HEURISTIC METHODS FOR PIPELINE NETWORK DESIGN

Christopher Yeates[1*], Cornelia Schmidt-Hattenberger[1], David Bruhn[1,2]
[1]Geoenergy Department, Geoforschungzentrum (GFZ), Potsdam, Germany
[2] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Netherlands
*cyeates91@gmail.com

Abstract

We showcase geospatial heuristic methods for network design and optimization. We propose and adapt graph algorithms to achieve optimal (or close to optimal) fluid transportation networks meeting quantifiable criteria (such as minimizing cost for example). Typically, these are used on pipeline infrastructure design, for CO_2 collection or H_2 distribution for example. The pipeline cost functions involved in the optimization depend on both pipeline length and a concave function of pipeline capacity. As such, discrete optimization methods are required. We have extended the tool to integrate other known aspects of network design. A sink placement algorithm can identify the minimum-cost storage location (and in parallel construct the rest of the a priori unknown network structure). The tools have finally been adapted to allow the inclusion of pre-existing pipeline infrastructure at a lower cost. They can then propose networks that prioritize planning along pre-existing pipeline routes.

Keywords: *Pipelines, transportation, optimization, graphs*

1. Introduction

German net-zero carbon emission targets of 2050 still require a variety of new technologies as many key sectors of industry currently don't have decarbonization pathways. To meet this challenge, carbon capture and storage of CO_2 originating from industrial sources represents a technologically mature and acheivable solution and its potential warrants a reevaluation in the wider German climate discussion [1]. The German case for Underground Carbon Storage is notbaly assisted by the large potential storage capacity idendified on- and off-shore [2,3].

While finding a capable, safe and acceptable storage sites is often the foremost subject of investigation in planning a large-scale storage operation, a key aspect of its success involves the judicious clustering of sources and transportation network design in a way that meets all the constraints (geological, social, technical) but also does so with the lowest cost possible.

For large, long-term CO_2 sources, pipeline transport is understood to be both safest and most economical method [4].

Due to specific technical pipeline requirements for CO_2 transport, the case for "from scratch" (i.e. new pipelines) networks is reasonable. This further opens the possibilities of network designs available.

To find minimum-cost networks with a broad range of system constraints, heuristic graph methods can be considered due to their intuitiveness, transparence, and adjustability. Heuristic methods try to make improvements by applying calculated small changes to the current best solution. They do not always result in optimal solutions but usually perform well and within reasonable time. Furthermore, exact methods (such as Branch & Bound [5]) often rely on having the best possible starter solution, usually obtained from a heuristic method.

We provide an overview of our contributions to the topic, through descriptions and rudimentary pseudo-code where possible. These include a new network topology (i.e., structure) optimization algorithm based on transferring edges of high valency nodes, that achieves (as far as we know) above state-of-the-art performance for finding minimum-cost pipeline networks. In particular for larger networks, it overcomes local minima and achieves optimal solutions in a significant number of cases within reasonable time.

In addition, we provide an optimal storage location algorithm, which attempts to identify the minimum-cost location to place a sink in a series of sources with an undefined network structure. While we use the example of a storage node, in a distribution network, this may take the role of a supply node. In a multi-level network, the algorithm could equivalently place an intermediate relay node optimally.

Finally, we provide adapted versions of our methods that integrate prior pipeline infrastructure or planning routes. In this way, priority is given to pipelines built on pre-existing pipeline routes, as a means of acknowledging the legal or social complexities of planning new pipeline routes.

Some overall examples are provided of real-world potential CO_2 collection networks. The point source emission volumes were established from taking the average of large emitters from the European Carbon Trading scheme registry for 2015-2018 [6]. Industrial CO_2 sources shown in Figure 1 are clustered in an infrastructure-aware way (detailed further down), and minimum-cost networks are calculated that place sink nodes optimally.

