006; Tsai *et al.*, 2010), other strategies (Sempe and Drogoul, 2003; Santana *et al.*, 2004; Chu *et al.*, 2007; Chen and Yum, 2010).

Greedy strategies. These use simple architectures to guide patrollers to attend places that have been visited less recently or are within shorter distances. In the literature, this is sometimes called simple pioneer strategy (Portugal and Rocha, 2013a). Such simple strategies achieve good performance in covering hotspots and coordinating patrollers. Moreover, in order to incorporate multiple factors that could influence decision-making, Portugal and Rocha (2013b) developed a Bayesian-based framework to include multiple factors in deciding which target to patrol in the next step; however, most of the greedy strategies have been developed in the context of robot patrols, and include no consideration of the special aspects of police patrolling, such as the unpredictable routes to prevent offenders from becoming familiar with the patrol, and being robust to the need for an emergency response.

Operation research strategies. These use graph theory tools to compute low-cost cycles and efficient routes for each patroller. The tools include the Travelling Salesman or Hamiltonian cycle (Elmaliach, Agmon and Kaminka, 2009; Smith and Rus, 2010; Pasqualetti, Franchi and Bullo, 2012), spanning trees (Gabriely and Rimon, 2001; Fazli and Davoodi, 2010) and graph partitioning (Sak, Wainer and Goldenstein, 2008; Stranders *et al.*, 2013). For instance, Elmaliach et al. (2009) developed a Hamiltonian cycle-based strategy for patrolling that guaranteed that each target point was covered at the same optimal frequency. The Multilevel Subgraph Patrolling strategy, presented by Portugal and Rocha (2010) uses balanced graph partitioning to assign sub-regions for each patroller and subsequently computes an effective path for each patroller based on Hamiltonian cycles, or similar approaches.

However, such strategies are naturally deterministic, which would more easily allow intelligent criminals to predict the patrol routes and take advantage of the idle time between the visits of the officers. Additionally, Hamiltonian cycles, among other algorithms, have a high computational complexity and are difficult to generalise to include large numbers of targets, compared to the greedy strategy. Moreover, such strategies would have to re-compute the patrol routes if the number of patrollers was to change because of an emergency response.

Interaction strategies. These firstly formulate interaction models between officers and offenders, and then use agent-based simulations or game theory models to obtain the optimal strategy for deterring offenders. For example, Reis et al. (2006) designed patrol routes based on genetic algorithms and a multi-agent-based simulation, wherein a set of criminals frequently tried to commit crimes while officers tried to prevent crimes. Notably, Tambe et al. (2013) have been leading the research field in applying computational game theory to security, which includes, but is not limited to, police patrolling. Their basic goal is to derive strategies for safe guards to protect an area, based on modelling the interactions between the two sides as an attacker-defender Stackelberg game, where a player always predicts their opponent's behaviour and chooses the best response. The details of their work are described in Tambe et al. (2013).

Other strategies. These use complex approaches such as task allocation (Sempe and Drogoul, 2003), reinforcement learning (Santana *et al.*, 2004), cross-entropy (Chen and Yum, 2010) and swarm intelligence (Chu *et al.*, 2007). For example, reinforcement learning is adopted to solve the patrolling problem by automatically adapting the agents' strategies to the topology of the environment (Santana *et al.*, 2004). Chen and Yum (2010) developed an algorithm for patrol route planning based on a cross-entropy method. However, strategies like reinforcement learning and cross-entropy are very complex in nature, and may be suitable for designing patrol routes for a single agent but struggle to extend to cooperative patrolling among multiple agents.

Although such strategies provide a novel perspective on the patrol routing problem, the validity of the interaction model remains to be discussed further. It requires substantial knowledge and behavioural research to create and verify such models, which is difficult.

Moreover, it is difficult to generalise such strategies to effectively guide police patrols in large areas and prevent crimes of multiple types.

2.3.3.3 Summary of the RPRP

In summary, existing approaches for solving RPRPs are not applicable in guiding police patrols, as they omit the peculiarities and challenges specific to police patrols, such as the unpredictable patrol routes to prevent offenders from being familiar with the patrolling, and being robust to the influence of emergency response. The issues of police patrolling need to be specified and formulated using clear guidelines, and require to be quantified by appropriate measures, in order to define a new patrol routing strategy.

2.4 **Incorporating Street Networks into Police Patrolling**

The street network structure is an essential part of the urban space, fundamentally shaping the human movement patterns. According to recent findings in criminology, street networks can influence long-term crime patterns (Davies and Johnson, 2015; Summers and Johnson, 2017), as well as the short-term dynamics of crime (Davies & Bishop 2013; Johnson & Bowers 2014). Therefore, it is of essential importance to incorporate street networks into the design of police patrol strategies. In this section, relevant street network representation and metrics are first introduced, followed by a description of street network analysis in Geographic Information Science (GIScience) and criminology. After that, a review of street networks in police districting models is presented.

2.4.1 Street Network Representation and Metrics

In the analysis of street network, the first step is to represent the structure of a street network as a network using the terminology of network theory, which is based on the mathematical field of graph theory (for an introduction, Bollobás 2002). In general, a network $G = (V, E)$ consists a set of nodes, V, and a collection of links, E, between pairs of nodes. If a link exists between two nodes i and j, the nodes are said to be adjacent and the link can be denoted by the unordered pair of node (i, j).

There are multiple ways to translate a street network to these terms, and the choice is dependent on the analysis task. One natural and popular representation is the ‘primal representation’ (Porta, Crucitti and Latora, 2006b), in which each junction of the street network is represented by a node, and a link is added between any two nodes which are connected by a street. In contrast, another choice is the ‘dual representation’ (Porta, Crucitti and Latora, 2006a), in which the whole streets (sets of street links which are associated based on street name or geometry) are represented as nodes, and links are placed between any two nodes that share a junction. Although this representation leads to the loss of any reference to geographic distance, it allows for "conceptually countless" edges for each node in the dual graph that is not dependent on the availability of geographic space (Porta, Crucitti and Latora, 2006b), thereby making the dual graph of a street network comparable in its structure with most other networks in society and biology.

In the study of street networks, and spatial networks more generally, some of the classical metrics used in network analysis are of little relevance. Here, two metrics are reviewed, which include the degree of a node and the betweenness centrality.

The degree of a node is defined as the number of links coincident with it. For many networks in biology or society, this metric can take a wide range of values and offers great insight, including the study of scale-free property of the network. However, in the analysis of street networks, the physical constraints of geographic space mean that it can take a very limited number of values, as it is very rare to find a junction at which 6 or more segments meet (Davies and Johnson, 2015). The vast majority of nodes have degree below 4. This issue is less severe for a dual graph, as it allows for "conceptually countless" edges for a node.

A more meaningful metric for street networks is the betweenness centrality of a street segment, which is concerned with travel through network. To explain this, the concept of a *path* is introduced. A path in a network is an ordered sequence of nodes such that every consecutive pair of nodes is connected by a link. The length of a path can be defined in two ways: using topological distance as the number of links it comprises, or in metric terms, as the sum of the physical length of all constituent links. These two path lengths have different interpretations: metric distance is a true measure of cost, but topological distance is more representative of distance as perceived by a human being navigating the network (Hillier and Iida, 2005). Clearly, given any pair of nodes i and j, it can be determined whether a path between them exists, and the shortest path which features minimal length, if such path exists. The shortest path length is denoted as d_{ij}.

If σ_{ij} denotes the total number of shortest paths between i and j, and $\sigma_{ij}(e)$ denotes the total number of shortest paths between i and j that traverses the link e, then the betweenness centrality B_e (Freeman, 1977) of a link e can be defined as:

$$B_e = \sum_{i,j \in V, i \sim j} \frac{\sigma_{ij}(e)}{\sigma_{ij}} \tag{2.1}$$

, where ~ represents the relation “there exists a path between *i* and *j*”. Betweenness has a clear interpretation in real-world terms as an estimate of the use of a link by traffic (either pedestrian or vehicular) traversing the network. The metric value of a given link reflects the role in the network as a whole.

2.4.2 Street Network Analysis in GIScience

In GIScience, spatial assembly or aggregation is a necessary step when modelling the geographical environment; in fact, it is fundamentally important in selecting basic units and spatial resolutions in building geo-related models. While some studies on mobile phone data have used Voronoi polygons, defined by base towers (Shi *et al.*, 2015; Lai, Cheng and Lansley, 2017), most studies have aggregated data based on grids (Reades, Calabrese and Ratti, 2009; Sun *et al.*, 2011; Adepeju, Rosser and Cheng, 2016). In other studies, census blocks have been adopted as the spatial units. While areal units are common in varying applications, the definition of such units can be subjective and arbitrary. The sensitivity of spatial analysis results to the choice of zoning system in which data are aggregated is known as the modifiable areal unit problem (MAUP) (Openshaw and Taylor, 1979; Openshaw, 1984). More precisely, the MAUP contains two related, but distinctive components; the scale effect represents variations in results that may be obtained when the same areal data are combined into sets of larger areal units of analysis, while the zoning effect represents variations in results due to alternative units of analysis when the number of units is fixed. As a matter of fact, the challenge of selecting the most appropriate scales or zoning structures has been a long-standing issue for geographers and practitioners. Although great efforts have been made to understand and eliminate MAUPs, there is still no consensus on the best unit to use for spatial analysis and modelling.

In many geographical applications, streets represent a promising and meaningful unit of spatial analysis and modelling, in order to characterise human activity and understand human movement pattern. For many people, streets and paths are the predominant elements in their perception of a city, and allowing them to observe the city as they move through them (Yin and Wang, 2016). Empirical research reveals that physical movement in the urban space is often constrained by a road network. Therefore, in comparison to traditional areal units, the street network represents an approximate decomposition of a block or an area, which is capable of minimising the effect of the MAUP during spatial analysis and modelling.

There is limited research on street network analysis in GIScience. Okabe and his collaborators extended spatial analysis models, involving planar space, to the street network, as exemplified by the point distribution and kernel density estimation on a network (Okabe, Yomono and

Kitamura, 1995; Okabe, Satoh and Sugihara, 2009), as well as the market area analysis and the demand of retail stores on a street network (Okabe and Kitamura, 1996; Okabe and Okunuki, 2001). They also formulated a theoretical framework called uniform network transformation, to transform a non-uniform network into a uniform one before conducting a points pattern analysis (Okabe and Satoh, 2006). These works have laid the foundations for more advanced spatial analysis and applications based on street networks.

Meanwhile, space syntax provides a theoretical and methodological framework for the analysis of street network configurations with respect to human space organisation (Hillier, 2007). The general idea of space syntax is that spaces can be broken into components, analysed as networks of choices, and then represented as graphs that describe the connectivity and integration of those components. It rests on three basic conceptions of space, namely *isovist*, *axial space*, and *convex space*. An isovist is the field of view from any particular point. An axial space is a straight sight-line and possible paths. A convex space is a space where, if imagined as a wireframe diagram, no line between any pair of points goes outside its perimeter. In other words, all points within the polygon of the convex space are visible to all other points in the polygon.

Space syntax provides various ways of analysing a street network, the most popular ones being *integration*, *choice*, and *depth distance*. First, integration measures how many turns are needed from a street segment to reach all other streets in the network via the shortest path. It shows the cognitive complexity of reaching a street, and arguably represents the pedestrian popularity of a street. Second, the choice measures the number of intersections that need to be crossed before reaching a street. It can be understood as a 'water-flow' in the street network, where each street is given one unit of water, and the water moves to all connected streets and is divided equally amongst the splitting streets when an intersection occurs. Third, the depth distance measures the linear distance from the centre point of each street to those of all other streets, and the streets with the lowest depth distance values are the nearest to all other streets.

Space syntax has been utilised to describe how easily navigable a space is in the design of museums and hospitals, among other settings. It has also been adopted to predict the correlation between spatial layouts and social effects, such as crime (Jones and Fanek, no date; Hillier and Sahbaz, 2007) and traffic flow (Croxford, 1999; Lerman, Rofe and Omer, 2014).

2.4.3 Street Network Analysis in Criminology

In criminology, the role of urban configurations in shaping crime patterns is a recent research topic, and is an essential component of environmental criminology (Brantingham, Brantingham and others, 1981; Reyns, 2016). Here, a brief review of the environmental criminology is given, followed by an introduction to the role of the street network in criminology.

Environmental criminology is an approach that focuses on the criminal act and the circumstances that lead to it, instead of the personal characteristics of the offender. Routine activity theory (Cohen and Felson, 1979) is central to environmental criminology, focusing on the conditions under which a crime occurs. In particular, a crime can only take place under the convergence of three factors in space and time - a motivated offender, a suitable target, and the absence of a capable guardian (Cohen and Felson, 1979). Consequently, the question shifts to how and how much such a concurrence might arise in a realistic setting. A popular and somewhat simplistic explanation is premised on the assumption that the majority of crime is opportunistic. Specifically, offenders encounter targets during non-criminal activities, and crime patterns are simply a manifestation of heterogeneities in the target distribution and awareness (Davies and Bishop, 2013). Hence, the conditions suitable for crime might be a result of the cumulative activity patterns of the public.

Going further, pattern theory (Brantingham and Brantingham, 1993) is aimed at adding more detailed geographic considerations, by situating the aforementioned concepts in a realistic urban environment. In detail, it considers both how activity patterns are shaped by configuration of space (or the *urban form*) and how the physical features of certain areas influence the decision making of a possible offender. The term 'urban backcloth' has been adopted to refer to the layout of the built environment, with stress on the elements relating to routine activities (including homes and workplaces). The implied activity patterns can be reconciled with types and levels of criminality. As a natural step, the notions of "nodes, paths and edges" (Brantingham and Brantingham, 1993) have been introduced into this theory, as a means of encoding the urban backcloth and its characteristics.

Street network-based models play an important role in the description of crime patterns in many aspects. First, street-level analysis is well-aligned with a trend towards the use of micro-units within criminology (Brantingham *et al.*, 2009), and crime variability among different areas can be explained at fine spatial scales including the street level. Steenbeek and Weisburd (2016) demonstrated that 58-69% of the variability of crime could be explained in terms of street segments. This highlights the capability of street-level analysis in reducing or eliminating the so-called ecological fallacy (Robinson, 1950). This assumption asserts that (crime) risk will be uniform across an area, which is unlikely to be true when streets that are co-located with an areal unit (e.g., grid cells) experience very different risks. Examining crime rates at the street level can help in identifying high-risk segments, which has clear implications for accurate crime prediction.

Second, the features of the street network could potentially influence overall or long-term levels of crime. Specifically, many studies have shown an association between the relevant attributes of street segments (e.g. permeability, betweenness, street type, geometric linearity) and the crime risk, which is well predicted by crime pattern theory. Beavon et al. (1994) found out that the risk of varying types of crime in Ridge Meadows (UK) was positively correlated with the number of roads that a street segment is connected to. Lasley (1996) showed that homes located on cul-de-sacs experience a lower risk of violent crime than those located on through roads in Los Angeles (USA), and a similar relationship has been reported for burglary risk (Armitage, 2007). Using data from Merseyside (UK) and several simple graph theoretic metrics, Johnson and Bowers (2010) showed that the burglary risk on streets was associated with the street type and the number of different types of streets that the given street was connected to. The more major roads a street was connected to, the greater the risk on that street. This study also confirmed that cul-de-sacs were at significantly lower risk. More recently, Davies and Johnson (2015) used the graph theory metric of betweenness to estimate the possible usage by people of street segments. As predicted by crime pattern theory, they found that the measure of betweenness of a street is a highly significant predictor of burglary risk. A similar pattern was discovered for incidents of serious outdoor violence in London (UK), based on the space syntax theory (Summers and Johnson, 2017).

Third, the street network also plays a role in the short-term dynamics of crime, and influences the diffusion of crime risk in the physical space. It has been observed that crime clusters in both time and space, such that when an event occurs at one location, there is a short-term increase in the probability that other events will occur nearby, known as near-repeat victimisation (Johnson and Bowers, 2014). Considering the movements of offenders and their awareness spaces, it is reasonable to suggest that networks serve as a means of the risk diffusion. In particular, such risk might spread not just to neighbours, but also to locations situated in the directly connected segments (Davies *et al.*, 2013). Moreover, in the urban space, the street network determines what it is for two places to be ‘near’, and determines the physical pathways along which offenders can travel and develop their awareness, and thus along which the risk diffuses (Rosser *et al.*, 2017).

In line with the role of the street network in the description of crime patterns, it is reasonable to use the street network in the design of police patrolling strategy and police deployment. In particular, the use of the street network has practical implications for policing from various aspects.

First, in the prediction of crime hotspots, the network-based kernel density estimation (KDE) model compares favourably with a grid-based model in terms of predictive accuracy. Rosser et al. (2016) demonstrated that the network-based model identified nearly 20% more crime at a coverage level of 5% (meaning that the identified hotspot area occupies 5% of the total area) using property crimes data from the UK. Another notable finding was that this improved accuracy is statistically significant at coverage levels from 1 to 10%. Figure 2.1 shows the mean hit rate of the two prediction approaches for varying levels of coverage from 0% to 20%. The hit rate is defined as the proportion of crimes captured relative to the total crime number in the testing data. The background shading indicates coverage levels when the network KDE hit rate is significantly higher than the grid KDE at a significance level of 5%, tested using Wilcoxon's signed-rank test (Rosser *et al.*, 2017). This implies that network-based methods for crime forecasting are likely to outperform grid-based ones, and should be used in police operations.

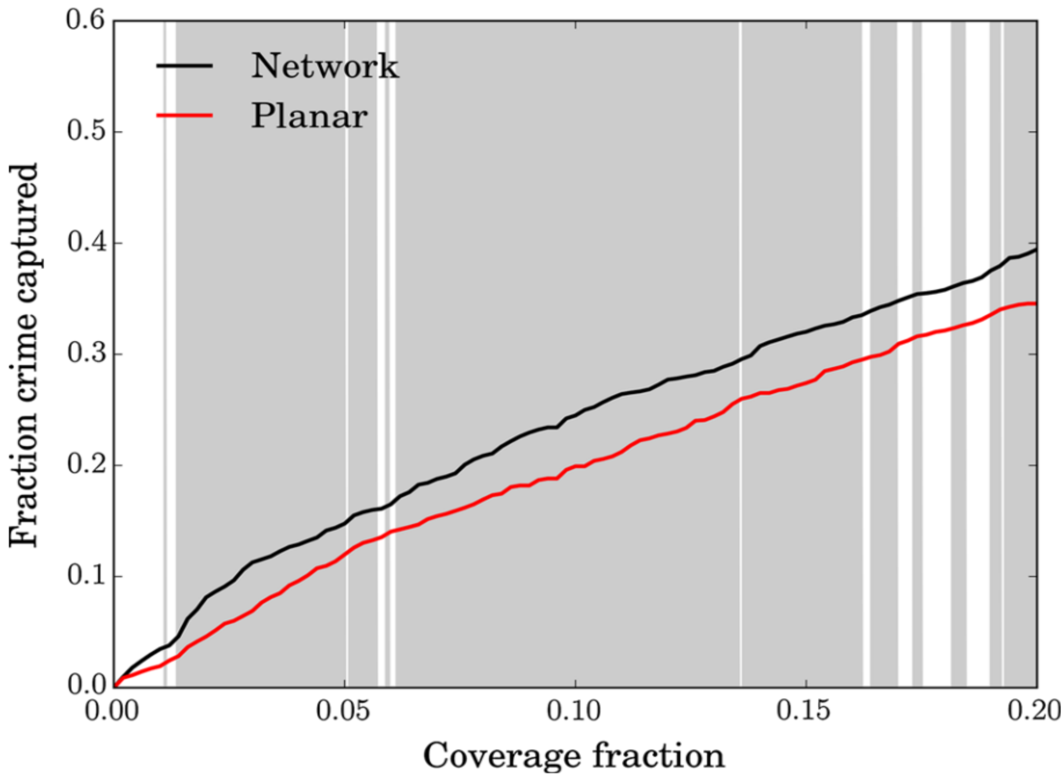


Figure 2.1 Mean hit rate versus coverage for network and grid-based KDE prediction approaches. The results were computed over 90 consecutive days. (Rosser *et al.*, 2017)

Second, street segments are well-defined targets for operational policing, as they are naturally more connectable and more permeable than areal units, such as grids. Generally, grid cells many intersect physical features (e.g., lakes) or barriers (e.g., railway tracks and gardens), and might include several unconnected streets. These make patrol plans unclear and difficult to implement (Rosser *et al.*, 2017). It seems reasonable to assume that when faced with a crime hotspot map that defines areas rather than streets, the patrolling officers would stick to the physically reachable routes within the squares, thus making large areas of the map redundant. Therefore, a patrol plan based on a street network is more physically viable and usable than the one built on grid cells.

Third, the street network plays a fundamental role in route design and performance evaluation for police patrols. As the street network constrains physical movement in the urban space, using the network distance would be a means of making the patrol routes usable and accessible. This applies to not only foot patrols but also vehicular patrols. Moreover, planning routes based on the street network enables performance evaluation of patrolling officers. Put simply, the question regarding whether a patrolling officer actually followed a planned route can be addressed in quantitative terms, by comparing the planned route with routes reconstructed from the Global Navigation Satellite System (GNSS) traces. Subsequently, the evaluation result could serve as feedback for strategic planning, thereby improving the operational deployment.

Because of the important role of the street network in police patrolling, there has been a growing body of research that has incorporated the street network into the design of police districting and routing models.

2.4.4 Street Networks in Police Districting Models

In formulating a PDP, the first step is to choose the basic units that can be consolidated into districts. Most PDP models use areal units or zones (e.g., grids, census blocks) as basic units (Mitchell, 1972; Zhang and Brown, 2013). These units are simple to define and readily usable. Moreover, using census blocks allows for the consideration of demographic data, and such created police sectors are thus suitable for use by other agencies (Sarac et al. 1999).

Alternatively, street segments can be used as the basic units in PDPs; there are a number of reasons why this may be advantageous. First, they are well aligned with the trend of street-level analysis in criminology, and could incorporate the theoretical and methodological progress on crime pattern descriptions and crime prediction. Second, according to Section 2.4.3, using street network segments as basic units creates districting plans with better usability, in comparison to grid cells. Third, a street level districting plan is able to mitigate the effects of the MAUP (see Section 0), which is inherent in grid-based districting models (Camacho-Collados, Liberatore and Angulo, 2015). Specifically, a districting plan may largely depend on the size of the grid cells, meaning that an alternation in cell size might lead to a substantial change in the resulting districting plan. However, in practice, the size of the cell unit is often determined empirically, leaving the districting plan unstable. In contrast, as street segments are enduring and well-defined, using them as units could reduce, or prevent, the effects of MAUPs and ecological fallacy (Robinson, 1950)

Despite the advantages of using street units in PDPs, few models have sought to use streets as basic units, although several relevant studies have used network distance. Recently, the Police Patrol Area Covering (PPAC) model (Curtin *et al.*, 2005; Curtin, Hayslett-McCall and Qiu, 2010) aimed at an efficient distribution of police patrol centres for maximal coverage (or backup coverage) of historical incidents, and new beats were formed by aggregating sectors, based on incidents served by the patrol centre. Figure 2.2 illustrates the new beat designs (i.e., districts). Although the network distance between the patrol centre and the incident is adopted, the district design is still based on existing sectors, failing to derive patrol districts based on the underlying street network. This direction was one of the objectives of the present study.

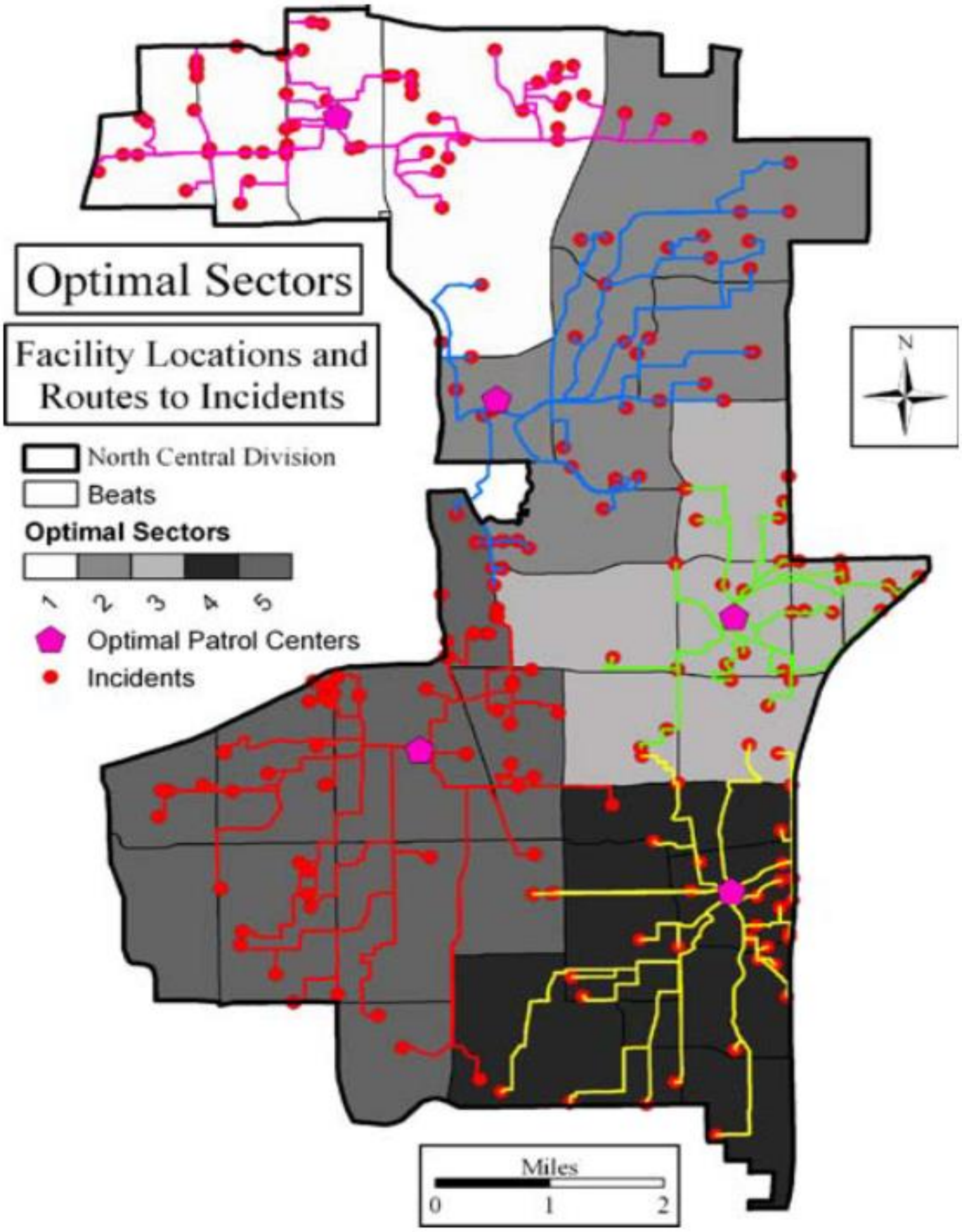


Figure 2.2 Example of optimal design of beats based on sectors and routes to incidents

(Curtin, Hayslett-McCall and Qiu, 2010)

2.5 **Chapter Summary**

This chapter provides a critical literature review of research fields relevant to police patrol strategies.

The PDP requires the grouping of basic spatial units into contiguous and balanced geographic clusters (i.e., districts) so as to guarantee efficient patrol coverage. Most of the existing studies on the PDP are based on areal units, giving very limited consideration to the underlying street network.

In designing routes for police patrols, it is necessary to take into account the unique characteristics of the police patrol task, such as multiple patrol officers, multiple depots, and the need to balance the route length. Moreover, depending on the objective, the PRP can be categorised into repeated and infrequent patrol routing, which should be considered separately.

The street network is a key determinant of the urban space configuration, and is the substrate on which crime and policing take place. Although street network analysis has received considerable attention in GIScience and criminology, more work needs to be done to incorporate the street network into the design of police patrol strategy, including districting and routing strategies.

Chapter 3

METHODOLOGY

3 METHODOLOGY

The literature review demonstrated several critical research gaps regarding police patrol strategies based on the urban street network. First, to my knowledge, very few existing police districting models use streets as their basic units, which neglects the effects of streets on the design of police districts. Second, the existing routing strategy that is relevant to police patrolling fails to incorporate the unique characteristics of police patrolling such as that officers start from multiple police stations. Third, different strategies for police patrolling have been designed separately, and there is a lack of a framework to combine the districting and routing strategies of police patrols.

This chapter presents the methodology that underpins this study. In Section 3.1, a framework for police patrol strategies is proposed. Section 3.2 contains a brief description of the three specific strategies proposed as a result of in this study, as well as the corresponding approaches for deriving the solutions. Section 3.3 provides an introduction to the main research methods used in this thesis. Finally, Section 3.3 is a summary of this chapter.

3.1 A Framework for Police Patrol Strategies

As mentioned in Section 1.2, the main aim of this research was to propose a framework for developing police patrol strategies based on street networks, and the framework should include strategies that are well suited for varying conditions and requirements. The methodological framework introduced here contains only the integration of the different modules that have contributed to realising this aim. The complete dataset, methodology and detailed algorithms employed during each module of the work are further discussed across the following chapters (Chapters 4, 5, 6, 7, and 8) of the thesis. Among these, Chapter 4 includes a description of the datasets used in this thesis. Chapter 5 contains a description of the police districting model. Chapter 6 discusses the infrequent coverage routing in police patrolling and describes a balanced route design. Chapter 7 includes an online patrol routing strategy for repeatedly covering the crime hotspots and the street network. Following that, Chapter 8 discusses the coupling of different police patrol strategies.

The methodological framework is illustrated in Figure 3.1.

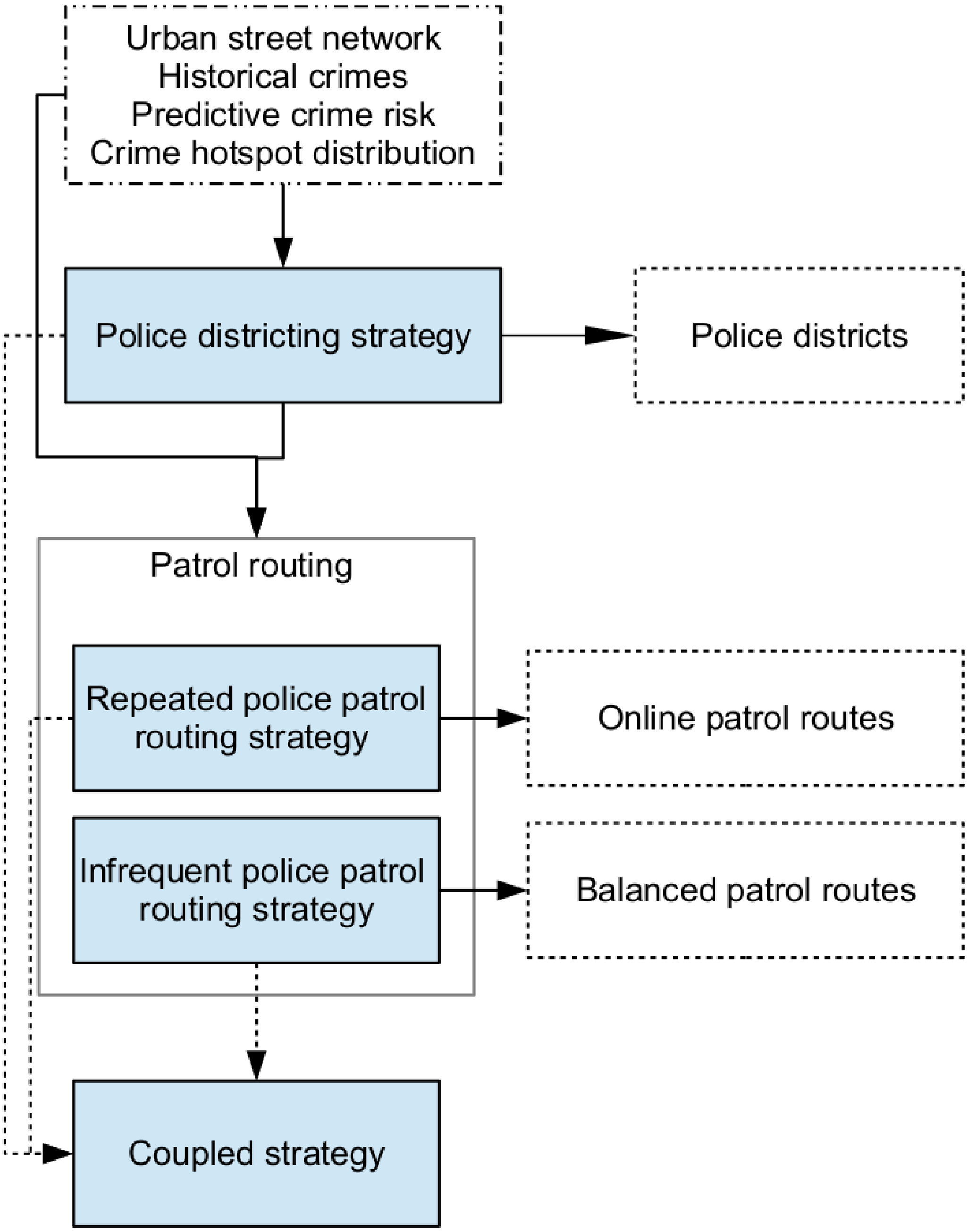


Figure 3.1 The methodological framework for police patrol strategies used in this study

In this framework, the districting and routing strategies have been combined to form an integrated police patrol strategy. Input to this framework includes the urban street network, historical crime records, the predictive crime risk on the streets, and the spatial distribution of crime hotspots. Given the territory to patrol, the first step is to design the police districts. If there is an existing system of police districts, the districting strategy can be used to adjust or redesign those districts. Then, for daily policing operations, patrol route planning is an essential procedure for guaranteeing the coverage of crime hotspots, and to maintain the efficiency of the patrol service. Depending on the objectives behind covering the crime hotspots, the patrol routing strategy can further be categorised into repeated and the infrequent patrol routing strategies. While the repeated patrol aims to patrol the targets (hotspots) repeatedly in a given time interval, the infrequent patrol is more concerned with covering each target at least once in the time interval using a limited patrol force.

These patrol strategies can be coupled to form new strategies. The coupling approach is promising, as it is capable of involving greater complexity in the patrolling procedure in the strategy design. Moreover, such a new strategy may be more efficient, and would bring about new features. There are a number of ways to couple the strategies. One possibility is to incorporate patrol districts into the routing of officers for repeated coverage, with each patrol officer being given a patrol district, and being restricted to that district. This leads to a number of advantages. First, less time is spent on travelling between different areas or districts, so that an officer can spend more time on patrolling. Second, it is easier for an officer to return to their station to deal with paperwork or other tasks that are associated with the patrolling, if necessary.

3.2 **Approaches for Deriving Police Patrol Strategies**

In this section, a brief introduction to the three patrol routing strategies used in this study is presented. For each strategy, a description of how the problem is formulated, and the corresponding approaches for deriving the strategy, are provided.

3.2.1 Police Districting Strategy

For the police districting, a street network PDP (SNPDP) is proposed that explicitly uses streets as basic units. This model defines the workload as a combination of street length, predicted crime risk, and district diameter. The outcome of the SNPDP is a balanced design of the districts.

This problem poses enormous computational challenges, as such costs normally increase rapidly with the number of basic units. For this reason, the exact methods are not applicable for practical use, as they fail to generate the optimal solutions in an acceptable time. In this study, a two-step heuristic method is proposed to derive a near optimal solution for this problem. The first step generates an initial solution, and the second step improves the solution via a tabu search process.

3.2.2 Infrequent Patrol Routing Strategy

The aim of infrequent patrol routing is to cover given crime hotspots at least once, and to balance the route lengths of patrollers from different stations. In this study, this problem was formulated as a special RPP, with a min-max objective, which attempts to minimise the length of the longest route of each different patroller.

This problem is computationally challenging because of the constraint of multiple stations and the objective of minimising the length of the longest route. The number of candidate solutions increases exponentially with problem size. To generate high-quality routes in an acceptable amount of time, a three-stage heuristic algorithm is proposed. The first stage generates an initial solution, and the second and third stages improve the solution via a tabu search procedure.

3.2.3 Repeated Patrol Routing Strategy

For the routing of repeated patrols, a set of measures to define a good routing strategy is proposed that considers the characteristics of the problem. Accordingly, an online patrol routing strategy was developed. This strategy considers multiple impact factors, including the patrolling history of crime hotspots and the movements of all officers.

This strategy was built on a Bayesian computational framework that computes the probability of visiting each crime hotspot and selects the hotspot with the largest probability as the next patrol target. Route planning is conducted by the control centre, and the planned routes are sent to the patrollers. As a validation of this routing strategy, an agent-based modelling environment was built, based on real-world datasets, that simulates the movements and interactions of the patrollers and the control centre.

3.3 Research methods

In this section, a brief introduction to the main research methods used in this thesis is presented. These methods include tabu search, Bayes' theorem, agent-based modelling, and Gurobi.

3.3.1 Tabu search

As introduced in Section 2.2.4.3, a meta-heuristic is a high-level procedure that is designed to find a heuristic that can provide a sufficiently good solution to an optimisation problem, especially with incomplete or imperfect information or limited computational capacity. One of the commonly used meta-heuristics is the tabu search, which is a local search-based method used for mathematical optimisation. Tabu search was proposed by Glover (1986), and a detailed description of tabu search is given by Glover and Laguna (1997).

Tabu search starts from a feasible solution and iteratively attempts to improve the solution quality by checking its immediate neighbours, which are only slightly different from the current solution. To avoid being stuck in suboptimal regions or areas with many equally fit solutions, tabu search uses a memory to guide the search, and discourages coming back to previously visited solutions. At each iteration, worsening moves can be accepted if no acceptable and improving move exists.

Tabu search has been applied to a wide range of optimisation problems, including districting (Liberatore and Camacho-Collados, 2016) and routing problems (Ahr and Reinelt, 2006; Willemse and Joubert, 2012). In this thesis, tabu search is used to solve the districting problems in Chapter 5 and the infrequent patrol routing problems in Chapter 6.

3.3.2 Bayes' theorem

In probability theory and statistics, Bayes' theorem (also called Bayes' law or Bayes' rule) describes the probability of an event, based on prior knowledge of conditions that inform the event. For example, if cancer is related to age, using Bayes' theorem, a person's age can be used as a condition to accurately assess the probability that they have cancer, which is superior to the assessment of the probability of cancer without knowledge of one's age.

The Bayes' theorem can be stated mathematically as follows (P. 29, Wasserman 2004): let $A_1, A_2, \ldots, A_k$ be a partition of an event, such that $P(A_i) > 0$ for each i and $\sum_i P(A_i) = 1$. If another event B has $P(B) > 0$, then for each i,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} \tag{3.1}$$

where $P(A_i|B)$ is a conditional probability, meaning the likelihood of event A_i occurring given that B is true. $P(B|A_i)$ is also a conditional probability. $P(A_i)$ and $P(B)$ are the probabilities of observing A_i and B independently, which are known as marginal probability. $P(A_i)$ is also called the prior probability of A_i and $P(A_i|B)$ the posterior probability of A_i.

