Lett Spat Resour Sci (2008) 1: 159–169
DOI 10.1007/s12076-008-0015-3

ORIGINAL PAPER

Efficient identification of geographic restriction conditions in anti-covering location models using GIS

Alan T. Murray · Hyun Kim

Received: 30 May 2008 / Accepted: 30 September 2008 / Published online: 16 October 2008

Abstract Spatial separation in located services and activities is often essential. Examples include homeland security, military asset defense, impacts on the environment, franchise outlet location, and promoting public wellbeing. When planning and management is supported by mathematical modeling, a difficulty has been efficient representation of spatial separation conditions. This paper reviews an optimization model, the anti-covering location problem, used to support planning and management problems where spatial separation must be ensured between sited services/activities. An approach is presented for the efficient and effective identification and use of spatial separation conditions called cliques in this model based upon the use of a geographic information system (GIS). Results highlight the significance of the developed methodology in terms of computational requirements, tractability and effectiveness. This research enhances capabilities for addressing important practical planning problems.

Keywords Dispersion · Node packing · r-separation · Integer programming · Cliques

JEL Classification C61

1 Introduction

An important class of location models involves maximizing the total value of facilities sited while also ensuring spatial separation between located facilities. Such opti-

A.T. Murray (✉)
School of Geographical Sciences, Arizona State University, Tempe, AZ 85287-5302, USA
e-mail: atmurray@asu.edu

H. Kim
Department of Geography, University of South Florida, Tampa, FL 33620, USA
e-mail: hkim1@cas.usf.edu

mization problems are commonly referred to as anti-covering or node packing models, and have been essential in supporting forest management, nature reserve design, telecommunications equipment siting, defense, water conservation, social service provision, zoning policy development, franchise outlet location, cartographic design, and a range of other planning contexts (see Moon and Chaudhry 1984; Murray 1999; Goycoolea et al. 2005; Downs et al. 2008; Grubesic and Murray 2008). Further, they are related to minimum separation and dispersion models (see Erkut et al. 1996; Church and Murray 2008).

Given the broad applicability and use of the anti-covering location problem (ACLP), researchers have been challenged by the complexity and computational difficulty in solving this basic model. Considerable effort, therefore, continues in developing improved approaches for solving it exactly and approximately (Murray and Church 1997; Goycoolea et al. 2005; Chaudhry 2006; Cravo et al. 2008). For exact approaches, the mathematical structure of formalized constraints has constituted the critical determinant in whether problem applications can be solved optimally in practice. However, the identification and implementation of an effective set of constraints imposing spatial separation between located facilities remains an issue because of combinatorial complexity.

This paper presents an approach for identifying and implementing a set of spatial restriction constraints for use in anti-covering problem applications based upon the use of a geographic information system (GIS). This enables spatial structure to be utilized and exploited in constraint identification, promoting computational efficiency. The remainder of the paper is organized as follows. The next section presents the basic model formulation and discusses related research. Cliques and their use in constraints are then detailed, and an approach for efficiently identifying cliques is developed through the use of GIS. Application results highlighting the utility of this approach are given, followed by concluding comments.

2 Background

The anti-covering location problem (ACLP) was discretely formalized as an integer program in Moon and Chaudhry (1984) to support spatial analysis. However, it is related to the well known node packing, vertex packing, stable set, independent set and r-separation problems (Padberg 1973; Nemhauser and Trotter 1975; Nemhauser and Sigismondi 1992; Barahona et al. 1992; Erkut et al. 1996). Considerable research continues to be devoted to this problem, particularly with respect to its use in applied work, variations to constraint structure and general solvability (see Church and Murray 2008). To facilitate discussion, a linear-integer formulation of the ACLP is given following the hybrid constraint structure proposed in Murray and Church (1997). The following notation is introduced:

$i =$ index of areas,

$k =$ index of cliques,

$\Psi_k =$ set of areas in clique k,

Φ_i = set of areas in conflict with area i,

β_i = benefit of selecting area i,

$$X_i = \begin{cases} 1 & \text{if area } i \text{ is selected,} \\ 0 & \text{otherwise.} \end{cases}$$

With this notation, the premise of the ACLP is to select areas (or sites or nodes) for some intended use. Selection examples include forest planning where a management area is to be harvested (or not) as well as social service provision where a parcel of land is considered (or not) for locating a center, like a clinic, rehabilitation facility, halfway house, etc. Thus, the variable X_i indicates a decision regarding whether or not area i will be selected for the specified activity or service. The benefit, β_i, could represent any metric of relative value of an area. In forestry, as an example, the benefit is the economic return to be realized if an area is harvested. In contrast, the benefit may be the same, so all could effectively equal one.