2. Technical contributions

2.1 A new topology optimization heuristic

Local network design heuristics make small modifications to initial solutions in order to find lower-cost configurations. Such methods are required as exhaustive solutions rapidly become impractical for large graphs. Indeed, Calyey's formula states that there exists n^{n-2} distinct spanning trees for n nodes. These heuristic methods often (but not exclusively) involve creating a cycle within the network and breaking the cycle in a different location. This simple transformation is named a local transformation. Heuristics then repeat such a procedure as soon as a better overall solution is found (for the case of first-descent heuristics). Sometimes, local network design heuristics get stuck in local minima. In this way, step-by-step local modifications of the network structure, in which one pipeline is replaced by another, do not lead to lower-cost solutions. Instead, multiple successive higher-cost jumps are required to find a lower-cost solution. These jumps are usually not permitted by local heuristics and require either random [7] or calculated moves [7] to other solutions. We describe a new metaheuristic, the High Valency Shuffle Metaheuristic, to guide local heuristics out of local minima by placing the edges of high valency nodes on their immediate neighbors and attempting a local heuristic on this tentative solution. Pseudocode is provided in Table 1:

Table 1: High Valency Shuffle Metaheuristic algorithm

```
Step 0:   Start from an initial local minimum
Step 1:   Initiate empty solution list
Step 2:   Identify all the nodes $n_{HV}$ with high valency
          (3 and above edges)
Step 3a:  For each node $n_{HVi}$ of these $n_{HV}$:
              - Identify closest nodes $n_C$
Step 3b:      For each node $n_{Ci}$ of these $n_C$:
                  - Transfer edges from $n_{HVi}$ to $n_{Ci}$
                  - Connect $n_{HVi}$ to $n_{Ci}$ if not already done
Step 3c:          If a cycle is detected in the graph:
                      For each of the edges in the cycle:
                          - Tentatively remove edge from cycle
                          - Run a lower-level local heuristic
                          - Add the solution to the solution list
                  Else:
                          - Run a lower-level local heuristic
                          - Add the solution to the solution list
Step 4:   Find minimal-cost solution from solution list
Step 5:   If this solution is better than current incumbent
          solution:
              - Set this new solution as incumbent
              - Restart algorithm from Step 1
          Else:
              - End algorithm and return incumbent
```

Detailed performance calculations and comparison to other literature solutions can be found elsewhere [8].

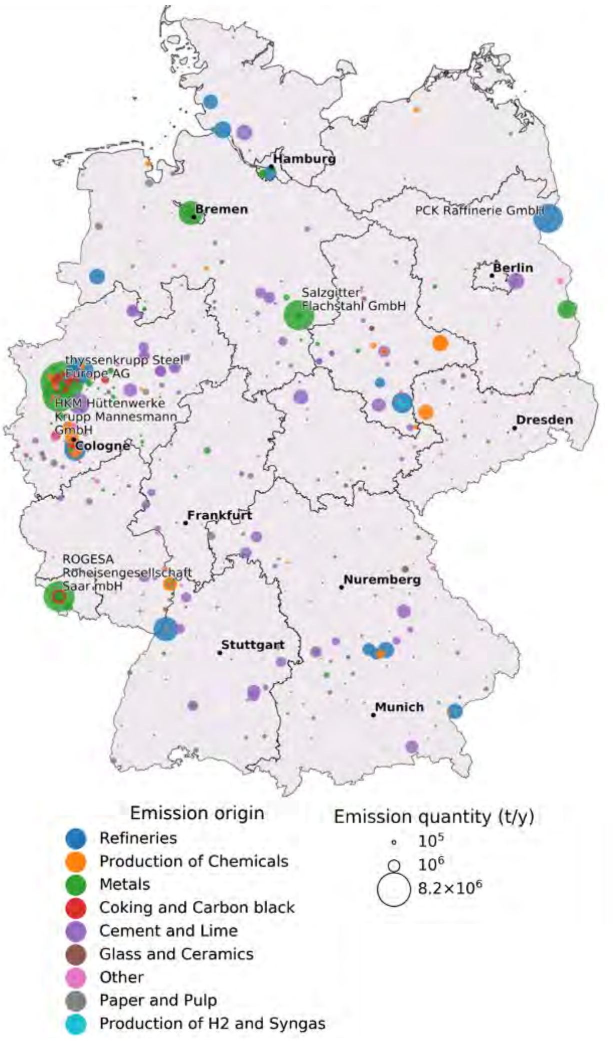

Figure 1: Map of Germany with emission origin, color-coded according to industrial category. The size of the circles displays the emission quantity.

