

# Placing Green Bridges Optimally, with Habitats Inducing Cycles

Maike Herkenrath, Till Fluschnik, Francesco Grothe and Leon Kellerhals

Technische Universität Berlin, Faculty IV, Algorithmics and Computational Complexity, Germany  
 {till.fluschnik, leon.kellerhals}@tu-berlin.de, {herkenrath, f.grothe}@campus.tu-berlin.de

## Abstract

Choosing the placement of wildlife crossings (i.e., green bridges) to reconnect animal species' fragmented habitats is among the 17 goals towards sustainable development by the UN. We consider the following established model: Given a graph whose vertices represent the fragmented habitat areas and whose weighted edges represent possible green bridge locations, as well as the habitable vertex set for each species, find the cheapest set of edges such that each species' habitat is connected. We study this problem from a theoretical (algorithms and complexity) and an experimental perspective, while focusing on the case where habitats induce cycles. We prove that the NP-hardness persists in this case even if the graph structure is restricted. If the habitats additionally induce faces in plane graphs however, the problem becomes efficiently solvable. In our empirical evaluation we compare this algorithm as well as ILP formulations for more general variants and an approximation algorithm with another. Our evaluation underlines that each specialization is beneficial in terms of running time, whereas the approximation provides highly competitive solutions in practice.

## 1 Introduction

Habitat fragmentation due to human-made infrastructures like roads or train tracks, leading to wildlife-vehicle collisions, a severe threat not only to animals, up to impacting biodiversity [Bennett, 2017; Sawaya *et al.*, 2014], but also to humans. Installing wildlife crossings like bridges, tunnels [Woltz *et al.*, 2008], ropes [Goldingay and Taylor, 2017], et cetera (we refer to those as *green bridges* from here on) in combination with road fencing (so as to ensure that the green bridges are being used) allows a cost-efficient [Huijser *et al.*, 2009] reduction of wildlife-vehicle collisions by up to 85% [Huijser *et al.*, 2008]. In this paper, we study the problem of finding the right positions for green bridges that keeps the building cost at a minimum and ensures that every habitat is fully interconnected. We focus on those cases in which the structure of the habitats is very simple and study the prob-

lem from both a theoretical (algorithmics and computational complexity) as well as an experimental perspective.

We follow a model recently introduced by Fluschnik and Kellerhals [2021]. Herein, the modeled graph can be understood as path-based graph [Urban *et al.*, 2009; Galpern *et al.*, 2011]: a vertex corresponds to a part fragmented by human-made infrastructures subsuming habitat patches of diverse animal habitats, and any two vertices are connected by an edge if the corresponding patches can be connected by a green bridge. The edges are equipped with the costs of building the respective green bridge (possibly including fencing) in the respective area. The goal is to construct green bridges in a minimum-cost way such that in the graph spanned by the green bridges, the patches of each habitat form a connected component. Formally, we are concerned with the following.

**Problem 1.** 1-REACH GREEN BRIDGES PLACEMENT WITH COSTS (1-REACH GBP-C)

**Input:** An undirected graph  $G = (V, E)$  with edge costs  $c: E \rightarrow \mathbb{N}$ , a set  $\mathcal{H} = \{H_1, \dots, H_r\}$  of habitats with  $H_i \subseteq V$  and  $|H_i| \geq 2$  for all  $i \in \{1, \dots, r\}$ , an integer  $k \in \mathbb{N}_0$ .

**Question:** Is there a subset  $F \subseteq E$  with  $c(F) := \sum_{e \in F} c(e) \leq k$  such that for every  $i \in \{1, \dots, r\}$  it holds that  $H_i \subseteq V(G[F])$  and  $G[F][H_i]$  is a connected graph?

In accordance with Fluschnik and Kellerhals [2021], we denote by 1-REACH GREEN BRIDGES PLACEMENT (1-REACH GBP) the unit-cost version of 1-REACH GBP-C.

**Our contributions.** Our study focuses on habitats which induce small cycles. This is well motivated from practical as well as theoretical standpoints. Small habitats, in terms of size and limited structures (as to trees and cycles), appear more often for small mammals, amphibians, and reptiles, among which several species are at critical state [Hamer and McDonnell, 2008; Hennings and Soll, 2010]. From a theoretical point of view, since the problem is already NP-hard in quite restricted setups [Fluschnik and Kellerhals, 2021], it is canonical to study special cases such as restrictions on habitat and graph structure. As the problem is polynomial-time solvable if each habitat induces a tree (Observation 1), studying habitats inducing cycles are an obvious next step.

Our theoretical results are summarized in Table 1. We prove that 1-REACH GBP remains NP-hard even if each habitat induces a cycle, even of fixed length at least three. Most of our NP-hardness results hold even on restricted input

Habitat family	NP-hard, even if	Upper
$\{P_2, C_3\}$	$G$ is a clique	$P^{\dagger\nu\dagger}$
$\{C_3\}$	(no further restrictions)	$P^{\dagger\nu\ddot{s}}$
$\{P_2, C_\ell\}, \ell \in \mathbb{N}_{\geq 4}$	$\Delta \geq 4$ or if $G$ is planar*	$P^{\dagger\nu\dagger}$
$\{C_\ell\}, \ell \in \mathbb{N}_{\geq 4}$	$\Delta \geq 10$ or if $G$ is planar*	$P^{\dagger\nu\dagger}$
$\{P_2\} \cup \mathcal{C}$	$\Delta \geq 3$ and $G$ is planar	$P^{\dagger\nu\dagger}$
$\mathcal{C}$	$\Delta \geq 9$ and $G$ is planar	$P^{\dagger\nu\dagger}$

Table 1: Our NP-hardness (refer to Theorem 1) and upper bound results regarding our habitat families.  $\mathcal{C} = \bigcup_{\ell \in \mathbb{N}_{\geq 3}} \{C_\ell\}$  denotes the class of all cycles  $C_\ell$  of length  $\ell$ . “P” is short for “polynomial-time solvable”,  $\Delta$  for  $\Delta(G)$ . \* ( $\ell \neq 5$ )  $\dagger$  (if every edge is in at most two habitats (Thm. 2))  $\ddot{s}$  (if  $\Delta \leq 2$  (Obs. 2))  $\ddot{s}$  (if  $\Delta \leq 3$  (Cor. 1))

graphs, i.e., planar graphs of small maximum degree. On the positive side, we prove that for cycle-inducing habitats we can reduce the problem to maximum-weight matching in an (auxiliary) multi-hypergraph. From this we derive a polynomial-time special case: If every edge is shared by at most two habitats, we can reduce the problem to maximum-weight matching, a well-known polynomial-time solvable problem.