The Bayes' theorem is capable of combining multiple conditions in the calculation of posterior probability. In this thesis, this theorem is used to calculate the probability of the next hotspot to visit by combining multiple factors, in the repeated patrol routing strategy in Chapter 7.

3.3.3 Agent-based modelling

Agent-based modelling (ABM) is a simulation technique that seeks to capture how individual behavioural units interact with each other and with the environment, which allows higher-order behaviours and structures to emerge from these interactions (Epstein and Axtell, 1996). In an agent-based model, an agent can represent any unit that is capable of behaviour (e.g. a person, a household, a car), and often many different types of agent exist in the same simulation. Agents are situated in surroundings that influence and constrain their behaviours, and these surroundings may include spatial spaces such as grids, street networks, or social networks, and are known to be the simulation's environment.

As a methodology, ABM has been applied to explore questions in varying fields such as geography, archaeology, economics, ecology, among others (Castle and Crooks, 2006). ABM represents a promising tool for investigating complex and important processes regarding human behaviours. It allows for spatially and temporally explicit individuals to move, perceive, observe and decide on actions based on their personal characteristics and surroundings, permitting simulations to capture all these variables (Heppenstall *et al.*, 2012).

In this thesis, ABM is utilised to test the effectiveness of the repeated patrol routing strategy, which is developed in Chapter 7.

3.3.4 Gurobi

The Gurobi Optimiser (Gurobi Optimization, 2016) is a commercial optimisation solver that was utilised to obtain the optimal or sub-optimal objective value of varying mathematical programming or optimisation problems, including linear programming, quadratic programming, mixed integer linear programming, among others. It is recognized as the state-of-the-art solver for mathematical programming, and has been designed from the ground up to exploit modern computer architectures, using the latest implementations of the latest algorithms. Therefore, Gurobi allows users more flexibility in how they model the problems, and increases their ability to add more complexity to the models, to better represent the real-world problems they are solving.

In this thesis, Gurobi is used to solve the districting problems in Chapter 5 to compare with the proposed methods, and is also applied to obtain a lower bound of the patrol routing problems in Chapter 6.

3.4 Chapter Summary

In this chapter, a framework for police patrol strategy was proposed. In accordance with the aims and objectives of this thesis, the framework includes two basic types of strategy, namely districting and routing. Further, these can be coupled to form new strategies that are able to deal with greater complexity of the patrolling procedure in the strategy design, and so can enhance efficiency. In the following chapters, each strategy in this framework is described in detail, including their formulation, implementation and validation.

Chapter 4

INFORMATION ABOUT THE DATA

4 INFORMATION ABOUT THE DATA

4.1 Chapter Overview

In this chapter, the case study areas and all the datasets used in this research are described. In Section 4.2, the major case study area – the London Borough of Camden (UK) – is briefly introduced. Three datasets contributed to the case study and the experiments: 1) the Integrated Transport Network (ITN) layer of the Ordnance Survey, which provided the fundamental and underlying street network (Section 4.3); 2) the crime records data, and the derived crime risk maps, which were used as inputs into the development of the police patrolling strategies (Section 4.4); and 3) the Corine land cover (CLC) dataset, which was used to reveal the signature of the identified patrol districts, and to validate the districting plan (Section 4.5). In Section 4.6, a brief introduction to another case study area, the South Side of Chicago (US), is provided. Section 4.6 summarises the data's characteristics.

4.2 The Case Study Area of Camden

This study was mainly conducted in London (UK). Greater London is organised into 33 local government districts, including 32 London boroughs and the City of London. Each borough has been assigned a BOCU by the MPS. Each BOCU has warranted police officers and special constables who patrol and respond to emergencies. The City of London was not included in this study of this thesis, as law enforcement in that is not administrated by the MPS. Instead, this area is the responsibility of the separate City of London Police.

The specific focus of this study was on the police patrols of the 12 BOCUs that correspond to the 12 Inner London boroughs, namely City of Westminster, Camden, Greenwich, Hackney, Hammersmith and Fulham, Islington, Kensington and Chelsea, Lambeth, Lewisham, Southwark, Tower Hamlets, and Wandsworth. Inner London is officially the wealthiest area in Europe, and contains the most expensive street in Europe. According to the statistical office of the European Union, Eurostat, in 2010, its GDP per capita was more than €80,000, while the UK GDP per capita was €27,500 and the European GDP per capita was €24,500 (European Commission, 2010). The population density in Inner London is more than twice that of Outer London.

Due to it being difficult to show all of the data and for the confidentiality of police and crime datasets, the results from all 12 Inner London BOCUs are not displayed herein. Rather, a study area covering the London Borough of Camden was chosen, to illustrate the mechanism and results of the proposed police patrol strategies. Camden was chosen because it is one of the boroughs with the highest amounts of urban activities and also the most criminal offences. According to MPS, between October 2015 and October 2017, the number of crimes in the borough of Camden was 69,338, ranking second in all 32 London boroughs (Met Police, 2017).

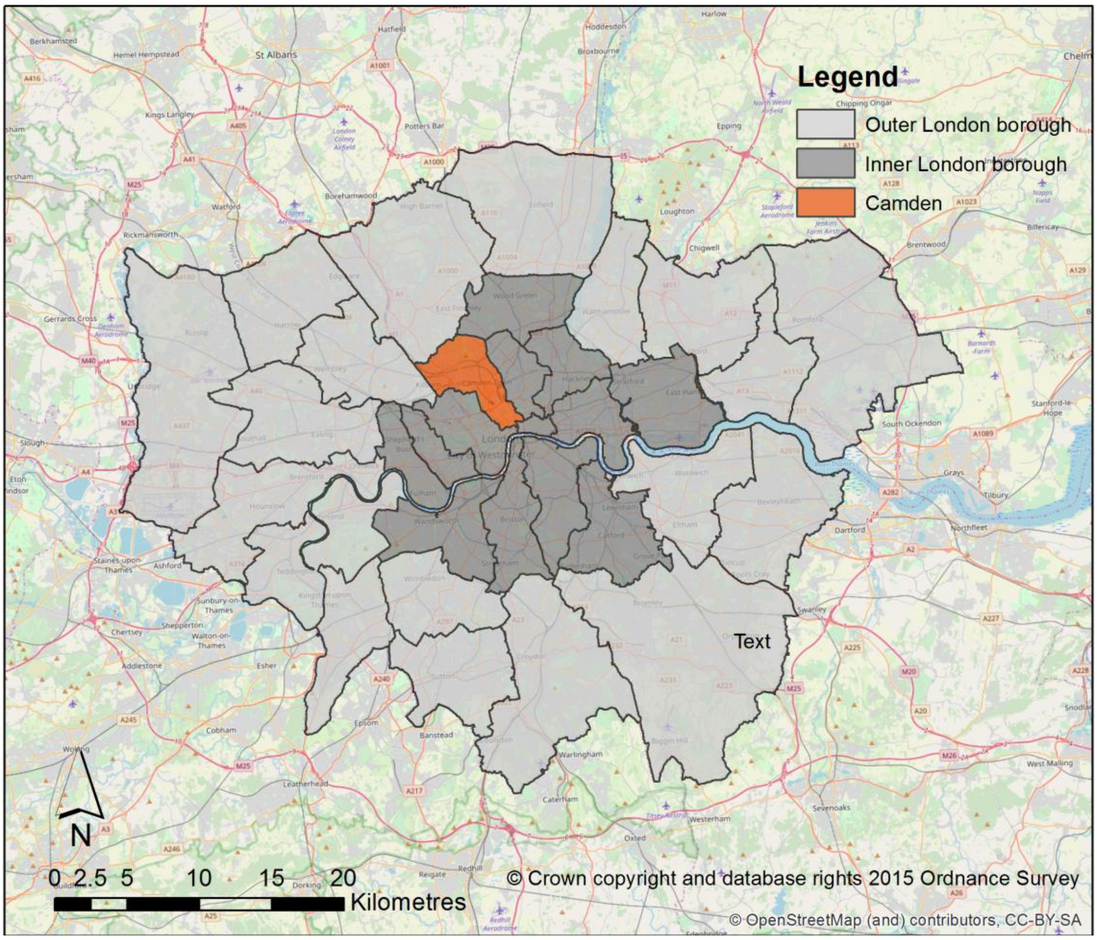


Figure 4.1 The London (UK) Borough of Camden study area, used in this thesis to exemplify the formulation of police patrol strategies

4.3 Street Network Data

The Ordnance Survey MasterMap ITN Layer Urban Paths Theme dataset (Ordnance Survey, 2015) was used, to represent the London street network. The ITN layer is a dataset of Great Britain’s road network from motorways to urban paths, including traffic restrictions, such as one-way systems and low bridges.

The topological structure of the ITN consists of Road Link features (edges) and their intersections (nodes). An edge is a line or curve segment connecting two nodes at its ends. The ITN data contains the identification and location information of each node, as well as the length, connectivity, and type of each edge. A named or numbered street corresponds to one or more edges, with such edges being given the name of the street. The simple ITN network structure illustrated in Figure 4.2 shows the relationship between nodes and edges.

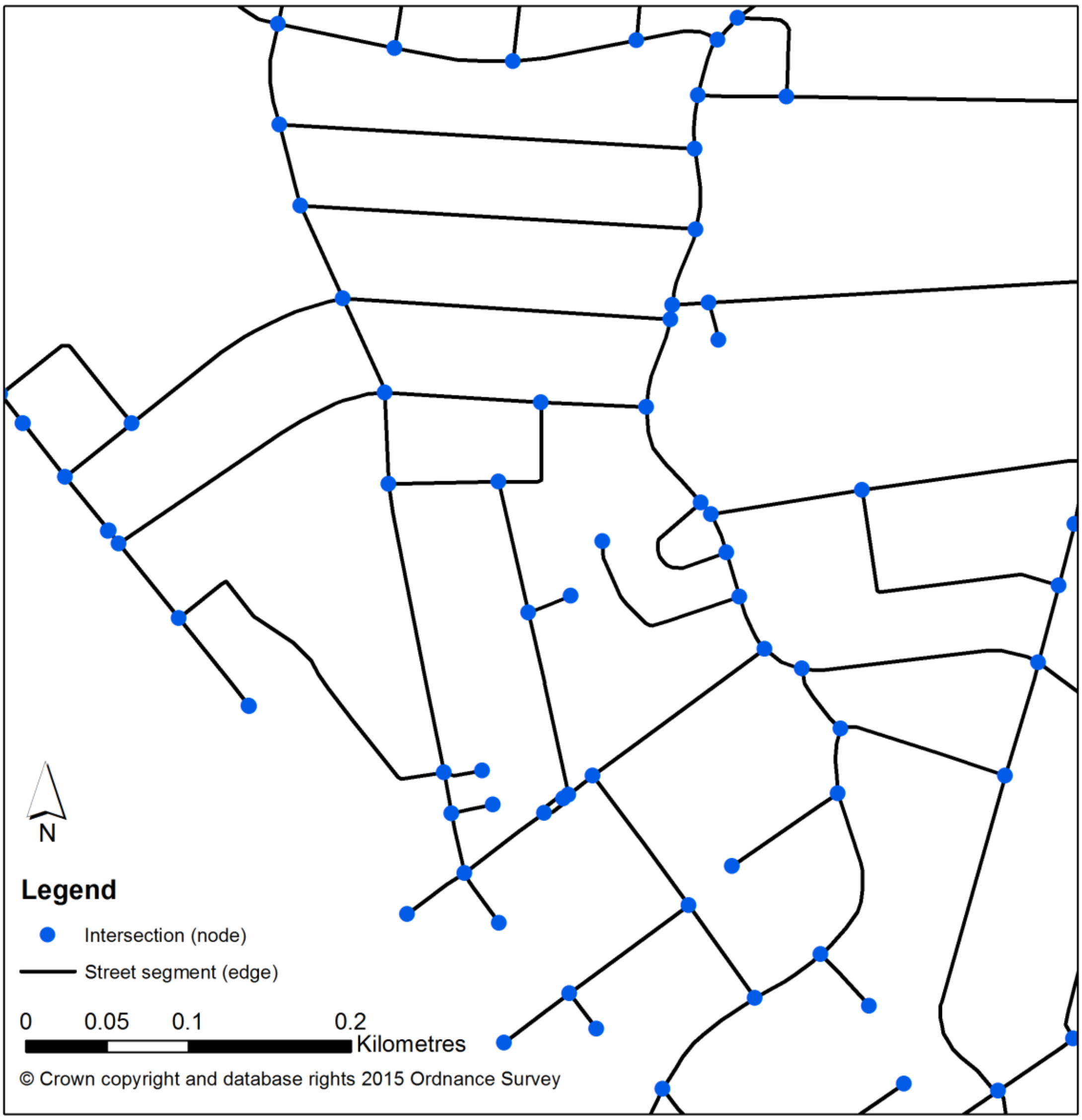


Figure 4.2 An example of the ITN network, illustrating nodes and edges

4.4 Crime Data and Crime Risk Maps

4.4.1 Crime Records Data

The historical crime records were acquired from two databases provided by the MPS, including the Computer Aided Dispatch (CAD) and the Crime Reporting Information System (CRIS).

The CAD database records all report incidents, which may be made by police officers or members. When a call is first received, the CAD operator categorises the nature of the incident and grades the call by importance. The call is also given a unique incident number. There may be multiple entries with the same incident number due to the way in which the CAD data are extracted from the source database. Furthermore, multiple unique incident numbers may actually refer to the same event if multiple people reported it.

The geographical location of the person reporting the incident is recorded. Due to change of CAD operation, there are two different ways of recording the location. CAD records before 14th July 2011 are geolocated using a 250m square grid, and after this time, locations are recorded as accurately as the location of the report can be determined. This may be an estimated location. For example, if an incident is reported at a range of house number (e.g., 1-50 High Street), the CAD entry will be snapped to the closest Ordnance Survey reference to the midpoint of this range (i.e., 25 High Street). There is no available indication to describe the potential error in this process.

No all entries in CAD are ‘real’ crimes. The available CAD dataset is associated with an attribute called *CRIS reference* to reflect the respondent’s description of the incident. If an incident is confirmed as a crime, it is added to the CRIS database, and the CRIS reference is included in the CAD entry. Otherwise, the CRIS reference is empty or marked as NOT CRIMES.

To extract ‘real’ crimes and remove duplicated entries, the CAD dataset is cleaned before use. First, it is filtered by CRIS entry by removing empty and NOT CRIMES records. Second, the data is de-duplicated by only counting each incident number once, regardless of the number of times it appears.

The crime record datasets for Camden are specifically introduced. Between 1st March 2011 and 31st March 2012, a total of 28,686 records of crimes of different types were recorded in the database for the Borough of Camden. The crime types include theft, burglary, homicide, battery, arson, motor vehicle theft, assault and robbery.

4.4.2 Crime Risk Maps

For this study, two types of crime risk maps were generated and used. The reason for using different types of crime risk map is to demonstrate that the proposed strategies are not limited to specific input and can accommodate varying types of crime risk.

The first type of crime risk map was computed using network-time KDE (NTKDE) (Rosser *et al.*, 2017). This method is conceptually similar to the original grid-based ProMap algorithm developed by Bowers et al. (2004). Given a target time, *t*, and location, *s*, the prospective risk level (a relative measure) at the location is calculated by summing up the risk contributions from all preceding crimes that are deemed sufficiently close in space:

$$\lambda_{grid}(t,s) = \sum_{0<c_i\le\tau;e_i\le v} \left(\frac{1}{c_i}\right)\left(\frac{1}{1+e_i}\right) \tag{4.1}$$

, where c_i is the time interval (e.g., number of weeks) that elapsed since crime i, e_i is the distance (e.g., number of 50-metre grid cells) between location s and the location of crime i, τ is the maximum time lag (i.e. the temporal bandwidth) and v is the maximum spatial lag or bandwidth. Each crime is represented as a point in space and time (t_i, s_i), and crimes included in the calculation are denoted ‘source crimes’.

This method has been generalised to cover a continuous two-dimensional planar space, and modified by normalising the function within the sum. Therefore, the resulting method becomes a form of space-time kernel density estimation (STKDE). In this method, two kernel components are defined, $f(\Delta s)$ and $g(\Delta t)$, which represent the spread of risk from a single source to a target over a distance, Δs, and a time lag of Δt. The planar STKDE at target time t and location s is given by:

$$\lambda_{planar}(t,s) = \sum_{i:t_i<t} f(\|s-s_i\|)g(t-t_i) \tag{4.2}$$

, where s_i and t_i represent the location and time of the i-th source, and $\|s-s_i\|$ denotes the Euclidean distance between s and s_i.

For the above form, we choose an exponentially decaying function of time and a linearly decaying function of distance:

$$f(\Delta s) = \begin{cases} \dfrac{h_s-\Delta s}{h_S^2} \; if\ \Delta s \le h_s \\ 0 \;\; otherwise \end{cases} \tag{4.3}$$

$$g(\Delta t) = \frac{1}{h_T}\exp(-\frac{\Delta t}{h_T}) \tag{4.4}$$

, where the parameters h_T and h_S are the temporal and spatial bandwidths, respectively.

The STKDE was then adjusted for the network space, to form the NTKDE. In this new form, the risk density decays linearly along an edge, and ‘splits’ at nodes (road junctions). The total risk contribution of the source at the target is calculated by summing over all non-cyclic paths between them, the length of which is less than the spatial bandwidth, h_S. Figure 4.3 illustrates the equal 'split' of kernel at node B.

For a given path p between s and s_i, assuming there are $m(p)$ vertices on p, then n_1, …, $n_{m(p)}$ denote the degrees of the $m(p)$ vertices on the path. The spatial component of the NTKDE is calculated by summing over all non-cyclic paths between s and s_i.

The NTKDE is given by

$$\lambda_{net}(t,s) = \sum_{i:t_i<t} g(t-t_i)k_{s_i}(s)$$

$$= \sum_{i:t_i<t} \frac{1}{h_T}\exp\left(-\frac{t-t_i}{h_T}\right)\left(\sum_{p:s_i\to s} \frac{h_S - \Delta s_i^{(p)}}{h_S^2(n_1-1)\dots\left(n_{m(p)}-1\right)}\right) \quad (4.5)$$

$$\Delta s_i^{(p)} < h_S$$

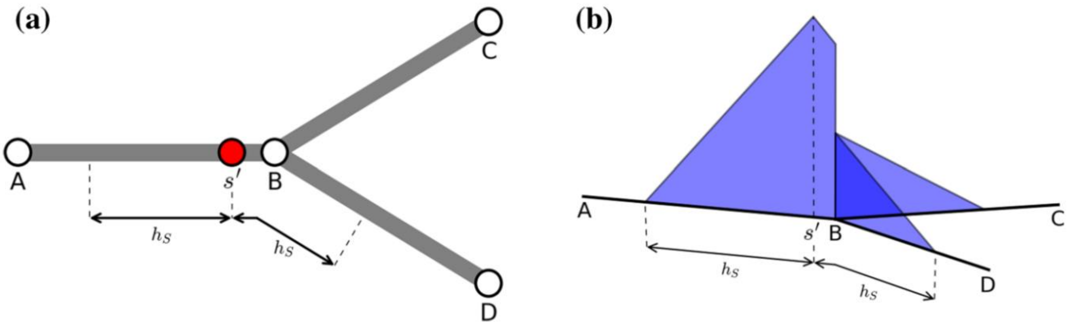


Figure 4.3 Kernel calculation on networks (Rosser *et al.*, 2017)**. (a) Given a kernel centred at s', the one-dimensional kernel function should be adjusted to apply to each of the branches. (b) The 'equal-split' approach, where the remaining kernel density at a junction is divided equally between the outgoing branches**

In the second type of crime risk map, crime density on the street is simply defined as the ratio of the number of crime incidents on the street during the study period to the street length. The streets with the highest crime densities, and covering a given proportion (e.g., 5%) of the total street length, are defined as crime hotspots.

These two types of risk map can be used interchangeably, as in either map each street has a non-negative risk value, through which the crime hotspots can be identified. In the following sections, the first type of crime risk map is used in the design of patrol districts (Chapter 5), while the second type is adopted in the design of the patrol routing strategies (Chapters 6 and 7).

4.5 Land Cover Dataset

The CLC 2012 dataset was used to interpret the characteristics of the planned patrol districts from the districting model. This covers the UK, Jersey and Guernsey (Cole *et al.*, 2015), and is published by the Natural Environment Research Council Environmental Information Data Centre.

The CLC dataset was produced within the framework of the Initial Operations of the Copernicus programme on land monitoring. The CLC inventory was initiated in 1985, and then it established a time series of land cover information with updates in 2000, 2006, and 2010. CLC products are based on the analysis of satellite images by national teams from participating countries (Cole *et al.*, 2015), following a standard methodology and nomenclature, and using the following base parameters: a minimum mapping unit (MMU) for status layers of 25 hectares; a minimum width of linear elements of 100 metres; and 44 classes in a three-level hierarchy. The five main (level-one) categories are artificial surfaces, agricultural areas, forests and semi-natural areas, wetlands, and water bodies. As an example, Figure 5.4 shows the CLC map for Camden, which consists of five different land cover types.

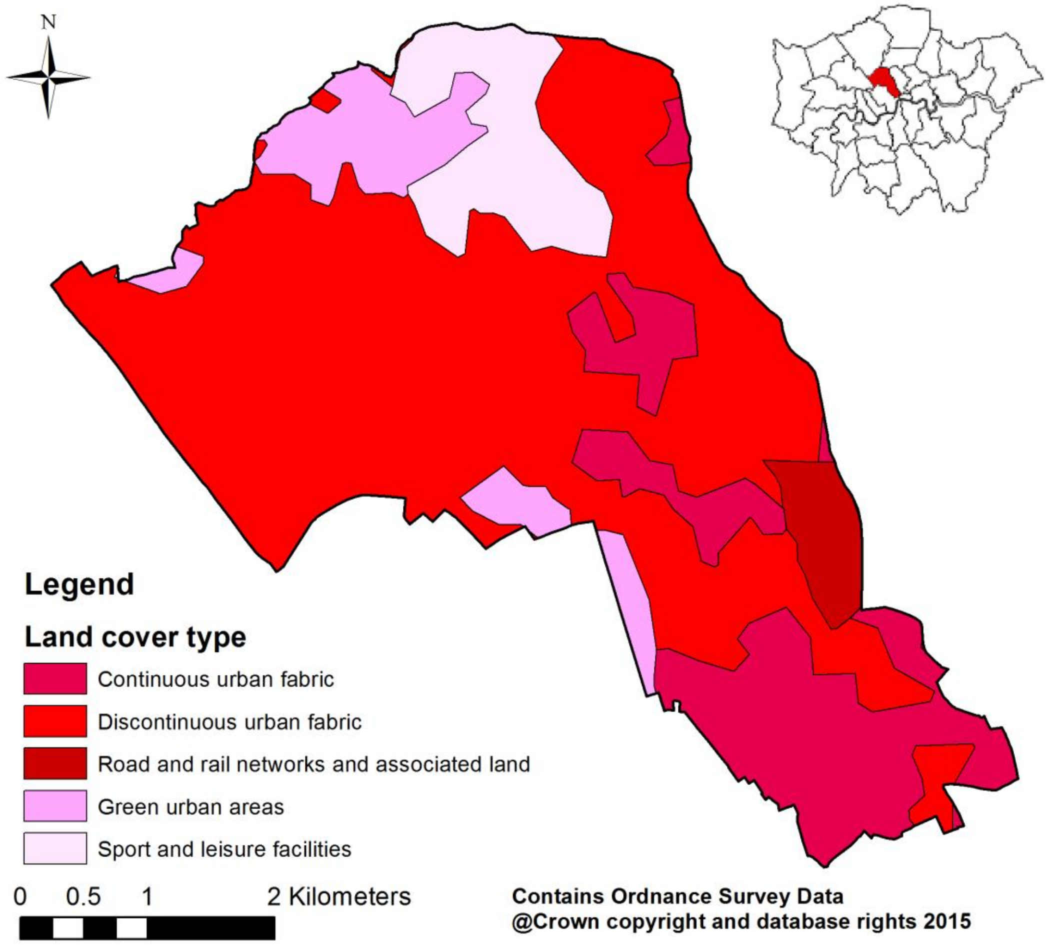


Figure 4.4 The 2012 CLC map for the London Borough of Camden

4.6 The Case Study Area of South Side

Another case study area used is the South Side, Chicago, which is located in the Illinois State in the United States of America. The South Side is an area of the city of Chicago, and is the largest of the three sides of the city that radiate from downtown, with the others being the North Side and the West Side. For convenience, the South Side is referred to as South Chicago. A map of South Chicago is shown in Figure 4.5.

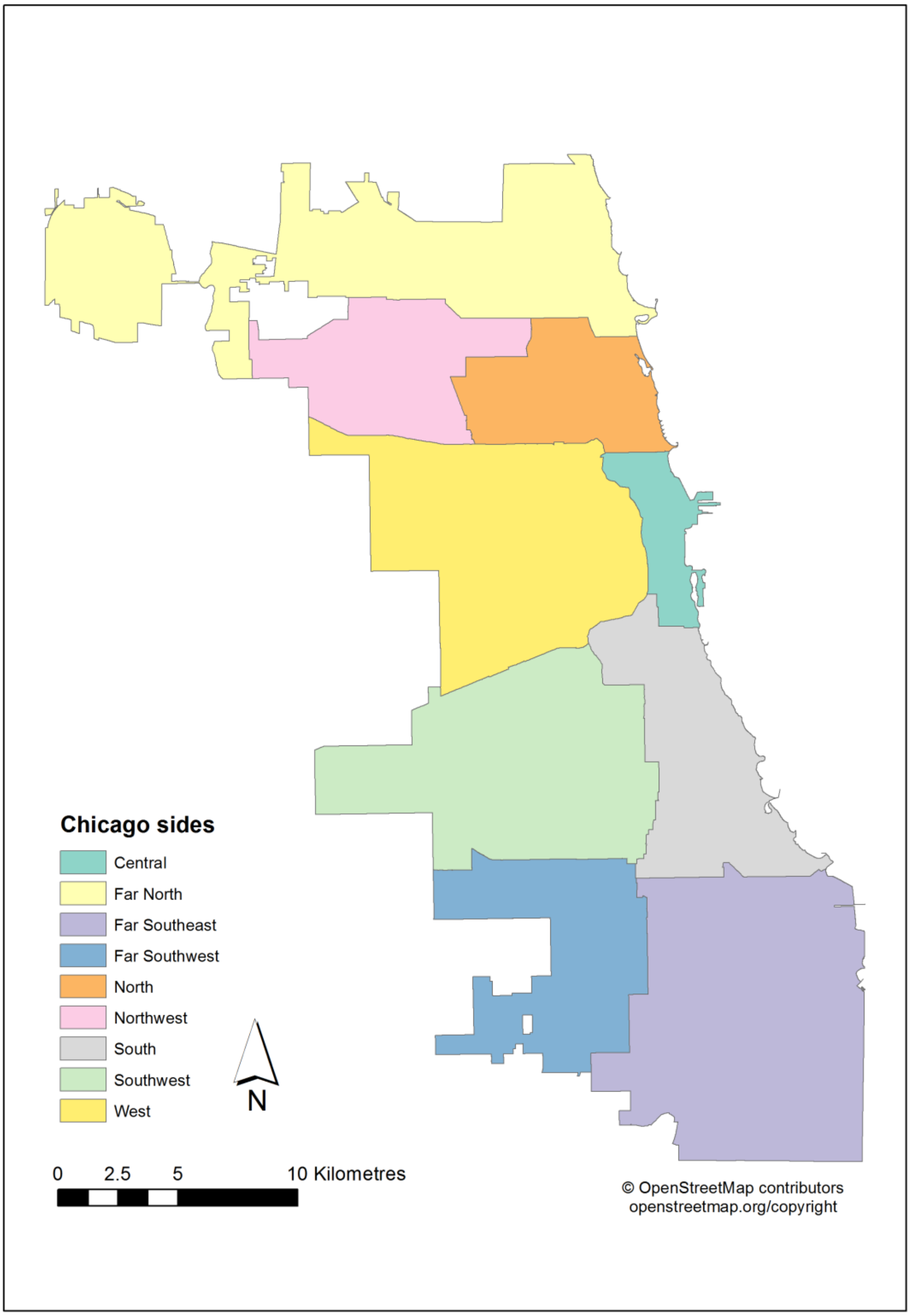


Figure 4.5 The South Chicago study area (denoted by 'South'), used in this thesis to exemplify the formulation of police patrol strategies

The case study of the South Chicago is described in Chapter 7, to validate the repeated patrol routing strategy. The datasets of South Chicago have been provided by various agencies. The street network data is provided from OpenStreetMap (http:// www.openstreetmap.org). The locations of police stations and the crime records are downloaded from the City of Chicago data portal (https://data.cityofchicago.org). The downloaded crime records range from the year 2011 to 2015, and the records used in this thesis are between March 1st, 2011 to March 1st, 2012. The crime types include theft, burglary, homicide, battery, arson, motor vehicle theft, assault, and robbery.

The computation of crime risk maps in South Chicago is the same as in the Camden case. Due to the lack of emergency call data or police despatch data of South Chicago, hypothetical emergency calls in South Chicago are generated using a uniform distribution in the given area and time period.

4.7 Chapter Summary

In this chapter, two study areas were introduced, namely Inner London and South Chicago. In the case study of Inner London, for the purpose of demonstration, the case study is focused on the London Borough of Camden, which is chosen because it is characterised by the busiest urban activities and has the most crime offences of all the London boroughs.

The design of police patrol strategies is based on crime records data and crime risk maps. The crime dataset used was obtained by the MPS system, and includes eight types of crime. Two types of crime risk maps were adopted for the case study, namely the NTKDE and simple historical crime risk maps.

The police patrol strategies are based on the underlying street network, specifically the Ordnance Survey MasterMap ITN Layer Urban Paths Theme dataset. The CLC dataset was introduced to characterise the signature of the identified patrol districts, and to validate the districting plan. This dataset consists of 44 land cover classes in a three-level hierarchy, and was produced via the analysis of satellite images by national teams in participating European countries.

In the case study area of South Chicago, the datasets were obtained from various agencies, and the computation of crime risk maps in South Chicago is the same as in the Camden case.

Chapter 5

A POLICE DISTRICTING MODEL BASED ON STREET NETWORKS

5 A POLICE DISTRICTING MODEL BASED ON STREET NETWORKS

5.1 Chapter Overview

In this chapter, a police districting model based on street networks is described. This chapter is organised as follows: in Section 5.2, a brief introduction to the PDP, and the objectives of this research, is provided; Section 5.3 describes the formulation of the SNPDP and, in Section 5.4, a two-step heuristic method to solve this problem is proposed; a grid-based aggregation model is outlined in Section 5.5, which is an alternative approach to deriving patrol districts based on streets. In Section 5.6, a case study based on the London Borough of Camden is reported on, which demonstrates the application and efficiency of the SNPDP model. The chapter is summarised in Section 5.7.

5.2 Introduction

The PDP concerns the optimal partitioning of territory into several patrol sectors with respect to performance attributes, such as workload and response time. This problem is motivated by various purposes, including patrol planning, emergency response, reporting, and demographic police research (Mitchell, 1972; Sarac *et al.*, 1999). In terms of patrol planning and emergency response, one of the most relevant criteria is the workload balance, which ensures the timeliness of emergency response and the effectiveness of patrol activities (Camacho-Collados et al. 2015). In contrast, an uneven workload distribution leads to an imbalance in staffing, morale problems, and span-of-control issues (Kistler 2009). In particular, some districts are overstaffed, while others are understaffed, and officers with excessive workloads can easily become dissatisfied. For these reasons, some police departments redesign the division boundaries after a certain period in order to equalise the workload (Kistler 2009).

Conventionally, the design of police districts is based on areal units, including grids (Mitchell, 1972; Bodily, 1978; Carroll and Laurin, 1981; Zhang and Brown, 2013; Camacho-Collados, Liberatore and Angulo, 2015), census blocks (Liberatore and Camacho-Collados, 2016), and r-districts (D'Amico *et al.*, 2002), among others. However, there are several limitations associated with such units. First, the grid structure is often selected by planners and administrators, and such decisions are empirical and arbitrary. The district configuration may be sensitive to the unit type and unit size, which is an instance of the MAUP (Openshaw and Taylor, 1979; Openshaw, 1984). More precisely, the MAUP includes two components: scale and aggregation problem. The scale problem is defined as a variation in results when the same areal data are aggregated into larger areal units for analysis. In contrast, the aggregation problem concerns variations in results due to alternative units of analysis where the number of units is constant. Second, the connectivity within and between these units in street networks is barely taken into account. Areal units could intersect physical features (e.g., lakes) or barriers (e.g., railway tracks), and are not well-defined targets for operational deployment in policing. Therefore, patrol sectors based on these units tend to consist of unconnected parts, making patrolling and emergency response difficult and leading to excessive and unbalanced workloads.

Street segments represent a promising unit for patrol sectors, as the street structure fundamentally influences human movement patterns. Recent findings in criminology suggest that street networks shape the long-term pattern of crimes (Davies and Johnson, 2015; Summers and Johnson, 2017), as well as the short-term dynamics of crime behaviour (Davies and Bishop, 2013; Johnson and Bowers, 2014). This implies that street-based models are appropriate for the description and prediction of crime risk (Rosser *et al.*, 2017). Thus, the use of street networks in crime prevention efforts, and particularly in the design of patrol districts, is well-motivated, having two possible advantages: 1) it may amplify the effect of targeted policing; 2) it may enhance the usability of the patrol district and patrol plans. Despite these advantages, few existing studies on the PDP have attempted to use street segments as basic units.

In summary, while it is essentially important to incorporate the street network into the PDP, there is a lack of police districting models that explicitly use street segments as basic units. This study aimed to explore the potential of street segments in police district design, and to develop efficient methods for large PDPs.

Formulating and solving the street-based PDP presented new challenges. The first challenge lies in the representation of streets in a policing context. Although several methods exist that can represent street networks as graphs (Jiang and Claramunt, 2004; Porta, Crucitti and Latora, 2006a, 2006b), it is unclear which method is most appropriate. Another challenge is that a street-based PDP with a large number of street units is computationally demanding, hindering the use of existing algorithms for PDPs. This required the development of efficient algorithms to generate high-quality districting plans in an acceptable time.

To overcome the challenges, in this study, an efficient and balanced district design was developed, based on urban networks. First, a multi-criteria police districting model was formulated that uses streets as basic units. The model incorporates several relevant workload attributes, and a trade-off between average workload and workload balance among districts. Second, an efficient heuristic method for large-sized PDPs was proposed, by integrating a graph partitioning algorithm and a tabu search method. Third, this model was compared with an alternative model that aggregated the patrol demand into a grid. The results indicate that the SNPDP is superior to the aggregation model, with respect to the workload balance and district connectivity.

5.3 A Street-Network-based Police Districting Problem (SNPDP)

In this section, the SNPDP is illustrated. The goal of this model is to partition the street network, denoted by *SN*, into a given number, *m*, of patrol districts to optimise regarding a set of given criteria. A solution for the PDP is denoted as $D = \{d_1, d_2, \ldots, d_m\}$. The workload of a district is evaluated by a weighted sum of three attributes: crime risk, area size, and diameter.

Here, an overview of the SNPDP model is also provided. First, there is a discussion on how to represent the street network, then the model constraints, workload attributes and objective functions are described. Finally, an exact formulation of the SNPDP is presented.

5.3.1 Street Network Representation

The first step in this model is to represent the street network as a graph. As introduction in Section 2.4.1, there are several methods for obtaining such a representation, including the primal and dual methods. The dual representation allows for a street-oriented computational analysis of the properties of an urban street network. As the PDP model in this study focuses on the partition of the streets, this network representation was adopted. The underlying street network is modelled by an undirected graph, $G = (V, E)$, with V corresponding to the named street links and E to the intersections of links (see Figure 5.1).

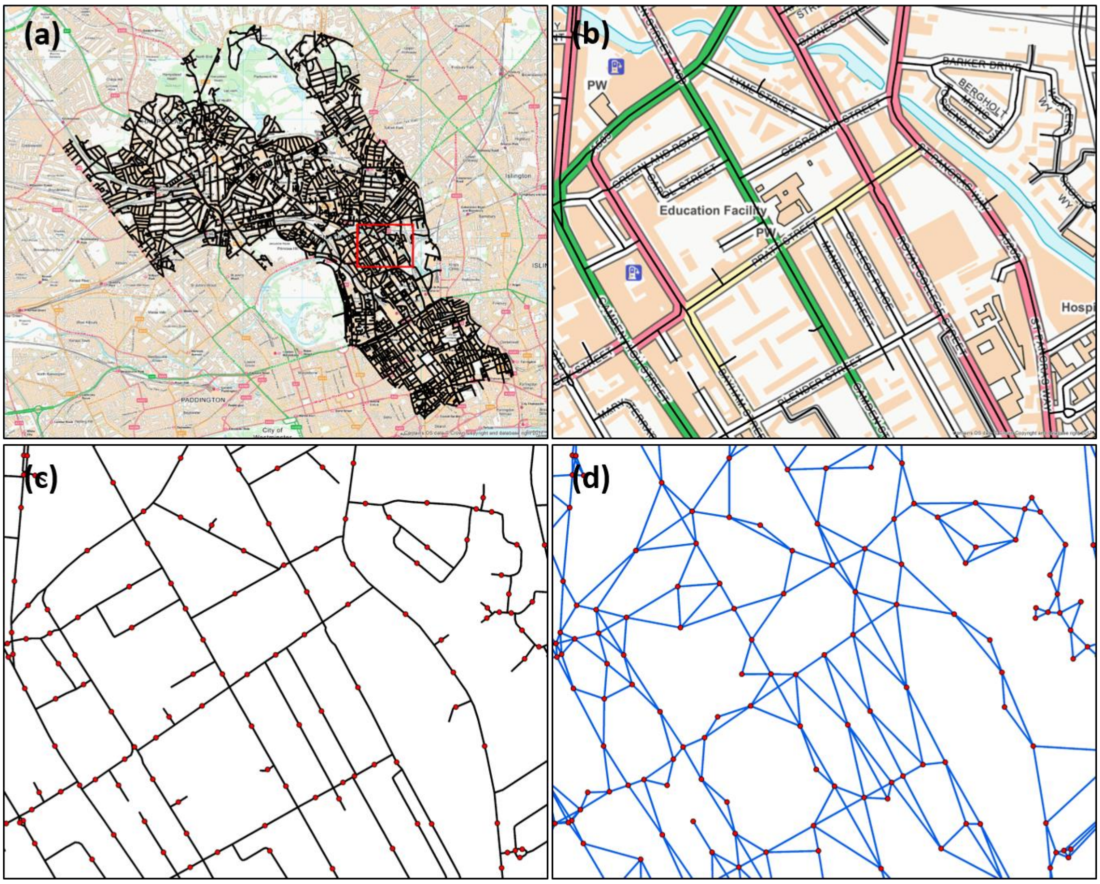


Figure 5.1 Construction of the dual representation of a street network. This example shows a section of the Camden network. a) The original street network map, as obtained from the Ordnance Survey's ITN dataset. b) Map zoomed to the section highlighted in red in (a), with the background map image removed. c) Nodes placed at the midpoints of streets. d) Links added between all pairs of connected streets, with the original street network removed

The distance between two streets, $dist(i, j)$, in graph G is defined as the shortest path distance between their midpoints (She *et al.*, 2015). Accordingly, the distance between all pairs of streets can be efficiently computed using the Floyd–Warshall algorithm (Floyd, 1962). A different definition was adopted by Butsch et al. (2014), which defined $dist(i, j)$ as the minimum distance between their end nodes.