An important notion in the ACLP formulation detailed here is a clique. A collection of areas that are all simultaneously in conflict with each other form a clique. The spatial separation requirement can therefore be tracked using the set of all cliques, or through the use of Φ_i. In either case, these sets identify all areas j that would conflict with, or be too close to, area i for additional facilities. With respect to Φ_i, this means that i and $j \in \Phi_i$ cannot simultaneously be selected. The formal ACLP model is as follows:

$$\text{Maximize} \sum_i \beta_i X_i \tag{1}$$

$$\text{Subject to} \sum_{i \in \Psi_k} X_i \leq 1 \quad \forall k, \tag{2}$$

$$|\Phi_i| X_i + \sum_{j \in \Phi_i} X_j \leq |\Phi_i| \quad \forall i, \tag{3}$$

$$X_i = \{0, 1\} \quad \forall i. \tag{4}$$

The objective of the ACLP (or node packing problem), (1), is to maximize the total weighted value of selected areas to accommodate the intended service or activity. Constraints (2) and (3) impose stipulated spatial proximity requirements between sited facilities. These conditions have been referred to as a hybrid constraint set in Murray and Church (1997) (see also Church and Murray 2008). Finally, constraints (4) impose integer restrictions on decision variables.

Of particular interest here is imposing spatial separation requirements between located (or selected) facilities/areas. In this formulation of the ACLP, spatial separation is ensured through the use of constraints (2) and (3). Constraints (2) structure a restriction within each clique k that at most one area could be selected. Constraints (3) structure a condition where if area i is selected, then no areas in conflict with area i could be selected. From a technical standpoint, it could be possible to use only constraints (2) or constraints (3) as theoretically all intended restrictions could be represented. Clique constraints, (2), have been proposed and/or relied upon in ACLP based models by Padberg (1973), Nemhauser and Trotter (1975), Jones et al. (1991), Nemhauser and Sigismondi (1992) and Barahona et al. (1992), among others. Clique are appealing in mathematical models because they are facet inducing, aiding in the solvability of integer programs. However, as noted by Nemhauser and Sigismondi (1992), there are possibly an exponential number of cliques in general cases, though Jones et al. (1991) detail a specific case when partial enumeration is feasible. Alternatively, the so called neighborhood constraints, (3), have been proposed and/or applied in ACLP based models by Moon and Chaudhry (1984), Torres-Rojo and Brodie (1990) and Murray and Church (1996), among others. Neighborhood constraints are appealing in practice because they are finite in number, requiring one constraint for each area. However, they are not facet inducing, making their use in practice problematic, as illustrated in Moon and Chaudhry (1984) and Murray and Church (1995). While potentially individually valid, Murray and Church (1997) demonstrated that there are benefits to using both constraint constructs, (2) and (3), in an ACLP based model.

A challenge remains, then, to identify and implement a meaningful hybrid constraint set in the ACLP. On one hand, there are an exponential number of potential cliques, so associated combinatorial complexity to identify them is significant. On the other hand, model solvability is paramount, so enhancing mathematical structure is key for addressing encountered problem applications. Thus, cliques are important for ACLP application because of their desirable facet inducing structure, yet due to the combinatorial complexity to identify them (combined with the fact that an exponential number exist) only a select finite set is feasible in practice.

3 Cliques and GIS

As suggested above, a clique is the condition where a set of areas are all simultaneously in conflict with each other. In the ACLP, this means that no two, or more, members of a clique k, Ψ_k, could be selected, because doing so would represent a violation. In a spatial context, there is a stipulated minimum separation between selected areas, so the clique reflects that set members are all within the minimum spatial separation requirement of each other. The "common conflict area" is the spatial manifestation of a clique.