We perform an experimental evaluation of several algorithms, including the two mentioned above, the approximation algorithm given by Fluschnik and Kellerhals [2021], as well as an ILP formulation for the case of general (i.e., not necessarily inducing cycles) habitat structures. Our evaluation shows that each more specialized algorithm for the cycle-inducing habitats perform much better than the next more general one. Moreover, we show that the approximation algorithm is fast with very small loss in solution quality. Finally, we underline the connection between solution quality and running time on the one side, as well as the way of how habitats intersect on the other side.

**Further related work.** Ament *et al.* [2014] gives an informative overview on the topic. The placement of wildlife crossings is also studied with different approaches [Clevenger *et al.*, 2002; Downs *et al.*, 2014; Loraamm and Downs, 2016; Bastille-Rousseau *et al.*, 2018]. A related problem is STEINER FOREST (where we only need to connect habitats, without requesting connected induced graphs), which (and an extension of it) is studied from an algorithmic perspective [Lai *et al.*, 2011; Jordán and Schlotter, 2015].

## 2 Preliminaries

Let  $\mathbb{N} (\mathbb{N}_0)$  be the natural numbers without (with) zero. For a set  $X \subseteq \mathbb{N}_0$  and  $y \in \mathbb{N}_0$ , let  $X_{\geq y} := \{x \in X \mid x \geq y\}$ .

**Graph theory.** For a graph  $G = (V, E)$ , we also denote by  $V(G)$  and  $E(G)$  the vertex set  $V$  and edge set  $E$  of  $G$ , respectively. For an edge set  $E' \subseteq E$  and vertex set  $V' \subseteq V$ , we denote by  $G[E'] = (\bigcup_{e \in E'} e, E')$  and by  $G[V'] := (V', \{e \in E \mid e \subseteq V'\})$  the graph induced by  $E'$  and by  $V'$ , respectively. The graph  $G[E'][V']$  is the graph  $G'[V']$  with  $G' = G[E']$ . By  $\Delta(G)$  ( $\delta(G)$ ) we denote the maximum (minimum) vertex degree of  $G$ . Denote

by  $N_G(v) := \{w \in V(G) \mid \{v, w\} \in E(G)\}$  and  $N_G[v] := N_G(v) \cup \{v\}$  the open and closed neighborhood of  $v$  in  $G$ .

**Basic observations.** We can assume that every vertex in our graph is contained in a habitat. We say that an edge  $e$  is *shared* by two habitats  $H, H'$  if  $e \subseteq H \cap H'$ . We say that a set  $F' \subseteq E$  *satisfies* a habitat  $H \in \mathcal{H}$  if  $H \subseteq V(G[F'])$  and  $G[F'][H]$  is connected. We have the following.

**Observation 1 (★<sup>1</sup>).** 1-REACH GBP-C where each habitat induces a tree is solvable in  $O(|\mathcal{H}| \cdot |G|)$  time.

**Observation 2 (★).** 1-REACH GBP-C on graphs of maximum degree two is solvable in  $O(|\mathcal{H}| \cdot |G|)$  time.

## 3 Lower Bounds

In this section we show that the NP-hardness of 1-REACH GBP persists even if the habitats induce simple structures and the graphs are restricted. We prove the following.

**Theorem 1.** 1-REACH GBP is NP-hard even if:

- (i) each habitat induces a  $P_2$  or a  $C_3$  and  $G$  is a clique, or each habitat induces a  $C_\ell$  for any fixed  $\ell \geq 3$ .
- (ii) each habitat induces (a  $P_2$  or) a  $C_\ell$  for any fixed  $\ell \in \mathbb{N}_{\geq 4} \setminus \{5\}$  and  $G$  is planar.
- (iii) each habitat induces a  $P_2$  or a  $C_\ell$  for any fixed  $\ell \geq 4$  and  $\Delta(G) \geq 4$ , or each habitat induces a  $C_\ell$  for any fixed  $\ell \geq 4$  and  $\Delta(G) \geq 10$ .
- (iv) each habitat induces a  $P_2$  or a cycle,  $G$  is planar, and  $\Delta(G) \geq 3$ , or each habitat induces a cycle,  $G$  is planar, and  $\Delta(G) \geq 9$ .

For each case we provide a polynomial-time reduction from the following NP-hard [Garey and Johnson, 1979] problem.

**Problem 2.** CUBIC VERTEX COVER (CVC)

**Input:** An undirected, cubic graph  $G = (V, E)$  and an integer  $p \in \mathbb{N}_0$ .

**Question:** Is there a set  $V' \subseteq V$  with  $|V'| \leq p$  such that  $G[V \setminus V']$  contains no edge?

We first provide constructions for the four cases before presenting the (exemplary) correctness for case (i). The correctness proofs for cases (ii)–(iv) are deferred to the full version.

**Constructions.** We next provide the constructions for the base cases (i.e., small cycle lengths and habitats inducing  $P_2$ s) of Theorem 1(i)–(iv). The results can be extended by employing a central gadget which we call *crown*. The crown allows us to exclude  $P_2$ s from the habitat family and to extend cycle lengths all while preserving the planarity of the reductions. See Figure 1(a) for a crown.

**Definition 1.** Let  $G$  be a graph with two distinct vertices  $a, b \in V(G)$  and habitats  $\mathcal{H}$ . When we say we  $(p, q)$ -crown  $a$  and  $b$ , then we connect  $a$  and  $b$  by a so-called base path  $P^0$  of length  $p + 1$  and two crown-paths  $P^1, P^2$ , each of length  $q + 1$ , and add two crown-habitats  $H_i = V(P^0) \cup V(P^i)$  for  $i \in \{1, 2\}$ .

