Analytical Mapping of Academic Research on *Hubs*

Imane Essaadi[1], Vincent Giard[1,2], Bernard Grabot[1,3]

1. Mohammed VI Polytechnic University – EMINES School of Industrial Management, Benguerir, Maroc

2. Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, 75016, Paris, France

3. Toulouse University, LGP/INP-ENIT

imane.essaadi@emines.um6p.ma, Vincent.giard@dauphine.fr, bernard.grabot@enit.fr

Abstract: *Hub location and design issues as well as hub use have attracted much academic interest over the last two decades. A review of literature shows the diversity of approaches and of underlying assumptions. This paper proposes a mapping of these aspects to draw up a general analytical grid for these papers. The idea is to help identify the relative positioning of the research papers and to highlight the contingency of each one of them, as their authors rarely set forth all of their assumptions. Our analysis also serves to identify potentially valuable research themes. Our grid is structured into nine dimensions: scope of the paper, decision-making framework, hub features, logistics network structure, transport organization, physical data, economic evaluation system, decision evaluation criteria, application field.*

Key words: *hubs, logistics network, analytical model, typology*

I. Introduction

For more than 25 years, the question of *hub* localization and design and of their use has been dealt with in many research papers belonging to a variety of fields (operational research, geography, telecommunications/internet, transport, industrial economy). Such academic interest reflects the fast growth of the transportation industry of goods, passengers and information in what has become a highly competitive environment where players are constantly seeking efficiency gains and improved efficacy. Direct transport between a starting point and a destination point often generates a host of small flows, and only makes sense under rarely met conditions of volume and frequency. The opposite option, involving concatenation of elementary transportations between an initial sender and a final addressee defining a type of circulation, may be plotted on a graph. In such a graph, at intermediary node level (end nodes being called *spokes* in specialized literature), a series of transport batch *rebundling* operations are performed to achieve sufficient massification to optimize transportation cost and to increase pick up frequency at the cost of multiplying handling operations

In this context, hubs may be seen as specific network nodes among other nodes. Qualifying hubs as nodes, however, is not a straightforward: objective operation as setting up a hub to achieve node flow consolidation implies flow concentration to, or from, the hub by eliminating other possible network connections. One therefore notes a strong interdependency between the definition of *hubs* and that of the relevant transportation network. This conceptual ambiguity is rarely pointed out in the literature.

This paper proposes an analytical grid that enables to better characterize scientific papers on hub network design while identifying their assumptions and relaxed constraints. Furthermore, this grid reveals some hub network design problems overlooked in the literature with the aim of bridging the gap between real-life problems and academic research. In particular, our analytical grid addresses the hub network design issue of variants and extensions. Moreover, our grid emphasizes the ambiguity of some recurring concepts (e.g. demand related to terminal node) and points to the limitations of some papers and/or the debatable character of some of their conclusions.

II. Analytical Grid

We analyzed a large sample of papers to draw up an analytical map structured into nine dimensions designed to highlight the multiple (often implicit) assumptions of these papers and to help circumscribe our intended future research on hubs. Our documentary research started with papers published over the last twenty years using a number of key words (*Hub-and-spoke network design problem; hub location-allocation problem; intermodal hub location problem; intermodal hub-and-spoke network design; hub network design problem; hub location problem in transportation network; many-to-many hub location routing problem*), which led to a pre-selection of some 150 papers, including seven literature surveys.

The typologies proposed in these survey articles (Alizadeh (2009), Daskin and al. (2005), Melo and al. (2009), Alumur and Kara (2008), Campbell and al. (2002), Campbell and O'Kelly (2012), Farahani and al. (2013), Kara and Taner (2011), O'Kelly and Bryan (1998), O'Kelly and Miller (1994)) were enriched by additional considerations suggested by other papers (as we kept reviewing the material) as well as by reference to standard analytical grids applied in industrial management. Each characteristic dimension of a paper comprises several items, not all of which are mutually exclusive: there is a certain degree of correlation between items belonging to other dimensions and we retained an open-ended approach to the number of sub-items as we feel that the diversity of angles enhances meaning. The order of presentation of our analytical dimensions is somewhat arbitrary and we rather focus on their complementarity.