A challenge in relying on cliques is efficiently identifying a good set, with "efficiently" and "good" being the operable words here. As noted previously, there are potentially an exponential number of cliques. For example, assume that areas A, B, C and D are all simultaneously in conflict with each other. There are 11 potential clique sets: $\{A, B\}$, $\{A, C\}$, $\{A, D\}$, $\{B, C\}$, $\{B, D\}$, $\{C, D\}$, $\{A, B, C\}$, $\{A, B, D\}$, $\{A, C, D\}$, $\{B, C, D\}$, and $\{A, B, C, D\}$. Further, some cliques may be better to use in the model. Thus, the idea is to find enough cliques to provide facet inducing structure in the model, but avoid enumeration or excessive computational search. Often a set of maximal cliques is noted, meaning that the cliques are the largest possible. However, their identification remains a difficult task as well. At any rate, with a good set of cliques, it is possible to supplement them with modified neighborhood constraints, as suggested in Murray and Church (1997), giving a complete and valid set

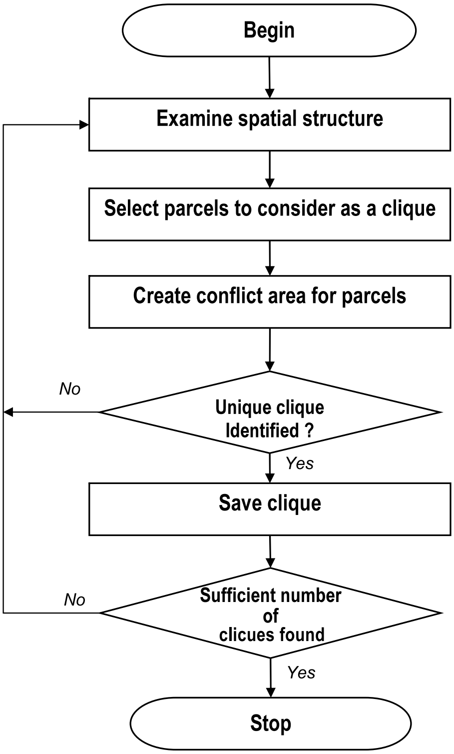

Fig. 1 Clique identification process

of constraints in the ACLP. The hope is that the ACLP will be solvable, but this is not guaranteed.

A geographic information system (GIS) can be used to exploit spatial structure to find cliques for the ACLP in order to enhance model solvability. An iterative process is suggested in Fig. 1 for identifying cliques, involving three major steps: (1) Examine spatial structure; (2) Select parcels to consider as a clique; and, (3) Create conflict areas for parcels. Each of these major steps is now discussed in more detail, highlighting the capabilities of GIS to provide efficient support.

The first major process identified in Fig. 1, *Examine spatial structure*, involves finding a collection of areas (or nodes or sites) having no current clique, thereby increasing the likelihood of deriving a unique maximal, or large size, clique. GIS is essential in this spatial search process because it enables areas with currently identified cliques in a region to be known, as well as where they are not. Specifically, the spatial proximity between identified cliques is examined. Therefore, it is possible to find the "most" isolated area where a unique clique might exist, as this is the area largest in size without a clique. The general evolution of common conflict areas, and therefore cliques, is illustrated in Fig. 2, showing the addition of common conflict areas in six iterations for this example. Initially, there are only parcels (Fig. 2a), then a common conflict area is found (Fig. 2b). This common conflict area represents a clique, or rather a collection of parcels that cannot simultaneously be selected given the imposed spatial separation distance (1320 ft in this case). The next step (Fig. 2c)