**Observation 3.** The minimum number of edges to satisfy both crown habitats of a  $(p, q)$ -crowning is  $p + 2q + 1$  (every edge from the base path and in each crown-path, all but one edge).

<sup>1</sup>Proofs of results marked with ★ are deferred to a full version, available at <https://arxiv.org/abs/2201.12273>.

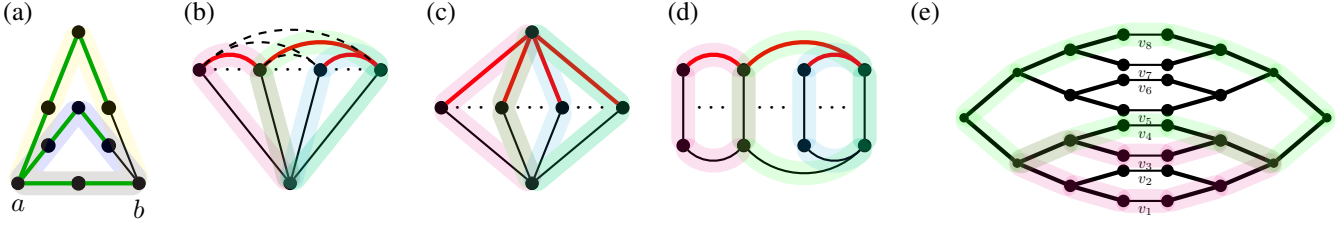


Figure 1: (a) A crowning and (b)–(e) four construction types. (a) The graph introduced by a (1, 3)-crowning with its two crown-habitats (blue and yellow). Thick (green) edges form a minimum-cardinality solution for the crown-habitats with 8 edges. (b)–(d) The construction are exemplified for a graph containing the edges  $\{v_1, v_i\}$ ,  $\{v_i, v_n\}$ , and  $\{v_j, v_n\}$ , but not the edges  $\{v_1, v_j\}$ ,  $\{v_1, v_n\}$ , and  $\{v_i, v_j\}$ , (b)+(c) Three habitats corresponding to the edges  $\{v_1, v_i\}$  (magenta),  $\{v_i, v_n\}$  (green), and  $\{v_j, v_n\}$  (blue) are depicted. Endpoints of thick (red) edges are candidates for  $(p, q)$ -crowning. In (b), dashed edge can be added without changing the correctness. (d) Thick (red) edges are candidates for subdivisions. (e) Habitats for the edges  $\{v_1, v_3\}$  and  $\{v_4, v_8\}$  are depicted. Each of the thick edges forms a habitat.

**Construction 1.** Let  $I = (G, p)$  be an instance of CVC with  $G = (V, E)$ ,  $V = \{v_1, \dots, v_n\}$ , and  $E = \{e_1, \dots, e_m\}$ . Construct an instance  $I' = (G', \mathcal{H}, k)$  as follows (see Figure 1(b)). Let  $G' = (V', E')$  with  $V' := V \cup \{x\}$  and  $E' := E \cup \bigcup_{i=1}^m \{\{x, v_i\}\}$ . Let  $\mathcal{H} = \{H_1, \dots, H_m\} \cup \{H'_1, \dots, H'_m\}$ , where  $H_i := e_i$  and  $H'_i := e_i \cup \{x\}$  for all  $i \in \{1, \dots, m\}$ . Let  $k := m + p$ .  $\diamond$

**Construction 2.** Let  $I = (G, p)$  be an instance of CVC with  $G = (V, E)$ ,  $V = \{v_1, \dots, v_n\}$ , and  $E = \{e_1, \dots, e_m\}$ . Construct an instance  $I' = (G', \mathcal{H}, k)$  as follows (see Figure 1(c)). Let  $G' = (V', E')$  with  $V' := V \cup \{x, y\}$  and  $E' := \bigcup_{i=1}^n \{\{x, v_i\}, \{y, v_i\}\}$ . Let  $\mathcal{H} = \{H_1, \dots, H_n\} \cup \{H'_1, \dots, H'_m\}$ , where  $H_i := \{v_i, x\}$  for all  $i \in \{1, \dots, n\}$  and  $H'_i := e_i \cup \{x, y\}$  for all  $i \in \{1, \dots, m\}$ . Let  $k := n + p$ .  $\diamond$

**Construction 3.** Let  $I = (G, p)$  be an instance of CVC with  $G = (V, E)$ ,  $V = \{v_1, \dots, v_n\}$ , and  $E = \{e_1, \dots, e_m\}$ . Construct an instance  $I' = (G', \mathcal{H}, k)$  as follows (see Figure 1(d)). Let  $G' = (V', E')$  with  $V' := V \cup V^*$  where  $V^* = \{v_1^*, \dots, v_n^*\}$  and  $E' := E \cup E^* \cup \bigcup_{i=1}^n \{\{v_i, v_i^*\}\}$ , where  $E^* = \{e_\ell^* = \{v_i^*, v_j^*\} \mid e_\ell = \{v_i, v_j\} \in E\}$ . Let  $\mathcal{H} = \{H_1, \dots, H_m\} \cup \{H_1^*, \dots, H_m^*\} \cup \{H'_1, \dots, H'_m\}$ , where  $H_i := e_i$ ,  $H_i^* := e_i^*$ , and  $H'_i := e_i \cup e_i^*$  for all  $i \in \{1, \dots, m\}$ . Let  $k := 2m + p$ .  $\diamond$