5.3.2 Constraints

Below, the two constraints that any feasible districting plan must satisfy are addressed.

5.3.2.1 Complete and Exclusive Assignment

The districts cannot overlap and must cover the entire patrol area. Moreover, each district should be non-empty.

5.3.2.2 Contiguity

A district must contain basic units that are connected to each other within the district. This implies that an officer cannot be assigned to a patrol sector that is composed of two or more separate parts. Moreover, unconnected districts would lead to excessive response time and inefficient resource deployment. In this study, the terms 'contiguity' and 'connectivity' are used interchangeably.

The contiguity constraint is satisfied in the methods for the SNPDP in two ways. First, in the exact formulation, the contiguity constraint was designed by borrowing concepts from graph partitioning. Specifically, the constraint is an extension of the ordered-area assignment conditions that are developed to enforce contiguity in the site design (Cova and Church, 2000). Second, in the proposed heuristic, the initial solutions are constructed by the graph partitioning method, and are then iteratively modified via moving a unit from one district to another, on condition that the district contiguity is maintained.

5.3.3 Workload Attributes

Here, three relevant workload attributes of a district are described. To be comparable, each attribute is formulated as a relative and dimensionless ratio between 0 and 1.

5.3.3.1 Risk

The future crime risk on a street can be estimated using historical data and a predictive crime mapping approach (Rosser *et al.*, 2017). The crime risk in a district, d_k, is the sum of the risks associated with the contained streets, and the risk measure, R_k, is expressed as the ratio of the risk of d_k to the entire area risk, as in equation (5.8).

5.3.3.2 Area

The area factor quantifies the size of the territory that an officer should patrol. The area measure, A_k, is defined as the ratio of the total length of streets in the district, d_k, to the entire network, as in equation (5.9).

5.3.3.3 Diameter

During patrolling, officers have to travel within districts to ensure security and respond to calls for service. Intuitively, a large (average or maximum) internal travel distance makes a district difficult to patrol. Therefore, it is favourable to reduce the day-to-day travel distance within districts. Here, the diameter measure is adopted to represent the maximum internal travel distance within a district, or the travel distance in the worst case.

The diameter of a district is defined as the maximum distance between any pair of points in it. It approximates the maximum distance an officer has to travel in the case of a call for service. The diameter measure, Dmt_k, is the ratio of the district diameter to the diameter of the entire graph, as in equation (5.10).

5.3.4 Objective Function

The workload is a convex combination of the three relevant attributes. The weights can be determined according to the preference of police experts. Given the weights, the district workload, L_k, can be computed using equation (5.1).

The objective of the SNPDP is to generate districting plans that are both efficient and balanced. First, a districting plan is said to be efficient if the average workload is low. Second, a districting plan is perfectly balanced if the workload of all districts is equal. As perfectly balanced districts cannot always be achieved, a common measure of workload balance is to use the mean absolute deviation of the workloads.

Both the workload average and deviation should be minimised to achieve a good districting plan. However, there might be a trade-off between these two objectives, meaning that improving on one objective may worsen the other. For example, a decrease in the workload deviation could lead to a higher average workload. Hence, two objectives are combined via a convex combination, as in equation (5.1). This objective allows the experts to express their preferences concerning the two objectives, and examine the trade-off between these. By varying the coefficient α, a range of different models from efficiency ($\alpha = 1$) and balance ($\alpha = 0$) can be observed.

5.3.5 Problem Formulation

Here, an exact mixed integer programming (MIP) formulation for the SNPDP is provided.

Parameters:

n = number of basic units

i, j = index of basic units; I = set of basic units, $I = \{1,\dots,n\}$

m = number of districts

k = index of district; K = set of districts, $K = \{1, \dots, m\}$

c = index of contiguity order, $c \in \{0, \dots, q\}$, $q = n - 1$

$$w_{ij} = \begin{cases} 1, & \text{if units i and j are adjacent to each other, with } i, j \in I \text{ and } i \neq j \\ 0, & \text{otherwise} \end{cases}$$

$N_i = \{j | w_{ij} = 1\}$, the set of units that are adjacent to unit i

r_i = the risk of unit i

a_i = the area of unit i

d_{ij} = the shortest distance between unit i and j

wt_{att} = the weight of an atttribute att in the workload, $0 \leq wt_{att} \leq 1, att \in \{A, R, Dmt\}$

L_k = the workload of district $k, k \in \{1, \dots, m\}$

α = the coefficient of the workload average in the objective function, $0 \leq \alpha \leq 1$

Decision variables:

$$t_{ijk} = \begin{cases} 1, & \text{if both unit } i \text{ and j belong to district } k, with\ i < j \\ 0, & \text{otherwise;} \end{cases}$$

$$x_i^{kc} = \begin{cases} 1, & \text{if unit } i \text{ is assigned to district } k \text{ in order } c \\ 0, & \text{otherwise;} \end{cases}$$

Minimise
$$obj = \alpha \frac{\sum_{k=1}^{m} L_k}{m} + (1-\alpha) \frac{\sum_{k=1}^{m} AD_k}{m} \tag{5.1}$$

Subject to
$$\sum_{i=1}^{n} x_i^{k0} = 1 \quad \forall k = 1, \ldots, m; \tag{5.2}$$

$$\sum_{k=1}^{m} \sum_{c=0}^{q} x_i^{kc} = 1 \quad \forall i = 1, \ldots, n; \tag{5.3}$$

$$x_i^{kc} \le \sum_{j \in N_i} x_j^{k(c-1)} \quad \forall\ i = 1, \ldots, n; \forall\ k = 1, \ldots, m;\ \forall\ c = 1, \ldots, q; \tag{5.4}$$

$$t_{ijk} \ge \sum_{c=0}^{q} x_i^{kc} + \sum_{c=0}^{q} x_j^{kc} - 1 \quad \forall\ i, j = 1, \ldots, n | i < j; \forall\ k = 1, \ldots, m \tag{5.5}$$

$$t_{ijk} \le \sum_{c=0}^{q} x_i^{kc} \quad \forall\ i, j = 1, \ldots, n | i < j; \forall\ k = 1, \ldots, m \tag{5.6}$$

$$t_{ijk} \le \sum_{c=0}^{q} x_j^{kc} \quad \forall\ i, j = 1, \ldots, n | i < j; \forall\ k = 1, \ldots, m \tag{5.7}$$

$$R_k = \frac{\sum_{i=1}^{n} \sum_{c=0}^{q} x_i^{kc}\, r_i}{\sum_{i=1}^{n} r_i} \tag{5.8}$$

$$A_k = \frac{\sum_{i=1}^{n} \sum_{c=0}^{q} x_i^{kc}\, a_i}{\sum_{i=1}^{n} a_i} \tag{5.9}$$

$$Dmt_k = \frac{\max_{i,j}\{t_{ijk} d_{ij}\}}{\max_{i,j}\{d_{ij}\}} \tag{5.10}$$

$$L_k = wt_R R_k + wt_A A_k + wt_{Dmt} Dmt_k \qquad \forall\ k = 1, \ldots, m \tag{5.11}$$

$$AD_k \geq L_k - \frac{\sum_{j=1}^{m} L_j}{m} \; \forall \, k = 1, \ldots, m \tag{5.12}$$

$$AD_k \geq \frac{\sum_{j=1}^{m} L_j}{m} - L_k \; \forall \, k = 1, \ldots, m \tag{5.13}$$

$$x_i^{kc} \in \{0,1\} \;\; \forall \, i = 1, \ldots, n; \forall \, k = 1, \ldots, m; \; \forall \, c = 0, \ldots, q; \tag{5.14}$$

$$t_{ijk} \in \{0,1\} \;\; \forall \, i, j = 1, \ldots, n | i < j; \forall \, k = 1, \ldots, m \tag{5.15}$$

This model is formulated as a minimisation problem with an objective function (5.1), and the two terms represent the workload average and workload deviation, respectively. They are merged into one single value by a weight, α.

Constraint equation (5.2) establishes that a district, *k*, should be non-empty, and have exactly one root unit. A root unit for a district has a contiguity order of 0. Constraint equation (5.3) requires that each unit should be assigned to exactly one district and one contiguity order. Constraint (5.4) enforces that a unit, *i*, can be assigned to a district, *k*, with order *c* if, and only if, a unit *j* exists in the neighbourhood of *i*, and that *j* is assigned to district *k* in order (*c*-1).

Constraint equations (5.5), (5.6) and (5.7) select the pairwise distance that must be considered for calculating the diameter of any distance. Therefore, the binary variable $t_{ijk} = 1$, if, and only if, both units i and j are assigned to district k, regardless of the order they are assigned in.

Constraint equations (5.8), (5.9) and (5.10) define the workload attributes, including risk, area, and diameter. Constraint (5.11) defines the workload of a district, k, as the linear combination of three attributes. Constraint equations (5.12) and (5.13) describe the absolute difference, AD_k, between the workload of district k and the average workload as the maximum of their differences. Constraint equations (5.14) and (5.15) guarantee that both x_i^{kc} and t_{ijk} are binary variables.

This formulation of the SNPDP model is computationally complicated. It has $\frac{3}{2}(mn^2 - mn)$ variables and $4m + n + \frac{5}{2}mn(n-1)$ constraints, which quickly make the model intractable as the number of units and districts increases.

5.3.6 Difference between the SNPDP and PPAC model

It is worth clarifying the differences between the SNPDP and the PPAC model (Curtin et al. 2010). The PPAC model optimises the delineation of the police sectors by assigning the existing beats to new sectors, following two steps. First, the optimal locations of the sector centres are identified by maximising the number of historical incidents that are within the acceptable service distance of a sector centre. Second, the sector boundaries are determined by assigning the beats to new sectors based on the incidents served by the sector centres.

The SNPDP is different from the PPAC in several aspects. First, the problem definition is different. While the PPAC seeks the maximal coverage of historical incidents, the SNPDP aims at a district design with a low average and deviation of the workload. Second, the basic units are different. In the PPAC, beat centroids are used as candidate locations for sector centres, and new sectors are generated by assigning the beats in an optimal way. In comparison, the SNPDP uses streets as the basic units for police districts, the output being a partition of the street network into several districts. Third, the role of the street network is different. In the PPAC, the optimal coverage is based on the street network distance between sector centres and historical incidents. In the SNPDP, the function of the street network is two-fold: the streets comprise the model units, and the network distance between the units is used.

5.4 The Heuristic Method

The problem discussed in Section 5.3 poses enormous computational challenges, especially when the problem size is large. Here, a two-step heuristic method for the PDP is proposed that incorporates the workload balance between districts.

This is partly inspired by the random-greedy-and-tabu-search (RG-TS) heuristic (Camacho-Collados, Liberatore and Angulo, 2015; Liberatore and Camacho-Collados, 2016), which is proposed for PDPs based on grids and census blocks. It consists of two steps: RG and TS. In the RG step, an initial solution is produced by randomly selecting a first node for each district and expanding the districts in a greedy way while keeping their contiguity. In each iteration of expansion, the algorithm extends the current solution by assigning a node to a district, and by choosing the node and the districting that results in the best feasible solution. The algorithm ends when all the nodes have been assigned. In the TS step, the initial solutions are improved by a tabu search method, which is introduced in Section 3.3.1.

Although the RG-TS proves efficient for PDPs based on grids and census blocks, it has several limitations. First, in the RG step, the initial solution expands gradually to involve one unit in one step, which is slow for a large problem. Second, the original tabu search procedure is also slow for a large problem, and the efficiency requires improvement.

Here, an efficient heuristic graph-partition-and-tabu-search (GP-TS) is proposed to approximately solve the SNPDP. First, a graph-partitioning method is used to generate a set of feasible solutions, and second, the solutions are improved using an improved TS procedure.

In the GP step, the Karlsruhe Fast Flow Partitioner Evolutionary algorithm (KaFFPaE) (Sanders & Schulz 2012) was used to generate the initial partition of the area. The KaFFPaE is a distributed evolutionary algorithm for tackling the graph partitioning problem, and it is well-suited for the SNPDP for the following reasons. First, it enforces the contiguity of each district by imposing a penalty on incontiguous partition. Second, the KaFFPaE can generate a balanced partition, in which the maximum block (or district) weight is constrained to $(1 + \epsilon)$ times the average block weight, and the block number and deviation tolerance (ϵ) are provided by the user. Moreover, the KaFFPaE includes a random component, meaning that it can generate a set of feasible and varying solutions to guarantee substantial exploration of the solution space; however, KaFFPaE only accepts node attributes as inputs, and it is unable to implement a district diameter. Therefore, it is used to generate balanced districts only with respect to area and risk.

The second step uses a TS algorithm to improve the solutions. The TS algorithm iteratively modifies the current solution by moving a node from one district to a neighbouring district, in the hope of finding a better solution. The advantage of the TS lies in its ability to avoid local optimality by putting the recently visited solutions in the tabu list ('tabu' meaning forbidden). In each iteration, the best move is chosen if it outperforms the best solution ever explored. Otherwise, the best non-tabu move is chosen. The TS terminates when any of three criteria are met: 1) the running time exceeds the predefined maximum run time; 2) no improvement over the current solution is found and all moves are tabu; and 3) the number of iterations since the last improvement over the best solution exceeds the threshold.

As the TS involves numerous times of evaluating a new solution, it is extremely slow for large problems. Here, the efficiency of the TS was improved by introducing two new features: a fast delta evaluation for new solutions, and rapid identification of feasible moves, as explained below.

Delta evaluation is a technique used in evaluating a modified solution quickly. As the modification involves only one or two districts, the delta evaluation updates only the attribute of the affected districts, which makes it more efficient than a complete evaluation following the definitions. The cost of the delta evaluation is the storage of the attributes of each district in the current solution. In each iteration of the TS, delta evaluation was used to evaluate all feasible moves.

A move generates a feasible solution if it does not disconnect any district. Since a move involves an origin district and a destination district, this condition means that both districts must not be disconnected by the move. According to Ricca & Simeone (2008), the migrating node: (a) must be adjacent to a node in the destination-district, and (b) must not be a cut-node for the origin district. A brute force method to identify all feasible moves would test every possible move and the connectivity of the two involved districts, which is very inefficient. Here, a fast identification of feasible moves is proposed. To test condition (a) efficiently, for a given district, the boundary is the subset of nodes that are adjacent to a node in a different district. In this implementation, the boundary of a district is stored and dynamically updated. In order to test condition (b) efficiently, for a migrating node, x, a sub-graph of the origin district is constructed that comprises the first- and second-order neighbours of x. Clearly, x is not a cut-node of the origin district if, and only if, either this sub-graph is connected or the graph of the origin district excluding x is connected, and the test can stop if the sub-graph is connected. Since the sub-graph is small, and the typical urban street network is well connected, the sub-graph is very likely to be connected, resulting in considerable time-savings by avoiding testing the entire district.

5.5 A Grid-based Aggregation Model of the SNPDP

This section addresses an alternative grid-based approach to deriving a street-based districting plan. This approach initially aggregates the policing workload into grids, then solves a grid-based PDP, ultimately transforming the solution into a street network. Using this model, one can use the formulation and algorithms for the grid-based districting problem. Moreover, this model alleviates the computational cost, as the grid-based PDP contains fewer units.

This approach is illustrated in Figure 5.2. In the first step, a spatial grid is created, which consists of uniform, square cells of a predefined size. The grid uses the Rook’s neighbourhood relation, in which any two cells that share a boundary line are neighbours. The distance between two neighbouring cells is equal to the grid size. Each cell is intersected with the street network, and its area size (or crime risk) is set as the sum of the intersections with the streets. The cells with no intersections with the network are removed from the grid.

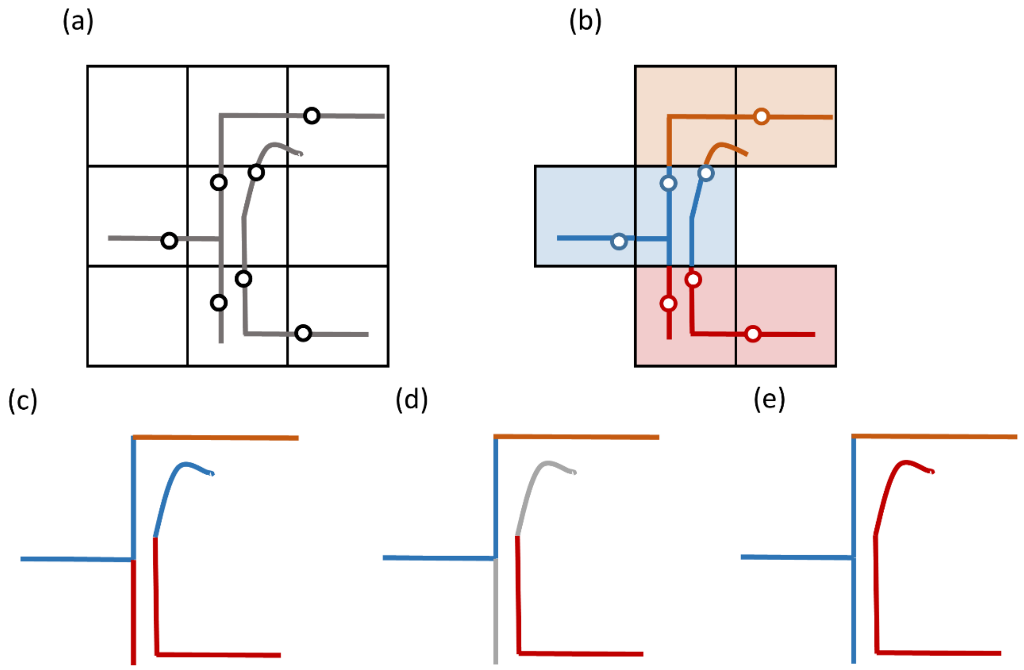


Figure 5.2 Illustration of the aggregation model for producing street-based districts (n=6, m=3). (a) The street network is overlaid with a spatial grid. The dots represent the street midpoints. (b) Cells with no intersections with streets are removed. A grid-based PDP is formulated and solved, resulting in three districts. (c) The grid solution is transformed into a street solution. The blue and red districts are unconnected and subject to refinement. (d) Two streets are marked as unassigned (coloured grey), as they are outside the main connected component of the district. (e) The partial solution is improved by greedy expansion, leading to a connected street solution

In the second step, a grid-based districting problem is formulated, in a similar way to that of the SNPDP in Section 5.3, except that the units represent cells. This model can be solved by a MIP solver, or the heuristic proposed in Section 5.4.

In the third step, the districting solution on grids is transformed into a solution based on the street network, as follows. A street is assigned to the same district as the cell in which the midpoint of the street is located. After all the streets are assigned, the street solution is produced and is then refined to enforce the contiguity of each district. Specifically, for each unconnected district that consists of more than one connected component, the nodes outside the main connected component are marked as unassigned. A partial solution is obtained in which all districts are connected. Then the districts expand in a greedy fashion, which is similar to the RG step in Section 5.4.

The size of the cells influences both the computational demand of all the steps and the quality of the final solution. In particular, it has a three-fold effect. First, it affects the assignment of the attributes of the streets to the cells. A small size implies a large number of cells and thus increases the computational cost of the intersection. Second, it affects the computational cost of solving the grid-based PDP. Similarly, the smaller the cell size, the higher the computational demand. Third, it influences the quality of the final solution. At a finer scale, it is more likely to achieve a high-quality solution. Because of the manifold effect, cell size should be carefully tested and selected.

This aggregation model results in an aggregation error, as the original street units are replaced by grids. To my knowledge, there has been no previous measurement of the aggregation error in this problem. Here, a straightforward measure to evaluate the aggregation error from the grid-based model is proposed. The measure is the relative gap between the solution, D', by the aggregation model and the solution, D, by the SNPDP, as follows:

$$gap(\%) = \frac{obj(D') - obj(D)}{obj(D)} \times 100\% \tag{5.16}$$

The gap is positive if the solution from the SNPDP outperforms that of the aggregation model. A gap close to 0 would indicate an insignificant difference between the two models.

5.6 **Experiments**

In this section, the solution quality of the proposed heuristic is demonstrated on several simulated SNPDP datasets, and then the problem-solving capability and scalability of the SNPDP model and the heuristic is illustrated in the Camden case study. The algorithms are implemented in Java 1.8, and were tested on a machine with two i7-4790 Intel CPUs and 32.0 GB DDR3 memory. Gurobi (version 7.5.2) is a commercial optimisation software package that was utilised to obtain the optimal or sub-optimal objective value of the MIP formulation.

5.6.1 Optimality Analysis

Here, the solutions of the GP-TS are compared with the solutions generated by Gurobi on a set of simulated datasets. A number of instances are used that were extracted from the Camden street network, with different numbers of units (n = 25, 50, 100, 200) and districts (m =3–9). The Camden street network is introduced in Chapter 4 and in Section 5.6.2. Each instance was generated by randomly selecting a node as the seed and gradually expanding the graph using the breadth-first search until the desired instance size was reached. The computational complexity of the SNPDP was discussed in Section 5.3, and the above problem size was selected to guarantee that at least some of the instances could be solved optimally within the 24-hour time limit for Gurobi.

The experimental settings were: for the parameters in the SNPDP, equal weights were adopted for both the workload function, equation (5.11), and the objective function, equation (5.1), meaning that wt_R, wt_A, and wt_{Dmt} are set to 0.333, and α is set to 0.5. For each instance, the GP-TS was conducted for 10 runs and, in each run, the maximum runtime of the TS was set to 0.5 hour, which is a tradeoff between solution quality and computation time. Hence, the time limit of the GP-TS is approximately five hours. Two parameters of TS, namely tabu length (i.e., size of the tabu list) and max iteration (i.e., maximum iteration since the last improvement), were both set to the number of units. It was noted that, in all the experiments, the TS did not terminate until the maximum runtime had elapsed. The combination of GP and Gurobi was also analysed, in which the GP solution was used as the initial solution of Gurobi. This led to the same result as the pure Gurobi solver. A list of parameter values is given in Table 5.1.

Table 5.1 Parameter values for the optimality analysis

Parameter	Meaning	Value
n	number of units	25, 50, 100, 200
m	number of districts	3, 4, 5, 6, 7, 8, 9
wt_R	weight of risk	0.333
wt_A	weight of area	0.333
wt_{Dmt}	weight of diameter	0.333
α	weight of workload average in SNPDP objective	0.5
time limit of Gurobi		10 h
number of runs using heuristic		10
tabuLength	size of the tabu list	equals number of units
maxIteration	maximum iteration since the last improvement	equals number of units
maxRuntime	maximum runtime	0.5 h

Table 5.2 shows the results of the instance using 25 units, as Gurobi failed to identify feasible solutions for the other instances within the time limit. The benchmark solution is the best

solution by Gurobi. The gap (%) is the relative quality of the solution from the heuristic compared to that of the benchmark solution.

$$gap(\%) = \frac{obj\ value\ by\ heuristic - obj\ value\ of\ benchmark}{obj\ value\ of\ benchmark} \tag{5.17}$$

Table 5.2 Computational results for the simulated datasets using Gurobi and the GP-TS

		Gurobi			GP-TS	
n	m	Solution time (h)	Objective value	Optimal?	Best gap (%)	Average gap(%)
25	3	0.03	0.1871	Y	0.00	0.00
25	4	0.04	0.1397	Y	0.00	0.00
25	5	0.24	0.1082	Y	12.80	12.80
25	6	11.3	0.1003	Y	0.89	0.89
25	7	15.6	0.0828	Y	0.00	11.55
25	8	24	0.0723	N	0.00	0.99
25	9	24	0.0605	N	9.11	18.58

As indicated in Table 5.2, Gurobi identified the optimal solutions for five problems (m = 3–7), and the GP-TS found the optimal solutions for three of them (m = 3, 4, 7). Note that for m = 7, the solution time needed by the GP-TS is five hours, which is significantly less than the 15.6 hours required by Gurobi. For all the problems in Table 1, the best gap and average gap are within 13% (with one exception), implying that most of the solutions found by the GP-TS are fairly close to the benchmark solutions. In addition, for problems with n > 25 (not shown in Table 1), while Gurobi failed to find a feasible solution within 24 hours, the GP-TS was capable of deriving a solution in a short time. This highlights the capability of the GP-TS to identify high-quality solutions in an acceptable time.

5.6.2 The London Borough of Camden Example

Camden, a borough in northwest Inner London, is policed by the MPS. The total area of Camden is 21.8 km^2, and the total street length is 347.0 kilometres. Recently, the MPS has been implementing and testing a street-level predictive policing system. By integrating street-level crime risk prediction into the design of patrol districts, the SNPDP model has practical implications and would improve the efficiency of policing operations.

In this study, Ordnance Survey ITN data was provided by via its MasterMap product. This borough includes 5575 street segments, with an average length of 62 metres, the maximum 800 metres, and 82% of the lengths less than 100 metres. Although link segmentation (that divides the links into approximately equal sizes) is common in network-based analyses (Xie and Yan, 2008; Yamada and Thill, 2010), it was not conducted as part of this research, as equal link size is not required for the districting task, and most of the streets have acceptable lengths. The street-level crime risk data is computed using the NTKDE model (Rosser *et al.*, 2017) and the historical burglary crime records (see Figure 5.3). Figure 5.4 shows the CLC map of Camden (Cole *et al.*, 2015), which reveals the characteristics of different parts of the borough, and is used to interpret the signatures of different patrol districts.

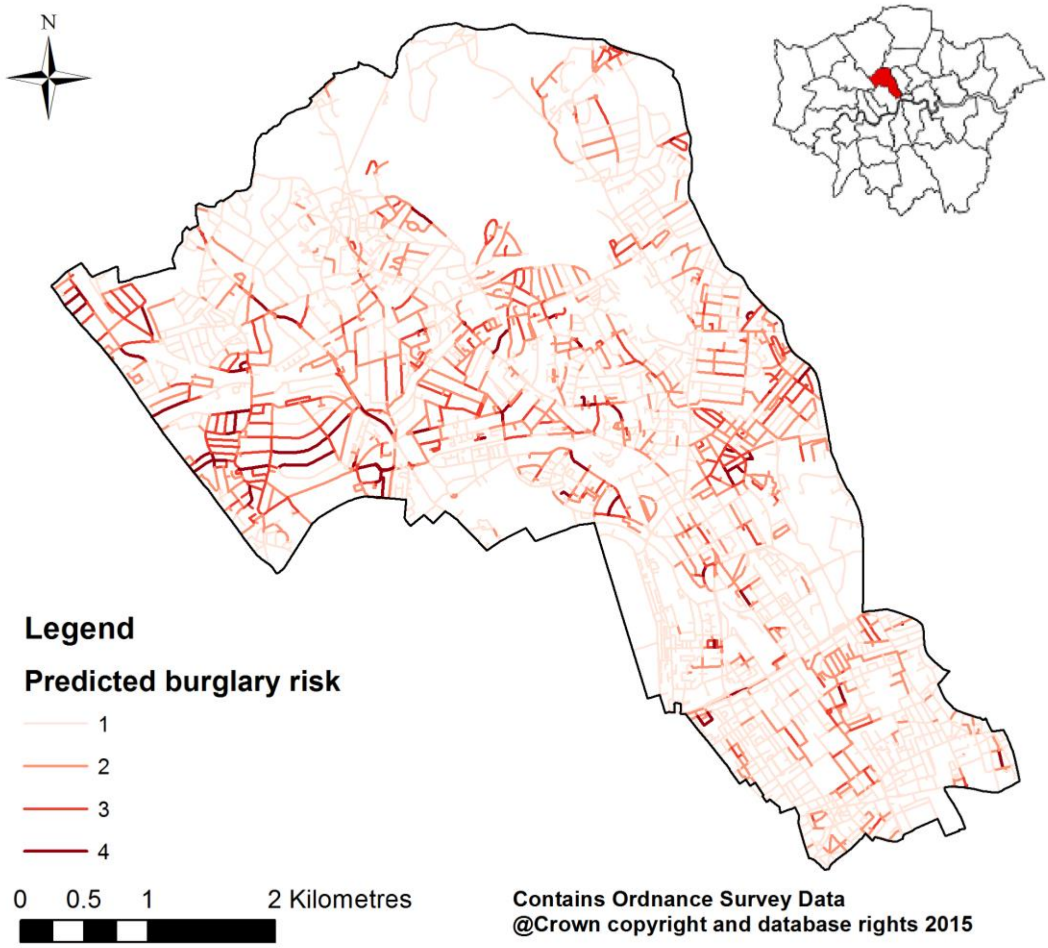


Figure 5.3 The predicted burglary crime risk in Camden.

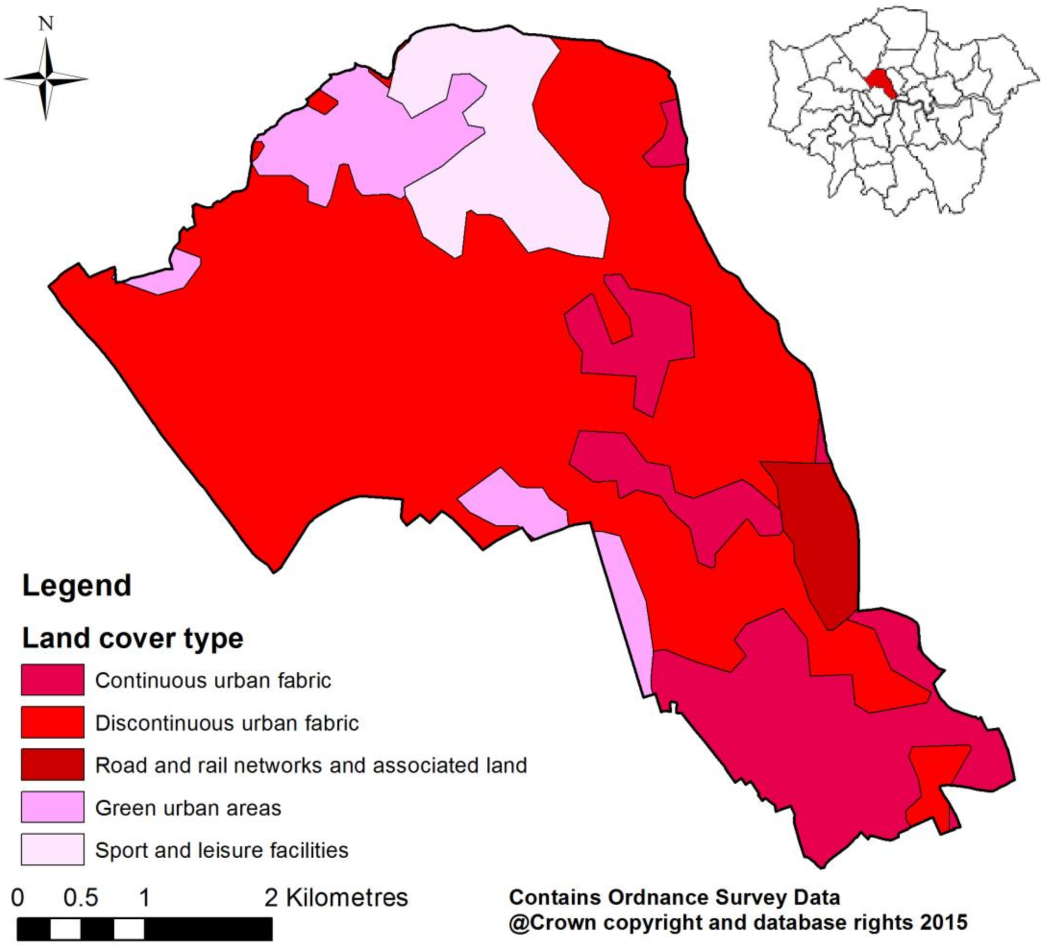


Figure 5.4 The 2012 CLC map of Camden

5.6.3 Results and Discussion

First, the GP-TS and RG-TS were used to produce districts in Camden, and to compare their efficiency and solution quality. Second, the best districting solution were focused on, and the characteristics of each district were interpreted. Finally, the aggregate grid-based model was formulated, using different cell sizes, and the aggregation error was discussed.

The configuration of the TS is the same for the optimality analysis. As with the optimality analysis, wt_R, wt_A, and wt_{Dmt} are set to 0.333, and α is set to 0.5. In operations, the weight values would be determined by the practitioners and experts. The number of districts is set to six so that each district is of moderate size.

Table 5.3 shows the computational results from the GP-TS and RG-TS for the SNPDP in Camden. To produce an initial solution, it took 30.0 and 298 seconds by the GP and RG procedures, respectively, meaning that the GP is considerably more efficient than the RG. The best and the average objective values obtained by GP-TS are significantly superior to those by RG-TS. The results demonstrate the advantages of GP-TS over RG-TS for large problems.

Table 5.3 Computational results of GP-TS and RG-TS for the SNPDP in Camden

	Time for an initial solution (s)	Best objective value	Average objective value
GP-TS	30.0	0.128	0.133
RG-TS	298	0.138	0.151

Figure 5.5 shows the district map with the best districting plan, and Table 5.4 presents the value of the attributes and workloads for each district. The six districts have different characteristics, depending on the ambient crime risk and land use types. In south Camden, the dominant land cover is ‘Continuous urban fabric’, meaning that more than 80% of the total surface is covered by buildings, roads and artificially surfaced areas. This part is largely used for education, transport, and business. Accordingly, Districts 1 and 2 in this area feature the largest accumulative street lengths. In contrast, north Camden is dominated by ‘Green urban areas’ and ‘Sport and leisure facilities’, and this is where the large green space, Hampstead Heath, is situated. Consequently, District 6 includes sparse streets and has the largest diameter. Moreover, in northwest Camden, District 5 features the highest crime risk score, which can be explained by a cluster of high crime risk streets (see Figure 5.3). The different signatures of the districts may have practical implications for policing operations.

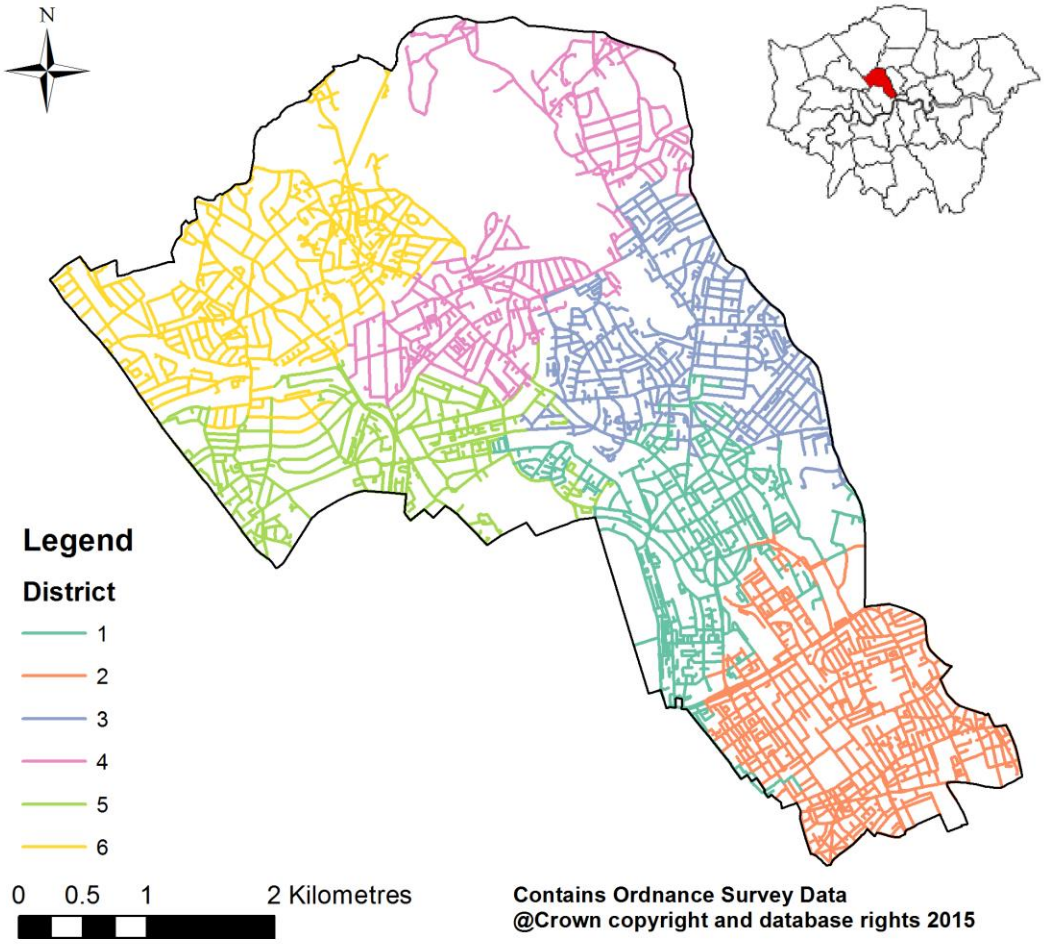


Figure 5.5 The best districting solution for Camden

Table 5.4 A summary of the best solution for Camden using GP-TS

District	Number of streets	Area	Risk	Diameter	Workload
1	1121	0.192	0.152	0.394	0.246
2	1110	0.181	0.131	0.339	0.217
3	965	0.159	0.196	0.383	0.246
4	682	0.132	0.118	0.549	0.267
5	847	0.165	0.260	0.339	0.255
6	850	0.170	0.142	0.426	0.246

The grid aggregation model was also used to generate the districting plan for the streets. The cell size was selected as 100, 150, 200, and 250 metres, respectively, in order to study the effects of cell size. The cell size was chosen to reflect a reasonable patrol unit for an officer, as well as having manageable computational demands. The chosen sizes are on the same scale as a case study on crime hotspot prediction in the same region (Adepeju, Rosser and Cheng, 2016). For each cell size, the grid-based districting problem was formulated and solved by the GP-TS heuristic to produce 10 solutions, which were then transformed and refined to obtain contiguous street solutions.

Table 5.5 shows the computational results of the SNPDP and the aggregation model using varying cell sizes. The benchmark solution is the best solution from the SNPDP. As the cell size increases, the unit number decreases dramatically. For all selected cell sizes, none of the street solutions that were directly transformed from the grid solutions is contiguous, meaning that refinement is needed. After refinement, the relative gap from the contiguous street districting solutions to the best solution obtained by solving the SNPDP was computed. The quality of the best solution decreases as the cell size increases, indicating that a small cell size leads to a small aggregation error. The best solution obtained by grid size of 100 metres is close to that provided by the SNPDP, with a gap of 1.68%; however, the aggregation model requires several extra steps of data pre-processing and solution refinement. The results confirm the advantages of the SNPDP over the aggregation model.

Table 5.5 Computational results of the two modelling approaches with varying cell size

Cell size(m)	Unit number	No. of contiguous street solutions before refinement	Average gap (%)	Best gap (%)
street	5575	10	4.18	0.00
100	1946	0	2.73	1.68
150	944	0	7.91	6.44
200	562	0	10.32	6.88
250	373	0	9.30	8.11

5.7 Chapter Summary

In this study, the SNPDP model is proposed. This is a novel approach to incorporating the street network structure and street-level predictive crime risk into the design of police districts. This model is multi-criteria-based, in that the objectives include the efficiency and balance of the district workload, and the workload is a combination of the crime risk, area size, and district diameter.