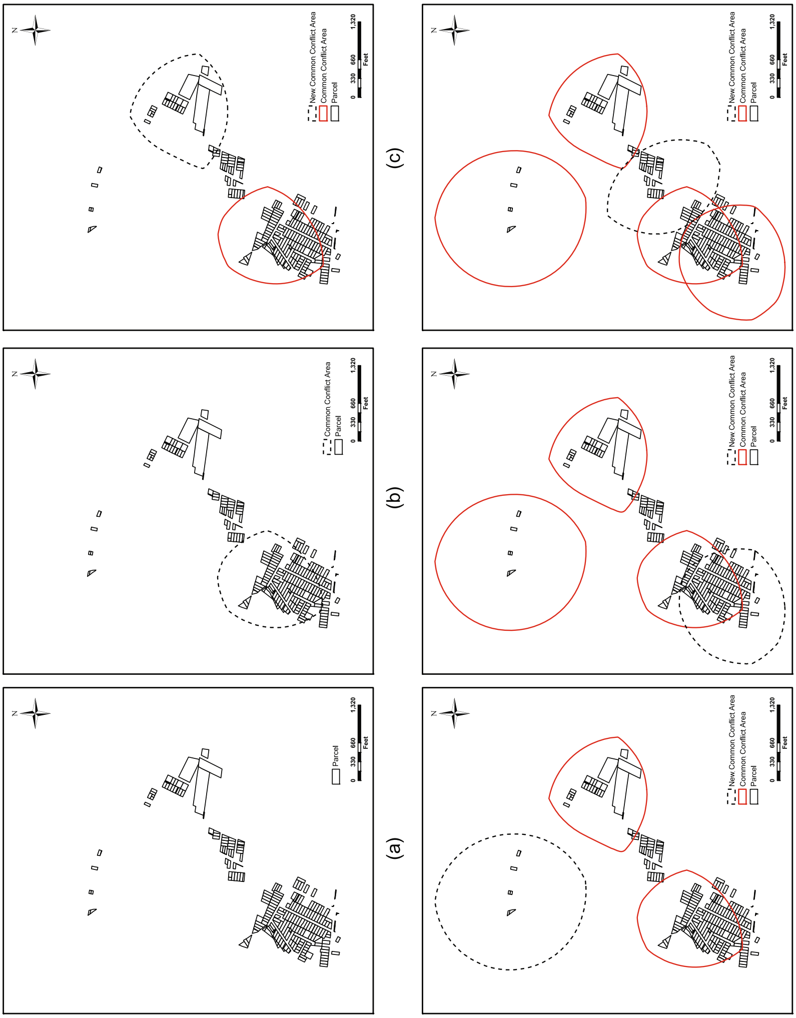

Fig. 2 Incremental addition of common conflict areas

finds another common conflict area away from the existing clique. This continues (Figs. 2d–2f) until five areas corresponding to cliques have been found in this case. In general, the process continues until user defined criteria are encountered regarding clique saturation and uniqueness.

Once spatial structure has been examined, it is then necessary to *Select parcels to consider as a clique* in the process outlined in Fig. 1. This amounts to identifying parcels, or areal units, to evaluate in clique formation. This selection is a function of the stipulated spatial separation standard. In our case, a distance of half the standard is utilized in selecting parcels to consider, but other criteria could be employed.[1] Thus, an irregular polygon centered on a location is produced (e.g., the polygon in Fig. 2b), and all parcels intersecting this polygon are selected. Basic functions in GIS enable creation of a polygon as well as selection of parcels that intersect the polygon. Again, identified are parcels to consider for a "maximal" clique as all are not necessarily within the spatial separation distance of each other, and therefore do not form a clique.

A final major step in Fig. 1 is to *Create conflict area for parcels*. This involves sub-set identification of the parcels found in the previous step, but doing so with the intent of finding the largest common conflict area (or largest clique). The approach is illustrated in Fig. 3. Given a set of parcels (Fig. 3a), the process requires that the associated conflict area for each parcel is found (Fig. 3b), and may be regular or irregular in shape. In GIS terms this is a regular or irregular buffer around a parcel, indicating that the selection of another parcel within this buffer would represent a spatial separation violation. Given all of the buffers, a common overlapping or intersecting area results, as shown for this example in Fig. 3c. Of course this assumes that the parcels in Fig. 3a are selected such that an overlapping area exists. The significance of the intersecting area is that we can use it to select parcels that in fact form a clique, as shown in Fig. 3d. Thus, we obtain the *common conflict area* by finding all parcels that are within the overlap/intersection area given in Fig. 3c.

4 Application results

The ACLP (anti-covering location problem) is applied here in the context of evaluating public policy regarding sex offender residency restrictions in Hamilton County, Ohio. Grubesic and Murray (2008) evaluate residency impacts of legislation oriented at convicted sex offenders. In particular, such laws aim to limit where offenders can live in relation to minors, and seek to ensure that inequitable concentrations of offenders in any local area do not result. Thus, communities are enacting laws the prohibit offenders from living within some pre-specified distance of each other. A typical standard is a quarter of a mile, or 1,320 ft. Grubesic and Murray (2008) use the ACLP to assess the maximum number of offenders a community should or could have, given housing conditions and other legislative mandates.