**Construction 4.** Let  $I = (G, p)$  an instance of CVC with  $G = (V, E)$ ,  $V = \{v_1, \dots, v_n\}$ , and  $E = \{e_1, \dots, e_m\}$ . W.l.o.g. we assume  $n$  to be a power of two as we can add isolated vertices. Construct an instance  $I' = (G', \mathcal{H}, k)$  as follows (see Figure 1(e)). Let  $T$  be a full binary tree of height  $\log_2(n)$  with root  $w$ .<sup>2</sup> Denote by  $w_1, \dots, w_n$  the leaves in the order provided by a depth-first search starting at  $w$ . Let  $T'$  be a copy of  $T$ , and denote by  $w'_i$  the copy of leaf  $w_i$ . Add  $T$  and  $T'$  to  $G'$  and for each  $i \in \{1, \dots, n\}$ , add the edge  $\{w_i, w'_i\}$ . For the construction of the habitats, denote the non-leaf vertices of  $T$  by  $w_C$  where  $C \subseteq \{1, \dots, n\}$  is the maximal subset of leaf indices in the subtree of  $T$  rooted at  $w_C$  (analogously for  $T'$ ). For each edge  $e \in E(T) \cup E(T')$ ,  $\mathcal{H}$  contains the habitat  $H_e = e$ . Now, for each edge  $e_\ell =$

<sup>2</sup>A full binary tree  $T$  with root  $w$  of height  $t$  is a tree with exactly  $2^t$  leaves each of distance exactly  $t$  to  $w$  and every inner node has exactly two children.

$\{v_i, v_j\} \in E$ ,  $\mathcal{H}$  contains the habitat  $H'_\ell := V(T_\ell) \cup V(T'_\ell)$ , where  $T_\ell$  is the subtree of  $T$  rooted at  $w_C$  with  $C$  being the smallest set with  $\{i, j\} \subseteq C$  (analogously for  $T'_\ell$ ). Finally, let  $k := |E(T)| + |E(T')| + p$ .  $\diamond$

**Correctness.** We next prove Theorem 1(i)–(iv), by using Constructions 1 to 4 and extending it with the crown (see Definition 1).

*Proof of Theorem 1(i).* Let  $I = (G, p)$  be an instance of CVC, and let  $I' = (G', \mathcal{H}, k)$  be an instance of 1-REACH GBP obtained from  $I$  using Construction 1. We claim that  $I$  is a **yes**-instance if and only if  $I'$  is a **yes**-instance.

( $\Rightarrow$ ) Let  $V' \subseteq V$  be a vertex cover of size at most  $p$ . We claim that  $F = E \cup \bigcup_{v_i \in V'} \{\{x, v_i\}\}$  is a solution to  $I'$ . Note that  $|F| \leq m + p$  and  $H_i \in F$  for all  $i \in \{1, \dots, m\}$ . Suppose the claim is false, that is, there is a habitat  $H'_\ell \in \mathcal{H}$  that is not connected. Since  $e_\ell = \{v_i, v_j\} \subseteq F$ , neither  $\{x, v_i\}$  nor  $\{x, v_j\}$  is in  $F$ . Hence,  $V' \cap e_\ell = \emptyset$ , a contradiction.

( $\Leftarrow$ ) Let  $F$  be a solution to  $I$ . We know that  $H_i \in F$  for all  $i \in \{1, \dots, m\}$ . We claim that  $V' = \{v_i \mid \{x, v_i\} \in F\}$  is a vertex cover of  $G$ . Note that  $|V'| \leq p$  since  $|F \setminus E| \leq p$ . Suppose the claim is false, that is, there is an edge  $e_\ell = \{v_i, v_j\} \in E$  with  $e_\ell \cap V' = \emptyset$ . By construction of  $V'$ , we have that each of  $\{x, v_i\}$  and  $\{x, v_j\}$  are not in  $F$ . Hence the habitat  $H'_\ell$  is not connected, a contradiction.

Note that adding any edge to  $G'$  does not change the correctness. For the NP-hardness for only  $C_\ell$ -habitats,  $\ell \geq 3$ , replace every edge  $e = \{v, w\}$  in  $E$  by a  $(\ell - 3, 1)$ -crowning and adjust  $\mathcal{H}$  and  $k$  accordingly.  $\square$

## 4 Upper Bounds

This section is devoted to instances of 1-REACH GBP-C in which every habitat induces a cycle. We will first show that this case can be reduced to the following problem.

**Problem 3.** MAXIMUM-WEIGHT HYPERGRAPH MATCHING (MWHM)

**Input:** A hypergraph  $G = (V, E)$  with edge weights  $w: E \rightarrow \mathbb{N}$ .

**Task:** Find a set  $M \subseteq E$  of maximum weight such that for all  $e, e' \in M$  holds that  $e \cap e' = \emptyset$ .

MWHM is NP-hard [Garey and Johnson, 1979], but if every hyperedge is of cardinality at most two, it is equivalent to

the well-known MAXIMUM-WEIGHT MATCHING (MWM) problem which is solvable in  $O(|V|(|E| + \log |V|))$  time [Gabow, 1990]. We make use of this to prove that some special cases of 1-REACH GBP-C are polynomial-time solvable.

#### 4.1 The General Case for Cycles

In this subsection we show the following.

**Proposition 1 (★).** *1-REACH GBP-C where every habitat induces a cycle can be decided by solving MWHM where the largest hyperedge is of size of the largest number of habitats intersecting in one edge.*

*Remark.* MWHM admits an ILP formulation with linearly many variables and constraints (used for our experiments):

$$\begin{aligned} \max \quad & \sum_{e \in E(H)} c(e) \cdot x_e \\ \text{s.t.} \quad & x_e \in \{0, 1\} \quad \forall e \in E(H) \\ & \sum_{e \in E(H): v \in e} x_e \leq 1 \quad \forall v \in V(H) \end{aligned} \quad (1)$$

Central for the translation to MWHM is the following graph.

**Definition 2** (Habitat graph). Let  $G = (V, E)$  be a graph with edge cost  $c: E \rightarrow \mathbb{N}$  and  $\mathcal{H}$  be a set of habitats each inducing a cycle. The multi-hypergraph  $B = (V_B, E_B)$  with edge weights  $w_B: E_B \rightarrow \mathbb{N}$  and bijection  $f: E \rightarrow E(B)$  are obtained as follows.  $B$  contains a vertex  $b_i$  for each habitat  $H_i$ . For every edge  $e \in E$  shared by at least two habitats  $H_{i_1}, \dots, H_{i_j}$ , add a hyperedge  $e' = \{b_{i_1}, \dots, b_{i_j}\}$  and set  $f(e) := e'$  and  $w_B(e') := c(e)$ . Finally, for every edge  $e \in E$  induced by only one habitat  $H_i$ , add a vertex  $b_e$  and the edge  $e' := \{b_e, b_i\}$ , and set  $f(e) := e'$  and  $w_B(e') := c(e)$ .