2.2 A sink location algorithm

Locating the lowest cost potential sink location in an undefined network adds significant complexity to the network design problem. Starting from the observation that a sink node will only be connected to a certain subset of nodes (named the "housing nodes", forming a polygon P) of the graph, we can then place the sink optimally for the given subset via the weighted geometric median. Indeed, for a given set of node weights (here given by pipeline costs per unit length for the given pipeline widths of pipelines flowing towards the sink), there exists a unique location for the placement of a new node that minimizes the total cost of the new pipelines. The difficulty here lies in the fact that the weights upon the housing nodes are yet undefined as they depend on the specific network topology of the remaining network. A combined process of simultaneous topology optimization and minimal-cost sink location is the algorithmic solution we propose. Starting from an initial guess of housing polygon, the algorithm is based around a back-and-forth between optimizing the topology of the network (allowing edges to change, while the sink node location is fixed), then optimizing the placement of the sink node (allowing only the sink node placement to change, while the edges do not change). In this way, an initial bad guess of housing nodes can potentially lead to a final good solution as the algorithm can "drift" gradually through the network to a better configuration, finding lower-cost solutions at each step. Despite permitting the drift, local minima can still be reached, and the choice of initial guess determines to a reasonable degree the extent of the drift. Further research is currently underway to establish a model for good initial guesses for varying cost functions. Within the scope of this paper, we take a Delaunay triangulation of the network, and use the 10 triangles with the largest total amount of node capacity as initial input guesses.

Table 2: Optimal Sink Placement with drift algorithm

Step 1:	Make initial guess of nodes P connected to sink
Step 2:	Initialize network around P (without changing edges to sink) – with an adapted topology optimization heuristic
Step 3:	Place sink optimally for initialized weights
Step 4a:	***While*** a better solution is found (drift loop):
Step 4b:	Use topology heuristic (allowing all edges to change) – e.g. the algorithm given in 2.1
Step 4c:	Reposition sink optimally for new weights on sink (potentially new set of nodes P')

2.3 Integrating established pipeline routes

Planning new pipeline routes is suspected to be highly dependent on land use due to technical or legal obstacles to creating pipeline routes on new areas. In this way, it is supposed that greater priority should be given to stretches of land with pre-existing pipeline channels in proposing a network design. As mentioned previously, due to CO_2 pipeline specificities, this does not necessarily entail recycling pre-existing pipelines themselves (although the methods could permit it with case-specific cost functions) but instead globally lowering the cost of individual pipelines along known pipeline corridors by a given amount. Some further modifications are made to local topology optimization heuristics and to the optimal sink placement algorithm. To integrate known pipeline routes to a new local network of CO_2 sources, a considerable portion of the pre-existing pipeline network is considered for potential use. Nodes of the preexisting network are also added to the new network graph object, despite not carrying any supplementary capacity. The combined network graph is therefore a single entity composed of 3 different types of edges:

1. Pre-existing routes which carry no flow. These have no cost but can still be part of the combined network and be considered to re-route flow.
2. Pre-existing routes which carry flow. These have reduced cost.
3. New routes which carry flow. These have full cost.

Topology optimization heuristics are therefore simplified if the breaking and recombing of the network is done in a way that flow carrying routes only are considered for re-routing (edges 2 and 3), as other modifications have no effect on cost.

The sink placement algorithm can be modified to accommodate for the reduced cost of pre-existing pipeline routes. The calculation of the geometric median presupposes that all potential sink locations will require pipelines to a given housing node following the same cost function. The reduced cost function for pre-existing pipeline routes then invalidates this assumption. Indeed, the geometric median often places the sink node on the highest flow-carrying node of the housing polygon. The geometric median calculation of a later step might then have to choose between replacing the sink node via constructing new pipelines from each housing node or leaving its current "sub-optimal" configuration with a reduced cost along a preexisting pipeline route. In this scenario, replacing the node may be denied if it leads to a lower-cost solution. This therefore results in a local minimum which terminates the algorithm.