[1]This threshold was arrived at through experimentation, enabling the identification of relatively large cliques in reasonable computational effort.

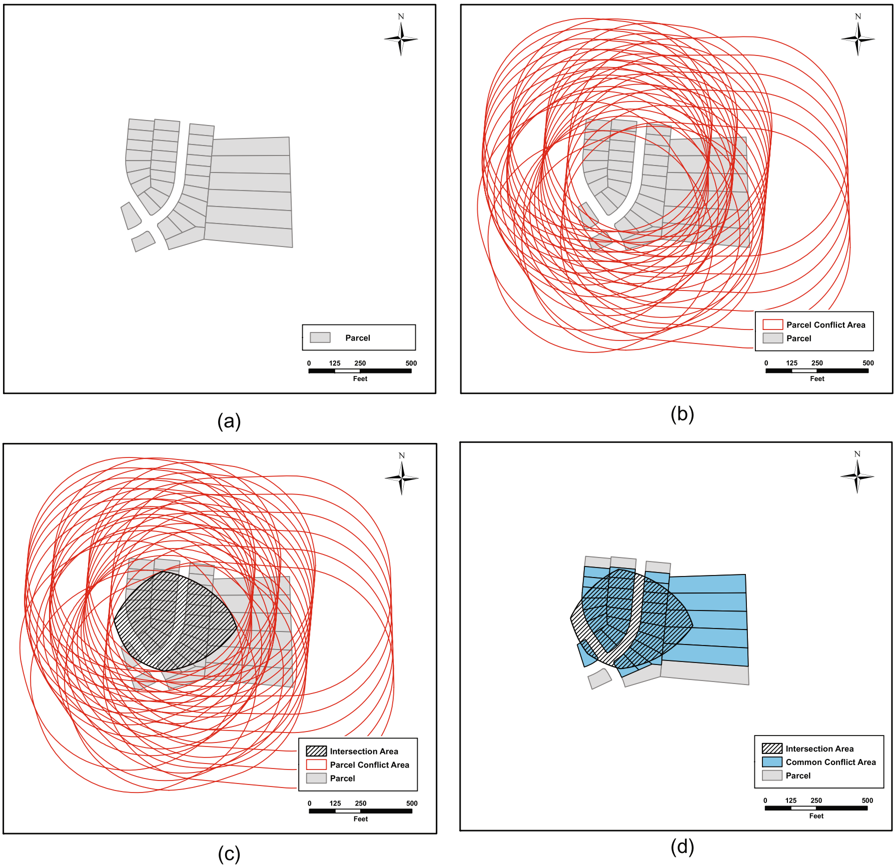

Fig. 3 Common conflict area creation

In the research reported here, two communities are examined, Reading and Norwood. The population of Reading is approximately 11,200 with some 12 convicted sex offenders as of June 2005 living in the community. Given limits on residential possibilities, including a ban on residing within restricted areas around schools and parks, there are 1,180 residentially zoned parcels in Reading. The population of Norwood is approximately 21,700 with 34 convicted sex offenders living within the city limits (again, as of June 2005). There are 3,252 possible parcels in Norwood for offenders to reside. The analysis undertaken here therefore assesses the implications of a 1,320 ft. minimum spatial separation limit between residences of convicted sex offenders in both communities.

The spatial information was managed and processed using ArcGIS (ver. 9.2) on an Intel Xeon based personal computer (2.99 GHz and 2 GB RAM) running Windows XP. A programming language within ArcGIS, ArcObjects, was used to produce text files of the associated ACLP. Cplex (ver. 10.1) was then utilized for model solution, with subsequent analysis on solutions conducted in ArcGIS.