The following connection between 1-REACH GBP-C and MWM proves Proposition 1.

**Lemma 1.** *Let  $G = (V, E)$  be a graph with edge cost  $c: E \rightarrow \mathbb{N}$ , let  $\mathcal{H}$  be a set of habitats each inducing a cycle, and let  $B$  denote the habitat graph with edge weights  $w_B: E_B \rightarrow \mathbb{N}$  and function  $f$ . (i) If  $M$  is a matching in  $B$ , then  $G[E \setminus f^{-1}(M)][H]$  is connected for every  $H \in \mathcal{H}$ ; (ii) If  $G[F][H]$  is connected for every  $H \in \mathcal{H}$ , then  $E(B) \setminus f(F)$  is a matching in  $B$ .*

*Proof.* (i) Let  $M$  be a matching in  $B$  and let  $F := E \setminus f^{-1}(M)$ . Then for every  $b_i \in V(B)$  there is at most one edge in  $M$  that is incident with  $b_i$ . Thus, for every  $H \in \mathcal{H}$ ,  $|E(G[H]) \cap F| \geq |H| - 1$ , and hence  $G[F][H]$  is connected.

(ii) Let  $G[F][H]$  be connected for every  $H \in \mathcal{H}$  and let  $M := E(B) \setminus f(F)$ . Suppose there are two edges in  $M$  that are both incident to some  $b_i \in V(B)$ . Then there are two edges in  $E(G[H_i])$  not contained in  $F$ , and hence  $G[H_i][F]$  is not connected—a contradiction. Thus,  $M$  is a matching.  $\square$

*Remark.* We can simplify the habitat graph to a simple hypergraph: If there are multiple edges with the same vertex set, then it is enough to keep exactly one of maximum weight. We will make use of this in our experiments.

#### 4.2 Polynomial-Time Solvable Subcases

If every habitat induces a cycle and every edge is in at most two habitats, then the habitat graph is a hypergraph with edges of cardinality at most two. We have the following.

**Theorem 2.** *1-REACH GBP-C where every habitat induces a cycle is solvable in  $O(|V| \cdot |E| \cdot |\mathcal{H}|)$  time when every edge is in at most two habitats.*

We next present two special cases of 1-REACH GBP-C that become polynomial-time solvable due to the above.

**Habitats inducing faces.** Suppose our input graph is a plane graph (that is, a planar graph together with an crossing-free embedding into the plane). If every habitat induces a cycle which is the boundary of a face, then clearly every edge is shared by at most two habitats since every edge is incident with exactly two faces. Thus, we get the following.

**Corollary 1.** *1-REACH GBP-C where every habitat induces a cycle is solvable in  $O(|V| \cdot |E| \cdot |\mathcal{H}|)$  time on plane graphs when every habitat additionally induces a face.*

**Habitats inducing triangles in graphs of maximum degree three.** Suppose our input graph has maximum degree three and each habitat induces a triangle. Observe that every vertex of degree at most one cannot be contained in a habitat. Moreover, every degree-two vertex is contained in at most one habitat. For degree-three vertices we have the following.

**Lemma 2 (★).** *If a vertex  $v$  is contained in three habitats, then  $N_G[v]$  is a connected component isomorphic to a  $K_4$ .*

We immediately derive the following data reduction rule.

**Reduction Rule 1.** *If a vertex  $v$  is contained in three habitats, then delete  $N_G[v]$  and reduce  $k$  by the minimum cost of a solution for  $G[N_G[v]]$ .*

If the reduction rule is inapplicable, then every vertex is contained in at most two habitats. Consequently, every edge is shared by at most two habitats. We obtain the following.

**Corollary 2.** *1-REACH GBP on graphs of maximum degree three is solvable in  $O(|V| \cdot |E| \cdot |\mathcal{H}|)$  time when every habitat induces a triangle.*

## 5 Experiments

In this section, we present and discuss our experimental and empirical evaluation.

### 5.1 Data

**Graphs.** Our experiments are conducted on planar graphs only. We used data freely available by Open Street Maps (OSM). For each state of Germany (except for the city states Berlin, Bremen and Hamburg; abbreviated by ISO 3166 code), we set a bounding box and extracted the highways within. For each area encapsulated by highways, we created a vertex. Two vertices are connected by an edge whenever they are adjacent by means of a highway. Additionally, we generated five artificial graphs which are relative neighborhood graphs [Toussaint, 1980] of sets of  $500 + i \cdot 1125$  points in the plane, placed uniformly at random, with  $i \in \{0, \dots, 4\}$ . To all graphs we randomly assigned edge costs from  $\{1, \dots, 8\}$ . Table 2 provides an overview over some instances' properties.

	$ V $	$ E $	faces	$\delta$	$2 E / V $	$\Delta$
BB	129	293	165	2	4.543	11
BW	224	508	285	1	4.536	15
BY	284	651	368	1	4.585	11
HE	113	245	133	1	4.336	13
MV	74	159	86	2	4.297	9
NI	141	321	181	2	4.553	11
NW	455	1035	581	1	4.549	14
RP	138	295	158	1	4.275	13
SH	49	98	50	1	4	8
SL	65	133	69	1	4.092	12
SN	96	211	116	1	4.396	10
ST	133	302	170	1	4.541	10
TH	94	205	112	1	4.362	9
A500	500	629	130	1	2.516	4
A1625	1625	2038	414	1	2.508	4
A2750	2750	3458	709	1	2.515	4
A3875	3875	4871	997	1	2.514	4
A5000	5000	6315	1316	1	2.526	4

Table 2: Properties of our real-world and artificial graphs.

**Habitats.** We created multiple instances from every graph above by equipping them with different types and numbers of habitats. We created three types of instances: *face* instances, *cycle* instances, and *walk* instances. Given a plane graph  $G$ , a number  $r$  of habitats and, in the case of cycle and walk instances, a habitat size  $q$ , the instances were created as follows.