3. Results

In this section we give an example of a simple network constructed from combining a CO_2 source emission cluster from the South of Germany with a preexisting pipeline network and placing sinks in a minimum-cost configuration. The sources whose primary function is energy production (essentially coal plants) were omitted due to their planned phase out as well as sources that emit less than 50000 tons per year to decrease network complexity. The cost of a new pipeline is given by the function $LC^{0.6}$ where L is the pipeline length and C is the required pipeline capacity. A concave dependence of the capacity integrates the economy of scale of larger pipelines (0.6 is a typical exponent for CO_2 networks [9]). The pipeline network data was obtained from a recent data supplement that carefully aggregates multiple decentralized sources for the German national pipeline grid [10]. Source clustering was performed in a "network-aware" manner. A DBSCAN clustering algorithm was used and input with a precalculated distance matrix, whose values correspond to the pairwise shortest path along the pre-existing network between the sources considered. As the CO_2 sources are not connected to the pre-existing pipeline network, direct links are added between each source and its 2 closest network nodes. In our network design procedure, the relative cost factor of building a pipeline on a pre-existing route, C_{eq}, was set to 75%, 50%, 25% and 0% of the equivalent cost of constructing an identical pipeline elsewhere. In the 0% scenario, utilizing established pipeline routes then adds no cost to the network. This simple choice led to final networks that vary greatly in structure and give distinct minimum-cost sink locations. The scenario with no reduction in cost for routing on pre-existing pipelines (i.e., $C_{eq} = 100\ \%$) is also provided for comparison.

To fasten the optimization process, we decrease the complexity of the pre-existing pipeline network by only selecting pipelines within the convex hull of the sources, further extended by 30 km.

We show the sources with the considered established pipeline routes in Figure 2.

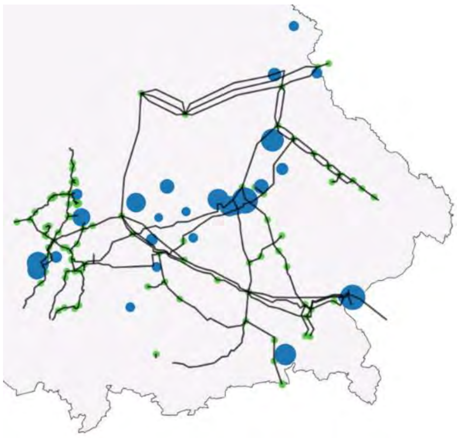

Figure 2: Considered CO2 sources (blue) and pre-existing pipeline infrastructure (black lines) and network nodes (green) for the South of Germany.

The final optimized networks for values of C_{eq} of 0%, 25%, 50 %, 75% and 100% are given in Figure 3 and 4. Calculation time was fast for all examples, taking roughly an hour to repeat the entire algorithm for the 10 different initial guesses of housing polygons. The intermediate topology optimization heuristic used in step 4b of the algorithm described in Table 2 was the metaheuristic described in Table 1.

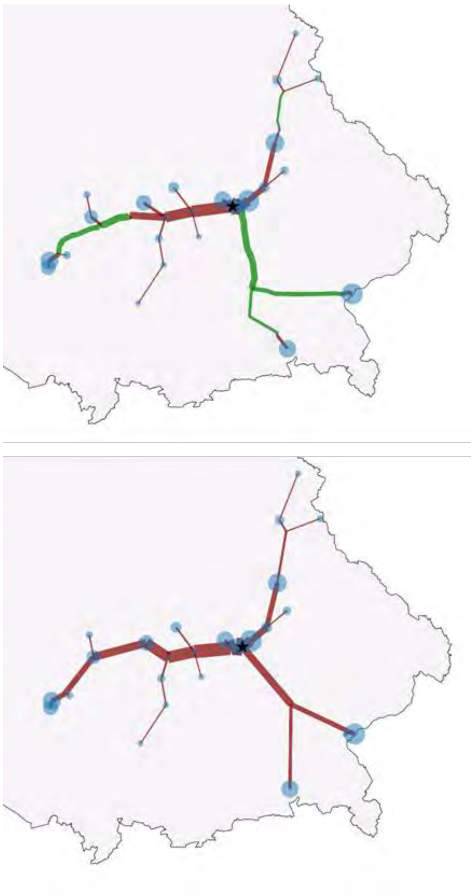

Figure 3: Minimum-cost networks and sink location utilizing pre-existing pipeline routes for C_{eq} of 75% (top) and 100% (bottom), the 100% case being the reference case without cost reduction for pipelines over pre-existing routes.