Table 1 Anti-covering location problem results for Reading

Constraint construct	Objective	Time (s)	Branches	Iterations	Number of constraints	Number of variables
Pairwise	24	27.34	0	4,659	237,371	1,180
Neighborhood	24*	22,465.61	1,715,535	16,252,312	1,180	1,180
Hybrid	24	1.80	0	530	1,225	1,180

*Optimality not verified due to memory limit termination (639.6% optimality gap at termination)

The analysis focuses on the viability of solving the ACLP for the above cases (Reading and Norwood). Three problem instances of the each case are considered: pairwise, neighborhood and hybrid. Specifically, each instance represents a valid ACLP with a unique spatial separation constraint construct. The pairwise instance uses only constraints (2), and not constraints (3), where cliques are exactly two in size.[2] The neighborhood instance used only constraints (3), and not constraints (2). Finally, the hybrid instance is the model presented previously, (1)–(4), with the neighborhood constraints, (3), modified to eliminate all conditions represented in identified cliques, as suggested in Murray and Church (1997).[3]

The pre-processing to set up the model instances was comparable, requiring only minutes in ArcGIS for all three approaches. Model solution, however, differs in a number of ways for the various instances. For Reading, a summary of Cplex solution findings is reported in Table 1, giving details for each of the model structure instances. In all three instances there are 1,180 decision variables, one for each potential residential parcel. A significant difference is found in the number of necessary constraints, with the pairwise approach needing 237,371 to impose all spatial separation requirements in the ACLP. In contrast, the hybrid instance needed only 1,225 constraints (124 of which are cliques identified using the proposed method). In terms of supporting policy evaluation, the ACLP finds that based on a spatial separation standard of a quarter of a mile between the residences of offenders it would be possible to have at most 24 offenders living in the community. Thus, given this policy, the current number of offenders residing in the area (12) could possibly increase. Derivation of this finding using the ACLP, however, varies depending on the spatial constraint construct utilized. Solution via the neighborhood approach failed to confirm an optimal solution after 6 hours of processing time. That is, the optimal solution could not be verified due to memory limit termination, though in this case an optimal solution was found. In contrast, both the pairwise and hybrid approaches enabled the ACLP to be solved in less than a minute (27.34 seconds and 1.80 seconds, respectively).

In the case of Norwood, comparative results are reported in Table 2. The hybrid instance again required the fewest number of constraints, relying on 170 identified

[2]This is a special case of the clique having the following form: $X_i + X_j \leq 1\ \forall i, j \in \Phi_i$.

[3]The structure of the modified neighborhood constraint is: $|\hat{\Phi}_i| X_i + \sum_{j \in \hat{\Phi}_i} X_j \leq |\hat{\Phi}_i|\ \forall i | \hat{\Phi}_i \neq \emptyset$. By definition, $\hat{\Phi}_i \subseteq \Phi_i$ due to removal of set members already represented in identified cliques, Ψ_k. Such removal is possible when the proximity restriction between sites i and j, $j \in \Phi_i$, is already represented in an imposed clique set, Ψ_k, where $i, j \in \Psi_k$. This is effectively a type of constraint lifting.

Table 2 Anti-covering location problem results for Norwood

Constraint construct	Objective	Time (s)	Branches	Iterations	Number of constraints	Number of variables
Pairwise	32	185.83	10	7,126	1,075,558	3,252
Neighborhood	31*	401,349.53	652,000	27,328,617	3,252	3,252
Hybrid	32	8.36	46	1,066	2,911	3,252

*Optimality not verified due to memory limit termination (645.23% optimality gap at termination)

cliques using the proposed method. The general statistics and computational findings are similar to those found for Reading, with the exception that the neighborhood approach was not able to identify the optimal solution that 32 offenders could potentially reside in the community. Interestingly, with respect to the policy implications, the ACLP finds that the spatial separation standards would result in fewer than the current 34 offenders in Norwood. Using the neighborhood constraint construct would not enable this upper bound to be identified as the model could not be solved. In contrast, both the pairwise and hybrid instances required only modest computational effort (185.83 seconds and 8.36 seconds, respectively).

While the pairwise and hybrid approaches both seem to perform reasonably, it is worth highlighting some important implications. Obviously the number of constraints necessary using the pairwise approach is substantially higher in both problem instances (237,371 for Reading and 1,075,558 for Norwood), in contrast to the number of constraints needed using the hybrid approach (1,225 for Reading and 2,911 for Norwood). This means the pairwise approach requires nearly 20,000% more constraints for Reading and 37,000% more constraints for Norwood than the hybrid approach. The reality is that number of constraints remains an important element in the solvability of linear-integer models using exact methods, especially commercial optimization packages. In fact, a slightly older version of Cplex (8.1) could not actually solve the problems using the pairwise constraint construct, but could using the hybrid constraints.