*Face instances:* Out of those faces of  $G$  that induce cycles, randomly choose  $r$  faces as habitats.

*Cycle instances:* List all induced cycles of length  $q \pm 1$ . For each such  $q$ , randomly choose  $r$  of the cycles as habitats.

*Walk instances:* Compute  $r$  self-avoiding random walks on  $q' = q \pm 1$  vertices, where  $q'$  is chosen uniformly at random. Add the vertices of each walk to a habitat.

For each instance type, for each graph above, for each  $r \in \{50, 100, 150, 200\}$ , for each  $q \in \{5, 7, \dots, 13\}$  in the case of cycle and walk instances, we generated 5 instances.

We remark that the real-world graphs MV, SH, and SL did not have sufficiently many cycles of length  $q \pm 1$  for  $q \in \{5, 7\}$ ,  $q \in \{5, 7, 9\}$ , and  $q = 5$ , respectively. In these cases, every cycle was chosen to be a habitat.

## 5.2 Algorithms

We implemented three exact solvers and one approximate solver. Two of the three exact solvers can be run only on some types of habitats

For the face instances we implemented the MWM-based algorithm (Corollary 1) which we will denote by  $\mathcal{A}_{\text{mwm}}$ . The habitat graph generation is implemented in Python 3. The matching is computed using Kolmogorov’s [Kolmogorov, 2009] C++ implementation of the Blossom V algorithm.

For the cycle instances generated the habitat graph in Python 3 and used Gurobi 9.5.0 to solve ILP formulation (1). We call the respective solver  $\mathcal{A}_{\text{mwhm}}$ .

The generic solver  $\mathcal{A}_{\text{gen}}$  can solve all instances of 1-REACH GBP-C and uses Gurobi 9.5.0 to solve the following

ILP formulation with an exponential number of constraints.

$$\begin{aligned} \min \quad & \sum_{e \in E(G)} c(e) \cdot x_e \\ \text{s.t.} \quad & x_e \in \{0, 1\} \quad \forall e \in E(G) \\ & \sum_{e \in \delta_H(S)} x_e \geq 1 \quad \forall \emptyset \neq S \subsetneq H, \forall H \in \mathcal{H} \end{aligned}$$

By  $\delta_H(S) := \{e \in E \mid e \cap S \neq \emptyset \wedge e \cap (H \setminus S) \neq \emptyset\}$ , we denote for a graph  $G = (V, E)$ , vertex set  $H \subseteq V$ , and subset  $\emptyset \neq S \subsetneq H$ , the set of edges between  $S$  and  $H \setminus S$ .

The approximate solver  $\mathcal{A}_{\text{apx}}$ , implemented in Python 3, is a weighted adaption of the  $\tilde{O}(r)$ -approximation algorithm for 1-REACH GBP given by Fluschnik and Kellerhals [2021]. Their algorithm computes for every habitat a spanning tree, and then combines the solutions. The weighted adaption has the same approximation guarantee. Further, for induced cycles it has an additive approximation guarantee that depends on the number of habitats and the maximum cost of any edge.

**Observation 4 (★).**  $\mathcal{A}_{\text{apx}}$  is an additive  $(r \cdot c_{\text{max}})$ -approximation for 1-REACH GBP-C with each habitat inducing a cycle, where  $c_{\text{max}} = \max_{e \in E} c(e)$ .

## 5.3 Results

We compared the implementations on machines equipped with an Intel Xeon W-2125 CPU and 256GB of RAM running Ubuntu 18.04. For the ILP-based solvers, we set a time limit of 30s for the solving time (not the build time). For 43 of the 100 artificial faces instances  $\mathcal{A}_{\text{gen}}$  was not able to compute any feasible solution. For all remaining instances,  $\mathcal{A}_{\text{gen}}$  provided an optimal solution in the given time limit.

All material to reproduce the results is available online.<sup>3</sup>

**Comparison of the optimal solvers.** Our experiments underline that each specialized solver outperforms the next less specialized solver (see Figure 2(a)). For instance, on real-world instances with face habitats,  $\mathcal{A}_{\text{mwm}}$  is on average 1.7 times faster than  $\mathcal{A}_{\text{mwhm}}$ , and on artificial instances with cycle habitats,  $\mathcal{A}_{\text{mwm}}$  is on average 82 times faster than  $\mathcal{A}_{\text{gen}}$  (see Table 4). Moreover,  $\mathcal{A}_{\text{mwm}}$  is 1.5 times faster than  $\mathcal{A}_{\text{mwhm}}$  on 80% of the face instances, and  $\mathcal{A}_{\text{mwhm}}$  is 10 times faster than  $\mathcal{A}_{\text{gen}}$  on 76% of the cycle instances.

**Approximate solver.** On real-world instances,  $\mathcal{A}_{\text{apx}}$  is 2 times faster than  $\mathcal{A}_{\text{mwm}}$  on face instances and 4 times faster than  $\mathcal{A}_{\text{mwhm}}$  on cycle instances, whereas the approximation factor never exceeds 1.348. The additive error is significantly better than the theoretical bound in Observation 4. See Table 3 and Figure 2(b). On the artificial instances, the approximation ratios are even better on average. This may be due to the fact that these instances are rather sparse (see Table 2).