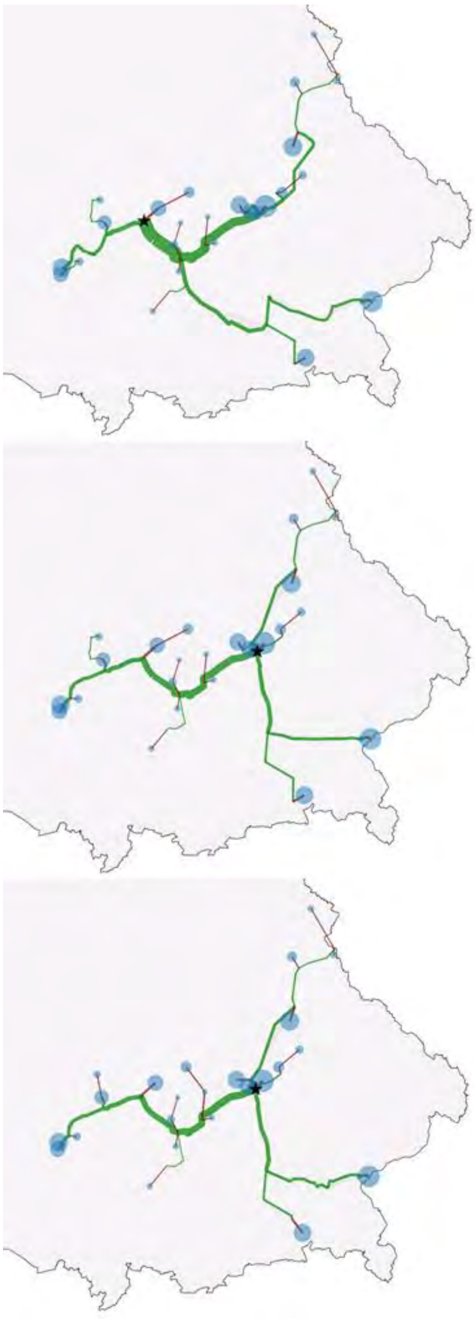

Figure 4: Minimum-cost networks and sink location utilizing pre-existing pipeline routes for C_{eq} of 0% (top) and 25% (middle) and 50% (bottom). The optimal sink location is shown as a black star in both plots.

In these figures, solution pipelines constructed upon pre-existing pipeline routes are shown in green whilst new routes are given in red. CO_2 sources are shown as blue circles with a size proportional to the emission volume. The pipeline thickness is proportional to the capacity required for flow material balance, which is furthermore unique if there is a match between total source and sink capacity.

We notably observe that as C_{eq} increases, overall length of the networks decreases as sinuous detours utilizing pre-existing routes no longer become economical. For lower values of C_{eq} seen in Fig. 4, the minimum-cost networks naturally tend to be built on a larger amount of pre-existing pipeline routes, whereas at the opposite extreme of full cost (Fig. 3, bottom), the minimum-cost network is constructed exclusively upon new routes (all pipelines shown in red).

The optimal locations of the sink (shown as black stars) appear also appear to change in each scenario, although chosen the locations are remarkably similar for the 25%, 50%, 75% and 100% C_{eq} parameters, situated within a dense cluster of 4 emitters and 4 preexisting network nodes. The chosen locations are identical for the C_{eq} = 25% and C_{eq} = 50%. The similarity of final optimal sink locations is observed despite a series of distinct initial guesses of housing nodes. The same final solution is often obtained for different initial guesses via the algorithm drift.