To efficiently solve large cases with SNPDP, a combined heuristic GP-TS was developed, which combines a graph partitioning algorithm and a tabu search procedure. In comparison with solutions found by an exact solver (Gurobi), the GP-TS is capable of producing high-quality solutions quickly for the SNPDP.

The SNPDP model was successfully tested in the case study of Camden, generating efficient and balanced patrol districts. The results confirm the advantages of the GP-TS over an existing heuristic RP-TS, with respect to the computation time and solution quality. Moreover, the superior capability of the SNPDP model is supported by comparison with an aggregation grid-based model, in terms of solution quality. The different characteristics of the patrol districts can be identified through a combination of the crime risk distribution and the land use map, and the results can guide the planning of policing strategy.

Putting a street-based district design into the police operations, it has both several advantages and certain limitations. The main advantage is that it incorporates the street network and the street-level predictive crime risk, and ensures a workload balance. Moreover, each district is guaranteed to be well connected in the street network. Therefore, it is expected to enhance the efficiency of patrol planning and emergency response. However, several limitations exist regarding the incorporation of census data and police activities that occur off the streets. First, as the street segments are incompatible with the current census units (e.g., census block, output area), it is difficult to carry out demographic research on a district consisting of streets. Second, while most of the police activities take place on, or near, the street, there are situations in which police should deviate from the street network, such as carrying out tasks in an area, such as a large green space, that has no streets.

This study opens up avenues for future research. First, this study proposes street-based and grid-based models to solve districting problems with edge demands, and offers a simple approach to comparing the solution quality. It would be useful to compare the two models systematically and to quantify the associated aggregation errors. Second, the street-based district design could be improved in certain ways so as to incorporate census data and off-street police activities. One possible approach would be to transform the districts into area-based districts whilst preserving the workload balance and connectivity of the new districts.

Chapter 6

INFREQUENT PATROL ROUTING: A BALANCED ROUTE DESIGN

6 INFREQUENT PATROL ROUTING: A BALANCED ROUTE DESIGN[3]

6.1 Chapter Overview

In this chapter, the infrequent patrol routing problem and its application in police patrolling are described. Firstly, in Section 6.2, an introduction to the infrequent patrol routing problem, and the modelling of this problem as a MMMDRPP is provided. In Section 6.3, the formulation, algorithm and evaluations for the MMMDRPP are described. Then, in Ssection 6.4, demonstrations of how to model the route designs for police patrolling using the MMMDRPP, and how to choose routes for implementation using a case study in Camden, are provided. In Section 6.5, the proposed algorithm is further tested on several benchmark problems. Finally, the chapter is summarised in section 6.6.

6.2 Introduction

The Infrequent Patrol Routing Problem (IPRP) concerns covering a set of target locations at least once in a patrolling interval, given a set of patrolling agents (Wolfler Calvo and Cordone, 2003; Willemse and Joubert, 2012). In police patrolling, the agents are patrol officers, and these officers may start from different locations. The route design can be implemented within a certain district generated by the PDP (Chapter 5), or in the whole territory. As this study focusses on the patrol routing problem on the street network, in which the patrol targets are street segments, the corresponding IPRP can be modelled as an arc routing problem (ARP). A detailed overview of the ARP and its variants can be found in Section 2.3.2.

More specifically, IPRP can be modelled as a Min-Max Multiple-Depot Rural Postman Problem (MMMDRPP) for the following reasons. First, the min-max objective function, which is used to minimise the length of the longest route, ensures a similar route length and workload balance among different patrol officers. This function not only ensures the fair scheduling and balanced routes but also increases the overall route efficiency and serves the demands as early as possible (Ahr and Reinelt, 2002). Second, the MMMDRPP takes into account the fact that officers may start from multiple police stations (or depots).

[3] Part of this chapter was presented in the following publication: Chen, H., Cheng, T. & Shawe-Taylor, J., 2018. A Balanced Route Design for Min-Max Multiple-Depot Rural Postman Problem (MMMDRPP): a police patrolling case. International Journal of Geographical Information Science, 32(1), pp.169–190.

The aim of this study is to formulate and solve the Min-Max Multiple-Depot Rural Postman Problem (MMMDRPP), and to apply this problem to route design of police patrolling. Specifically, the contributions from this chapter include:

- The introduction of a new arc routing problem, namely the MMMDRPP, which can generate balanced routes for postmen from different depots;
- The proposal of a routing algorithm TABU-PATROL that approximately solves the MMMDRPP.
- The approximation of an optimal solution value for the MMMDRPP, with the proposal of three lower bounds;
- A model for the route design of police patrolling based on the MMMDRPP, with a demonstration of how to design and choose routes for police patrolling in practice using the proposed algorithm; and
- Further testing of the TABU-PATROL algorithm with multiple benchmark routing problems, demonstrating that TABU-PATROL is robust enough to generate quality solutions in different instances.

6.3 A TABU-PATROL Algorithm for MMMDRPP

In this section, the MMMDRPP is first formulated, then the TABU-GUARD algorithm for the MMRPP is briefly reviewed, and a TABU-PATROL algorithm for MMMDRPP is proposed, based upon TABU-GUARD. After that, solution evaluations of the MMMDRPP are described.

6.3.1 Problem Definition

In the formulation of the MMMDRPP, a transportation network is represented as a connected graph consisting of nodes and edges. Each node represents a street intersection in the geographic space, and each edge between two nodes represents a street segment with a travelling distance. The set of required edges consists of segments that must be traversed at least once by a route. The MMMDRPP is defined as follows:

Input: A connected and undirected graph, G, with:

- a set of nodes V, and a set of edges E;
- a set of required edges $R \subseteq E$;
- a non-negative distance function for the edges, f: $E \rightarrow \mathbb{R}_{\geq 0}$; and
- a number of postmen, k, and a list of k depots D. The depots may be different.

Output: A set of k closed routes, T.

Objective: To minimise the distance of the longest route of T, denoted by $Lmax$.

Constraints:

- The closed route constraint: the *k*-th route of *T* should be a closed route (or a cycle) and traverse the *k*-th depot of *D*; and
- The traversal constraint: each edge in *R* should be traversed at least once by T.

A mathematical model of the MMMDRPP is formulated as follows, which is adapted from the MMRPP (Ahr and Reinelt, 2002):

$$\min Lmax \tag{6.1}$$

subject to:

$$Lmax \in \mathbb{R}^{+} \tag{6.2}$$

$$\sum\nolimits_{e\in E} w(e)x^{i}(e) + w(e)y^{i}(e) \leq Lmax, \forall\ i = 1, \dots, k \tag{6.3}$$

$$x^{i}(e) \in \{0,1\}, y^{i}(e) \in \mathbb{N}_0, \forall\ e \in E, i = 1, \dots, k \tag{6.4}$$

$$\sum\nolimits_{i=1}^{k} x^{i}(e) = 1\,, \forall\ e \in \mathrm{R} \tag{6.5}$$

$$\sum\nolimits_{e\in\delta(v)} x^{i}(e) + y^{i}(e) \equiv 0\ (mod\ 2), \forall\ v \in V, i = 1, \dots, k \tag{6.6}$$

$$x^{i}(\delta(S)) + y^{i}(\delta(S)) \geq 2 * x^{i}(e), \forall\ S \subseteq V \backslash d_i, e \in E(S), i = 1, \dots, k \tag{6.7}$$

The binary decision variable, $x^{i}(e)$, equals 1 if edge e is serviced by tour C_i and $e \in R$. The integer variable, $y^{i}(e)$, represents the number of times that edge e is traversed by C_i without being serviced. $\delta(v)$ denotes the set of edges that are incident to the node, v, and $\delta(S)$ denotes the set of edges that have one, and only one, endpoint belonging to the node set, S. $E(S)$ denotes the set of required edges both endpoints of which belong to the node set S. *Lmax* is defined as the positive length of the longest patrol route, and is expressed by constraint equations (6.2) and (6.3). Constraint equation (6.5) guarantees that each required edge is serviced. Constraint equation (6.6) enforces that each single route must be a closed walk. It requires that the number of times any node is traversed by any route must be even. If the number is odd, the route cannot form a closed walk. Constraint equation (6.7) ensures that the depot is traversed. For any route and any sub-graph (S, $E(S)$) excluding the depot node, if any edge in the sub-graph is serviced ($x^i(e) = 1$), this route must "enter" or "leave" the sub-graph at least twice. Otherwise, the route will probably fail to include its depot. For example, a route would violate constraint equation (6.7) if it traversed only one required edge and excluded the depot.

The integer program of the MMMDRPP is extremely difficult to solve directly. Assuming that there are m required edges ($|R|=m$) and k postmen, there would be approximately 2^m sub-graphs and more than $2^m \times k$ inequalities from constraint equation (6.7).

The MMMDRPP is an NP-hard problem, which can be proved in two steps. First, the MMRPP is NP-hard, as the RPP reduces to the MMRPP with a single postman, and the RPP is known to be NP-hard (Lenstra and Kan, 1976). Second, the MMRPP reduces to the MMMDRPP with only one depot.

A feasible solution for the MMMDRPP is a set, T, of k tours $T = \{C_1,\ldots,C_k\}$. The length of a route, C_i, which is the total length of its traversed edges, is denoted by $w(C_i)$. For a feasible solution T, $wmax(T) = max(w(C_i))$ denotes the length of the longest route. The objective of the MMMDRPP is to find a solution, T^*, with the minimum $wmax(T)$.

$SP(u,v)$ denotes the shortest path between two nodes, u and v. This can be calculated by Dijkstra's algorithm (Dijkstra, 1959). The shortest path between two edges $e_1 = (u_1, u_2)$ and $e_2 = (v_1, v_2)$, which is called $SP(e_1, e_2)$, is defined as the shortest path out of $\{SP(u_1, v_1), SP(u_1, v_2), SP(u_2, v_1), SP(u_2, v_2)\}$. Note that $SP(e_1, e_2)$ does not include e_1 and e_2. Furthermore, $SPR(e_1, e_2)$ denotes the set of required edges that are contained in $SP(e_1, e_2)$.

6.3.2 TABU-GUARD for the MMRPP

TABU-GUARD was proposed to tackle the MMRPP (including the MMCPP) by Willemse and Joubert (2012). This algorithm has three stages: 1) it generates multiple random initial solutions by a PI method (see Section 6.1); 2) it uses IMPROVE-SOLUTION, which iteratively attempts to shorten the longest route by exchanging or reallocating edges between different routes; and 3) it employs the TABU-SEARCH procedure to further improve the solutions.

This TABU-SEARCH procedure has several components and multiple different choices. First, in terms of neighbourhood exchange procedures to create neighbouring solutions, there are three choices: 1) REMOVE-INSERT-NEIGHBOURHOOD (RIN); 2) EXCHANGE-NEIGHBOURHOOD (EN); and 3) a combination of RIN and EN. Second, there are three tabu list strategies for determining whether a move is forbidden: 1) COMPLEX-TABU; 2) SIMPLE-TABU; and 3) SIMPLE-AGGRESSIVE-TABU. Third, TABU-SEARCH has two parameters – the maximum number of iterations with improvement, *tmax*, and the tabu tenure, α.

In the application of TABU-GUARD, the first step is to manually determine the component combination and select the parameter values. Given a set of problem instances, nine different combinations of neighbourhood exchange procedures and tabu list strategies are tested over all instances, with α ranging from 1 to 16 and *tmax* fixed at 500. Then, the best three combinations with appropriate α ranges, are chosen and used.

TABU-GUARD is not readily applicable to the MMMDRPP, as it does not consider multiple depots. Specifically, if TABU-GUARD was to be used for the MMMDRPP, the first step would be to generate k routes by assigning each depot with the farthest required edge. However, the resulting routes would be excessively long, and would overlap considerably with each other. For example, assuming there are two depots, d_1 and d_2, as well as a required edge e_1, where e_1 is the farthest required edge to d_1 and $distance(d_1, e_1) > distance(d_2, e_1)$, intuitively, it would be more efficient to assign e_1 to d_2 rather than to d_1. Therefore, it is necessary to develop new algorithms that are well-suited to the MMMDRPP.

6.3.3 TABU-PATROL for the MMMDRPP

Here, the TABU-PATROL algorithm for the MMMDRPP is proposed. TABU-PATROL is influenced by TABU-GUARD, but it is different from TABU-GUARD in three respects. Firstly, it includes two novel constructive methods to generate the initial routes, which are better suited for the MMMDRPP than the PI of TABU-GUARD. Secondly, it contains only the most efficient neighbourhood exchange procedures and tabu list strategy, which makes the algorithm simpler than TABU-GUARD. Thirdly, in contrast to the manual parameter selection in TABU-GUARD, TABU-PATROL uses an automatic and systematic parameter selection.

TABU-PATROL works in three stages (Figure 6.1). First, it produces a number of random initial solutions using GENERATE-INITIAL-SOLUTION. Second, the initial solutions are improved by using IMPROVE-SOLUTION. Third, the solutions are further improved by a SIMPLE-TABU-SEARCH procedure, which includes automatic parameter selection.

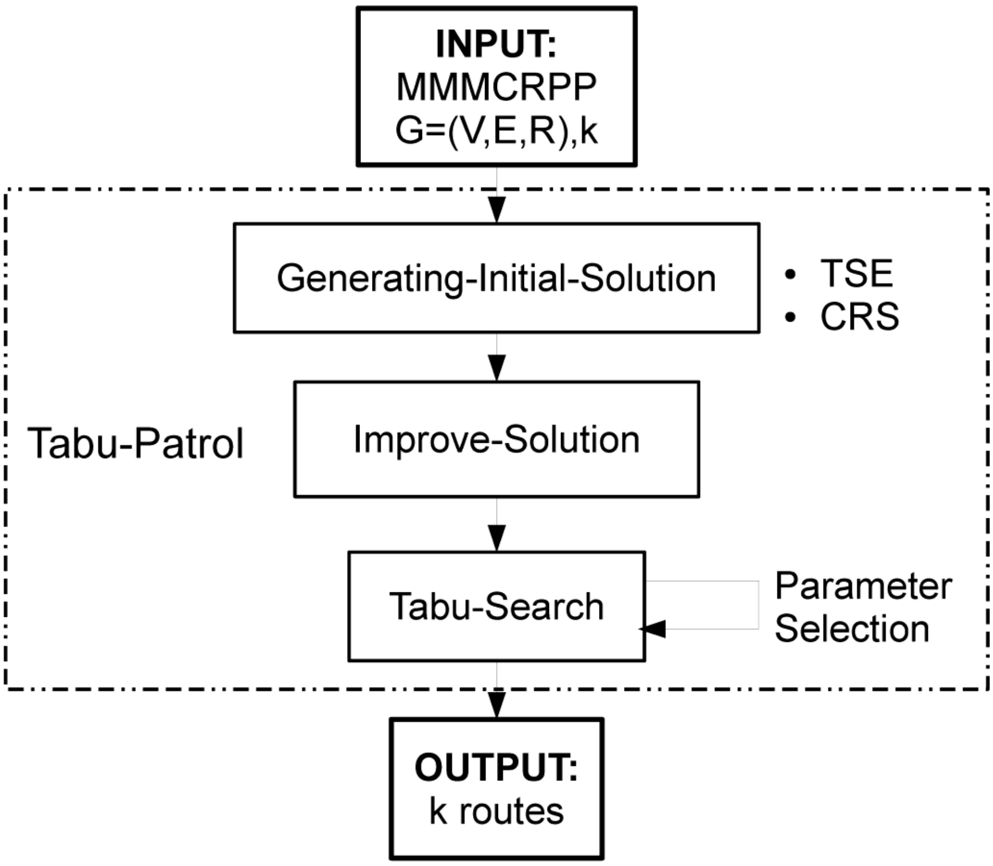


Figure 6.1 Flowchart illustrating the steps involved in TABU-PATROL

6.3.3.1 Stage 1: Generating Initial Solutions

The first stage of TABU-PATROL is to generate a number of different initial solutions using two new construction methods. The first method – TOUR-SPLIT-EXCHANGE (TSE) – follows the route-first-cluster-second paradigm (Prins, Lacomme and Prodhon, 2014), creating a giant tour to traverse all required edges, decomposing the tour into a given number of sub-tours, and then improving the routes by exchanging stations. The other method – CLUSTER-ROUTE-SPLIT (CRS) – follows the cluster-first-route-second paradigm (Prins, Lacomme and Prodhon, 2014). It firstly forms clusters of required edges based on the distance to the stations, then computing the tour to traverse the required edges within each cluster, and further splits the tours if necessary.

TSE is illustrated in Figure 6.2. Firstly, it creates a single tour, *1-tour*, that traverses all required edges at least once. Constructing such a tour is an undirected RPP, which can be near-optimally solved by the heuristic of Christofides et al. (1981). The length of *1-tour* is *w*(*1-tour*). The tour is encoded by the scheme of Willemse and Joubert (2012), in which only required edges are explicitly represented per route and it is assumed that the shortest path is used between consecutive required edges. Thus, *1-tour* is denoted as $RPPT=\{r_1, SPR(r_1, r_2), r_2, \ldots, r_{n-1}, SPR(r_{n-1}, r_n), r_n, SPR(r_n, r_1)\}$, where $r_i \in R$. Then, *1-tour* is simplified by only retaining the first position of any duplicated required edge.

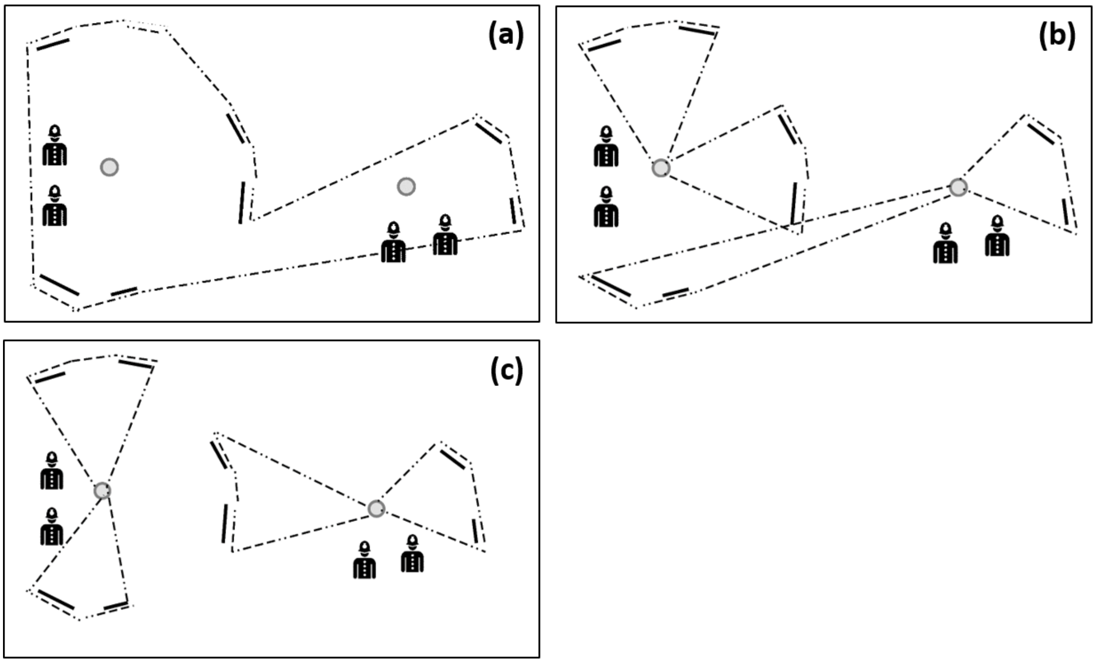


Figure 6.2 Illustration of TSE. (a) Creating a single tour traversing all required edges. (b) Spliting the tour into k parts with similar lengths, and linking these parts with the nearest depot. (c) Exchanging the depots among different tours so as to minimise the length of the longest tour

Secondly, the obtained *1-tour* is split into k partitions with similar lengths. Here D' denotes the list of unused depots and, initially, $D' = D$. The first splitting node is selected randomly in *1-tour*, and then (k-1) splitting nodes are determined in order to divide the tour into k segments $\{P_1, P_2, \ldots, P_k\}$ that have similar lengths. The i-th route $\{d, P_i, d\}$ is formed by finding the depot, d, from D', such that the length of the route, $\{d, P_i, d\}$, is (one of) the shortest, and then d is removed from D'. This step is repeated k times to form k routes.

The resulting solution, $T = \{C_1, \ldots, C_k\}$, is feasible for the MMMDRPP, as all required edges are traversed. It is then further improved by exchanging the depots between any pair of tours. If the exchange reduces a longer tour length, the depots are exchanged to form two new tours. This process is repeated until no improvements can be gained.

Since the first splitting node is chosen randomly and is different each time that TSE is invoked, the TSE is capable of generating different initial solutions.

The second method CRS is illustrated in Figure 6.3. Given the list of stations, D, the set of unique stations is denoted by $DU=\{d_1,\ldots,d_m\}$. For each $i\in[1,m]$, k_i is the number of occurrences of d_i in D. Then, the set of required edges is divided into m clusters $\{CL_1, \ldots, CL_m\}$ based on the shortest distance to the depots, and the i-th cluster corresponds to the i-th depot d_i. Afterwards, for each cluster, CL_i, k_i tours should be formed, such that each tour traverses d_i and each edge in CL_i should be traversed at least once. This sub-problem can be solved by invoking the TSE procedure. As the random factor of TSE is brought into CRS, CRS is also capable of generating different initial solutions.

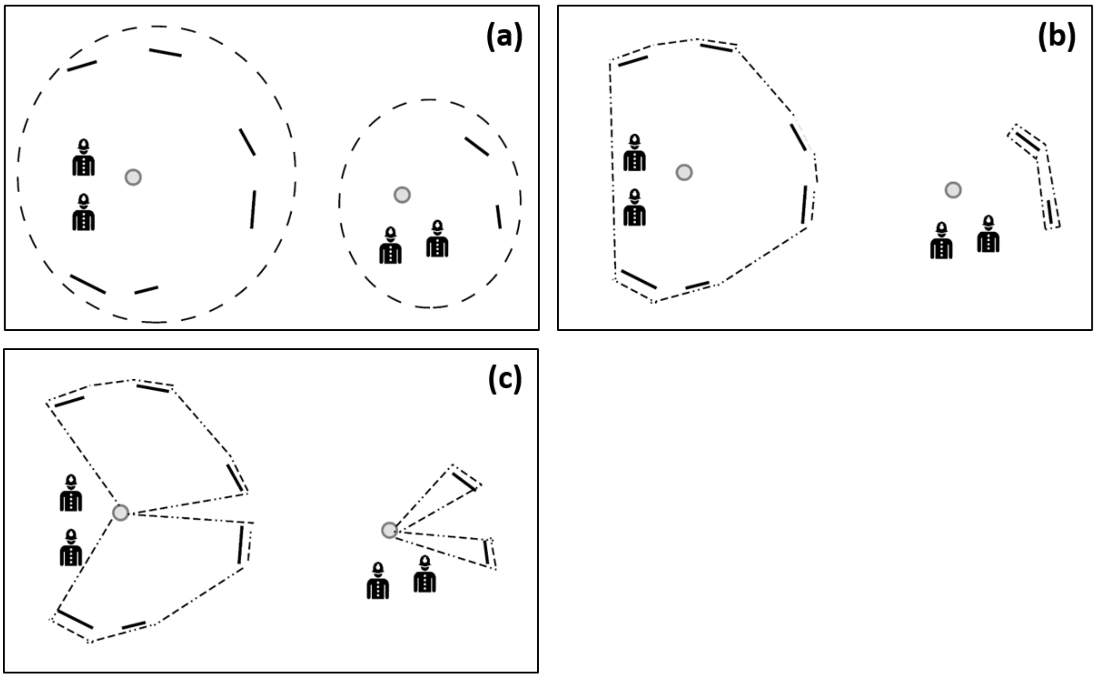


Figure 6.3 Illustration of CRS. (a) Clustering the required edges based on their nearest depots. (b) Generating the routes for each cluster of required edges. (c) Spliting each tour into certain parts and linking the parts with the corresponding depots

Both TSE and CRS take into account the different depots for the postmen. TSE generates routes from different depots by assigning each partition of the *1-tour* to the closest depot, and by exchanging the depots of certain routes to further shorten the tour length. In this way, the generated routes may have similar lengths, and contain edges that are close to each other, as they are from a partition of the *1-tour*. However, the route lengths are affected by the depot-edge distance. Meanwhile, CRS deals with different depots by assigning each required edge to the nearest depot before generating routes. It is guaranteed that the generated routes contain edges in close proximity, but they do not necessarily have similar lengths.

As TSE and CRS follow different paradigms, it cannot easily be judged which method is better. Therefore, both of them were used in TABU-PATROL, and the one that generated a better solution was selected.

6.3.3.2 Stage 2: Improving the Initial Solutions

The initial solutions from TSE and CRS can be excessively long and unbalanced, and are improved in Stages 2 and 3. For Stage 2, the IMPROVE-SOLUTION procedure developed by Willemse and Joubert (2012) was directly used without adjustment. IMPROVE-SOLUTION iteratively attempts to improve a solution, *T*, by exchanging or reallocating edges between different tours such that the longest tour is shortened. It terminates when no improvement is obtained in the current iteration.

6.3.3.3 Stage 3: SIMPLE-TABU-SEARCH

In this stage, SIMPLE-TABU-SEARCH is used to find new solutions by reallocating the edges in the current solution. SIMPLE-TABU-SEARCH, as used here, is different from the TABU-SEARCH algorithm in TABU-GUARD in two respects. First, SIMPLE-TABU-SEARCH is simpler and easier to use than TABU-GUARD, as it includes only one tabu list strategy (SIMPLE-TABU) and only one exchange neighbourhood (EN) procedure. SIMPLE-TABU and EN were selected, as they performed the best in the tests of Willemse and Joubert (2012) and the pilot tests in this study. Second, the parameters of SIMPLE-TABU-SEARCH are selected by an automatic algorithm, which leads to a non-biased parameter configuration and solutions of higher quality. The manual parameter selection in TABU-GUARD is biased, as it is usually guided by the personal experience of the user, and is influenced by random factors in experiments.

SIMPLE-TABU-SEARCH begins with constructing neighbourhood solutions by modifying the initial solution. Then the neighbourhood solutions are examined and the best is chosen for the next iteration. The chosen new solution may be better or worse than the previous solution. However, the search may cycle by continuously moving back to the visited solutions such as the local optimum, which hinders further progress. To tackle this problem, SIMPLE-TABU-SEARCH uses a tabu list to keep track of the recent neighbourhood moves, known as tabu moves. Each tabu move is stored for a certain number of iterations, and then ‘expires’ and is removed from the list. If a new neighbourhood solution comes from a move that coincides with any tabu move, it cannot be chosen. However, if no non-tabu neighbourhood solutions exist, the algorithm can move to a neighbourhood solution that comes from a tabu move, but has better performance than the best solution found. The best solution found is called the incumbent solution, and is tracked in the search. SIMPLE-TABU-SEARCH ends when no non-tabu neighbourhood solution is available and no neighbourhood solution that outperforms the incumbent solution exists. Finally, SIMPLE-TABU-SEARCH returns the incumbent solution. Below, details on the essential components of SIMPLE-TABU-SEARCH are presented.

Constructing neighbourhood solutions

SIMPLE-TABU-SEARCH constructs neighbourhood solutions using EN, which removes an edge from one tour and adds it to another. EN is applied to two tours at a time – the longest tour, C_i, and any other tour, C_j. It removes an edge, *u*, from C_i and another edge, *v*, from C_j, and then inserts *u* into C_j and *v* into C_i by using INSERT-ARC (Willemse and Joubert, 2012) to form a new solution. This process is repeated for any two-edge combination with *u* of C_i and *v* of C_j. Then, the overall best solution and the best non-tabu solution, obtained from this process is improved by IMPROVE-SOLUTION, and returned. SIMPLE-TABU-SEARCH moves to the overall best solution if it outperforms the incumbent solution, or otherwise moves to the best non-tabu neighbouring solution.

Tabu list and stopping criteria

The tabu list is updated in each iteration. After moving to a neighbourhood solution, T_{new}, SIMPLE-TABU-SEARCH adds to the tabu list the edges *u* and *v* that have been moved to create T_{new}. Any move that involves edges *u* or *v* leads to a tabu solution. This tabu list strategy is called TABU-SIMPLE (Willemse and Joubert, 2012). The edges *u* and *v* are removed from the tabu list after a predefined number of iterations (tabu tenure) have passed.

SIMPLE-TABU-SEARCH has two stopping criteria. First, it terminates if all neighbourhood solutions are tabu, and the overall best neighbourhood solution does not outperform the incumbent solution. Second, it terminates if the incumbent solution has not improved after a predefined number of iterations. After SIMPLE-TABU-SEARCH terminates, the incumbent solution is returned as the improved solution.

Parameter configuration

SIMPLE-TABU-SEARCH has two parameters – the maximum number of iterations with improvement, *tmax*, and the tabu tenure, α. As the initial solutions involve randomness, it is challenging to find the best parameters of SIMPLE-TABU-SEARCH to improve different initial solutions. Therefore, the parameter tuning requires repeated experiments and statistical testing. This task is an instance of the algorithm configuration problem and can be solved by iterated racing (Balaprakash, Birattari and Stuetzle, 2007; López-Ibáñez *et al.*, 2016).

Iterated racing is a method used for automatic configuration, which consists of three steps: 1) sampling new configurations according to a particular distribution; 2) selecting the best configurations from the newly sampled ones via racing; and 3) updating the sampling distribution in order to bias the sample towards the best configurations. These steps are repeated until a termination criterion is met: reaching a minimum number of surviving configurations, a maximum number of used instances, or a predetermined computational budget. More details of iterated racing can be found in López-Ibáñez et al. (2016).

6.3.4 Solution Evaluation for the MMMDRPP

Since the MMMDRPP is NP-hard, three lower bounds are proposed to approximate the minimum longest route length. Then, a solution for the MMMDRPP is evaluated in terms of its lower bound gap (LBG), in comparison with the tightest lower bound. Given TABU-PATROL and similar algorithms involving randomness, its efficiency must be evaluated based on different solutions from it. The algorithm was run multiple times, and the best LBG and mean LBG of the solutions, and the execution time, were used to evaluate the efficiency of the algorithm.

Here, the lower bounds are introduced. The first lower bound – SPT-LB – is adapted from the STP-LB for the MMRPP. For any required edge, e, there must be one nearest depot from D and the corresponding shortest path length, $SPT(e) = \min_{d \in D} w\big(SP(d,e), e, SP(e,d)\big)$. In an optimal solution, each required edge has to be traversed, and the route to traverse e is no shorter than $SPT(e)$. Thus, the longest route in an optimal solution must be no shorter than the $\max_{e \in E} SPT(e)$.

The second lower bound – the IPR-LB – is obtained by solving the relaxed integer programming problem of the MMMDRPP without the constraint equation (6.7). This lower bound is the same as the IPR-LB for the MMRPP, as the different depot nodes are not considered.

The third lower bound is called the Rural Postman Tour Lower Bound (RPT-LB). The idea is to create a RPT using the optimal solution $T^* = \{C_1, \ldots, C_k\}$ to the MMMDRPP. Let $wmax(T^*)$ be the length of the longest tour of T^*. Then, generate the shortest cycle to traverse all depot nodes at least once – the well-known Travelling Salesman Problem – which can be near-optimally computed by an approximation algorithm (JGraphT, 2017), with the approximation ratio of 2. That is, if the shortest cycle and the calculated cycle are C^* and C', respectively, it can be proved that $w(C^*) \le w(C') \le 2 \cdot w(C^*)$. Then, the tours of T^* and C^* can be combined to construct a connected and larger tour, T', with $w(T') = w(T^*) + w(C^*) \le k * wmax(T^*) + w(C')$. As T' traverses all required edges of E, it is a feasible solution to the RPP on $G = (V, E, R)$

with one postman (*1-RPP*). Moreover, this *1-RPP* can be near-optimally solved by the algorithm of Christofides et al. (1981), with an approximation factor of 3/2. Let $RPPT^*$ be the optimal RPP tour and $RPPT'$ the obtained RPP tour, and so $w(RPPT^*) \leq w(RPPT') \leq 3/2 * w(RPPT^*)$ exists. As T' is a feasible solution to the 1-RPP, then $w(T') \geq w(RPPT^*) \geq 2/3 * w(RPPT')$. It is not difficult to prove that $k * wmax(T^*) \geq w(T') - w(C') \geq 2/3 * w(RPPT') - w(C')$, hence $[2/3 \cdot w(RPPT') - w(C')] / k$ is a new lower bound for the MMMDRPP.

To evaluate the quality of a solution *T*, the LBG is used, which is defined as:

$$LBG(T) = \frac{wmax(T) - maxLB}{wmax(T)} \tag{6.8}$$

where *wmax(*T*)* is the length of the longest route in *T* and, *maxLB* is the tightest lower bound for the given problem. The range of *LBG*(*T*) is (0,1), and a lower value indicates higher solution quality.

Based on different solutions generated by TABU-PATROL or similar algorithms, the best LBG and Mean LBG can be calculated. An algorithm with a lower best LBG and lower mean LBG is more efficient than others. Moreover, as the time constraint is essential in such applications, an efficient algorithm should generate quality solutions within an acceptable time.

6.4 Case Study: Balanced Routes for Police Patrolling in Camden

In Section 6.4 and 6.5, different MMMDRPP cases are used to test and evaluate TABU-PATROL with three constructive methods – TSE, CRS, and PI. Here, the PI method was adjusted, and used as a baseline method, as follows: for each depot, $d \in D$, it creates a route $\{d, SPR(d,r), r, SPR(r,d), d\}$ by finding the untraversed required edge, *r*, that is furthest from *d*. Then, it adds the remaining untraversed required edges to an existing tour that results in the minimum route length increase.

Here, TABU-PATROL is used to develop patrol routes from multiple stations in Camden, which is policed by the MPS, and which had five major police stations between 2011 and 2012. The MPS employs different types of officers, including foot patrol officers, community support officers, and senior officers. The responsibilities and tasks of these officers vary, but the foot patrol officers were selected for this study, as their main task is to patrol and cover an area. The daily patrolling activities in a day are divided into three shifts (i.e., early, late and night), each lasting for about eight hours. In each shift, officers are dispatched to patrol the environment and to cover certain important locations.

The purpose of this case study was to generate routes for patrolling in Camden, using a group of officers from different stations in order to balance the route lengths amongst them. To illustrate the balanced patrol routes, the number of officers was assumed to be five and 10, equating to one and two patrollers from each police station, respectively. It was also assumed that the police patrollers are dispatched to cover the crime hotspots in Camden (see Figure 6.4). Herein, 69 street segments were identified as crime hotspots, having the highest crime densities and covering 1% of the total road length. The time interval of crimes used was from March 2011 to March 2012, and the types include theft, burglary, homicide, battery, arson, motor vehicle theft, assault, and robbery.

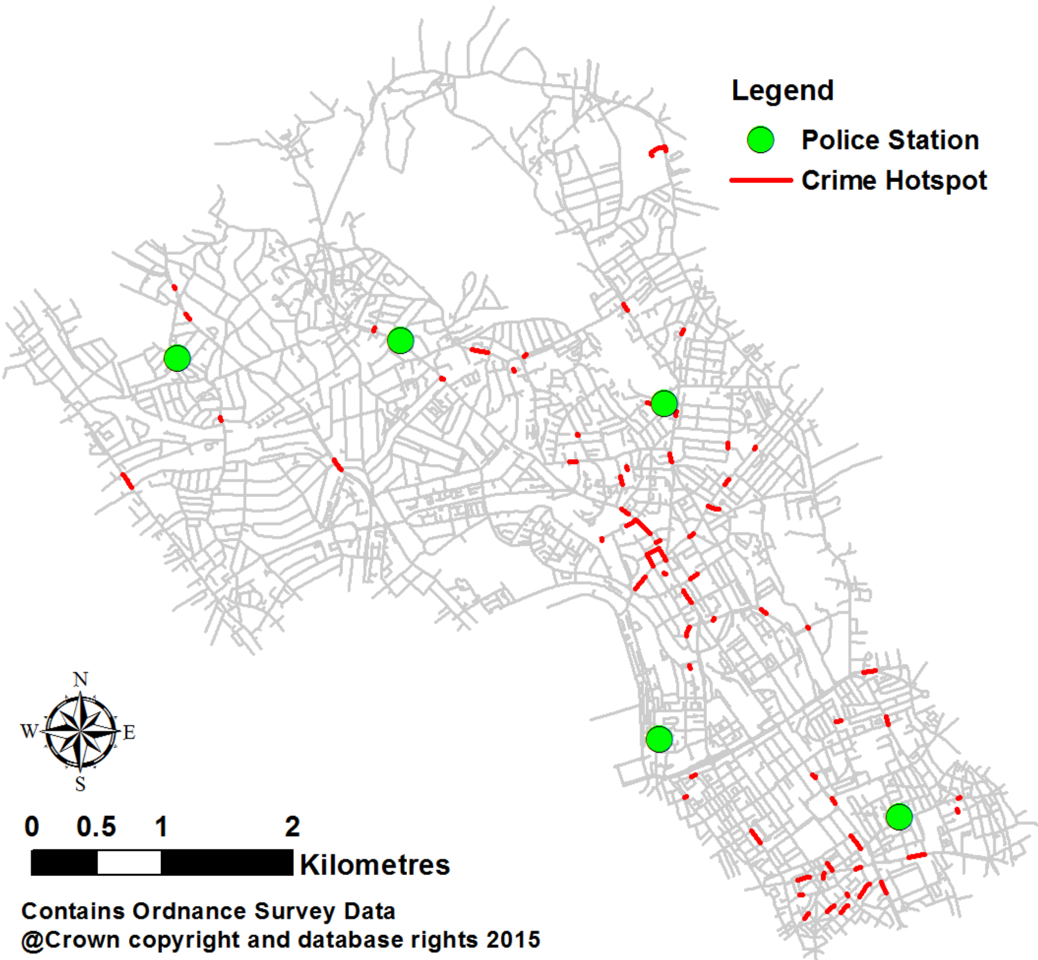


Figure 6.4 Crime hotspot map of Camden, London (UK)

All algorithms described in Sections 6.3 and 6.4 were coded in Java version 1.8.0_66, and run in the Java SE Runtime Environment, on two 3.60 GHz Intel® Core™ i7 CPUs with 32.0 GB of RAM. The integer programming of IPR-LB was solved by 64-bit Gurobi Optimizer version 6.5.2 (Gurobi Optimization, 2016). Parameter configuration was conducted in the R environment version 3.2.2 (R Core Team, 2015) and the IRACE package version 2.1.1662 (López-Ibáñez *et al.*, 2016). Data analysis was also performed in the R environment.

6.4.1 Computational Results

Here, how to use TABU-PATROL to generate patrolling routes is first discussed, and then the solutions using the best LBG, mean LBG, and execution are evaluated. Following this, the routes from TABU-PATROL are demonstrated on the street map.

TABU-PATROL generated 10 different initial solutions for five and 10 patrollers by using three constructive methods – TSE, CRS, and PI, and then improved the solutions using IMPROVED-SOLUTION and SIMPLE-TABU-SEARCH.