*Intersection.* We also considered the intersection rate  $\lambda := |\sum_{H \in \mathcal{H}} |E(G[H])| / |\bigcup_{H \in \mathcal{H}} |E(G[H])|$ , which measures the average number of habitats in each edge (see the full version for a comparison of the solution quality and the speedup factor of  $\mathcal{A}_{\text{apx}}$ ). For  $\lambda \geq 10$  the approximation quality improves slightly in the real-world cycle and walk instances. As the habitats lie more dense, it is more likely for an

<sup>3</sup><https://git.tu-berlin.de/akt-public/gbp-experiments>

	quality ratio				additive ratio				running time ratio			
	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd
Faces <sub>Art</sub>	1	1.134	1.039	0.036	0	0.575	0.125	0.178	0.586	10.877	5.142	2.884
Cycles <sub>Art</sub>	1	1.119	1.038	0.026	0	0.724	0.234	0.151	0.583	19.957	8.265	4.694
Walk <sub>Art</sub>	1	1.044	1.008	0.009	0	0.31	0.053	0.056	2.835	4652.877	737.436	1157.331
Faces <sub>Real</sub>	1.018	1.243	1.141	0.049	0.042	0.326	0.201	0.055	0.802	5.586	2.256	1.015
Cycles <sub>Real</sub>	1.026	1.313	1.16	0.041	0.03	0.853	0.29	0.147	1.515	22.083	4.648	2.105
Walk <sub>Real</sub>	1.016	1.348	1.174	0.054	0.007	1.27	0.34	0.233	5.456	5797.882	1266.843	1815.414

Table 3: Summary of our results regarding  $\mathcal{A}_{\text{apx}}$ . The quality ratio is  $\text{cost}(\mathcal{A}_{\text{apx}})/\text{OPT}$ . The additive ratio is  $(\text{cost}(\mathcal{A}_{\text{apx}}) - \text{OPT}) / (d \cdot |\mathcal{H}|)$ , where  $d$  is the average weight of an edge in an optimal solution. The running time ratio is  $\text{time}(\text{BEST})/\text{time}(\mathcal{A}_{\text{apx}})$ , where BEST is the exact solver with the best overall running time on the instance.

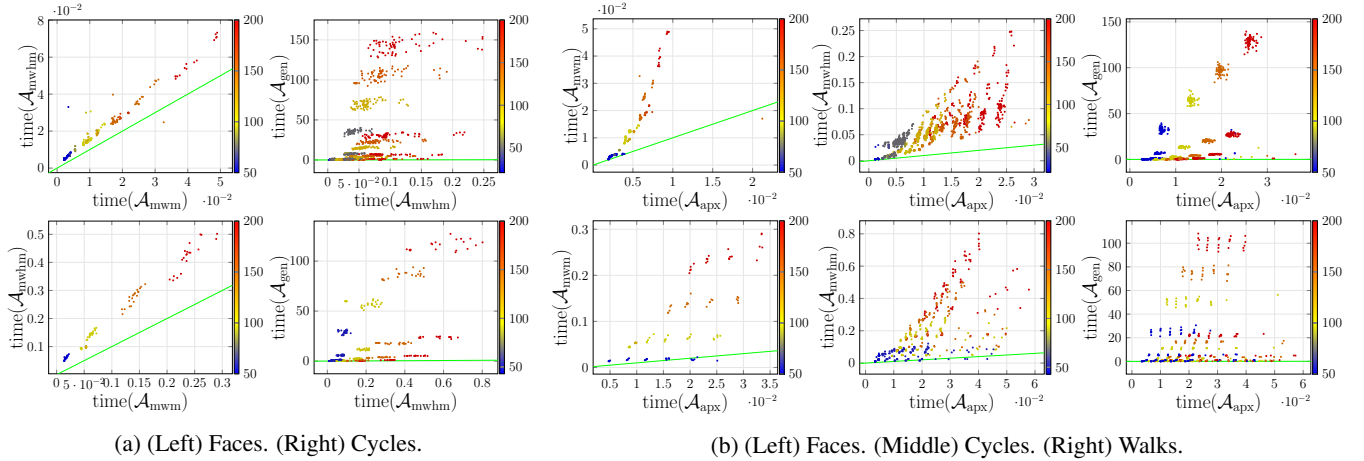


Figure 2: (Top) Real-World. (Bottom) Artificial. Data points are colored by the number of habitats. (a) Running times of the two best exact solvers on the respective instance type. (b) Running times of  $\mathcal{A}_{\text{apx}}$  against the best exact solver on the respective instance type.

	running time ratio			
	min	max	mean	sd
Faces <sub>Art</sub>	1.578	4.424	2.382	0.709
Faces <sub>Real</sub>	0.76	9.566	1.756	0.572
Cycles <sub>Art</sub>	1.35	645.87	82.933	137.271
Cycles <sub>Real</sub>	0.834	2266.312	315.597	483.523

Table 4: Summary of our results regarding the running time ratio of the respective two best exact algorithms.

edge to be in the solution. It thus seems plausible that  $\mathcal{A}_{\text{apx}}$  chooses fewer unnecessary edges.

As for the running time, the intersection rate seems only to have an effect on  $\mathcal{A}_{\text{mwhm}}$ . Especially on real-world cycle instances one can see that the running time quotient of  $\mathcal{A}_{\text{mwhm}}$  and  $\mathcal{A}_{\text{apx}}$  decreases with growing  $\lambda$ . This is likely due to the habitat graph  $B$ : If there are less edges that are in a unique habitat, then there are less vertices in  $B$  that are incident to a single cardinality-two edge, and hence less variables and constraints in (1). Note that one can deal with such vertices in a preprocessing routine as proposed i.e. by Koana *et al.* [2021]. This may significantly improve the running times of  $\mathcal{A}_{\text{mwhm}}$

and  $\mathcal{A}_{\text{mwhm}}$  for instances with low intersection rate.

## 6 Conclusion

While we prove that when every habitat induces a cycle, 1-REACH GBP-C remains NP-hard, even on planar graphs and graphs of small maximum degree, we provide an ILP-based solver that performs exceptionally well in our experiments. Moreover, when each habitat additionally induces a face in a given plane graph, the problem becomes solvable in polynomial time, with even faster practical running times. Lastly we observe that the approximation algorithm by Fluschnik and Kellerhals [2021] performs well in our experiments.

In a long version of this paper, we wish to address several theoretical and experimental tasks. On the experimental side, we plan to test our code on larger input instances. As mentioned in Section 5.3, we believe that the implementation of preprocessing routines may improve the running times of  $\mathcal{A}_{\text{mwhm}}$  and  $\mathcal{A}_{\text{mwhm}}$  significantly. On the theoretical side, we plan to settle the computational complexity of 1-REACH GBP-C for the following cases: for habitats in  $\{P_2, C_3\}$  in planar graphs or with constant maximum degree at least 4; for habitats in  $\{P_2, C_\ell\}$  in graphs of maximum degree at most 3. Future work may include restrictions on the habitats other than induced cycles.