5 Conclusions

The ability to identify and implement hybrid spatial separation constraints in the ACLP was shown to be possible through the use of GIS. In fact, GIS enables spatial structure to be taken into account in finding large cliques efficiently, thereby making the mathematical structure of the ACLP more amenable to exact solution using commercial optimization software. The application results highlight that the proposed GIS-based approach is computationally feasible, offering the capability to address large practical planning/policy problems.

Noted previously was that the pairwise and hybrid approaches enable the ACLP applications to be solved optimally in this research. However, the number of constraints using the pairwise approach is significant, leaving little optimism for addressing larger planning problems in practice. In order to address offender issues in Hamilton County, as an example, it would be necessary to consider some 45,000 residential parcels, each with an associated decision variable. The hybrid approach is likely the only viable way the ACLP could be solved exactly.

A final observation is that greater use and integration of GIS in spatial optimization is inevitable because of the ability to make better use of spatial information and geographic relationships. This was certainly the case with the dispersion model discussed in this paper.

References

Barahona, F., Weintraub, A., Epstein, R.: Habitat dispersion in forest planning and stable set problem. Oper. Res. **40**, S14–S21 (1992)

Chaudhry, S.S.: A genetic algorithm approach to solving the anti-covering location problem. Expert Syst. **23**, 251–257 (2006)

Church, R.L., Murray, A.T.: Business Site Selection, Location Analysis and GIS. Wiley, New York (2008)

Cravo, G.L., Ribeiro, G.M., Nogueira Lorena, L.A.: A greedy randomized adaptive search procedure for the point-feature cartographic label placement. Comput. Geosci. **34**, 373–386 (2008)

Downs, J.A., Gates, R.J., Murray, A.T.: Estimating carrying capacity for sandhill cranes using habitat suitability and spatial optimization models. Ecol. Model. **214**, 284–292 (2008)

Erkut, E., ReVelle, C., Ulkiisal, Y.: Integer-friendly formulations for the r-separation problem. Eur. J. Oper. Res. **92**, 342–351 (1996)

Goycoolea, M., Murray, A.T., Barahona, F., Epstein, R., Weintraub, A.: Harvest scheduling subject to maximum area restrictions: exploring exact approaches. Oper. Res. **53**, 490–500 (2005)

Grubesic, T.H., Murray, A.T.: Sex offender residency and spatial equity. Appl. Spatial Anal. Policy (2008, in press)

Jones, J.G., Meneghin, B.J., Kirby, M.W.: Formulating adjacency constraints in linear optimization models for scheduling projects in tactical planning. For. Sci. **37**, 1283–1297 (1991)

Moon, A.D., Chaudhry, S.: An analysis of network location problems with distance constraints. Manag. Sci. **30**, 290–307 (1984)

Murray, A.T.: Spatial restrictions in harvest scheduling. For. Sci. **45**, 45–52 (1999)

Murray, A.T., Church, R.L.: Measuring the efficacy of adjacency constraint structure in forest planning. Can. J. For. Res. **25**, 1416–1424 (1995)

Murray, A.T., Church, R.L.: Constructing and selecting adjacency constraints. INFOR **34**, 232–248 (1996)

Murray, A.T., Church, R.L.: Facets for node packing. Eur. J. Oper. Res. **101**, 598–608 (1997)

Nemhauser, G., Sigismondi, G.: A strong cutting plane/branch-and-bound algorithm for node packing. Oper. Res. **43**, 443–457 (1992)

Nemhauser, G.L., Trotter, L.E.: Vertex packings: structural properties and algorithms. Math. Program. **8**, 232–248 (1975)

Padberg, M.W.: On the facial structure of set packing polyhedra. Math. Program. **5**, 199–215 (1973)

Torres-Rojo, J.M., Brodie, J.D.: Adjacency constraints in harvest scheduling: an aggregation heuristic. Can. J. For. Res. **20**, 978–986 (1990)