The parameter configuration of SIMPLE-TABU-SEARCH was selected by the iterating-racing algorithm from the IRACE package. Both parameters, *tmax* and α, are treated as categorical variables, with ranges of {100, 200, 300, 400, 500, 600, 700} and {1, 2, ..., 15}, respectively. These ranges were selected based on the TABU-GUARD algorithm and pilot experiments. The evaluation measure used by iterated racing is LBG. Up to 300 experiments are performed in the iterated racing. The parameter configuration was conducted on each combination of constructive method and *k*. Table 6.1 shows the best configuration for each algorithm and *k*. The best configurations differ from case to case. Overall, the best performance of SIMPLE-TABU-SEARCH is obtained using a high value of *tmax* and medium value of α.

Table 6.1 Best parameter configuration for each combination of (method, k)

(method, *k*)	(TSE,5)	(TSE,10)	(CRS,5)	(CRS,10)	(PI,5)	(PI,10)
α	9	7	12	6	9	5
Tmax	500	700	700	600	600	700

Before evaluation, the three MMMDRPP lower bounds for patrolling Camden, with the patroller number *k* being five and 10, were calculated, as reported in Table 6.2. All results are given in metres, with the tightest bounds in bold; the column '*Tightest*' in Table 6.2 indicates the tightest lower bound of the three. SPT-LB does not consider *k*, and hence does not change with *k*, whereas IPR-LB and RPT-LB decrease as *k* increases. For both values of *k*, SPT-LB is the tightest lower bound, probably because of an edge that is significantly distant from all the depots. For clarity, here lower bound refers to the tightest lower bound for the given *k*.

Table 6.2 Lower bounds for Camden hotspot map (in metres)

k	SPT-LB	IPR-LB	RPT-LB	Tightest
5	**4748**	1554	2259	**4748**
10	**4748**	784	1129	**4748**

Table 6.3 and Table 6.4 show the results of the best and mean LBGs in Camden, using the three methods and after the three stages. For all methods, large LBGs were observed in the initial solutions. The solution quality was slightly improved after using IMPROVE-SOLUTION, and then considerably improved after SIMPLE-TABU-SEARCH. From Table 6.3, when $k = 5$, in the initial solutions, the mean LBG was the lowest with TSE and the highest with PI. In the final solutions, TSE performed the best, in terms of both Best LBG and Mean LBG, while PI produced the worst solutions with the highest best LBG and mean LBG. The performance of CRS is slighter better than PI. Large LBGs were still observed in the final solutions. This may be partly explained by the weak lower bound in this case, which may be much lower than the true value.

When k increases to 10, a different comparison of the three algorithms was observed. Overall, both TSE and CRS significantly outperformed PI. The final solutions of CRS were better than those of TSE, although the initial solutions of CRS were not the best.

Table 6.3 Summary of 10 solutions for *k* = 5 in the Camden case study, using the three methods

	TSE		CRS		PI	
	Best LBG	Mean LBG	Best LBG	Mean LBG	Best LBG	Mean LBG
Initial	0.600	0.631	0.643	0.648	0.765	0.781
Improved	0.579	0.611	0.642	0.645	0.764	0.779
Final	0.437	0.453	0.443	0.508	0.446	0.567

Table 6.4 Summary of 10 solutions for *k* = 10 in the Camden case study, using the three methods

	TSE		CRS		PI	
	Best LBG	Mean LBG	Best LBG	Mean LBG	Best LBG	Mean LBG
Initial	0.292	0.395	0.420	0.526	0.710	0.711
Improved	0.231	0.367	0.393	0.513	0.710	0.711

Final	0.128	0.188	0.086	0.121	0.196	0.513

Table 6.5 shows a comparison of the execution times for the three algorithms in the Camden case study. Both TSE and CRS are capable of generating 10 different solutions within 200 seconds, which is very time-efficient. Although PI is quicker than TSE and CRS, PI performs the worst in terms of solution quality.

Table 6.5 Summary of execution times (in seconds) for generating 10 solutions for the Camden case study using the three methods

	TSE	CRS	PI
k=5	193	191	131
k=10	84	49	42

To visualise the patrolling routes, the initial and final solutions by TSE and CRS, with 10 patrollers, are shown on the street map (Figure 6.5). Each colour represents the route of a patroller, and the colour of the route overlap is randomly selected from the overlapped routes. When using TSE, the routes in the initial solution constitute a giant tour to traverse all required edges, and the routes have considerable overlaps. In the final solution, the overlap of routes has been significantly reduced, and the route lengths have become much more balanced. Moreover, each new route is compact and is restricted to a small area, which is suitable for patrol operations. In contrast, in the initial solution by CRS, each route only traverses the required edges that are adjacent to the station, and visually the route lengths are distinct and unbalanced. In the final solution, the new routes become far more balanced and less overlapped, while the compactness of the routes is retained. The results reveal the differences between TSE and CRS in generating initial routes, and also acknowledges the effectiveness and efficiency of the SIMPLE-TABU-SEARCH algorithm in balancing the routes, alleviating route overlaps, and keeping the routes compact.

6.4.2 Applying TABU-PATROL in Practice

Here, applying TABU-PATROL in practice is discussed. In practice, patrol planners can use TABU-PATROL to generate efficient patrol routes. First, the route planners decide on the available patrollers and then use TABU-PATROL with TSE and CRS to generate a sufficiently large set of solutions. The number of solutions depends on the time constraint; as is shown in the case study, generating 10 solutions is a good trade-off between execution time and solution quality. Second, the generated solutions are evaluated by the LBG, and a subset of solutions with the lower LBG is chosen. Third, these chosen solutions can be visualised by GIS to show the spatial extent and the traversed area of each route. The route planners can evaluate the feasibility of these routes, based on their knowledge and experience of policing and the local area. Fourth, the route planners can now choose one or several feasible solutions, the sequence in which they are implemented, and which routes should follow a clockwise or anti-clockwise direction. If there are multiple patrolling shifts in the day, the route planners can choose a different solution for each shift, which makes the patrol routes unpredictable and flexible (Willemse and Joubert, 2012).

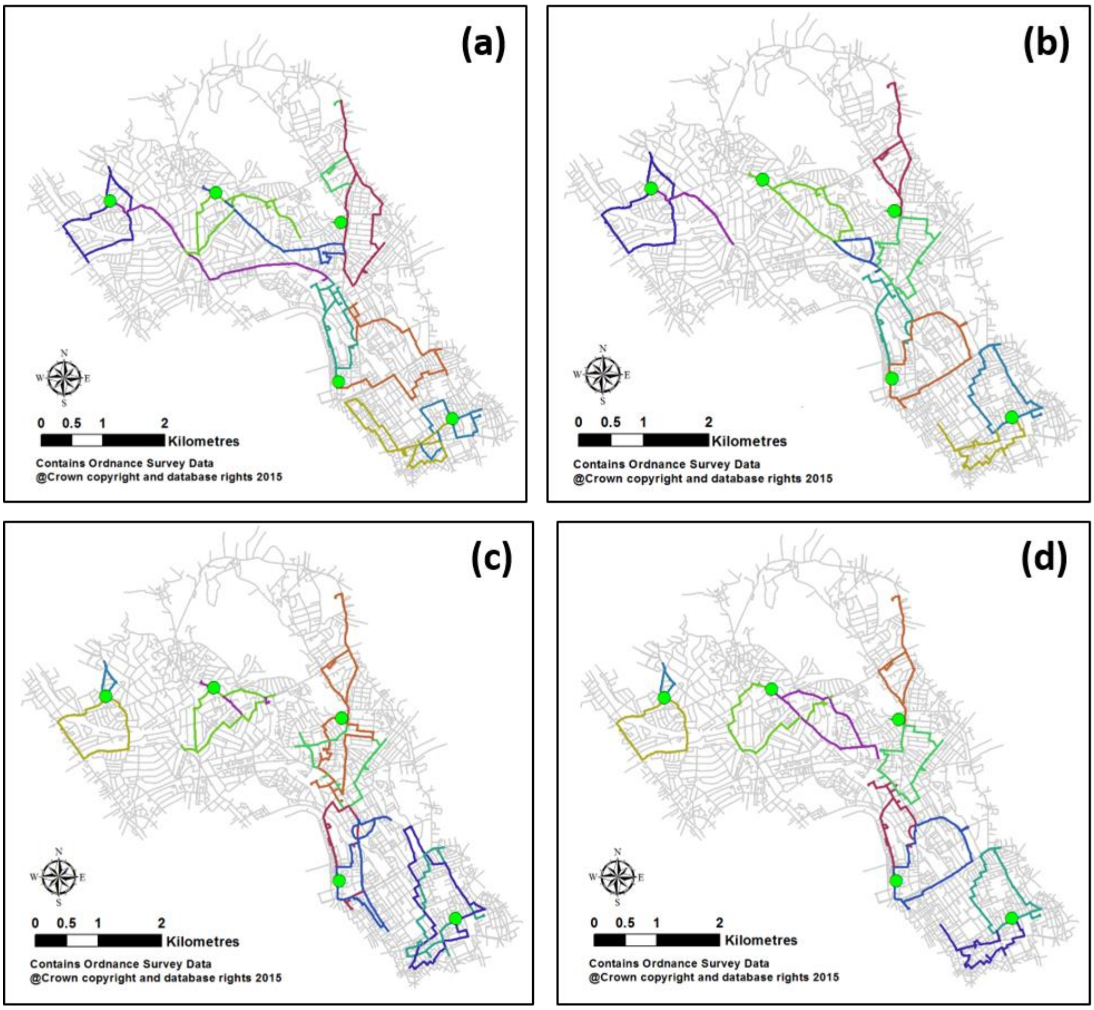


Figure 6.5 Comparison of the initial and final solutions of TSE and CRS with *k* = 10 for Camden police patrolling. a) an initial solution generated by TSE; b) improving the solution in (a) via TS; c) an initial solution generated by CRS; d) improving the solution in (c) via TS

6.5 A Case Study Using Benchmark Problems

For further testing of TABU-PATROL, the algorithm was run on eight problem instances that were originally proposed for the CARP by Li and Eglese (1996). These benchmark cases have been intensively adopted in the CARP (Belenguer *et al.*, 2006; Brandão and Eglese, 2008) and MMCPP/MMRPP (Ahr, 2004; Ahr and Reinelt, 2006; Willemse and Joubert, 2012). The original cases included demands on edges and only one depot, and were adapted for the

MMMDRPP as follows. First, as the MMMDRPP does not model edge demand, the edge demands are omitted. Moreover, a list of 10 depots is added to each instance to form the new MMMDRPP problems, which are called md-egl-**. The depot list of md-egl-e* and md-egl-s* instances is {1, 51, 31, 42, 12, 1, 51, 31, 42, 12}, and {1, 40, 32, 98, 3, 1, 40, 32, 98, 3}, respectively. The new depots are chosen to be evenly distributed in the network. To my knowledge, this is the first work on the MMMDRPP benchmark problems.

For each instance, TABU-PATROL generated and improved five different solutions for each k=2,…,10. The initial solutions are generated using TSE, CRS, and PI. Each method is coupled with the best SIMPLE-TABU-SEARCH parameter configuration for the Camden case and k = 10 (see Table 6.1). Table 6.6 reports the average of the mean LBGs and execution times over different values of k. For the complete results, please contact the author. For the benchmark problems, either TSE or CRS produced the best-quality solutions, except for the md-egl-e1 instance. The worst LBGs for TSE and CRS was 0.150 and 0.161, respectively, whereas the worst LBG for PI was 0.440. Overall, these methods have similar execution times on md-egl-e* problems, whereas on md-egl-s* problems, TSE is slower than CRS and PI. As CRS produces high-quality solutions within reasonable execution times, this method is ideal when good solutions have to be generated under the constraint of time. If there is no time constraint, it is recommended to test TSE and CRS to solve the MMMDRPP, and to choose the better results.

Table 6.6 Results of TABU-PATROL using three constructive methods on the MMMDRPP of eight adapted benchmark problems

				TSE		CRS		PI	
	V	E	R	Mean LBG	Time(s)	Mean LBG	Time(s)	Mean LBG	Time(s)
md-egl-e1	77	98	47	0.145	13.3	0.161	14.6	0.140	16.6
md-egl-e2	77	98	72	0.140	39.4	0.111	37.9	0.193	39.4
md-egl-e3	77	98	87	0.133	78.1	0.118	72.1	0.215	73.1
md-egl-e4	77	98	98	0.123	159	0.134	144	0.440	148
md-egl-s1	140	190	115	0.150	56.1	0.143	44.3	0.247	48.1
md-egl-s2	140	190	147	0.122	414	0.130	326	0.254	317
md-egl-s3	140	190	159	0.118	520	0.123	416	0.261	402
md-egl-s4	140	190	190	0.062	847	0.074	683	0.202	667

6.6 Chapter Summary

In this chapter, the infrequent coverage routing in police patrolling, and a balanced route design, were discussed. The infrequent coverage problem was formulated as a MMMDRPP, which seeks balanced routes from different depots to cover a street network. This problem is challenging because of the constraint of multiple depots and the objective of minimising the length of the longest route. To solve this problem, a novel TABU-PATROL algorithm was proposed, which follows three stages to generate and improve the routes using tabu search. To approximate the optimal solution value, three lower bounds were devised for this problem, namely SPT-LB, IPR-LB, and RPT-LB. Then, the solution quality was measured by using the LBG, and the algorithm efficiency was evaluated by using the best LBG and the mean LBG of its multiple solutions.

In the Camden case study, it was demonstrated how to model a route design for police patrolling as the MMMDRPP, using TABU-PATROL to generate balanced routes. The results show that TABU-PATROL can generate quality solutions within a reasonable time. In TABU-PATROL, three methods (TSE, CRS, and PI) were used to generate the initial routes, and it was concluded that TSE and CRS can generate better solutions than PI. How to apply TABU-PATROL in practice was also discussed.

Furthermore, the TABU-PATROL algorithm was tested on several adapted benchmark problems, with results showing that TABU-PATROL was robust enough to generate quality solutions for different cases.

Chapter 7

REPEATED PATROL ROUTING: AN ONLINE APPROACH

7 REPEATED PATROL ROUTING: AN ONLINE APPROACH [4][5]

7.1 Chapter Overview

In this chapter, the repeated coverage routing problem and its application in the police patrolling are described. This chapter is organised as follows: in Section 7.2, a brief introduction to the repeated patrol routing strategy and the objectives of this study are provided. In Section 7.3, the guidelines and evaluation measures are formulated for police patrol routing strategy, which is followed by a description of the BAPS in Section 7.4. In Section 7.5, an agent-based simulation environment of police patrols is described, in order to test the effectiveness of the proposed routing strategies. In Section 7.6, two case studies of police patrol are presented using realistic police and crime settings, so as to validate the BAPS. Finally, Section 7.7 is a summary of this chapter.

7.2 Introduction

The repeated patrol routing problem (RPRP) concerns patrolling or covering the targets repeatedly in a given time interval, using a fairly large patrol team. This problem is similar with the multi-agent patrolling problem (Almeida *et al.*, 2004) or multi-robot patrolling problem (Portugal and Rocha, 2011). A detailed review of these problems can be found in Section 2.3.3.

In the context of police patrolling, the RPRP is different from the IPRP described in Chapter 6 in many aspects. First, the definition is different. While the IPRP is defined as the routing problem to cover each target at least once, the RPRP concerns covering the targets repeatedly in the time interval. Second, the objectives are different. As is described in Chapter 6, the objective of IPRP is to balance the route length, which is represented by minimising the length of the longest route. However, this objective does not apply to the RPRP, as each patroller keeps moving in the patrolling process. The objectives of RPRP include efficiency, flexibility, scalability, unpredictability, and robustness, which are to be discussed in this chapter.

---

[4] Part of this chapter was presented in the following publication: Chen, H., Cheng, T. & Wise, S., 2015. Designing daily patrol routes for policing based on ant colony algorithm. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-4/W2, pp.103–109.

[5] Part of this chapter was presented in the following publication: Chen, H., Cheng, T. & Wise, S., 2017. Developing an online cooperative police patrol routing strategy. Computers, Environment and Urban Systems, 62, pp.19–29.

There are several challenges to solve the RPRP. First, there are usually tens of patrol officers that cooperate to cover tens of targets in the patrolling interval. Second, the routes are required to simultaneously satisfy multiple objectives. For example, the objective of efficiency concerns the minimisation of the average time length between two consecutive visits to a target, and the objective of flexibility requires more visits to the targets with a higher priority. Third, there are uncertain factors in the patrolling process. The number of patrol officers may change when some of them are required to stop patrolling and to respond to emergencies, and when and where the future emergencies occur are difficult to predict. These challenges make the RPRP difficult to solve.

In this study, the model is formulated in an online and multi-agent way. When a patrol officer completes her last patrolling task, a new target is generated to it immediately based on her current location, the status of the other officers, as well as other factors. This approach has two advantages. First, it enables the incorporation of multiple objectives and varying factors in the route generation, especially the real-time situation. Second, generating a target for each patroller one by one can largely reduce the action space, compared with that of generating actions for all patrollers at one time. This would substantially reduce the computational demand of the routing process.

7.3 Guidelines and Evaluation Measures

The major aim of police patrols is to prevent and reduce potential crime. A fundamental research question is: what makes a good police patrol routing strategy? Taking the characteristics and challenges of police patrolling into account, a good patrol routing strategy should follow the guidelines proposed in this study, including efficiency, flexibility, scalability, unpredictability, and robustness.

It is worth noting that several previous studies provide inspirations for the guidelines and measures. Two basic concepts, namely the idleness and global idleness of patrol targets, are firstly introduced by Machado et al. (2002) and directly used in this study. Portugal and Rocha (2013a) proposed the measure of team scalability to quantify the impact of team size in robot patrolling, and the measure of robustness to consider the influence of communication errors in robot patrolling. These measures are adapted here to account for the impact of team size and emergency response in police patrol. To our knowledge, this is the first study to use the measures for team scalability and robustness in police patrol. The other measures, including flexibility, unpredictability, and spatial scalability, are first proposed in this study.

7.3.1 Efficiency

Police patrol requires every important place or hotspot to be regularly and repetitively visited. Thus, efficiency is the foremost requirement for police patrol, which means patrols should minimise the time lag between two visits to a hotspot, for all hotspots. The measure for efficiency is based on idleness and global idleness of hotspots. The set of *n* hotspot segments is denoted as $H = \{h_1, h_2, \dots, h_n\}$. The instantaneous idleness (or idleness) of any hotspot $h_i \in H$ at time t is given by:

$$\mathrm{Idl}(h_i, t) = t - t_{lv}(h_i) \tag{7.1}$$

where $t_{lv}(h_i)$ corresponds to the last time when the hotspot h_i was visited by any patroller.

Thus, the average idleness of a hotspot h_i at time t is defined as:

$$AIdl(h_i, t) = (t - t_0)/(\mathrm{C}(h_i, t) + 1) \tag{7.2}$$

where t_0 is the starting time of the patrolling; $\mathrm{C}(h_i, t)$ is the accumulative number of visits to h_i at time *t*. Here, a value of one is added to the cumulative number to avoid the case where the denominator is 0, meaning that all the hotspots are visited at least once before the patrolling begins.

The global measure of idleness for all hotspots, the global average idleness (*GAI*), is given by:

$$GAI(\mathrm{t}) = \sum_{i=1}^{n} AIdl(h_i, t)/n \tag{7.3}$$

$GAI(t)$ is a variable dependent on time t and has different values at different time. The converged $GAI(t)$ is denoted as GAI, which is the measure of the efficiency.

7.3.2 Flexibility

Another aspect of an effective routing strategy is flexibility, which is the ability to cover hotspots with varied priority. Intuitively, the more important hotspots should have higher visiting frequency or lower average idleness. The flexibility can be measured using weighted global average idleness (*WGAI*) (Chen, Cheng and Wise, 2015), with $W(h_i)$ representing the weight of the hotspot h_i and the larger value representing the higher priority:

$$WGAI(t) = \sum_{i=1}^{n} W(h_i) \times AIdl(h_i, t) \Big/ \left[\sum_{i=1}^{n} W(h_i)\right] \tag{7.4}$$

The converged $WGAI(t)$ is denoted as $WGAI$, which is the measure of flexibility. This measure is a balance between two requirements, namely covering all hotspots and highlighting hotspots of greater importance.

7.3.3 Scalability

A promising patrol strategy should be applicable to different areas and with different numbers of patrollers, which is described as scalability. There are two types of scalability for a patrol routing strategy, namely *team scalability* and *spatial scalability*.

Team scalability is related to how well the given strategy performs as the number of patrollers increases (Portugal and Rocha, 2013b). A scalable patrolling strategy is able to adapt to different sizes of the team without severe performance degradation.

The team scalability can be evaluated based on a classical metric, called Balch's speedup measure (Balch and Arkin, 1994). In the patrolling problem, Balch's speedup measure reveals how much more efficient a group of patrollers are than just one patroller in completing the patrolling task, and is defined as:

$$v(R) = GAI(1)/[R \times GAI(R)] \tag{7.5}$$

where $GAI(R)$ is the GAI value of patrolling by R patrollers; $GAI(1)$ is the GAI value of patrolling by one single patroller.

If a group of R patrollers are more efficient in patrolling and achieve low GAI value, the resultant $v(R)$ would be greater than 1, and this performance is said to be superlinear. Linear performance is equal to 1, which means equal performance. Sublinear performance is less than 1, corresponding the lower efficiency of R patrollers in the task (Balch and Arkin, 1994). Since it is uncommon to patrol a large area using a single patroller, the measure is modified using a small size S rather than 1 as the reference, and the new measure is named as ST (Scalability of Team Size):

$$ST(R) = [S \times GAI(S)]/[R \times GAI(R)] \tag{7.6}$$

On the other hand, spatial scalability of a patrol routing strategy concerns its performance in different areas, including but not limited to the layout of the area or the road network, the density of crime hotspots and the distribution of police officers. Unlike the team scalability, the relevant factors of spatial scalability are various and difficult to quantify. Generally, the spatial scalability can be measured by comparing the efficiency of the designed strategy with the benchmark strategy. To our knowledge, no previous study has considered the spatial scalability of routing strategy in the context of police patrol.

Here, the influence of crime hotspot density level is considered, as one example of spatial scalability. For convenience, the notation *crime density level at x%* represents that the total length of crime hotspots covers x% of the total segment length of the road network. The measure of spatial scalability to compare the performance in two hotspot density levels is:

$$SS(L_i) = [GAI(L_i) - GAI(L_B)]/[GAI(L_B)] \tag{7.7}$$

Here $SS(L_i)$ refers to relative change of *GAI* performance due to the change of hotspot density from the baseline level L_B (e.g. level at 5%) to L_i (e.g. level at 10% or 15%).

7.3.4 Unpredictability

If the patrol routes or the visits to hotspots are easily deduced by potential criminals, they could commit a crime in the elapsed time between two visits, rendering police patrol ineffective (Sak, Wainer and Goldenstein, 2008). Therefore, it is important to keep the patrolling strategy unpredictable (Yin *et al.*, 2012; Sherman *et al.*, 2014). Specifically, the greater the uncertainty of police visits, the greater risk is perceived by potential criminals (Sherman, 1990). Therefore, it is essential to impart randomness or unpredictability in the patrolling strategy.

There are two kinds of randomness in a patrolling problem, namely randomness of patrol routes and randomness of visitations to hotspots. The former can be evaluated using entropy of a patrol strategy, as proposed by Chen and Yum (2010). However, the entropy quantifies the dissimilarity of multiple candidates of patrol routes and fails to measure the uncertainty of police arrivals at and departures from hotspots, or the randomness of visits to a given place. An experienced burglar waiting around a potential target for a time when no patrols are nearby would be more concerned about predicting the time of next visit rather than the routes of the patrol team. Here the randomness of visitations to hotspots is evaluated by the average value of the standard deviation of idleness of each hotspot:

$$ASDI(t) = \sum_{i=1}^{n} SDI(h_i, t)/n \tag{7.8}$$

where $SDI(h_i, t)$ is the standard deviation of idleness of hotspot h_i at time *t*, and $ASDI(t)$ is the average of the standard deviation of idleness of all hotspots at time *t*.

Similarly, the converged $ASDI(t)$ is denoted as $ASDI$. The higher value of $ASDI$ is favoured as it means higher unpredictability in the patrol routes and is more unlikely to be predicted by intruders. This is the first-time that the unpredictability is quantified for police patrol.

Apart from the random visiting time, there are other methods to impart unpredictability to patrol routes, such as accessing a long hotspot segment randomly from both ends. These random factors also confuse potential criminals and increase the deterrence of patrolling.

7.3.5 Robustness

Because officers are also responsible for dealing with emergency calls during patrolling, it is necessary to use a robust strategy, i.e., to remain effective even if some patrollers are dispatched for emergencies. The measure of robustness in this work is the relative increase of *GAI* value in the emergency scenario, in comparison with *GAI* in the non-emergency scenario, which is represented as:

$$RI\_GAI = (GAI\_Emerg - GAI\_Norm)/(GAI\_Norm) \times 100\% \tag{7.9}$$

where RI_GAI is the relative increase of *GAI*, GAI_Emerg and GAI_Norm are the *GAI* in the emergency scenario and normal scenario respectively. A routing strategy with low RI_GAI is preferable as it is less influenced by emergencies.

To summarise, these guidelines describe five requirements for an effective police patrol routing strategy to minimise the idleness globally and its change in different situations. They are efficiency (global idleness), flexibility (weighted hotspots), scalability (change of the number of hotspots or patrollers), unpredictability (randomness of idleness), and robustness (dealing with emergencies). It is difficult to define the general threshold values of each evaluation criteria to distinguish an effective strategy, as it may depend on the situation, as to be shown in the two case studies in Section 7.6.

7.4 The Bayesian Ant-based Patrolling Strategy (BAPS)

This section describes the BAPS. This is a greedy strategy, searching for the locally optimal choice at each stage. BAPS is designed according to the proposed guidelines, as shown in Figure 7.1.

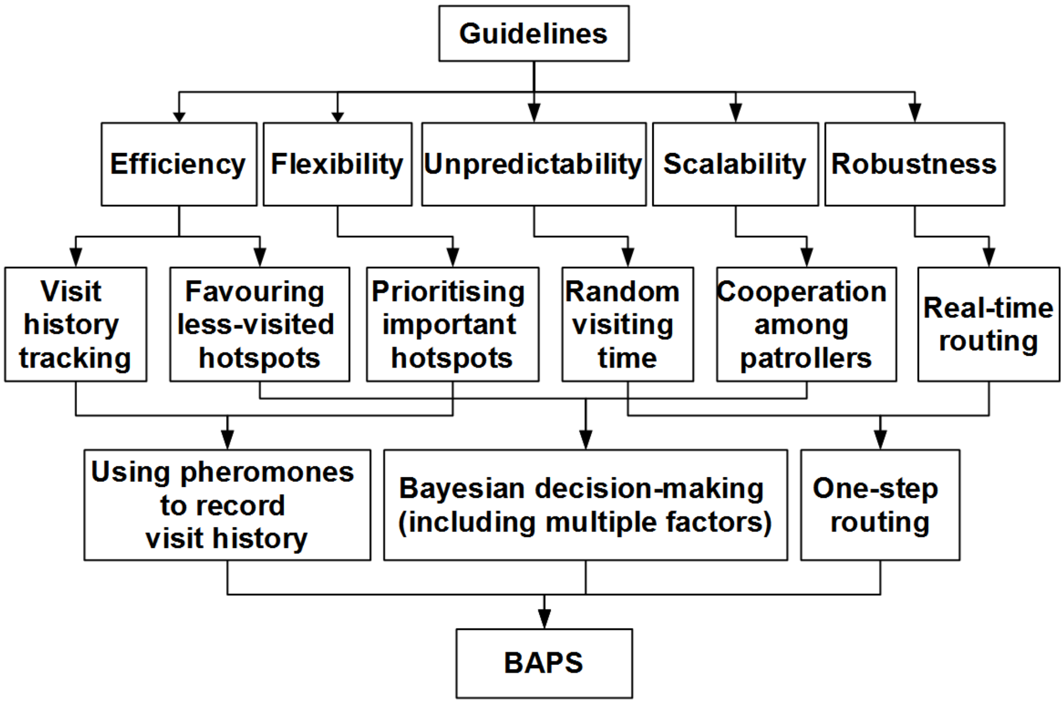


Figure 7.1 Schematic diagram of how to design BAPS according to guidelines

In particular, to achieve high *efficiency*, BAPS uses the pheromone variable to mark visiting history on hotspots and guide patrollers to visit less frequently visited places. To guarantee *flexibility*, a Bayesian framework is adopted to allow more prior knowledge and additional decision variables; the parameters of each hotspot can also be set to different values to account for varied priorities. Moreover, *scalability* is ensured when the movements and intentions of other patrollers are accounted for in choosing the next hotspot. The *unpredictability* of the patrol routes would result from the probabilistic nature of BAPS and the real-time manner to building the routes. Additionally, BAPS is *robust* to the influence of emergency calls because of the information sharing and the resultant coordination among different patrollers, which is a real-time system conducted in police practice. If one or several patrollers are busy with emergencies, the hotspots in their neighbourhood would be less patrolled and see an increase of instantaneous idleness in a short time, leading to the higher possibility of being visited by other patrollers.

In this strategy, patrol routes are built in an online and centralised manner. The routes are built according to the state of the system, and are adapt to the dynamic environment, i.e., each route is built step by step and in every decision, only the next step choice is made. Moreover, to guarantee better coordination, every next-step decision is made by the control centre based on the position and movements of all patrollers. Every time a patroller finishes patrolling a hotspot, they update their situation with the control centre and asks for the next hotspot to patrol.

Therefore, the core problem is the decision making of which hotspot to cover in the next step for every patroller. Intuitively, in order to achieve a regular visit to all hotspots, the less recently visited places are preferable to other places. To that end, a variable called *pheromone* is used to mark the visiting history of every hotspot, which is inspired by Fu and Ang (2009). The pheromone concept originated in the ant colony algorithm, which is a type of algorithm for solving various optimisation problems by simulating ants' behaviour in seeking food (Dorigo et al. 1991). The ant colony algorithm has been used in the multi-agent patrolling algorithm (Fu and Ang, 2009; Doi, 2013), which used pheromone deposition and decay to indicate the frequency and time of recent visits to a location. For this reason, the pheromone variable is used here.

In BAPS, every hotspot maintains a pheromone value. The pheromone level is affected by deposition and decay. The deposition is a process to increase the pheromone level by a certain amount after a visit by any patroller. If a hotspot is visited at time t, the deposition is as follows:

$$Ph_{h_i}(t+1) = Ph_{h_i}(t) + Ph\_inc \tag{7.10}$$

where Ph_inc is the increment of pheromone level in a deposition.

Decay occurs at each time step and at each hotspot as an exponentially decreasing process. The decay of pheromone level $Ph_{h_i}(t)$ after a duration of $(t_2 - t_1)$ is calculated as follows:

$$Ph_{h_i}(t_2) = Ph_{h_i}(t_1) \times \lambda_{h_i}^{(t_2 - t_1)} \tag{7.11}$$

where λ_{h_i} is the pheromone decay rate at hotspot h_i, and $\lambda_{h_i} \in (0,1)$, which could be defined based upon the risk level of crime (as shown in the case study in Section 4), as lower pheromone decay rate is needed for the higher risk hotspots to ensure relatively high frequency of visits compared with lower risk hotspots.

Moreover, the decision making needs to consider other decision variables in calculating the probability of the next hotspot. Besides the pheromone level, two other factors are considered important in the decision, namely the distance to the hotspot of the patroller and the targeted hotspots of the teammates. On one hand, moving to faraway hotspots is time-consuming and inefficient and should be discouraged. On the other hand, to achieve better coordination and scalability, the decision making should account for the targeted hotspot of other patrollers. Therefore, the Bayesian rule is applied to combine multiple decision variables in calculating the probability of next-step target.

Before describing the decision-making process, several terms need to be defined. The gain of a patroller A to patrol hotspot h_i is defined as:

$$G(A, h_i) = 1/\left[Ph_{h_i}(t) \times NORMdpatrol(A, h_i)\right] \tag{7.12}$$

where $Ph_{h_i}(t)$ is the pheromone level of h_i; $NORMdpatrol(A, h_i)$ is the rescaled value of $dpatrol(A, h_i)$, the distance from A's position to finish patrolling hotspot h_i. The rescaling on distances, which changes the range of distance into a fixed range using a linear transformation, is designed to avoid local optima where patrollers repeatedly visit vertices that are very close to each other.

The expected gain of patroller *A* given the probability of moving towards different hotspots is defined as $G(A)$ and calculated as follows:

$$G(A) = \sum_i P(move(h_i)) \times G(A, h_i) \tag{7.13}$$

where $P(move(h_i))$ is the probability of moving towards h_i.

If A moves towards h_i, $S(A, h_i)$ is used to represent the number of patrollers targeting at the same hotspot h_i as A. Thus, the expected number of teammates visiting the same target as A can be calculated as $S(A)$:

$$S(A) = \sum_i P(move(h_i)) \times S(A, h_i) \tag{7.14}$$

Based upon these definitions, the probability of a patroller A moving towards a hotspot h_i given the gain of patrolling h_i and the targeted hotspot of the teammates, is calculated using the following Bayesian-based formula (Portugal and Rocha, 2013a):

$$\begin{aligned} &P\big(\text{move}(h_i) \big| G(A), S(A)\big) \\ &\qquad = P(\text{move}(h_i)) \times P(G(A)|\text{move}(h_i))/P(G(A)) \\ &\qquad \times P(S(A)|\text{move}(h_i))/P(S(A)) \end{aligned} \tag{7.15}$$

$P(\text{move}(h_i))$ represents prior knowledge among hotspots. In the case study in Section 7.6, $P(move(h_i))$ is defined as uniform among all hotspots and thus this term is constant among different hotspots. According to the definition of $G(A)$ and $S(A)$, these terms are independent of the targeted hotspot. Therefore, $P(G(A))$ and $P(S(A))$ are regarded as normalisation factors to guarantee the resultant probability within the range of [0, 1], and are omitted for simplification in the calculation of the posterior probability, as the relative order of the posterior probability among hotspots is not influenced by omitting the constant terms.

It is necessary to define the probability distribution function for Gain *G*. As suggested in the State Exchange Bayesian Strategy (Portugal and Rocha, 2013a), the gain *G* is a continuous random variable with a probability density function *f(G)*. The cumulative distribution function of *G* is given by:

$$F(G = \mathrm{g}) = P(G \leq \mathrm{g}) = \int_0^{\mathrm{g}} f(G)\mathrm{d}G \tag{7.16}$$

To represent the idea that higher values of gain should be rapidly more influential in the decision-making, $F(G)$ is defined as an exponentially increasing function, as follows:

$$F(G) = \mathrm{a} \times \exp(\mathrm{b} \times G) \tag{7.17}$$

where a and b are constants. Based upon the two extreme conditions, namely $F(G = 0) = \mathrm{L}$ and $F(G = \mathrm{M}) = 1$, $F(G)$ can be reformatted as:

$$F(G) = \mathrm{L} \times \exp[\ln(1/\mathrm{L})/\mathrm{M} \times G] \tag{7.18}$$

where *L* and *M* are constants that control the distribution function, $\mathrm{L}, \mathrm{M} > 0$ and $G \leq \mathrm{M}$. L is the probability value for zero gain and M is the gain saturation or maximum gain. L is simply defined as a value close to 0 (e.g., 0.001), and M is calculated using the lower bound of the pheromone level and minimum rescaled distance from the experiments. Thus, the probability density function $f(G)$ is obtained by differentiating $F(G)$:

$$f(G) = \mathrm{L} \times \ln(1/\mathrm{L})/(\mathrm{M}) \times \exp[\ln(1/\mathrm{L})/(\mathrm{M}) \times G] \tag{7.19}$$

Thus $P(G(A)|move(h_i))$ is the probability of $G(A)$ given the movement towards h_i, and is calculated as:

$$\begin{aligned} P\big(G(A)\big|\mathrm{move}(h_i)\big) &= P\big(G(A, h_i)\big) \\ &= \mathrm{L} \times \ln(1/\mathrm{L})/\mathrm{M} \times \exp(\ln(1/\mathrm{L})/\mathrm{M} \times G(A, h_i)) \end{aligned} \tag{7.20}$$

$P(G(A))$ is the expected probability of $G(A)$ given the probability distribution of moving towards different hotspots. Based upon equation (7.13), $P(G(A))$ can be calculated as:

$$P\big(G(A)\big) = \sum_i P(\mathrm{move}(h_i)) \times P(G(A, h_i)) \tag{7.21}$$

It is seen that $P(G(A))$ is independent of the targeted hotspot, and is regarded as a normalisation factor (Jensen and Nielsen, 2007) to guarantee the resultant probability within the range of [0, 1], which can be omitted for simplification in the calculation of the posterior probability, as the relative order of the posterior probability among hotspots is not influenced.

Similarly, a distribution function has to be selected for *S* variable. Obviously the greater the number of teammates moving towards a hotspot h_i, the lower the probability for patroller A to move towards h_i. Thus the probability mass function of S, which describes the probability of moving towards a hotspot towards which other n teammates are moving, is defined as follows (Portugal and Rocha, 2013a):

$$P(S = \mathrm{n}) = 2^{R-(\mathrm{n}+1)}/(2^R - 1) \tag{7.22}$$

where R is the number of patrollers and R > 1. The above definition guarantees that the probability of all $n \in [0, R]$ sums up to 1.

Accordingly, $P(S(A)|\mathrm{move}(h_i))$ represents the probability of $S(A)$ given the movement to h_i, which can be calculated as:

$$P\big(S(A)\big|\mathrm{move}(h_i)\big) = P\big(S(A, h_i)\big) = 2^{R-(S(A,h_i)+1)}/(2^R - 1) \tag{7.23}$$

As with $P(G(A))$, $P(S(A))$ is independent of the targeted hotspot, and can be omitted in the calculation of the posterior probability.

Thus, the probability of moving to hotspot h_i is given as follows, with all constant terms omitted:

$$P\big(\mathrm{move}(h_i)\big|G(A), S(A)\big) \propto \exp[\ln(1/\mathrm{L}) \times G(A, h_i)/\mathrm{M}]\,/2^{S(A,h_i)} \tag{7.24}$$

The next hotspot to patrol is the one with the highest probability:

$$h_{next} = \underset{h_i}{\mathrm{argmax}}\, P(\mathrm{move}(h_i)|G(A), S(A)) \tag{7.25}$$

If more than one hotspot has the highest probability, the next hotspot is selected from these candidates with an equal probability.

To summarise, BAPS accounts for several factors in deciding which hotspot to visit next, including recent visit history on every hotspot, travel distance, and the movements of other patrollers.

7.5 Agent-based Modelling of Online Police Patrols

This section presents a multi-agent modelling framework of online police patrol, to test the effectiveness of the routing strategy. In this framework, the environment is a street network in the urban area, and crime hotspots are the street segments with high crime density. There are two types of agents, namely patrollers and the control centre. Foot patrols with uniform skills and speed are dispatched to patrol or respond to an emergency. They have a full knowledge of the environment and always travel to the destination via the shortest path. The control centre is responsible for recording the system state (idleness and visiting history of each hotspot, etc.), communicating with patrollers, calculating the patrol routes and sending tasks of patrolling or emergency to patrollers.