## Acknowledgements

Till Fluschnik and Francesco Grothe were supported by the DFG project MATE (NI 369/17).

## References

- [Ament *et al.*, 2014] R. Ament, R. Callahan, M. McClure, M. Reuling, and G. Tabor. Wildlife connectivity: Fundamentals for conservation action, 2014. Center for Large Landscape Conservation: Bozeman, Montana. 2
- [Bastille-Rousseau *et al.*, 2018] Guillaume Bastille-Rousseau, Jake Wall, Iain Douglas-Hamilton, and George Wittemyer. Optimizing the positioning of wildlife crossing structures using gps telemetry. *Journal of Applied Ecology*, 55(4):2055–2063, 2018. 2
- [Bennett, 2017] Victoria J Bennett. Effects of road density and pattern on the conservation of species and biodiversity. *Current Landscape Ecology Reports*, 2(1):1–11, 2017. 1
- [Clevenger *et al.*, 2002] Anthony P Clevenger, Jack Wierzechowski, Bryan Chruszcz, and Kari Gunson. Gis-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. *Conservation Biology*, 16(2):503–514, 2002. 2
- [Downs *et al.*, 2014] Joni A. Downs, Mark W. Horner, Rebecca W. Loraamm, James Anderson, Hyun Kim, and Dave Onorato. Strategically locating wildlife crossing structures for Florida panthers using maximal covering approaches. *Transactions in GIS*, 18(1):46–65, 2014. 2
- [Fluschnik and Kellerhals, 2021] Till Fluschnik and Leon Kellerhals. Placing green bridges optimally, with a multivariate analysis. In *Proceedings of the 17th Conference on Computability in Europe – Connecting with Computability (CiE '21)*, pages 204–216, 2021. 1, 2, 5, 6
- [Gabow, 1990] Harold N. Gabow. Data structures for weighted matching and nearest common ancestors with linking. In *Proceedings of the 1st Symposium on Discrete Algorithms (SODA '90)*, pages 434–443, 1990. 4
- [Galpern *et al.*, 2011] Paul Galpern, Micheline Manseau, and Andrew Fall. Patch-based graphs of landscape connectivity: a guide to construction, analysis and application for conservation. *Biological Conservation*, 144(1):44–55, 2011. 1
- [Garey and Johnson, 1979] M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. 2, 3
- [Goldingay and Taylor, 2017] Ross L Goldingay and Brendan D Taylor. Can field trials improve the design of road-crossing structures for gliding mammals? *Ecological Research*, 32(5):743–749, 2017. 1
- [Hamer and McDonnell, 2008] Andrew J. Hamer and Mark J. McDonnell. Amphibian ecology and conservation in the urbanising world: A review. *Biological Conservation*, 141(10):2432–2449, 2008. 1
- [Hennings and Soll, 2010] Lori A Hennings and Jonathan Andrew Soll. *Wildlife corridors and permeability: a literature review*. Metro Sustainability Center, 2010. 1
- [Huijser *et al.*, 2008] Marcel P. Huijser, Pat T. McGowen, Amanda Hardy, Angela Kociolek, Anthony P. Clevenger, Dan Smith, and Robert J. Ament. Wildlife-vehicle collision reduction study: Report to congress. 2008. 1
- [Huijser *et al.*, 2009] Marcel P. Huijser, John W. Duffield, Anthony P. Clevenger, Robert J. Ament, and Pat T. McGowen. Cost-benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the united states and canada: a decision support tool. *Ecology and Society*, 14(2), 2009. 1
- [Jordán and Schlotter, 2015] Tibor Jordán and Ildikó Schlotter. Parameterized complexity of spare capacity allocation and the multicost steiner subgraph problem. *Journal of Discrete Algorithms*, 30:29–44, 2015. 2
- [Koana *et al.*, 2021] Tomohiro Koana, Viatcheslav Korenwein, André Nichterlein, Rolf Niedermeier, and Philipp Zschoche. Data reduction for maximum matching on real-world graphs: Theory and experiments. *ACM Journal of Experimental Algorithmics*, 26:3.1–3.30, 2021. 6
- [Kolmogorov, 2009] Vladimir Kolmogorov. Blossom V: a new implementation of a minimum cost perfect matching algorithm. *Mathematical Programming Computation*, 1(1):43–67, 2009. 5
- [Lai *et al.*, 2011] Katherine J. Lai, Carla P. Gomes, Michael K. Schwartz, Kevin S. McKelvey, David E. Calkin, and Claire A. Montgomery. The steiner multi-graph problem: Wildlife corridor design for multiple species. In *Proceedings of the 25th Conference on Artificial Intelligence (AAAI '11)*. AAAI Press, 2011. 2
- [Loraamm and Downs, 2016] Rebecca W. Loraamm and Joni A. Downs. A wildlife movement approach to optimally locate wildlife crossing structures. *International Journal of Geographical Information Science*, 30(1):74–88, 2016. 2
- [Sawaya *et al.*, 2014] Michael A. Sawaya, Steven T. Kalinowski, and Anthony P. Clevenger. Genetic connectivity for two bear species at wildlife crossing structures in banff national park. *Proceedings of the Royal Society B: Biological Sciences*, 281(1780):20131705, 2014. 1
- [Toussaint, 1980] Godfried T. Toussaint. The relative neighbourhood graph of a finite planar set. *Pattern Recognition*, 12(4):261–268, 1980. 4
- [Urban *et al.*, 2009] Dean L. Urban, Emily S. Minor, Eric A. Treml, and Robert S. Schick. Graph models of habitat mosaics. *Ecology Letters*, 12(3):260–273, 2009. 1
- [Woltz *et al.*, 2008] Hara W. Woltz, James P. Gibbs, and Peter K. Ducey. Road crossing structures for amphibians and reptiles: Informing design through behavioral analysis. *Biological Conservation*, 141(11):2745–2750, 2008. 1