Details of the algorithm progression for each final lowest cost solution are given in Figure 5. We show how the overall network cost is decreased at each step of the algorithm, including notably multiple rounds of the drift loop for C_{eq} = 75% and C_{eq} = 25% scenarios.

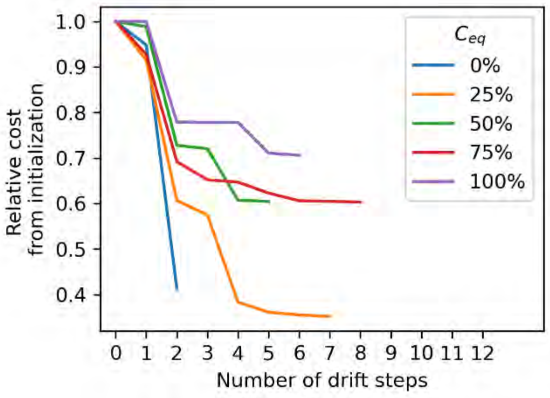

Figure 3: Algorithm progression (post-initialisation) for the final lowest-cost networks for each pre-existing pipeline route cost parameter.

4. Conclusion

Throughout this paper we explore a combination of methods enabling the design of minimum-cost pipeline networks considering features such as optimal sink placement or integration of prior pipeline routes. We give distinct network solutions established from clustering large industrial CO_2 sources of the South German region and utilize the prior regional gas pipeline network routes with varying cost of inclusion. While the networks shown here may not take into account other regional complexities, either geographical or geological (subsurface storage potential has nonetheless been established in the Southern region), the example demonstrates the functionality of the methods.

References

[1] Speech by Federal Chancellor Dr Angela Merkel at the Xth Petersberg Climate Dialogue, Berlin, 14 May 2019: https://www.bundesregierung.de/breg-en/news/speech-by-federal-chancellor-dr-angela-merkel-at-the-xth-petersberg-climate-dialogue-berlin-14-may-2019-1612592

[2] Müller, C & Reinhold, K (2011). Geologische Charakterisierung tiefliegender Speicher- und Barrierehorizonte in Deutschland – Speicher-Kataster Deutschland. – Schriftenr. dt. Ges. Geowiss., Vol. 74; Hannover. (ISSN 1860-1782)

[3] Höller, S, & Viebahn, P (2011). Assessment of CO2 storage capacity in geological formations of Germany and Northern Europe. *Energy Procedia*,*4*:4897-4904.

[4] Peletiri, SP; Rahmanian, N; Mujtaba, IM (2018). CO_2 Pipeline Design: A Review. *Energies*, *11*, 2184. https://doi.org/10.3390/en11092184

[5] Lawler EL, Wood, DE (1966). "Branch-and-Bound Methods: A Survey," *Operations Research*, INFORMS, vol. 14(4):699-719

[6] German Environment Federal Office (2019). Installations covered by ETS Germany in 2018, Deutsche Emissionshandelsstelle. Available at: https://www.dehst.de/SharedDocs/downloads/EN/installation_lists/2018.pdf (Accessed: 2 September 2020)

[7] Brimberg J, Hansen P, Lin K, Mladenovi N, Breton M, Brimberg, J (2003). An oil pipeline design problem. *Operations Research*, 51(2):228–239. https://doi.org/10.1287/opre.51.2.228.12786

[8] Yeates, C., Schmidt-Hattenberger, C, Weinzierl, W, & Bruhn, D (2020, December 17). Comparison of heuristic methods for achieving minimum-cost capacitated networks with a new metaheuristic based on node valency. https://doi.org/10.31224/osf.io/eqczh

[9] Kazmierczak T, Brandsma R, Neele F, Hendriks C (2009). Algorithm to create a CCS low-cost pipeline network. Energy Procedia. 1(1):1617–1623. https://doi.org/10.1016/j.egypro.2009.01.212

[10] Kunz, F, Kendziorski, M, Schill, W, Weibezahn, J, Zepter, JM, Hirschhausen, C, Hauser, P, Zech, M, Möst, D, Heidari, S, Felten, B, & Weber, C (2017). Electricity, heat, and gas sector data for modeling the German system.