The agent-based framework is used to model two patrolling strategies, BAPS and a benchmark strategy Christofides Cyclic Patrolling Strategy (CCPS), which is a deterministic and cyclic patrolling strategy based on graph theory (Chen et al. 2015). The reason for using CCPS as the benchmark is that the real-world patrol strategy is confidential and difficult to obtain and that the family of cyclic strategies are classical algorithms for the patrolling problem and perform well in different situations (Chevaleyre 2004).

In operation, CCPS is fundamentally different from BAPS. The first stage in CCPS is to compute the shortest cyclic route, which covers every hotspot at least once. The problem to find this route is known as Rural Postman Problem and can be solved using the Christofides Algorithm (Christofides et al. 1981). After the cyclic route is computed, the patrollers are distributed evenly on the route and begin to patrol following the same direction on the route, keeping even distribution on the route. In this way, CCPS strives a regular and fair visit to each hotspot, using the shortest cycle. In contrast, patrol routes in BAPS are built in real time, which requires patrollers to communicate with control centre every time they finish a patrol task, and to wait for the command of the next patrol target.

Moreover, in the emergency scenario, patrols in both BAPS and CCPS would be interrupted if an emergency occurs. According to the command from the control centre, the officers in the neighbourhood of the emergency would stop patrolling and head for the emergency site. They would resume patrolling after dealing with the emergency. To my knowledge, this is the first time that a cyclic patrolling strategy is tested in the emergency scenario.

The detailed processes of the two strategies are presented in Figure 7.2. The control centre consists of two schedulers, namely the patrol route scheduler and the emergency scheduler. The patrol route scheduler calculates the next patrol target for BAPS patrol (Figure 7.2a), or the patrol cycle for CCPS patrol (Figure 7.2b), while the emergency scheduler sends out the emergency task to patrollers close to the emergency sites. In both strategies, patrollers are assigned to patrol or respond to emergency calls. Whenever a patroller receives the emergency task, s/he would stop patrolling and respond to the emergency. In BAPS, whenever a patroller finishes the given task (emergency or patrol), s/he sends out “Task completed” to the control centre and then awaits the next task. However, in CCPS, patrollers follow the patrol cycle in their patrolling. The normal scenario is a simplification of the emergency scenario in which no emergency occurs and the emergency scheduler is removed.

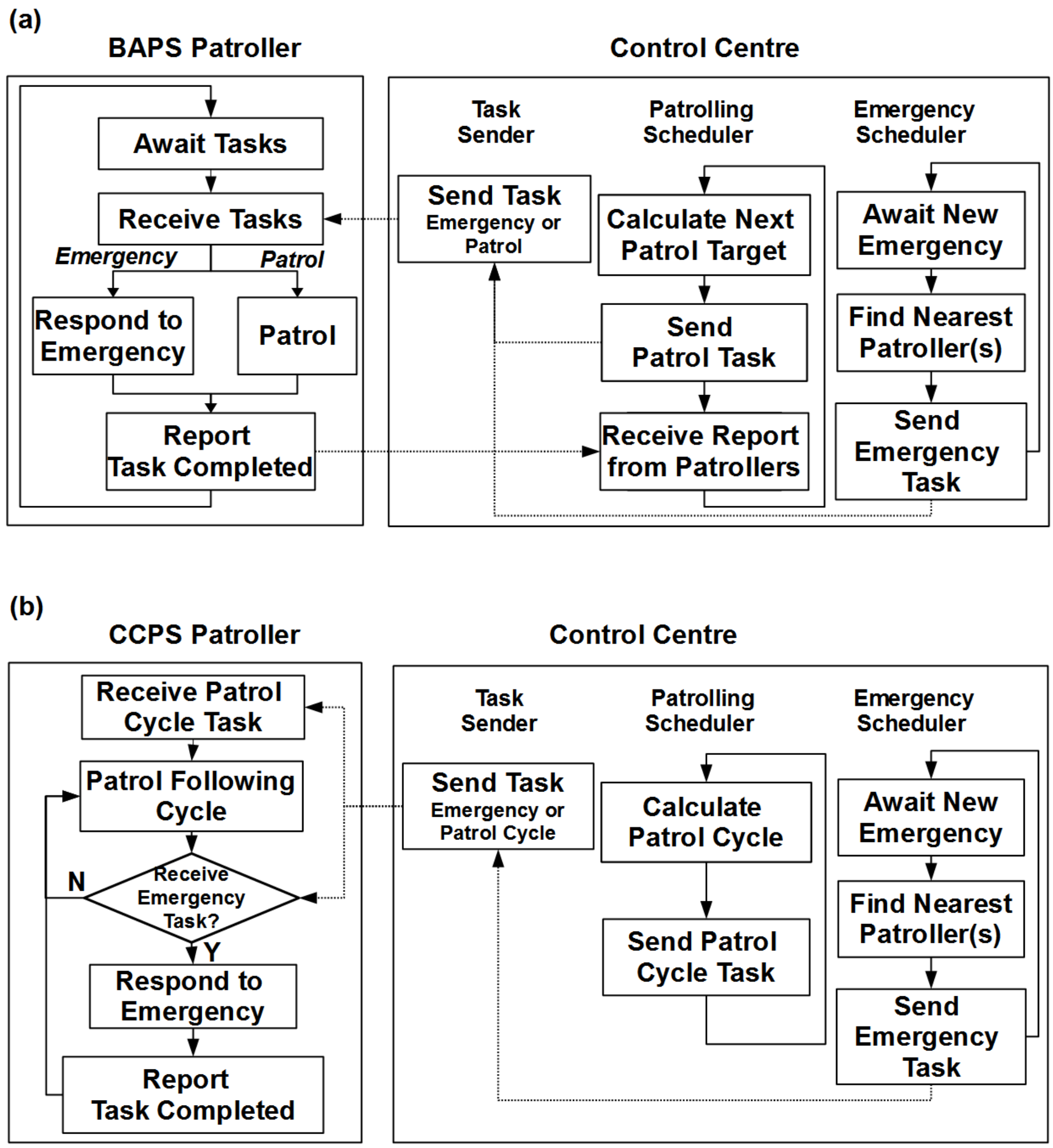


Figure 7.2 Schematic diagram of routing strategies in the emergency scenario. (a) BAPS; (b) CCPS (Chen, Cheng and Wise, 2017)

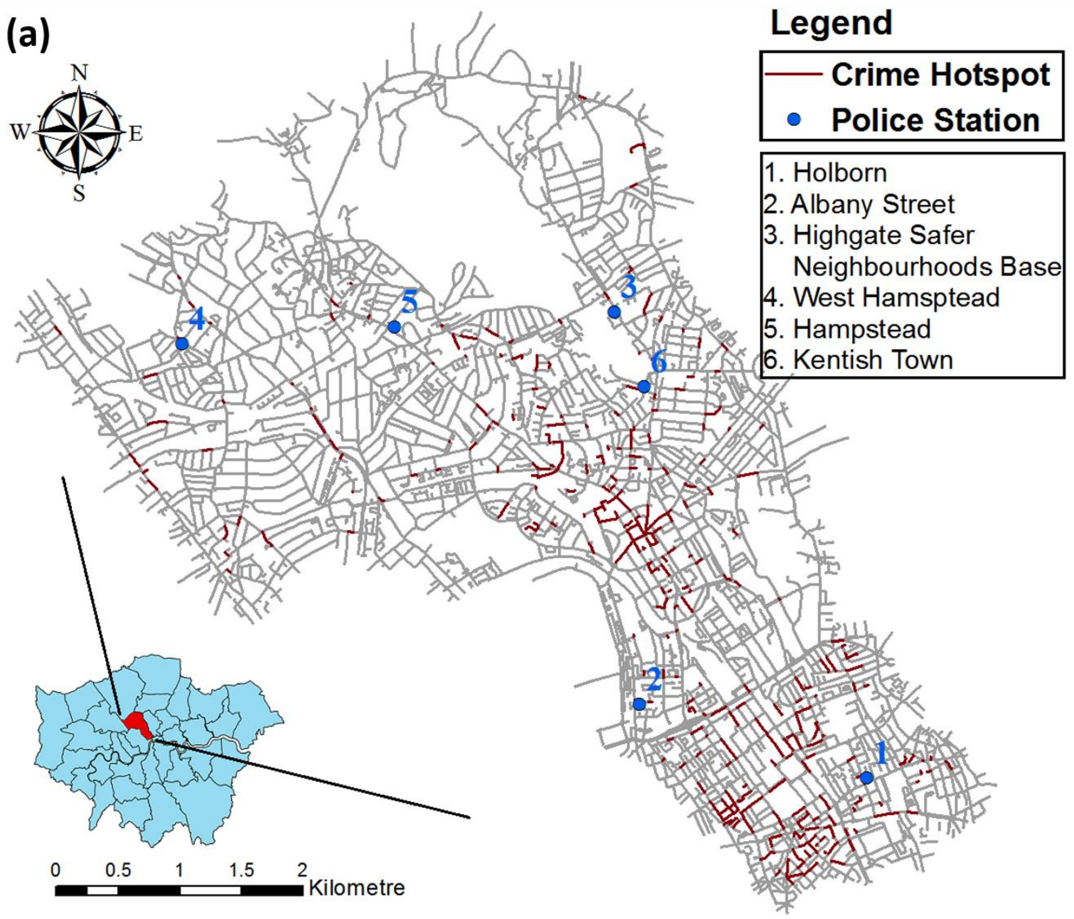
7.6 Case Studies

In order to test the applicability of the guidelines and the BAPS, two case studies were conducted, in the London Borough of Camden, London, United Kingdom and in South Side, Chicago, Illinois, the United States of America. For convenience, they are called Camden and South Chicago.

The data used are provided by various agencies. In the Camden case, some of the original crime incidents in Camden were aggregated to the centre of grids when they were recorded. The emergency calls data are added here to test the robustness. In the South Chicago case, the street network data is downloaded from OpenStreetMap. The locations of police stations in

South Chicago and the crime dataset from 2001 to present are downloaded from the City of Chicago data portal (https://data.cityofchicago.org). The time duration of crimes used is from March 1st, 2011 to March 1st, 2012, and the crime types used include theft, burglary, homicide, battery, arson, motor vehicle theft, assault, and robbery. The computation of crime counts for each segment is the same as in the Camden case. Due to the lack of emergency call data or police despatch data of South Chicago, the robustness is tested using hypothetical emergency calls, whose location and time is generated using a uniform distribution in the given area and time period.

Crime hotspots are identified as the street segments with the highest crime density and covering 5% of the total road length. The number of crime hotspots is 311 in Camden and 289 in South Chicago. Figure 7.3 shows the crime hotspot maps.



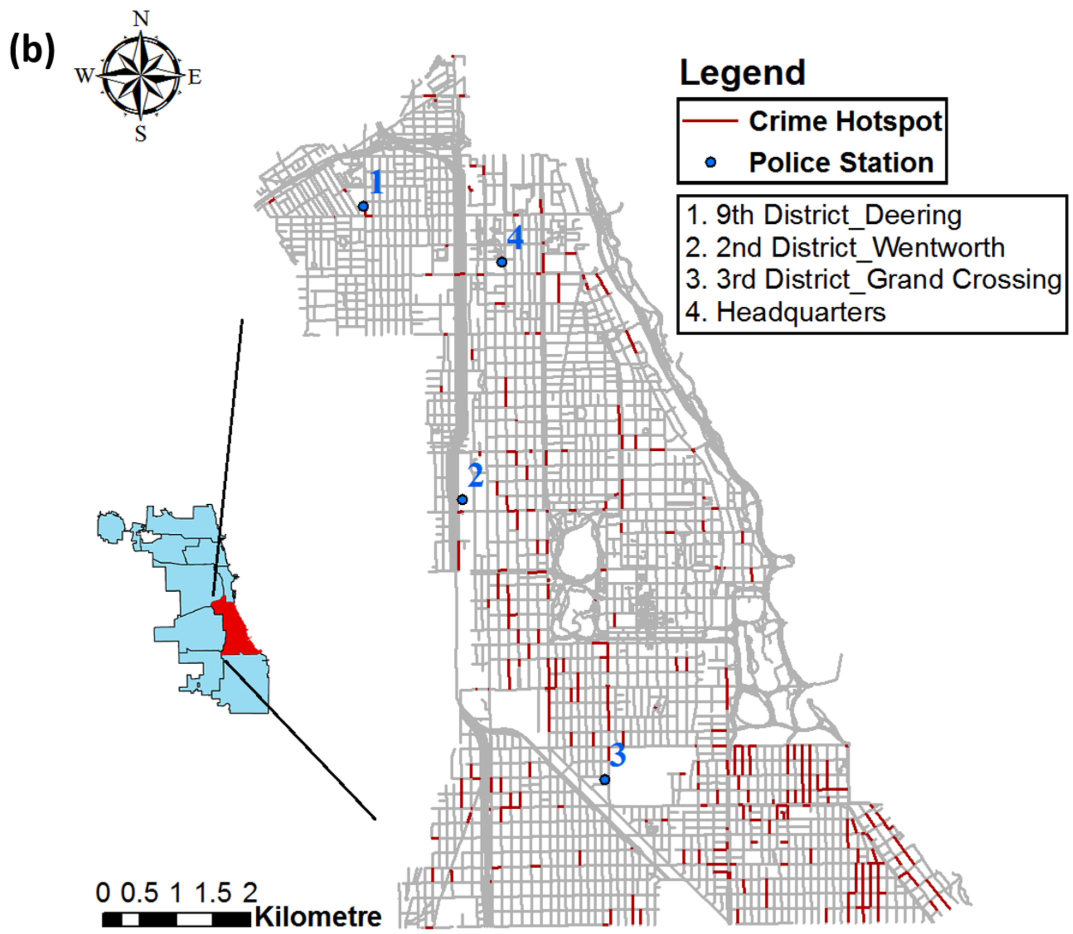


Figure 7.3 Crime hotspot map of top 5% crime. (a) in Camden; (b) in South Chicago. (Chen, Cheng and Wise, 2017)

The model framework used in this work is built in Java, using the MASON simulation toolkit (Luke et al. 2005). The simulation trial proceeds at a physical scale of 1m^2 resolution and is updated on a temporal scale of five seconds per step. The analysis of simulation result is done using R language and environment (R Core Team 2015).

Two patrolling strategies (BAPS and CCPS) were tested in the above environment with different sizes of patroller groups (2~8 officers per police stations). Each simulation ran for 11 patrol cycles, after which each hotspot had been visited at least 11 times. The GAI or WGAI is considered to have converged when its value after any patrol cycle converges with no more than 1% difference to that of the previous cycle. The number of patrol cycles was selected experimentally to guarantee the convergence of GAI or WGAI. The parameters of BAPS were selected to minimise the GAI in the trial experiments. For example, in the selection of λ, which represents the decaying rate of pheromone level, different λ values (0.9999, 0.99991, 0.99992, 0.99993, etc.) were tested in a typical simulation (Camden case, 30 patrollers, 311 hotspots) with other parameters fixed, and 0.99993 was selected as the final value as it resulted in the lowest GAI. The parameter settings in Camden were directly applied to the South Chicago case without further experimenting, leading to a good performance in South Chicago as well.

The computational efficiency of BAPS is tested. The simulations ran on a Dell machine, with a 3.60 GHz Intel Core i7-4790 processor, 32.0 GB RAM and 64-bit Windows 7 operating system. In the typical experiment of 48 officers covering 311 hotspots in Camden, the simulation lasted 82 seconds, with 7000 patrol target computations and thus a cost of 0.01 seconds for each computation. Note that the assumptions are that the hotspot distribution is known, the officers always use the shortest path, and the travelling time between any two locations is fixed. The computational efficiency in the large-scale problem or dynamic situations is subject to further experiments.

The evaluation and comparison of BAPS and CCPS is presented in order according to the proposed guidelines.

7.6.1 Efficiency

Efficiency is measured by GAI, which is the converged value of GAI(t). Table 7.1 and Table 7.2 present the result of GAI and the relative change (Bennett & Briggs 2005) of GAI:

$$\text{Relative Change} = (GAI\_BAPS - GAI\_CCPS)/(GAI\_CCPS)\ \times 100\% \qquad (7.26)$$

In Camden, for different patrol team sizes, BAPS has lower global average idleness and consequently better performance than CCPS. The relative change varies slightly with the team size, reaching the maximum at the size of 18 and 24. In South Chicago, the GAI values in BAPS are lower than CCPS by around 10%, except for the team size of 8. A possible reason is that with a small patrol team, officers have to travel long distances to cover the whole area, leading to the degeneracy of BAPS efficiency.

Table 7.1 Efficiency performance in Camden (values in seconds)

Team Size	12	18	24	30	36	42	48
GAI_CCPS	5079	3477	2605	2039	1709	1443	1254
GAI_BAPS	4407	2833	2128	1700	1401	1223	1075
Relative Change (%)	-13.2	-18.5	-18.3	-16.6	-18.0	-15.2	-14.3

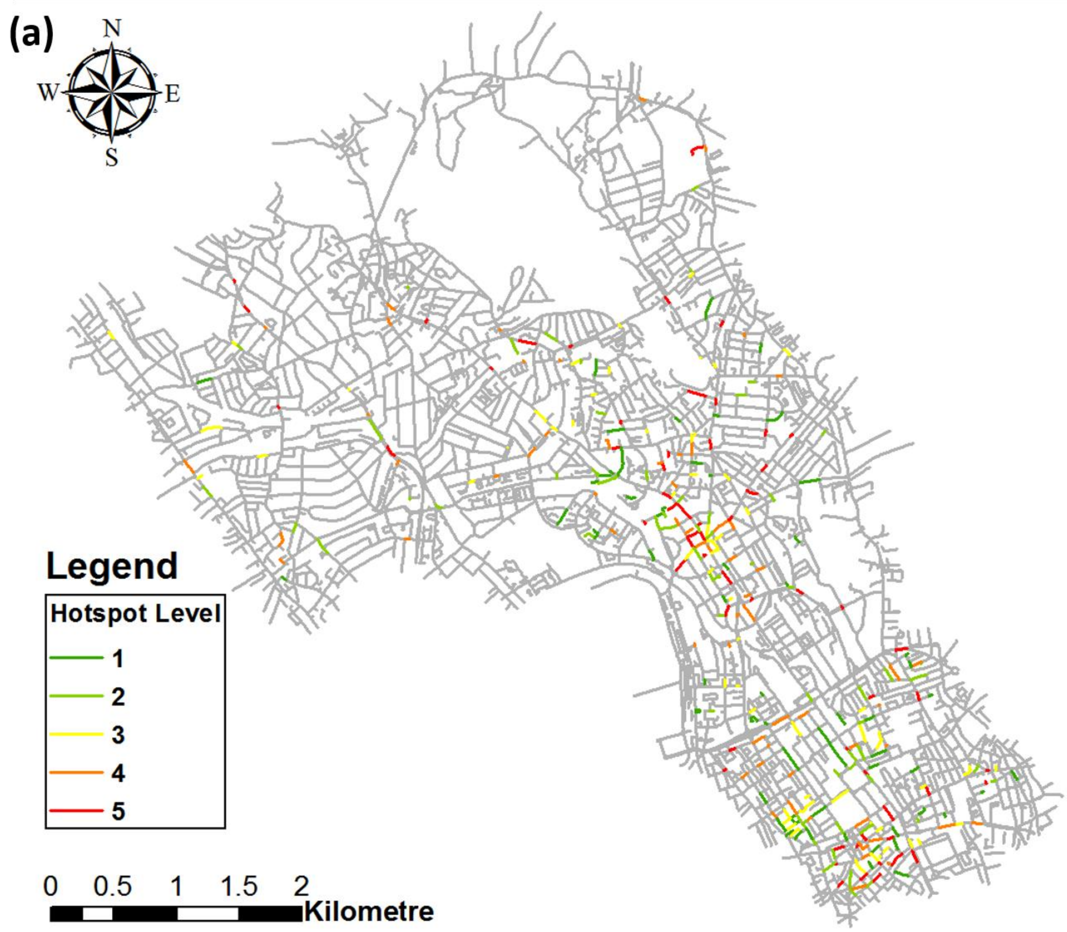
Table 7.2 Efficiency performance in South Chicago (values in seconds)

Team Size	8	12	16	20	24	28	32

GAI_CCPS	11133	7425	5613	4418	3711	3152	2762
GAI_BAPS	11725	6899	4981	3917	3253	2824	2478
RelativeChange (%)	5.32	-7.08	-11.26	-11.34	-12.34	-10.41	-10.28

7.6.2 Flexibility

The flexibility of routing strategies is measured by Weighted BAPS (WBAPS) which uses smaller decay rates for hotspots with higher crime density (risk). Here the hotspots are evenly divided into five levels 1-5, with crime density and weight decreasing from Level 5 to Level 1. Figure 7.4 shows the hotspot map with different risk levels in both cases.



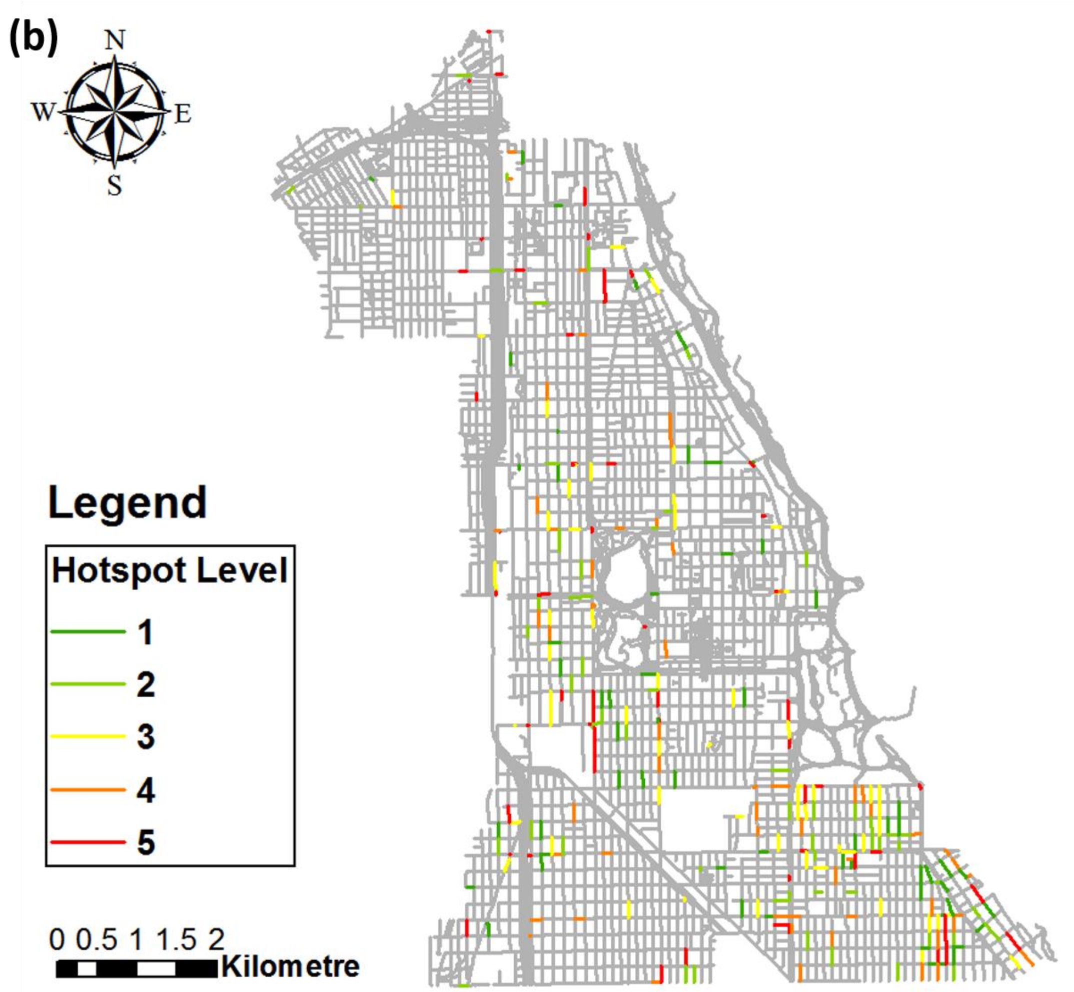


Figure 7.4 Crime hotspot maps with different risk levels. (a) in Camden; (b) in South Chicago (Chen, Cheng and Wise, 2017)

The decay rates are selected experimentally. The decay rates for these 5 levels (from 5 to 1) are: 0.99989, 0.99990, 0.99991, 0.99992, and 0.99993. Table 7.3 compares the performance of three strategies (BAPS, WBAPS, and CCPS) in patrolling hotspots of multiple levels using 30 patrollers in Camden. BAPS or WBAPS have superior performance than CCPS in terms of WGAI, GAI and the GAI of each risk level. In comparison with BAPS, WBAPS decreases the weighted global average idleness by about 1.8%, in the cost of a slight rise (2.4%) in GAI. Moreover, the GAI at Level 4 and Level 5 hotspots reduce moderately by 4.7% and 9.4% when WBAPS is used. WBAPS provides an easy and effective approach to targeting hotspots of higher importance, which shows the advantage of BAPS in that it can be tuned for specific aims. A similar trend is observed in Table 7.4 in the South Chicago case with 20 patrollers. WBAPS and BAPS have lower GAI and WGAI, compared with CCPS. WBAPS has slightly lower GAI and higher WGAI than BAPS, as well as lower GAI in the prioritised hotspots of

Level 4 and 5. The result verifies the flexibility of BAPS to patrol hotspots of varying levels by using varied decay rates.

Table 7.3 Flexibility performance in Camden (all values in seconds)

Strategy	WGAI	GAI	GAI of Each Level				
			1	2	3	4	5
CCPS	2030	2040	2075	2048	2088	2039	2004
BAPS	1653	1671	1718	1725	1697	1700	1607
WBAPS	1623	1712	2054	1872	1720	1620	1456

Table 7.4 Flexibility performance in South Chicago (all values in seconds)

Strategy	WGAI	GAI	GAI of Each Level				
			1	2	3	4	5
CCPS	4442	4361	4358	4380	4460	4346	4550
BAPS	3890	4032	4046	4035	4000	3869	3753
WBAPS	3888	4214	4762	4466	3999	3715	3554

7.6.3 Scalability

To test the team scalability of BAPS and CCPS patrols, the ST metric (See Equation (7.6)) is calculated for different sizes. Table 7.5 reveals that in Camden, for all tested team size, BAPS systems present superlinear performance as the speedup is greater than 1.0 while the performance of CCPS systems are sublinear when team size is between 18 and 36. On every tested team size, the speedup performance of BAPS outperforms that of CCPS, indicating better scalability of BAPS. Moreover, the scalability of CCPS is achieved by setting all patrollers evenly distributed in time and space (Smith & Rus 2010; Pasqualetti et al. 2012), which means the starting positions have to be calculated for each size, while in BAPS starting positions have little influence on its performance. Similarly, in South Chicago, BAPS has superlinear performance and outperforms CCPS on every tested team size (See Table 7.6).

Table 7.5 Performance of team scalability in Camden (GAI values in seconds)

Team Size	12	18	24	30	36	42	48
GAI_CCPS	5079	3477	2605	2039	1709	1443	1254

GAI_BAPS	4407	2833	2128	1700	1401	1223	1075
TS_CCPS	1.000	0.974	0.975	0.996	0.991	1.006	1.013
TS_BAPS	1.000	1.037	1.035	1.037	1.049	1.030	1.025

Table 7.6 Performance of team scalability in South Chicago (GAI values in seconds)

Team Size	8	12	16	20	24	28	32
CCPS	11133	7425	5613	4418	3711	3152	2762
BAPS	11725	6899	4981	3917	3253	2824	2478
ST_CCPS	1.000	1.000	0.992	1.008	1.000	1.009	1.008
ST_BAPS	1.000	1.133	1.177	1.197	1.201	1.186	1.183

The spatial scalability (SS) of patrol strategies was tested by changing the crime hotspot density level from 5% to 10% and 15%. The higher density level requires better cooperation between agents in order to cover all hotspots. The corresponding hotspot maps are shown in Figure 7.5, and the testing results are presented in Table 7.7. For example, Camden_10p_30 represents the experiment of covering the Camden hotspot map of 10% total length with 30 patrollers. The SS values in Table 7.7 use 5% as the baseline density level, with other factors fixed, including the total patrol area, strategy, and team size. Overall, BAPS outperforms CCPS in all hotspot density levels, except for the 15% South Chicago hotspot map with 20 patrollers. Furthermore, the SS value of BAPS is consistently larger than CCPS, indicating that BAPS is more affected by the hotspot density level. The summary on spatial scalability is that BAPS has better performance on different hotspot density levels, but it is more sensitive to the high hotspot density levels.

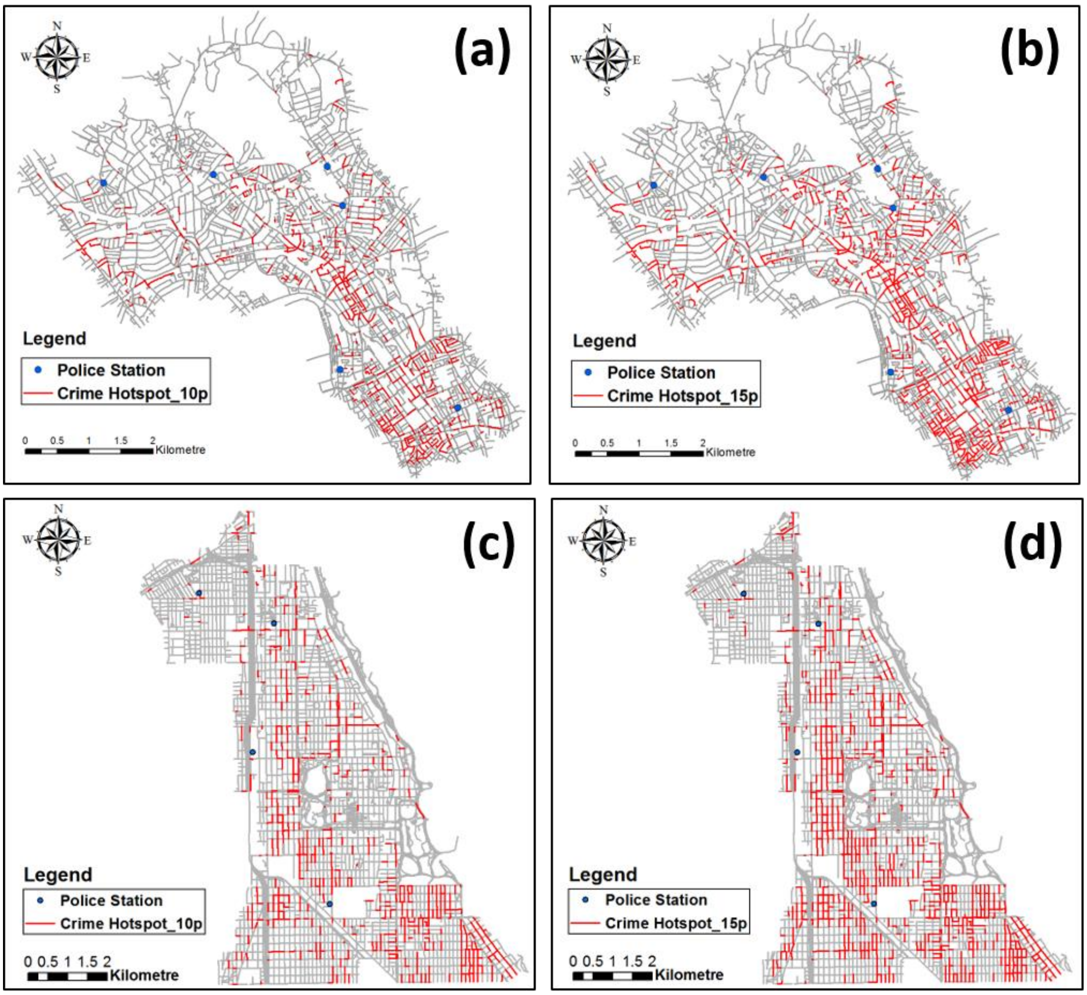


Figure 7.5 Crime hotspot maps with denser hotspots. (a) in Camden with 10% of total road length; (b) in Camden with 15% of total road length; (c) in South Chicago with 10% of total road length; (d) in South Chicago with 15% of total road length (Chen, Cheng and Wise, 2017)

Table 7.7 Performance of spatial scalability in denser hotspot maps (values in seconds)

HotspotMap_TeamSize	Camden _10p_30	Camden _15p_30	SouthChicago _10p_20	SouthChicago _15p_20
GAI_CCPS	2806	3415	6492	8035
GAI_BAPS	2489	3132	6358	8392
RelativeChange(%)	-11.30	-8.29	-2.06	4.44
SS_CCPS(%)	37.62	67.48	46.94	81.87
SS_BAPS(%)	46.41	84.24	62.32	114.25

7.6.4 Unpredictability

ASdIdl is measured to evaluate the unpredictability of patrolling. In Camden, for different team sizes, the ASdIdl values of BAPS are higher than CCPS (See Table 7.8). Low standard deviation in CCPS can be explained by the even distribution of patrollers on the cycle and the same patrolling cycle used by all patrollers. In contrast, the high deviation of idleness in BAPS and the high randomness of patrol routes would create a perceived "omnipresence" of the police that deters crime in crime hotspots (Sherman & Eck 2002). Likewise, in South Chicago, for every team size, BAPS has higher ASdIdl values, compared with CCPS (See Table 7.9).

Table 7.8 Measure of unpredictability in Camden (all values in seconds)

Team Size	12	18	24	30	36	42	48
ASdIdl_CCPS	939	1233	1599	497	753	583	493
ASdIdl_BAPS	4349	3368	2590	2055	1649	1422	1197

Table 7.9 Measure of unpredictability in South Chicago (all values in seconds)

Team Size	8	12	16	20	24	28	32
ASdIdl_CCPS	2218	1309	1843	1441	955	1132	1231
ASdIdl_BAPS	7836	5696	4735	4148	3696	3410	3010

7.6.5 Robustness

The experiments of robustness were conducted, using the real-world emergency records in Camden (March of 2011) and hypothetical emergency calls in South Chicago. When an emergency is reported, the nearest n patrollers have to stop patrolling and respond. Due to insufficient details about the emergency (time length, number of patrollers dispatched, etc.), different settings were attempted, including the total number of patrollers and the number of patrollers per emergency.

Table 7.10 and Table 7.11 present the robustness performance in Camden and South Chicago. For example, BAPS_18 represents the BAPS simulation with 18 patrollers. The percentages represent relative changes in comparison with non-emergency scenario (0 patrollers per emergency). In Camden (Table 7.10), the performance of both BAPS and CCPS degraded slightly or moderately as the number of patrollers that are required per emergency increased.

Evidently, the higher number of patrollers needed by an emergency, the greater the effect on the patrolling performance. Further, holding constant the patrolling strategy and the number of patrollers per emergency, the emergency response has more a prominent impact on the performance when the patrolling group is smaller. Comparatively, with the group size of 18, BAPS patrol was more affected by the emergency response than CCPS patrol. However, when the group size increased to 48, the influence on BAPS patrol was less prominent than that on CCPS patrol. Similar comparison exists in South Chicago (Table 7.11), where BAPS patrol was more influenced than CCPS patrol when the group size was 12 or 20, and was more robust than CCPS when the group size increased to 32. The result supports the robustness of BAPS against emergency response.

Table 7.10 Robustness performance of emergency scenario in Camden

Patrollers per Emergency	0	1	2	3	4
BAPS_18	0.0%	1.8%	4.5%	4.7%	10.2%
BAPS_30	0.0%	3.0%	3.3%	3.9%	5.4%
BAPS_48	0.0%	-0.2%	0.3%	1.0%	2.4%
CCPS_18	0.0%	1.1%	0.9%	1.6%	4.1%
CCPS_30	0.0%	3.2%	4.1%	3.4%	5.5%
CCPS_48	0.0%	4.7%	4.8%	5.7%	6.1%

Table 7.11 Robustness performance of emergency scenario in South Chicago

Patrollers per Emergency	0	1	2	3	4
BAPS_12	0.0%	1.7%	13.7%	21.2%	30.9%
BAPS_20	0.0%	4.8%	7.4%	11.4%	18.8%
BAPS_32	0.0%	-1.3%	0.3%	3.1%	5.2%
CCPS_12	0.0%	2.9%	7.3%	12.1%	31.2%
CCPS_20	0.0%	3.5%	2.5%	2.7%	0.7%
CCPS_32	0.0%	3.8%	4.1%	4.5%	5.5%

7.7 Chapter Summary

In this study, a set of guidelines was proposed for repeated police patrolling, along with an online cooperative police patrol routing strategy. Five quantitative measures were developed for the guidelines, namely efficiency, flexibility, unpredictability, scalability, and robustness. Under these guidelines, an online routing strategy BAPS was developed. This strategy accounts for multiple factors which impact patrol and adopts a probabilistic computational framework, resulting in effective patrolling. As demonstrated by two real-world case studies, BAPS generally outperforms a cyclic patrolling strategy CCPS in terms of multiple measures. Thus, BAPS has great potential for online cooperative police patrol and other related applications.

Chapter 8

COUPLING DIFFERENT STRATEGIES IN POLICE PATROL

8 COUPLING DIFFERENT STRATEGIES IN POLICE PATROL

8.1 Chapter Overview

In this chapter, the coupling of different strategies for police patrolling is discussed. First, in Section 8.2, the motivation for, and the possibilities of coupling different police patrol strategies are provided, with the focus on a combination of the districting and repeated coverage routing strategies. Then, in Section 8.3, a case study is presented that combines these strategies. The combined strategy is compared to the original BAPS, and the advantages and limitations of the combined strategy are revealed. Finally, in Section 8.4, the chapter is summarised.

8.2 Coupling Different Strategies

The districting and routing strategies for police patrolling were extensively discussed in Chapters 5, 6, and 7. In their application, the proposed strategies can be adopted in two different ways. The first option is to select the most suitable strategy according to the scenario requirement, and to apply it after essential adjustments. The other option is to couple different strategies, and to form an integrated patrol planning system. As each of the strategies are built on a scenario that represents a simplification of the realistic situation, the coupling of different strategies has the advantage of incorporating the greater complexities of a real situation into a single strategy. Moreover, the coupling of different strategies can lead to an improvement in performance, including lower operational costs or better coverage of the street network.

For these reasons, this chapter focuses on the coupling of different police patrol strategies. There exist several possibilities for strategy coupling. First of all, the districting strategy can be coupled with the routing strategy, in the way the route planning is conducted within each district. Second, a route planning system can be built to incorporate the two routing strategies that can switch between these two strategies in response to the real-time situation. For example, at the beginning of the patrol shift, the patrollers follow the balanced routes to monitor the crime hotspots. As calls for service occur in the territory, and the number of patrol officers changes, the system can switch to an online patrolling mode.

This chapter specifically focuses this study on the coupling of the districting strategy (Chapter 5) and the online repeated patrol routing strategy (Chapter 7). This coupling is advantageous for a number of reasons. First, as demonstrated in Chapter 5, the districting strategy is capable of deriving balanced and efficient patrol districts. However, it is unclear how the patrol districts are utilised in the policing operations. Second, in Chapter 7, the BAPS was described for the repeated coverage of crime hotspots. However, under this strategy, there is no boundary to the patrol area for each officer, and every officer can be asked to cover a large proportion of the territory. This is unfavourable, as an officer may spend too much of their time travelling between different areas. Assigning every officer to a patrol district may reduce the travel time and improve the performance of patrolling.

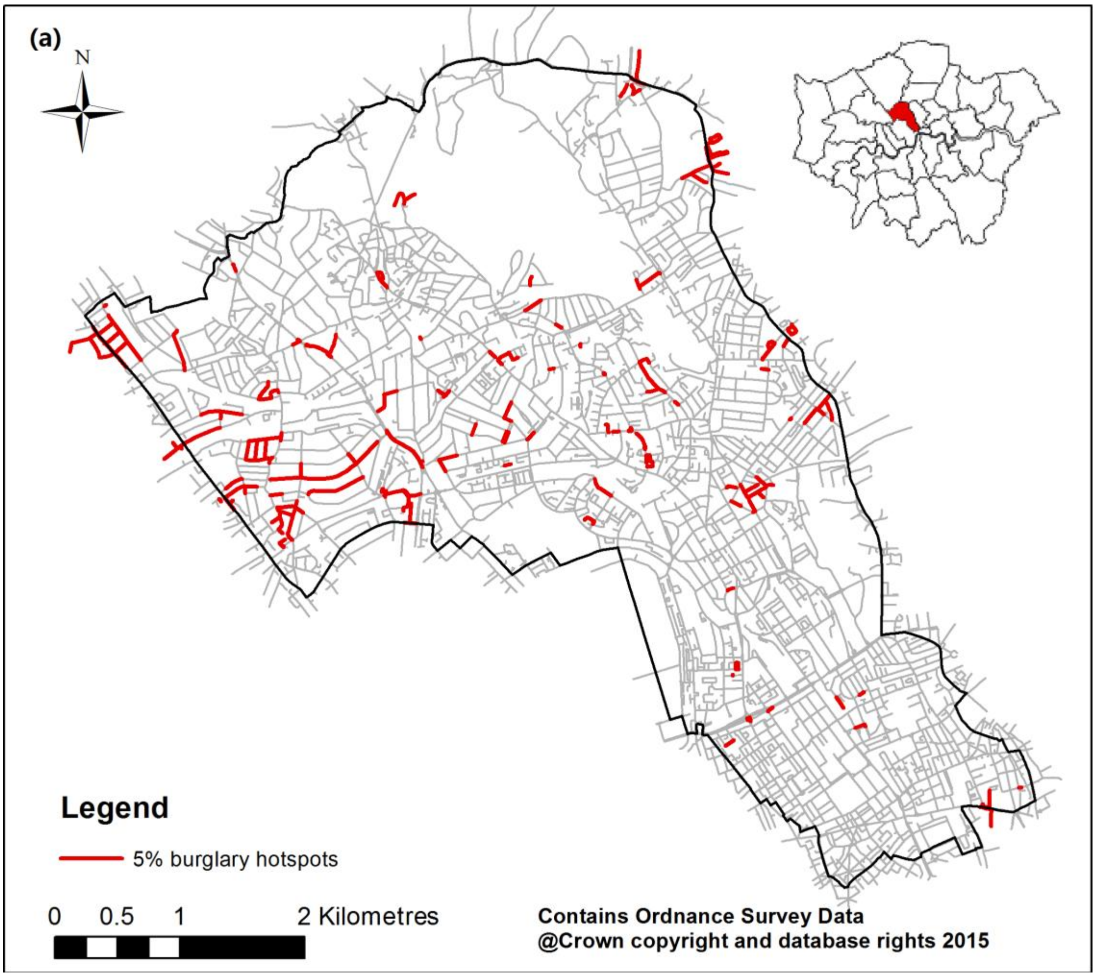
The coupling of the districting and online repeated patrol routing strategies is described as follows. The first step is to produce the patrol districts, based on the distribution of crime risk and crime hotspots. The second step is to assign the officers to the patrol districts, according to the workload of the patrol districts. In the third step, the officers are dispatched to monitor the crime hotspots within the patrol district, following the routing from BAPS.

8.3 Case Study

This section presents a case study that was conducted to demonstrate how the districting strategy and the online repeated patrol routing strategy (BAPS) were coupled. For clarity, the coupled strategy is called Districting-BAPS. The mechanism of the Districting-BAPS is very similar to BAPS (see Figure 7.2), and the only difference is that each officer stays in the corresponding district and patrols only the hotspots within that district.

This case study was based on the experiments reported in Chapters 5 and 7. The study area was the London Borough of Camden, and the street-level burglary risk map was computed using the NTKDE model (Rosser *et al.*, 2017) and the historical burglary crime records (see Figure 5.3). The districting solution, with six districts, was derived in Section 5.6 (see Figure 5.5), and is used in this case.

Here, the burglary hotspots were identified based on the burglary risk and coverage level. Given a group of crime hotspots, the level of coverage is defined as the proportion of the accumulative street length of the hotspots to the total length of the network (Rosser *et al.*, 2017). The first step in identifying the hotspots is to place all street segments in ranked order according to their predicted risk, starting from the highest. Second, given the level of coverage, the street segments are selected in order – from the highest ranked units up to the point where the cumulative coverage is equal to the selected level. In this case, the coverage levels of 0.05 and 0.10 were selected, and the hotspot maps are illustrated in Figure 8.1 (a)(b). Considering the district map (Figure 5.5), the number of crime hotspots located in each district is summarised in Figure 8.1 (c)(d). It is not surprising that the number of crime hotspots is not balanced in different districts, as the districts are designed to achieve workload balance regarding crime risk, street length and diameter, which does not include the number of crime hotspots.


(a)
N
Legend
5% burglary hotspots
0 0.5 1 2 Kilometres
Contains Ordnance Survey Data
@Crown copyright and database rights 2015

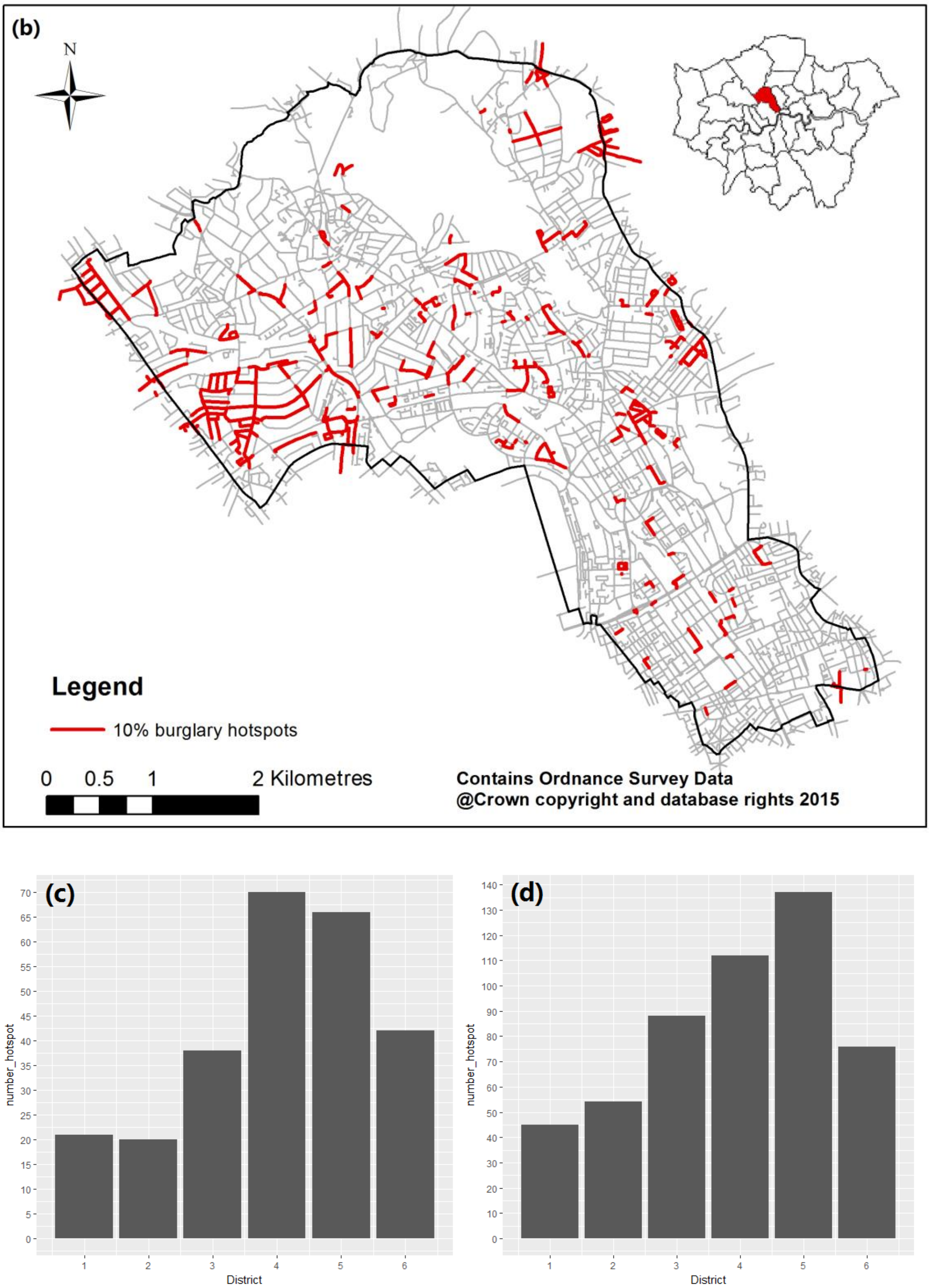


Figure 8.1 Two burglary hotspot maps and the summary of hotspots in each district. (a) 5% coverage burglary hotspot map with a total of 257 hotspots. (b) 10% coverage burglary hotspot map with a total of 512 hotspots. (c) Number of 5% coverage burglary hotspots in each district, (d) Number of 10% coverage burglary hotspots in each district. The district map is shown in Figure 5.5

The case study was carried out following the steps described in Section 8.2. First, the districting solution consisting of six patrol districts, is derived (see Figure 5.5). Second, given the number of patrol officers (12 or 18), the number of officers for each district is set approximately proportional to the hotspot number in each district. Intuitively, a larger patrol team would be dispatched to the district with relatively more crime hotspots. Third, the patrolling experiment was conducted in an agent-based simulation environment. The model framework used in this work is built in Java, using the MASON simulation toolkit (Luke et al. 2005). The simulation trial proceeds at a physical scale of $1m^2$ resolution and is updated on a temporal scale of five seconds per step.

In the simulation, each patrol officer starts from the police station they belong to, and if the Districting-BAPS strategy is used, each district is taken care of by the patrol officers from one station. Figure 8.2 shows the environment with the patrol districts. The large and small circles represent the police station and the patrol officers, respectively. The colour of a street indicates the district it belongs to, and the same colour is used for the corresponding station and officers. In comparison, Figure 8.3 shows the usage of street links in the same simulation experiment. The colour of the streets indicates the number of times of usage, and brighter colours represent more frequent usage.

The simulation environment provides different ways to examine patrol performance. First, by examining the location and movement of patrol officers and patrol districts (Figure 8.2), the user can check where the officers are patrolling the corresponding districts. Second, by combing the frequency of using street links (Figure 8.3) and hotspot maps (Figure 8.1), the user can check whether the officers have patrolled the hotspots. Third, the simulation environment saves the results as text files, which can be analysed by the user to evaluate performance.

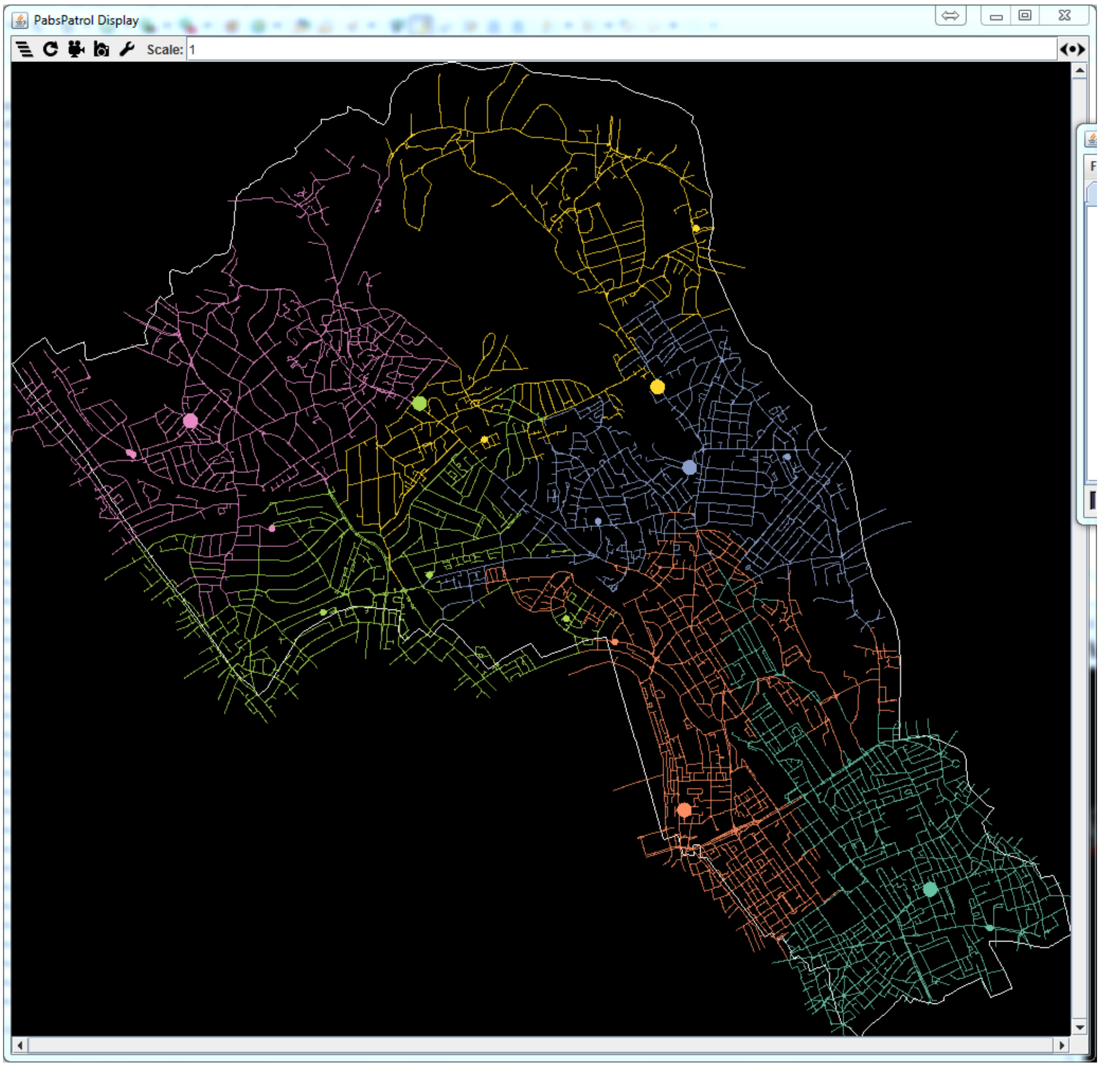


Figure 8.2 The simulation environment for police patrolling, including patrol districts, police stations and patrol officers

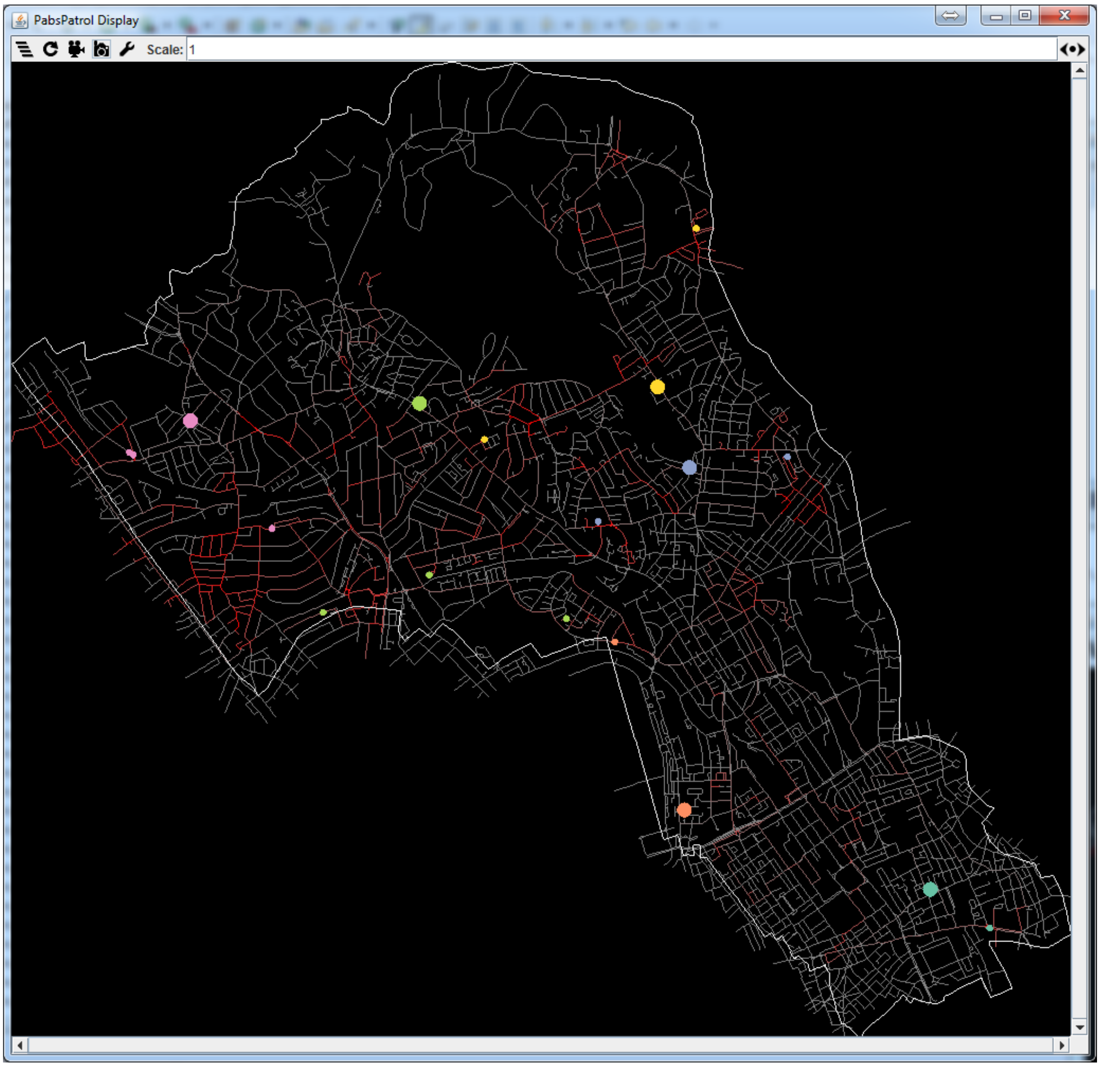


Figure 8.3 The simulation environment recording street usage by police. The environment includes the patrol officers, police stations, and the street network. The street colour indicates the number of times it is used, with brighter colours representing more frequent use

Several experiments have been carried out on two different routing strategies (Districting-BAPS and BAPS), two different crime hotspot maps, and two numbers of patrol officers (12 and 18). The results are evaluated using the Global Average Idleness (GAI, see equation (7.3)). Briefly, GAI measures the average duration between two consecutive visits to a crime hotspot. A smaller GAI value indicates a more efficient coverage of the hotspots. The relative gap is defined as follows:

$$gap(\%) = \frac{GAI(\text{Districting\_BAPS}) - GAI(\text{BAPS})}{GAI(\text{BAPS})} \times 100\% \tag{8.1}$$

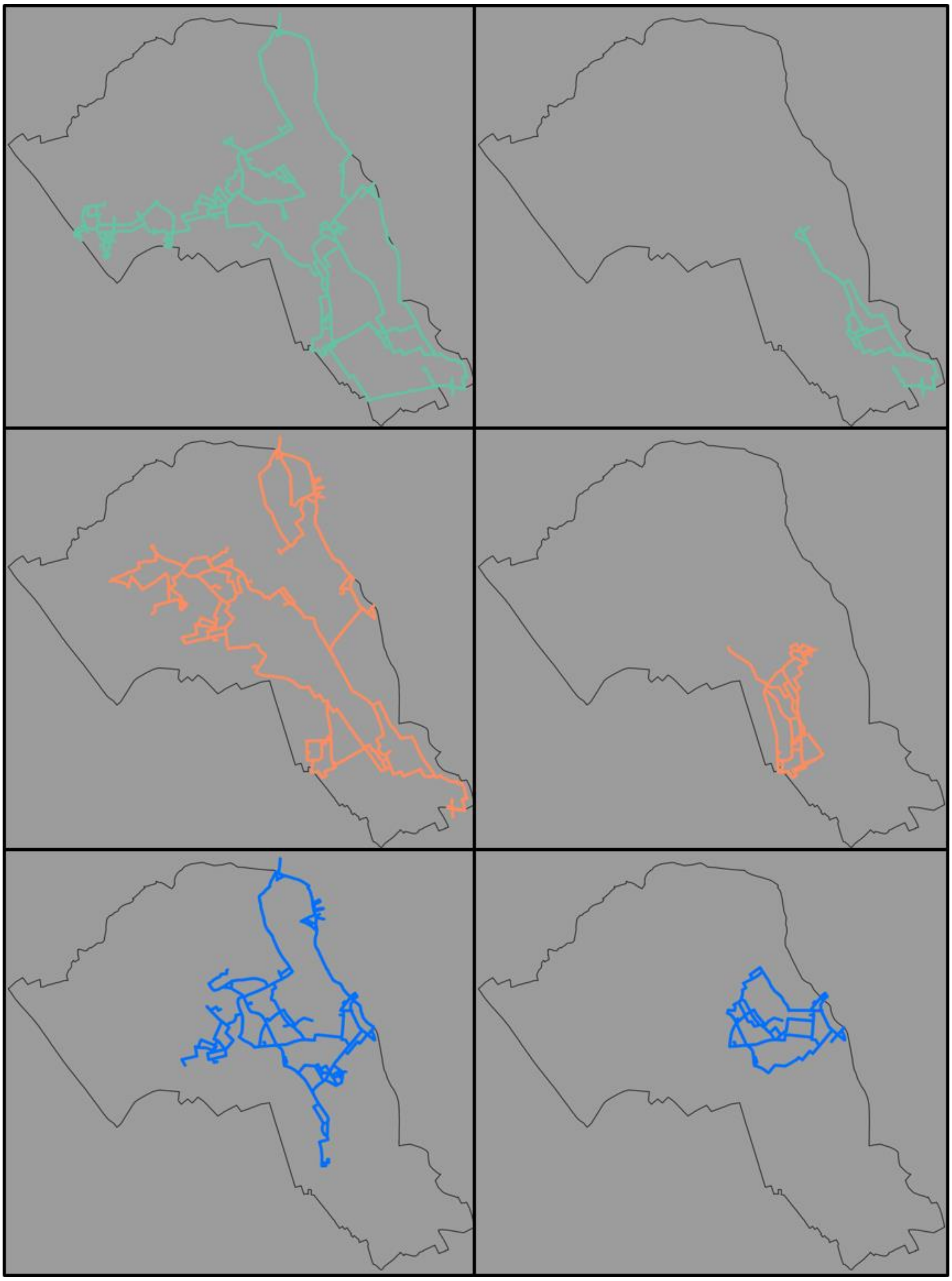
A negative relative gap indicates that Districting-BAPS is superior to BAPS. Otherwise it is less efficient than BAPS.

The experimental results are summarised in Table 8.1, in which it is shown that the GAI of the Districting-BAPS is slightly higher than that of BAPS, and the maximum relative difference is 3.32%. This implies that the Districting-BAPS can guide the routing of the repeated patrol coverage, although it is slightly less efficient than the BAPS.

Table 8.1 Experimental results of the Districting-BAPS and BAPS (the GAI is in seconds)

Hotspot map	5%	5%	10%	10%
No. of officers	12	18	12	18
GAI (BAPS)	2504	1627	4443	2770
GAI (Districting-BAPS)	2521	1662	4476	2862
relative gap (%)	0.68	2.15	0.74	3.32

A major difference between the two strategies is that the patrol officers have different activity spaces. Here, activity space refers to the physical area in which an officer moves and patrols. In the BAPS, the patrol officers may move across a large area, whereas in the Districting-BAPS, each officer only moves within the given district. Figure 8.4 shows a comparison between the activity spaces of officers using the two strategies. In this figure, each row corresponds to a district in Figure 8.2 and a station. The left and right columns show the range of activity of the officers using the BAPS and Districting-BAPS, respectively. In the BAPS, the patrol officers move across a large part of the territory, which is the natural result of globally searching the next patrol target in the strategy. In contrast, in the Districting-BAPS, the patrol officers move within the district. This makes it easier for the officers to return to their station to perform other tasks that must be done inside the station. Moreover, this also reduces the travel time across different districts, which can increase the efficiency of patrolling. Therefore, it is an advantage of the Districting-BAPS that patrol officers are restricted to their own district.



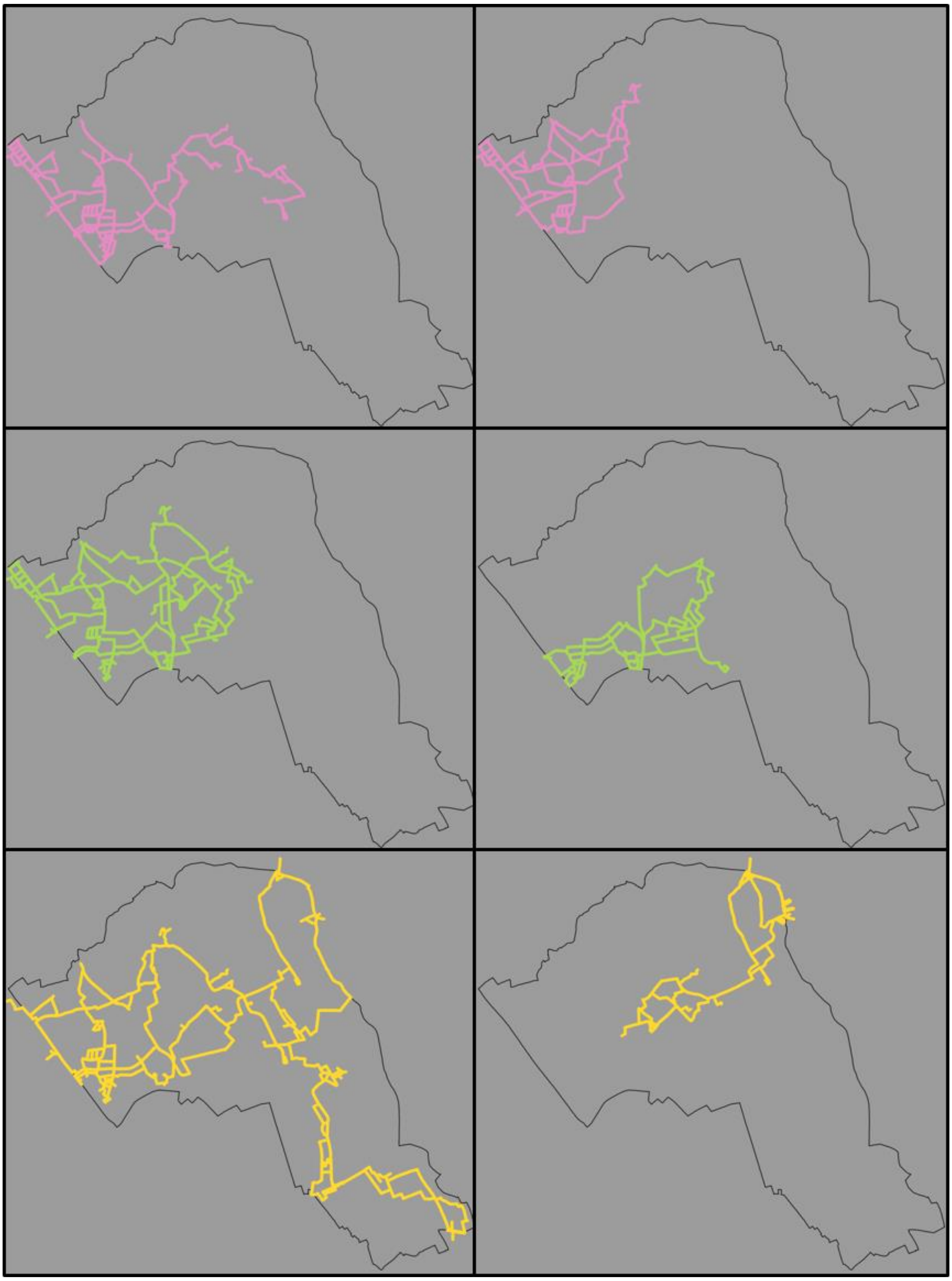


Figure 8.4 The activity space of patrol officers using different patrol strategies. The left and right columns correspond to the BAPS and Districting-BAPS, respectively. Each row represents a district and a police station. The colour scheme is consistent with Figure 8.2

8.4 Chapter Summary

In this chapter, the coupling of distinct police patrol strategies was discussed. Strategy coupling is well motivated, as it allows for greater complexities of the patrolling scenario, and it can increase the efficiency of police patrolling. Focus was placed on the coupling of the districting and the online repeated patrol routing strategies.

A case study was conducted to experiment with the combined strategy. It was demonstrated that, although the new strategy is slighted less efficient than the original routing strategy, in terms of global average idleness, the new strategy is applicable for guiding the patrol routes of officers. The advantage of the new strategy over the original one is that the patrol officers are restricted to their corresponding districts, which makes it easier for them to return to the police station for indoor tasks.

These patrol strategies can be coupled in other ways. For instance, a route planning system can integrate the repeated and infrequent routing strategies that can switch between these two strategies in response to the real-time situation. Particularly, at the beginning of the patrol shift, the patrollers follow the balanced routes to monitor the crime hotspots. As calls for service occur in the territory, and the number of patrol officers changes, the system can switch to an online patrolling mode. Another possibility is to couple the districting and infrequent patrol routing strategy, both of which contains the objective of balancing workloads. In this way, the workload balance at different levels can be achieved: 1) the districting strategy guarantees the workload balance among districts and in the long term; 2) the routing strategy for infrequent patrol routing strategy ensures the workload balance among the officers in a district, and in the daily operations. This coupled strategy can be also used for deliveries from multiple warehouses.

Chapter 9

CONCLUSIONS AND FURTHER WORK

9 CONCLUSIONS AND FURTHER WORK

9.1 Chapter Overview

This thesis is concluded in this chapter. First, in Section 9.2, a chapter-by-chapter summary of the thesis is given, in which the main results of each chapter are summarised. Second, in Section 9.3, the contribution of this thesis to the literature is expounded on. Third, in Section 9.4, the limitations of this study are considered. Finally, in Section 9.5, an outlook is provided for future research and applications in police patrol strategies.

9.2 Thesis Summary

Chapter 1 demonstrated the importance of police patrolling and efficient police patrol strategies. However, the current literature on police patrol strategies fails to systematically account for the effects of the urban street network that determines the configuration of the built environment and shapes human activity patterns relating to crime and policing. Therefore, the aim of this study was to develop a framework, consisting of different police patrol strategies, that was built on the urban street network, in order to cover crime hotspots and the entire network effectively.

Chapter 2 provided a thorough review of the relevant research. First, the police patrol models were categorised into two types, namely police districting models and patrol routing models. Second, the police districting models were introduced, including their definition, the attributes considered, and solution approaches. Previous research was found to fail to explicitly use streets as the basic units in PDP. Third, the police routing models were described and further categorised into RPRP and IPRP. While there is a body of research on repeated patrol routing, the approaches used were determined not to be applicable in guiding police patrol, as they omit the peculiarities of police patrols. In terms of infrequent police patrol routing, the existing research assumes that patrollers start from the same depot, ignoring the fact that patrol officers can start from different stations.

In Chapter 3, the methodological framework of this thesis was discussed. This framework includes three different police patrol strategies, and these strategies can be coupled to form new strategies. This chapter also provided a brief introduction to the formulation and solution approaches of the three strategies.

In Chapter 4, the case study areas and the datasets used in this study were introduced. Most of the case studies were conducted in the London Borough of Camden (UK). The Ordnance Survey ITN layer provided the fundamental street network. The crime record datasets and the derived crime risk maps were used as input to the police patrolling strategies. Moreover, the

CLC dataset was utilised to reveal the characteristics of the identified patrol districts, and to validate the districting plan. Another case study area is the South Chicago.

In Chapter 5, the SNPDP was introduced, which is a novel approach to incorporating the street network structure and street-level predictive crime risk into the design of police districts. A mathematical formulation of this model was proposed, which demonstrated that problem complexity increases rapidly with problem size. Therefore, a heuristic GP-TS was developed to solve this problem. This heuristic compared favourably to the exact solver Gurobi and the existing heuristics in the fictional and real-world districting problems, and it is capable of producing high-quality solutions quickly for the SNPDP.

In Chapter 6, the infrequent coverage routing in police patrolling and a balanced route design were discussed. Firstly, the infrequent coverage problem was formulated as a MMMDRPP, and a linear formulation of the MMMDRPP was given. This problem was challenging because of the constraints imposed by multiple depots and the objective of minimising the length of the longest route. Therefore, a tabu search algorithm TABU-PATROL was proposed to solve this problem. The case study demonstrated that TABU-PATROL can generate high-quality balanced routes for patrol officers from different depots within a reasonable time.

In Chapter 7, an online cooperative patrol routing strategy was introduced, in order to cover crime hotspots repeatedly. First, a set of five quantitative measures were proposed for the repeated police patrol routing – efficiency, flexibility, unpredictability, scalability, and robustness. In light of these guidelines, a real-time routing strategy BAPS was developed. This strategy considered multiple factors that impact patrolling, and adopted a probabilistic computational framework. The case studies in Camden and Chicago demonstrated that BAPS generally outperforms a cyclic patrolling strategy on the proposed measures.

In Chapter 8, the coupling of the proposed police patrol strategies was discussed. The strategy coupling was well motivated, as it accommodates greater complexities in the patrolling process, and it can increase the efficiency of patrolling. The focus was placed on coupling the districting strategy and the online repeated patrol routing strategy. The case study demonstrated that the new strategy is applicable in guiding the patrol routes of the officers, although it was slightly less efficient than the original strategy. The advantage of the new strategy is that the patrol officers are restricted to their corresponding districts.

9.3 Contribution to the Literature

In this section, the potential contributions of the research behind this thesis to the literature are summarised. The contributions can be divided into three categories: 1) contributions to the literature on police patrolling strategies; 2) contributions to the literature on routing problems; and 3) contributions to the literature on districting models.

9.3.1 Contributions to the Literature on Police Patrol Strategies

This study addressed the challenges of developing police patrolling strategies by:

1. Proposing the design of police districts based on street networks, which incorporates the underlying network structure and the predictive crime risk on the streets;
2. Formulating the RPRP, and proposing a set of guidelines and measures to evaluate the patrol performance;
3. Developing an online and cooperative routing strategy for effective foot patrol, by combining a pheromone-based algorithm and Bayesian decision-making; and
4. Developing a model and algorithm to design balanced routes for patrol officers from different stations, in order to cover each given crime hotspot at least once.

9.3.2 Contributions to the Literature on Routing Problems

This study contributes to the literature of routing problems by:

1. Extending the MMRPP to the multiple depot cases, which has great potentials in warehouse deliveries and other applications;
2. Formulating the MMMDRPP, which seeks balanced routes from multiple depots to cover a subset of a road network, and by analysing the problem complexity;
3. Proposing three lower bounds for this problem, in order to approximate the optimal solution value; and
4. Developing a three-step heuristic method for the MMMDRPP, which is capable of generating high-quality solutions in an acceptable time.

9.3.3 Contributions to the Literature on Districting Models

This study contributes to the literature on districting models (or regionalisation) by:

1. Proposing a PDP that uses the streets as basic units, and that seeks an efficient and balanced partition of the workload;
2. Proposing two different approaches to generate police districts, based on the street network. The first approach solves the districting problem on grids, and then transforms the results to the street network, while the second directly generates districts consisting of street segments; and
3. Developing an efficient heuristic to solve the districting problem, which consists of a graph partitioning algorithm and a TS procedure.

9.4 A Critique of the Limitations of the Study

The contributions of the research presented in this thesis to the current literature have been presented in the previous chapters. A number of limitations, however, were encountered during the research process. These limitations are highlighted here, and need to be considered in applying the research outcomes to police operations.

9.4.1 Limitation in the Cooperation between Foot Patrols and Vehicular Patrols

The research presented in this thesis focuses on the patrolling activity of officers of a single type, those involved in foot patrols. Another assumption of the research was that the patrollers moved at a constant speed. However, there is a lack of consideration for the cooperation of multiple types of patrollers. In operations, foot patrols and vehicular patrols may cooperate in certain tasks, and foot patrollers may occasionally use vehicles for long-distance travelling. This cooperation can have two-fold effect on the patrolling. First, the patrollers may be heterogeneous, having different speeds, movement ranges, and movement patterns. For instance, foot patrols can traverse all roads (except highways) in either direction, and even use paths not on the network links, especially in open areas, such as parks, grasslands and squares. In contrast, vehicular patrols are restricted to main roads that are accessible to cars and must obey the rules of the road, including traffic lights and the one-way systems, although they may neglect the rules when responding to emergencies. Second, the patroller may change their moving speed when necessary. This is obvious when patrol cars have to slow down in congested areas and stop for red lights, or when foot patrols take a vehicle for a long-distance trip.

9.4.2 Limitation in the Model Validation

The study defined several models concerning effective patrolling, and proposed the corresponding algorithms. The algorithms have been validated using real-world cases and were based on proposed quantitative measures. This validation, however, failed to consider the attitudes of police commanders and front-line officers. The successful application of a certain strategy depends on the willingness of commanders and executors. Will front-line officers accept change in their work routines that results from a new strategy? Given the suggested patrol route, do they fully implement it during the patrolling period, or do they deviate from it on occasion? Do they find the new patrol route usable and easy to follow? These questions should be investigated and answered as part of a comprehensive validation process for the proposed strategies.

9.4.3 Simplification of the Patrolling Tasks

This research focused on the task of covering crime hotspots (via patrol routing), or covering the entire street network (via the design of police districts). Although this is one of the central tasks of policing, other tasks exist that must be performed. Policing is a complex and culturally specific process, and officers may have many intersecting responsibilities. Apart from patrolling territory to increase their visibility and deter crime, officers from MPS in London are also expected to respond to calls for service, help with finding missing persons, and handle traffic accidents. Therefore, these tasks must be taken into account in order to derive a realistic and usable police patrol strategy. In particular, when and how often do police officers switch between different task modes? In the design of patrol routes, are there any other important tasks or locations that should be considered, apart from monitoring crime hotspots?

9.5 **Conclusion and Outlook**

In this thesis, a framework for police patrolling strategies on the street network has been proposed in order to ensure an efficient coverage of crime hotspots and a balanced workload among police officers. This framework consists of three strategies, including the districting model, the patrol routing strategy for repeated coverage, and the patrol routing strategy for infrequent coverage. These strategies consider the underlying structure of the street network, the distribution of crime hotspots, and the cooperation between patrol officers belonging to different stations. They were applied to patrolling of the street network in the London Borough of Camden, and exhibited strong performance compared to the corresponding benchmark strategies, suggesting that they may have considerable potential in police operations. Moreover, these strategies are applicable to other similar fields and scenarios. For example, the districting model could be used to design the fire emergency response areas, and the routing strategy for infrequent coverage could be applied to generating routes in fields such as deliveries from multiple warehouses and estate security guards.

As populations in the big cities continue to increase, and new types of crime emerge in the urban areas, police forces are under increasing strain to suppress and deter crime. In London, the number of total crimes increased by 4.56% in the year 2017 compared with the year 2016 (The Metropolitan Police, 2018). Such a situation is exacerbated by cuts into police funding, making it harder for officers to tackle crime. For example, the MPS has faced a cut of more than £700 million in cash terms or approximately 40% in real terms, in the last 10 years (Mayor of London, 2018). This implies that intelligent policing (Cheng *et al.*, 2016) is of immediate practical use, to ensure the efficient deployment of resources. The patrolling strategies proposed in this thesis constitute an essential part of intelligent policing, and are likely to improve the efficiency of patrolling and so strengthen the deterrence of crime.

To conclude, efficient police patrol strategies for different scenarios are an important approach in maintaining police deterrence and preventing potential crime. However, their successful application in policing operations relies on incorporating the complexity of policing activities as well as the preferences of police officers. The proposed police patrol strategies are likely to have considerable potentials in practice, and can possibly improve the efficiency of patrolling in operations.

BILIOGRAPHY

Abate, F. (1996) *The Oxford dictionary and thesaurus: The ultimate language reference for American readers*. Oxford: Oxford Univ. Press.

Adepeju, M., Rosser, G. and Cheng, T. (2016) 'Novel evaluation metrics for sparse spatio-temporal point process hotspot predictions - a crime case study', *International Journal of Geographical Information Science*, 30(11), pp. 2133–2154. doi: 10.1080/13658816.2016.1159684.

Ahr, D. (2004) *Contributions to Multiple Postmen Problems*, *Naturwissenschaftlich-Mathematische Gesamtfakultät*. Ruprecht-Karls-Universitat, Heidelberg. Available at: http://www.ub.uni-heidelberg.de/archiv/4963.

Ahr, D. and Reinelt, G. (2002) 'New Heuristics and Lower Bounds for the Min-Max k-Chinese Postman Problem', in Mohring, R. and Raman, R. (eds) *Algorithms — ESA 2002*. Heidelberg: Springer Berlin, pp. 64–74. Available at: http://link.springer.com/chapter/10.1007/3-540-45749-6_10.

Ahr, D. and Reinelt, G. (2006) 'A tabu search algorithm for the min-max k-Chinese postman problem', *Computers and Operations Research*, 33(12), pp. 3403–3422. doi: 10.1016/j.cor.2005.02.011.

Akyurt, I. Z., Keskinturk, T. and Kalkanci, C. (2015) 'Using Genetic Algorithm For Winter Maintenance Operations: Multi Depot K-Chinese Postman Problem', *Emerging Markets Journal*. University Library System, University of Pittsburgh, 5(1), p. 50.

Almeida, A. *et al.* (2004) 'Recent advances on multi-agent patrolling', in *Advances in Artificial Intelligence - SBIA 2004*. Heidelberg: Springer Berlin Heidelberg, pp. 474–483.

Applegate, D. *et al.* (2002) 'Solution of a Min-Max Vehicle Routing Problem', *INFORMS*

Journal on Computing, 14(2), pp. 132–143. doi: 10.1287/ijoc.14.2.132.118.

Arkin, E. M., Hassin, R. and Levin, A. (2006) ‘Approximations for minimum and min-max vehicle routing problems’, *Journal of Algorithms*, 59(1), pp. 1–18. doi: 10.1016/j.jalgor.2005.01.007.

Armitage, R. (2007) ‘Sustainability versus safety: confusion, conflict and contradiction in designing out crime’, in *Imagination for Crime Prevention. Crime Prevention Studies Volume 21*, pp. 81–110.

Balaprakash, P., Birattari, M. and Stuetzle, T. (2007) ‘Improvement Strategies for the F-Race algorithm: Sampling Design and Iterative Refinement’, in *Hybrid Metaheuristics, Proceedings*, pp. 108–122. doi: 10.1007/978-3-540-75514-2.

Balch, T. and Arkin, R. (1994) ‘Communication in reactive multiagent robotic systems’, *Autonomous Robots*, 1(1), pp. 27–52.

Beavon, D. J. K., Brantingham, P. L. and Brantingham, P. J. (1994) ‘The influence of street networks on the patterning of property offenses’, *Crime Prevention Studies, Vol 2*, pp. 115–148. Available at: http://www.popcenter.org/library/crimeprevention/volume_02/06beavon.pdf.

Belenguer, J.-M. *et al.* (2006) ‘Lower and upper bounds for the mixed capacitated arc routing problem’, *Computers & Operations Research*, 33(12), pp. 3363–3383. doi: 10.1016/j.cor.2005.02.009.

Benavent, E. *et al.* (2009) ‘Min-Max K-vehicles windy rural postman problem’, *Networks*, 54(4), pp. 216–226. doi: 10.1002/net.20334.

Benveniste, R. (1985) ‘Solving the combined zoning and location problem for several emergency units.’, *The Journal of the Operational Research Society*, 36(5), pp. 433–450. doi:

10.1057/jors.1985.70.

Bergey, P. K., Ragsdale, C. T. and Hoskote, M. (2003a) 'A decision support system for the electrical power districting problem', *Decision Support Systems*, 36(1), pp. 1–17. doi: 10.1016/S1344-6223(02)00033-0.

Bergey, P. K., Ragsdale, C. T. and Hoskote, M. (2003b) 'A Simulated Annealing Genetic Algorithm for the Electrical Power Districting Problem', *Annals of Operations Research*, 121(1–4), pp. 33–55. doi: 10.1023/A:1023347000978.

Blais, M., Lapierre, S. D. and Laporte, G. (2003) 'Solving a home-care districting problem in an urban setting', *Journal of the Operational Research Society*, 54(11), pp. 1141–1147. doi: 10.1057/palgrave.jors.2601625.

Bodily, S. E. (1978) 'Police sector design incorporating preferences of interest groups for equality and efficiency', *Management Science*. INFORMS, 24(12), pp. 1301–1313. doi: 10.1287/mnsc.24.12.1301.

Bollobás, B. (2002) *Modern graph theory*. 2nd Editio. New York: Springer.

Bowers, K. J., Johnson, S. D. and Pease, K. (2004) 'Prospective hot-spotting: The future of crime mapping?', *British Journal of Criminology*, 44(5), pp. 641–658. doi: 10.1093/bjc/azh036.

Bozkaya, B., Erkut, E. and Laporte, G. (2003) 'A tabu search heuristic and adaptive memory procedure for political districting', *European Journal of Operational Research*, 144(1), pp. 12–26. doi: 10.1016/S0377-2217(01)00380-0.

Braa, J. J. and Hedberg, C. (2002) 'The struggle for district-based health information systems in South Africa', *The Information Society*, 18(April), pp. 113–127. doi: 10.1080/01972240290075048.

Brandão, J. and Eglese, R. (2008) 'A deterministic tabu search algorithm for the capacitated arc routing problem', *Computers & Operations Research*, 35(4), pp. 1112–1126. doi: 10.1016/j.cor.2006.07.007.

Brantingham, P. J., Brantingham, P. L. and others (1981) *Environmental criminology*. Sage Publications Beverly Hills, CA.

Brantingham, P. L. *et al.* (2009) 'Crime Analysis at Multiple Scales of Aggregation: A Topological Approach', in *Putting Crime in its Place*. New York: Springer, pp. 87–107. doi: 10.1007/978-0-387-09688-9.

Brantingham, P. L. and Brantingham, P. J. (1993) 'Environment, routine, and situation: Toward a pattern theory of crime', *Routine Activity and Rational Choice: Advances in Criminological Theory*, (1789), pp. 259–294.

Bruce, C. (2009) 'Districting and resource allocation: A question of balance', *Geography & Public Safety*, 1(4), pp. 1–3. Available at: https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=247736 (Accessed: 20 March 2017).

Butsch, A., Kalcsics, J. and Laporte, G. (2014) 'Districting for arc routing', *INFORMS Journal on Computing*, 26(4), pp. 809–824. doi: 10.1287/ijoc.2014.0600.

Camacho-Collados, M., Liberatore, F. and Angulo, J. M. (2015) 'A multi-criteria Police Districting Problem for the efficient and effective design of patrol sector', *European Journal of Operational Research*, 246(2), pp. 674–684. doi: 10.1016/j.ejor.2015.05.023.

Caro, F. *et al.* (2004) 'School redistricting: embedding GIS tools with integer programming', *Journal of the Operational Research Society*, 55(8), pp. 836–849. doi: 10.1057/palgrave.jors.2601729.

Carroll, J. M. and Laurin, P. G. (1981) 'Using simulation to assign police patrol zones', *Simulation*, 36(1), pp. 1–12. doi: 10.1177/003754978103600102.

Castle, C. J. E. and Crooks, A. T. (2006) *Principles and concepts of Agent-Based Modelling for developing geospatial simulations*. London. Available at: https://www.bartlett.ucl.ac.uk/casa/pdf/paper110.pdf.

Chen, H., Cheng, T. and Wise, S. (2015) 'Designing daily patrol routes for policing based on ant colony algorithm', *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-4/W2, pp. 103–109. doi: 10.5194/isprsannals-II-4-W2-103-2015.

Chen, H., Cheng, T. and Wise, S. (2017) 'Developing an online cooperative police patrol routing strategy', *Computers, Environment and Urban Systems*, 62, pp. 19–29. doi: 10.1016/j.compenvurbsys.2016.10.013.

Chen, X. and Yum, T.-S. P. (2010) 'Cross entropy approach for patrol route planning in dynamic environments', *2010 IEEE International Conference on Intelligence and Security Informatics*. IEEE, pp. 114–119. doi: 10.1109/ISI.2010.5484767.

Cheng, T. *et al.* (2016) *CPC: Crime, Policing and Citizenship – Intelligent policing and big data*. London. Available at: https://www.researchgate.net/publication/303539868_CPC_Crime_Policing_and_Citizenship_-_Intelligent_policing_and_big_data.

Chevaleyre, Y. (2004) 'Theoretical analysis of the multi-agent patrolling problem', *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'04)*. doi: 10.1109/IAT.2004.1342959.

Christofides, N. *et al.* (1981) *An algorithm for the rural postman problem*, *Report IC.OR.81.5*. London. Available at:

https://www.researchgate.net/publication/239385320_An_algorithm_for_the_Rural_Postman_Problem (Accessed: 22 April 2015).

Chu, H. N. *et al.* (2007) ‘Swarm approaches for the patrolling problem, information propagation vs. pheromone evaporation’, in *19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007)*. IEEE, pp. 442–449. doi: 10.1109/ICTAI.2007.80.

Cirincione, C., Darling, T. A. and O’Rourke, T. G. (2000) ‘Assessing South Carolina’s 1990s congressional districting’, *Political Geography*, 19(2), pp. 189–211. doi: 10.1016/S0962-6298(99)00047-5.

Cohen, L. E. and Felson, M. (1979) ‘Social change and crime rate trends: A routine activity approach’, *American sociological review*. JSTOR, pp. 588–608.

Cole, B. *et al.* (2015) ‘Corine land cover 2012 for the UK, Jersey and Guernsey’. NERC Environmental Information Data Centre. doi: 10.5285/32533dd6-7c1b-43e1-b892-e80d61a5ea1d.

Cook, P. J. (1980) ‘Research in criminal deterrence: Laying the groundwork for the second decade’, *Crime and Justice*, 2, pp. 211–268. doi: 10.1086/449070.

Cova, T. and Church, R. (2000) ‘Contiguity constraints for single-region site search problems’, *Geographical Analysis*, 32(4), pp. 306–329. doi: 10.1111/j.1538-4632.2000.tb00430.x.

Croxford, B. (1999) ‘Can Space Syntax Predict Traffic Flows, Speeds and Mix?’, in *Space Syntax Second International Symposium*. Available at: http://www.spacesyntax.net/symposia-archive/SSS2/SpSx 2nd Symposium 99 -2003 pdf/2nd Symposium Vol 2 pdf/49 Croxford 300.pdf (Accessed: 28 October 2017).

Curtin, K. M. *et al.* (2005) ‘Integrating GIS and maximal covering models to determine optimal police patrol areas', in *Geographic Information Systems and Crime Analysis*. Hershey: IGI Global, pp. 214–235.

Curtin, K. M., Hayslett-McCall, K. and Qiu, F. (2010) 'Determining optimal police patrol areas with maximal covering and backup covering location models', *Networks and Spatial Economics*, 10(1), pp. 125–145. doi: 10.1007/s11067-007-9035-6.

D'Amico, S. J. *et al.* (2002) 'A simulated annealing approach to police district design', *Computers and Operations Research*, 29(6), pp. 667–684. doi: 10.1016/S0305-0548(01)00056-9.

Davies, T. and Johnson, S. D. (2015) 'Examining the relationship between road structure and burglary risk via quantitative network analysis', *Journal of Quantitative Criminology*, 31(3), pp. 481–507. doi: 10.1007/s10940-014-9235-4.

Davies, T. P. *et al.* (2013) 'A mathematical model of the London riots and their policing', *Scientific Reports*. Nature Publishing Group, 3. doi: 10.1038/srep01303.

Davies, T. P. and Bishop, S. R. (2013) 'Modelling patterns of burglary on street networks', *Crime Science*. Springer, 2(1), p. 10. doi: 10.1186/2193-7680-2-10.

Delgado, C. and Pacheco, J. (2001) 'Minmax Vehicle Routing Problems: Application to School Transport in the Province of Burgos', in *Computer-Aided Scheduling of Public Transport*, pp. 9–11. doi: 10.1007/978-3-642-56423-9.

Dijkstra, E. W. (1959) 'A note on two problems in connection with graphs', *Numerische Mathematik*, pp. 269–271. doi: 10.1007/BF01386390.

Doi, S. (2013) 'Proposal and evaluation of a pheromone-based algorithm for the patrolling problem in dynamic environments', in *2013 IEEE Symposium on Swarm Intelligence (SIS)*. IEEE, pp. 48–55. doi: 10.1109/SIS.2013.6615158.

Dorigo, M., Maniezzo, V. and Colorni, A. (1991) ‘Positive feedback as a search strategy’, *Technical Report 91-016*. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.52.6342 (Accessed: 8 April 2015).

Dror, M. (ed.) (2000) *Arc Routing: Theory, Solutions and Applications*. 1st edn. New York: Springer Science & Business Media.

Eck, J. *et al.* (2005) ‘Mapping crime: Understanding Hotspots’, *National Institute of Justice: Washington DC.* National Institute of Justice. Available at: http://discovery.ucl.ac.uk/11291/ (Accessed: 15 October 2017).

Elmaliach, Y., Agmon, N. and Kaminka, G. A. (2009) ‘Multi-robot area patrol under frequency constraints’, *Annals of Mathematics and Artificial Intelligence*, 57(3), pp. 293–320. doi: 10.1007/s10472-010-9193-y.

Epstein, J. and Axtell, R. (1996) *Growing artificial societies: Social science from the bottom up*. 1st edn. Washington, D.C.: Brookings Institution Press.

European Commission (2010) *Regional GDP per capita in the EU in 2010: eight capital regions in the ten first places*.

Fazli, P. and Davoodi, A. (2010) ‘Complete and robust cooperative robot area coverage with limited range’, *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 5577–5582.

Fleischmann, B. and Paraschis, J. N. (1988) ‘Solving a large scale districting problem: A case report’, *Computers & Operations Research*, 15(6), pp. 521–533. doi: http://dx.doi.org/10.1016/0305-0548(88)90048-2.

Floyd, R. W. (1962) ‘Algorithm 97: Shortest path’, *Communications of the ACM*, 5(6), p. 345. doi: 10.1145/367766.368168.

Freeman, L. C. (1977) 'A Set of Measures of Centrality Based on Betweenness', *Sociometry*. doi: 10.2307/3033543.

Fu, J. G. M. and Ang, M. H. (2009) 'Probabilistic ants (PAnts) in multi-agent patrolling', in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*. IEEE, pp. 1371–1376. doi: 10.1109/AIM.2009.5229880.

Gabriely, Y. and Rimon, E. (2001) 'Spanning-tree based coverage of continuous areas by a mobile robot', *Annals of Mathematics and Artificial Intelligence*, pp. 77–98.

Galvão, L. C. *et al.* (2006) 'A multiplicatively-weighted Voronoi diagram approach to logistics districting', *Computers and Operations Research*, 33(1), pp. 93–114. doi: 10.1016/j.cor.2004.07.001.

Glover, F. (1986) 'Future paths for integer programming and links to artificial intelligence', *Computers and Operations Research*, 13(5), pp. 533–549. doi: 10.1016/0305-0548(86)90048-1.

Glover, F. and Laguna, M. (1997) *Tabu search*. Boston: Kluwer academic publishers.

Golden, B. L. and Wong, R. T. (1981) 'Capacitated arc routing problems', *Networks*, 11(3), pp. 305–315. doi: 10.1002/net.3230110308.

Gurobi Optimization, I. (2016) *Gurobi Optimizer Reference Manual*. Available at: http://www.gurobi.com (Accessed: 1 September 2017).

Heppenstall, A. J. J. *et al.* (2012) *Agent-Based Models of Geographical Systems*, *Agent-Based Models of Geographical Systems*. Edited by A. J. Heppenstall et al. Dordrecht: Springer Netherlands. doi: 10.1007/978-90-481-8927-4.

Hillier, B. (2007) *Space is the machine: a configurational theory of architecture*, *Space*

Syntax. London, UK: Space Syntax. doi: 10.1016/S0142-694X(97)89854-7.

Hillier, B. and Iida, S. (2005) ‘Network and psychological effects in urban movement’, in *International Conference on Spatial Information Theory*. Springer, Berlin, Heidelberg, pp. 475–490. doi: 10.1007/11556114_30.

Hillier, B. and Sahbaz, O. (2007) ‘Beyond hot spots; using space syntax to understand dispersed patterns of crime risk in the built environment’, *Conference on crime analysis at Institute of Pure and Applied Mathematics, University of California*, pp. 1–36. Available at: http://discovery.ucl.ac.uk/18549/ (Accessed: 28 October 2017).

Iannoni, A. P., Morabito, R. and Saydam, C. (2009) ‘An optimization approach for ambulance location and the districting of the response segments on highways’, *European Journal of Operational Research*, 195(2), pp. 528–542. doi: 10.1016/j.ejor.2008.02.003.

Jensen, F. V. and Nielsen, T. D. (2007) *Bayesian Networks and Decision Graphs*. 2nd edn, *Bayesian Networks and Decision Graphs: February 8*. 2nd edn. New York: Springer New York (Information Science and Statistics). doi: 10.1007/978-0-387-68282-2.

JGraphT (2017) *JGraphT Library*. Available at: http://jgrapht.org/ (Accessed: 1 September 2017).

Jiang, B. and Claramunt, C. (2004) ‘Topological Analysis of Urban Street Networks’, *Environment and Planning B: Planning and Design*, 31(1), pp. 151–162. doi: 10.1068/b306.

Johnson, S. and Bowers, K. (2014) ‘Near Repeats and Crime Forecasting’, in *Encyclopedia of Criminology and Criminal Justice*. New York: Springer-Verlag New York Inc., pp. 3242–3254. Available at: http://link.springer.com/10.1007/978-1-4614-5690-2_210 (Accessed: 25 August 2015).

Johnson, S. D. and Bowers, K. J. (2010) ‘Permeability and burglary risk: Are cul-de-sacs safer?', *Journal of Quantitative Criminology*, 26(1), pp. 89–111. doi: 10.1007/s10940-009-9084-8.

Jones, M. A. and Fanek, M. F. (no date) 'Crime in the Urban Environment'. Available at: http://www.spacesyntax.net/symposia-archive/SSS1/SpSx 1st Symposium 97 -2003 pdf/1st Symposium Vol II pdf/4 - Urban themes/25 Fanek 300.pdf (Accessed: 28 October 2017).

Jozefowiez, N., Semet, F. and Talbi, E.-G. (2008) 'Multi-objective vehicle routing problems', *European Journal of Operational Research*, 189(2), pp. 293–309. doi: 10.1016/j.ejor.2007.05.055.

Kalcsics, J. (2015) 'Districting Problems', in *Location Science*. Berlin: Springer International Publishing, pp. 595–622. doi: 10.1007/978-3-319-13111-5.

Kalcsics, J., Nickel, S. and Schröder, M. (2005) 'Towards a unified territorial design approach — Applications, algorithms and GIS integration', *Top*, 13(1), pp. 1–56. doi: 10.1007/BF02578982.

Kistler, A. (2009) 'Tucson police officers redraw division boundaries to balance their workload', *Geography & Public Safety*, 1(4), pp. 3–5. Available at: https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=247736.

Lacomme, P., Prins, C. and Sevaux, M. (2006) 'A genetic algorithm for a bi-objective capacitated arc routing problem', *Computers and Operations Research*, 33(12), pp. 3473–3493. doi: 10.1016/j.cor.2005.02.017.

Lai, J., Cheng, T. and Lansley, G. (2017) 'Improved targeted outdoor advertising based on geotagged social media data', *Annals of GIS*. Taylor & Francis, pp. 1–14. doi: 10.1080/19475683.2017.1382571.

Larson, R. C. (1974) 'A hypercube queuing model for facility location and redistricting in urban emergency services', *Computers and Operations Research*, 1(1), pp. 67–95. doi: 10.1016/0305-0548(74)90076-8.

Lasley, J. R. (1996) *Using traffic barriers to design out crime: a program evaluation of LAPD's Operation Cul de Sac*. Washington, DC.

Lei, T. L., Church, R. L. and Lei, Z. (2015) 'A unified approach for location-allocation analysis: integrating GIS, distributed computing and spatial optimization', *International Journal of Geographical Information Science*, 8816(July), pp. 1–20. doi: 10.1080/13658816.2015.1041959.

Lenstra, J. K. and Kan, A. H. G. R. (1976) 'On general routing problems', *Networks*. Wiley Subscription Services, Inc., A Wiley Company, 6(3), pp. 273–280. doi: 10.1002/net.3230060305.

Lerman, Y., Rofe, Y. and Omer, I. (2014) 'Using Space Syntax to Model Pedestrian Movement in Urban Transportation Planning', *Geographical Analysis*, 46(4), pp. 392–410. doi: 10.1111/gean.12063.

Li, L. Y. O. and Eglese, R. W. (1996) 'An interactive algorithm for vehicle routeing for winter gritting', *Journal of the Operational Research Society*, 47(2), pp. 217–228. doi: 10.1057/palgrave.jors.0470202.

Liberatore, F. and Camacho-Collados, M. (2016) 'A comparison of local search methods for the multicriteria police districting problem on graph', *Mathematical Problems in Engineering*. Hindawi Publishing Corporation, 2016, pp. 1–13. doi: 10.1155/2016/3690474.

López-Ibáñez, M. *et al.* (2016) 'The irace package: Iterated racing for automatic algorithm configuration', *Operations Research Perspectives*, 3, pp. 43–58. doi: http://dx.doi.org/10.1016/j.orp.2016.09.002.

López-Pérez, J. F. and Ríos-Mercado, R. Z. (2013) ‘Embotelladoras ARCA uses operations research to improve territory design plans’, *Interfaces*, 43(3), pp. 209–220. doi: 10.1287/inte.1120.0675.

Machado, A. *et al.* (2002) ‘Multi-agent patrolling: An empirical analysis of alternative architectures’, in *3rd International Conference on Multi-agent-based Simulation II (MABS)*, pp. 155–170. doi: 10.1007/3-540-36483-8_11.

Mayer, A. (2009) ‘Geospatial technology helps East Orange crack down on crime’, *Geography & Public Safety*, 1(4), pp. 8–10. Available at: https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=247736.

Mayor of London (2018) *Mayor sets out reality of Government’s dangerous police funding cuts | London City Hall*. Available at: https://www.london.gov.uk/press-releases/mayoral/stark-reality-of-government-police-funding-cuts (Accessed: 30 January 2018).

Mehrotra, A., Johnson, E. L. and Nemhauser, G. L. (1998) ‘An Optimization Based Heuristic for Political Districting’, *Management Science*, 44(8), pp. 1100–1114. doi: 10.1287/mnsc.44.8.1100.

Met Police (2017) *Crime Data Dashboard*. Available at: https://www.met.police.uk/stats-and-data/crime-data-dashboard/ (Accessed: 13 November 2017).

Minciardi, R., Puliafito, P. P. and Zoppoli, R. (1981) ‘A districting procedure for social organizations’, *European Journal of Operational Research*, 8(1), pp. 47–57. doi: 10.1016/0377-2217(81)90028-X.

Mitchell, P. S. (1972) ‘Optimal selection of police patrol beats’, *Journal of Criminal Law, Criminology & Police Science*, 63(4), pp. 577–584. doi: 10.2307/1141814.

Muyldermans, L. (2003) ‘Routing, districting and location for arc traversal problems’, *4OR*, 1(2), pp. 169–172. doi: 10.1007/s10288-003-0015-5.

Muyldermans, L., Cattrysse, D. and Van Oudheusden, D. (2003) ‘District design for arc-routing applications’, *Journal of the Operational Research Society*, 54(11), pp. 1209–1221. doi: 10.1057/palgrave.jors.2601626.

Okabe, A. and Kitamura, M. (1996) ‘A Computational Method for Market Area Analysis on a Network’, *Geographical Analysis*, 28(4), pp. 330–349. doi: 10.1111/j.1538-4632.1996.tb00939.x.

Okabe, A. and Okunuki, K. (2001) ‘A Computational Method for Estimating the Demand of Retail Stores on a Street Network and its Implementation in GIS’, *Transactions in GIS*, 5(3), pp. 209–220. doi: 10.1111/1467-9671.00078.

Okabe, A. and Satoh, T. (2006) ‘Uniform network transformation for points pattern analysis on a non-uniform network’, *Journal of Geographical Systems*, 8(1), pp. 25–37. doi: 10.1007/s10109-005-0009-2.

Okabe, A., Satoh, T. and Sugihara, K. (2009) ‘A kernel density estimation method for networks, its computational method and a GIS-based tool’, *International Journal of Geographical Information Science*, 23(1), pp. 7–32. doi: 10.1080/13658810802475491.

Okabe, A., Yomono, H. and Kitamura, M. (1995) ‘Statistical Analysis of the Distribution of Points on a Network’, *Geographical Analysis*, 27(2), pp. 152–175. doi: 10.1111/j.1538-4632.1995.tb00341.x.

Openshaw, S. (1984) *The Modifiable Areal Unit Problem*. Norwich: Geo Books (Concepts and techniques in modern geography). Available at: https://books.google.co.uk/books?id=BJ5jQgAACAAJ.

Openshaw, S. and Taylor, P. J. (1979) ‘A million or so correlation coefficients: Three experiments on the modifiable areal unit problem’, in *Statistical methods in the spatial sciences*. London: The Royal Geographical Society (with the Institute of British Geographers), pp. 127–144.

Ordnance Survey (2015) *OS MasterMap Integrated Transport Network Layer*. Available at: https://www.ordnancesurvey.co.uk/business-and-government/products/itn-layer.html (Accessed: 14 November 2017).

Park, K. *et al.* (2000) ‘Telecommunication Node Clustering with Node Compatibility and Network Survivability Requirements’, *Management Science*, 46(3), pp. 363–374. doi: 10.1287/mnsc.46.3.363.12066.

Pasqualetti, F., Franchi, A. and Bullo, F. (2012) ‘On cooperative patrolling: Optimal trajectories, complexity analysis, and approximation algorithms’, *IEEE Transactions on Robotics*, 28(3), pp. 592–606. doi: 10.1109/TRO.2011.2179580.

Pereira, J. R. and Fernández, E. (2015) ‘The Multi-Depot Rural Postman Problem’, in *Workshop on Combinatorial Optimization, Routing and Location*. Salamanca(Spain). Available at: http://coral.confereasy.com/en/papers/639/the-multi-depot-rural-postman-problem.

Platz, T. T. and Hamers, H. (2013) ‘On Games Arising from Multi-Depot Chinese Postman Problems’, *SSRN Electronic Journal*. doi: 10.2139/ssrn.2205692.

Porta, S., Crucitti, P. and Latora, V. (2006a) ‘The network analysis of urban streets: a dual approach’, *Physica A: Statistical Mechanics and its Applications*, 369(2), pp. 853–866. doi: 10.1016/j.physa.2005.12.063.

Porta, S., Crucitti, P. and Latora, V. (2006b) ‘The network analysis of urban streets: a primal approach', *Environment and Planning B: Planning and Design*, 33(5), pp. 705–725. doi: 10.1068/b32045.

Portugal, D. and Rocha, R. (2010) 'Msp algorithm: Multi-robot patrolling based on territory allocation using balanced graph partitioning', in *Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM. Available at: http://dl.acm.org/citation.cfm?id=1774360 (Accessed: 27 December 2014).

Portugal, D. and Rocha, R. P. (2011) 'On the performance and scalability of multi-robot patrolling algorithms', in *9th IEEE International Symposium on Safety, Security, and Rescue Robotics, SSRR 2011*, pp. 50–55. doi: 10.1109/SSRR.2011.6106761.

Portugal, D. and Rocha, R. P. (2013a) 'Distributed multi-robot patrol: A scalable and fault-tolerant framework', *Robotics and Autonomous Systems*, 61(12), pp. 4759–4764. doi: 10.1016/j.robot.2013.06.011.

Portugal, D. and Rocha, R. P. (2013b) 'Multi-robot patrolling algorithms: Examining performance and scalability', *Advanced Robotics*, 27(5), pp. 325–336. doi: 10.1080/01691864.2013.763722.

Prins, C., Lacomme, P. and Prodhon, C. (2014) 'Order-first split-second methods for vehicle routing problems: A review', *Transportation Research Part C: Emerging Technologies*, 40, pp. 179–200. doi: 10.1016/j.trc.2014.01.011.

R Core Team (2015) *R: A language and environment for statistical computing*. Vienna, Austria. Available at: https://www.r-project.org/ (Accessed: 1 September 2017).

Rao, T. M., Mitra, S. and Zollweg, J. (2011) 'Snow-plow route planning using AI search', in *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, pp. 2791–2796. doi: 10.1109/ICSMC.2011.6084095.

Ratcliffe, J. (2010) 'Crime mapping: spatial and temporal challenges', in *Handbook of Quantitative Criminology*. 2010th edn. New York: Springer New York.

Reades, J., Calabrese, F. and Ratti, C. (2009) 'Eigenplaces: analysing cities using the space – time structure of the mobile phone network', *Environment and Planning B: Planning and Design*, 36(5), pp. 824–836. doi: 10.1068/b34133t.

Reis, D. *et al.* (2006) 'GAPatrol: An evolutionary multiagent approach for the automatic definition of hotspots and patrol routes', *Advances in Artificial Intelligence - IBERAMIA-SBIA 2006*, 4140, pp. 118–127.

Reyns, B. W. (2016) 'Environmental criminology: Evolution, theory and practice', *Security Journal*. Palgrave Macmillan UK, 29(3), pp. e1–e3. doi: 10.1057/sj.2014.37.

Ricca, F., Scozzari, A. and Simeone, B. (2013) 'Political Districting: from classical models to recent approaches', *Annals of Operations Research*, 204(1), pp. 271–299. doi: 10.1007/s10479-012-1267-2.

Ricca, F. and Simeone, B. (2008) 'Local search algorithms for political districting', *European Journal of Operational Research*, 189(3), pp. 1409–1426. doi: 10.1016/j.ejor.2006.08.065.

Robinson, W. S. (1950) 'Ecological Correlations and the Behavior of Individuals', *American Sociological Review*. American Sociological Association, 15(3), p. 351. doi: 10.2307/2087176.

Rosser, G. *et al.* (2017) 'Predictive crime mapping: arbitrary grids or street networks?', *Journal of Quantitative Criminology*, 33(3), pp. 569–594. doi: 10.1007/s10940-016-9321-x.

Sak, T., Wainer, J. and Goldenstein, S. K. (2008) 'Probabilistic multiagent patrolling', in *Lecture Notes in Computer Science*, pp. 124–133. doi: 10.1007/978-3-540-88190-2-18.

Santana, H. *et al.* (2004) 'Multi-agent patrolling with reinforcement learning'. IEEE Computer Society, pp. 1122–1129. doi: 10.1109/AAMAS.2004.180.

Sarac, A. *et al.* (1999) 'Reconfiguring Police Reporting Districts in the City of Buffalo', *OR Insight*. Palgrave Macmillan UK, 12(3), pp. 16–24. doi: 10.1057/ori.1999.12.

Schoepfle, O. B. and Church, R. L. (1991) 'A new network representation of a "classic" school districting problem', *Socio-Economic Planning Sciences*, 25(3), pp. 189–197. doi: 10.1016/0038-0121(91)90017-L.

Sempe, F. and Drogoul, A. (2003) 'Adaptive patrol for a group of robots', *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, 3, pp. 2865–2869. doi: 10.1109/IROS.2003.1249305.

Shafahi, A. and Haghani, A. (2015) 'Generalized Maximum Benefit Multiple Chinese Postman Problem', *Transportation Research Part C: Emerging Technologies*, 55, pp. 261–272. doi: 10.1016/j.trc.2015.01.017.

She, B. *et al.* (2015) 'Weighted network Voronoi Diagrams for local spatial analysis', *Computers, Environment and Urban Systems*, 52, pp. 70–80. doi: 10.1016/j.compenvurbsys.2015.03.005.

Sherman, L. (1990) 'Police crackdowns: Initial and residual deterrence', *Crime and Justice,* 12, pp. 1–48.

Sherman, L. W. *et al.* (1998) 'Preventing Crime: What Works, What Doesn't, What's Promising', *National Institute of Justice, Office of Justice Programs*, (July), pp. 1–19. doi: 10.1111/an.1986.27.3.11.5.

Sherman, L. W. *et al.* (2014) 'An integrated theory of hot spots patrol strategy: Implementing prevention by scaling up and feeding back', *Journal of Contemporary Criminal Justice*, 30(2), pp. 95–122. doi: 10.1177/1043986214525082.

Shi, L. *et al.* (2015) 'Human mobility patterns in different communities: a mobile phone data-based social network approach', *Annals of GIS*, 21(1), pp. 15–26. doi: 10.1080/19475683.2014.992372.

Smith, S. L. and Rus, D. (2010) 'Multi-robot monitoring in dynamic environments with guaranteed currency of observations', in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, pp. 514–521. doi: 10.1109/CDC.2010.5717132.

Steenbeek, W. and Weisburd, D. (2016) 'Where the Action is in Crime? An Examination of Variability of Crime Across Different Spatial Units in The Hague, 2001–2009', *Journal of Quantitative Criminology*, 32(3), pp. 449–469. doi: 10.1007/s10940-015-9276-3.

Stranders, R. *et al.* (2013) 'Near-optimal continuous patrolling with teams of mobile information gathering agents', *Artificial Intelligence*, 195, pp. 63–105. doi: 10.1016/j.artint.2012.10.006.

Summers, L. and Johnson, S. D. (2017) 'Does the configuration of the street network influence where outdoor serious violence takes place? Using space syntax to test crime pattern theory', *Journal of Quantitative Criminology*, 33(2), pp. 397–420. doi: 10.1007/s10940-016-9306-9.

Sun, J. B. *et al.* (2011) 'Exploring space–time structure of human mobility in urban space', *Physica A: Statistical Mechanics and its Applications*, 390(5), pp. 929–942. doi: 10.1016/j.physa.2010.10.033.

Tambe, M. *et al.* (2013) 'Computational Game Theory for Security: Progress and Challenges'.

Tavares-Pereira, F. *et al.* (2007) 'Multiple criteria districting problems: The public transportation network pricing system of the Paris region', *Annals of Operations Research*, 154(1), pp. 69–92. doi: 10.1007/s10479-007-0181-5.

Tavares Pereira, F. *et al.* (2009) 'Comparing two territory partitions in districting problems: Indices and practical issues', *Socio-Economic Planning Sciences*, 43(1), pp. 72–88. doi: 10.1016/j.seps.2007.04.001.

The Metropolitan Police (2018) *2016 - 2017 crime statistics | The Met*. Available at: https://www.met.police.uk/stats-and-data/year-end-crime-statistics/ (Accessed: 30 January 2018).

Tsai, J. *et al.* (2010) 'Urban security: Game-theoretic resource allocation in networked physical domains', *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10),* pp. 881–886.

Wang, F. (2012) 'Why police and policing need GIS: an overview', *Annals of GIS*, 18(3), pp. 159–171. doi: 10.1080/19475683.2012.691900.

Wasserman, L. (2004) *All of Statistics: A Concise Course in Statistical Inference, The American Statistician*. New York, NY: Springer New York (Springer Texts in Statistics). doi: 10.1007/978-0-387-21736-9.

Weisburd, D. (2005) 'Hot spots policing experiments and criminal justice research: Lessons from the field', *Annals of the American Academy of Political and Social Science*, 599(May), pp. 220–245. doi: 10.1177/0002716205274597.

Willemse, E. J. and Joubert, J. W. (2012) 'Applying min–max k postmen problems to the routing of security guards', *Journal of the Operational Research Society*, 63(2), pp. 245–260. doi: 10.1057/jors.2011.26.

Wolfler Calvo, R. and Cordone, R. (2003) 'A heuristic approach to the overnight security service problem', *Computers and Operations Research*, 30(9), pp. 1269–1287. doi: 10.1016/S0305-0548(02)00070-9.

Xie, Z. and Yan, J. (2008) 'Kernel Density Estimation of traffic accidents in a network space', *Computers, Environment and Urban Systems*, 32(5), pp. 396–406. doi: 10.1016/j.compenvurbsys.2008.05.001.

Yamada, I. and Thill, J.-C. (2010) 'Local indicators of network-constrained clusters in spatial patterns represented by a link attribute', *Annals of the Association of American Geographers*. Taylor & Francis Group, 100(2), pp. 269–285. doi: 10.1080/00045600903550337.

Yin, L. and Wang, Z. (2016) 'Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery', *Applied Geography*, 76, pp. 147–153. doi: 10.1016/j.apgeog.2016.09.024.

Yin, Z. *et al.* (2012) 'TRUSTS: Scheduling randomized patrols for fare inspection in transit systems using game theory', *AI Magazine*, 33(4), pp. 59–72.

Zhang, Y. and Brown, D. (2014) 'Simulation optimization of police patrol district design using an adjusted simulated annealing approach', *Spring Simulation Multi-Conference*. Available at: http://ptl.sys.virginia.edu/ptl/sites/default/files/zhang2013sa.pdf (Accessed: 10 November 2014).

Zhang, Y. and Brown, D. E. (2013) 'Police patrol districting method and simulation evaluation using agent-based model & GIS', *Security Informatics*, 2(1), p. 7. doi: 10.1186/2190-8532-2-7.

Zoltners, A. A. and Sinha, P. (2005) 'Sales Territory Design: Thirty Years of Modeling and Implementation', *Marketing Science*, 24(3), pp. 313–331. doi: 10.2307/40056963.