II.1 Scientific scope of the papers

Papers generally aim to expand the body of knowledge of the scientific community. They may be either purely theoretical and targeted at academics or have a practical application. The scope of a paper may be classified into three non-exclusive dimensions:

- **Description-oriented of complex situations**. Some papers analyze real life cases considered as representative in order to highlight actual problems deemed to be ill-understood by the scientific community. In this framework, the items of scope typology (section 9) are relevant.
- **Model-oriented.** The complexity of a problem under review, to do with the combinations of features that are less restrictive than those usually looked at, may lead to reformulating a problem (mathematical programming…), towards generally prescriptive ends (decision-making). This model-orientation may be backed by a case study deemed to be representative and followed by a straightforward numerical example (mandatory in certain reviews), to help readers grasp the proposed relations. One may add that a paper’s model could be equivalent to a previously proposed one though using a different formal approach to facilitate numerical resolution
- **Numerically-oriented.** A highly circumscribed class of problems may lead to proposing a numerical resolution technique that is more efficient than previous solutions or to show the limitations of an existing method.

II.2 Decision-making framework

The decision-making framework, a key focus in management, deals with the type of decision and the stakeholders concerned. The decision-making framework is a key defining dimension to assess the scope of model-based prescriptive papers. The framework tends to be implicit but we believe that explicitness is important to understand the contingent nature of a model.

Nature of the decision. For the last fifty years, a distinction is made between *strategic*, *tactical* and *operational* decisions. In our context, strategic decisions have to do with sizing and design of a logistics network, and this refers to a wide range of possible decisions as described under sections 3, 4 and 5. Decisions rely on factual information (section 6) and, very often, on an evaluation system (section 7). The tactical and operational decisions tend to focus on the use made of logistics networks and on the opportunity of recourse to additional resources to satisfy demand. Note that the assumptions of these decisions are based on choice made on the strategic level and conversely, strategic decisions lay on strong assumptions of network operation studied at tactical/operational level. The high degree of interdependence between these decisions, which are impossible to take simultaneously, represents one of the most difficult problems in designing a logistics network. The hierarchy of approaches used in many papers plays a major role in design as optimization made on local scale irrespective of constraints resulting from upper and/or lower decision levels ultimately leads to inefficient decisions. Thus, a consistency check of data is needed as the decisions yielded by a model implicitly rely on parameters and assumptions stemming from choices made in the context of the resolution of a model at a particular management level. For example the economic design of a logistics network depends on transportation and handling costs that in turn depend on the organization defined at tactical or operational level.

Stakeholders. Logistics network organizational models and supporting decision evaluation systems impact the *stakeholders* of the transportation system. The system is often implicitly considered as able to be managed centrally (*single stakeholder*), whether or not the resources belong to a single legal entity. The paper's proposed decisions are seen as being globally the best for all the logistics chain actors. However, this view is often misleading: while the actors collaborate they also seek to optimize their share of the added value of transportation services by leveraging pricing power (a distinct approach from the 'cost of exchanged services' that prevails in a single stakeholder situation). The *multiple stakeholder* approach is far more complex to figure out, particularly in connection with finding shared economic criteria to strike an acceptable compromise for all. Failing such compromise, the final say will be that of an actor able to impose his or her choices. Note that stakeholders also include the customers and the owners of network companies as well as the economic players who are impacted by the benefits and nuisance induced by network activities.

II.3 Hub Features

Hubs may be described with reference to four dimensions.

- **Number and location of hubs**. Hubs may be ***single*** or ***multiple*** (the latter is the most frequent case). The number may be ***predetermined*** or ***to be determined*** (involving a strategic decision). In the case of multiple hubs to be determined, the most frequent case is that of a choice in a finite set of eligible hubs of known geographic location; this allocation process is called *discrete localization*. Some papers deal with this problem by including the geographic coordinates of hubs as variables for orders, which is called *continuous localization*.
- **Hub capacity**. The ***definition*** of hub capacity generally focuses on storage volume or space. In reality, this concept is far more difficult to define than is generally acknowledged. Of course, physical space enters in the definition of capacity but other relevant factors for duration of warehousing stay include: frequency and regularity of incoming and outgoing flow, storage organization, human and material resources, flow management rules, etc. Additionally, capacity ***sizing*** may vary: it could be *fixed*, *extensible* relative to a basic known capacity or *to be determined*; in the latter case, the decision is strategic whereas in the former one it may be either strategic (investment) or tactical (ad hoc recourse to additional resources).
- **Hub types.** Hubs may play different roles depending on variant parameters such as type of management, location and organization. Concerning management types, hubs may be **private**

(managed by one or several private stakeholders (such as industrial firms or logistics service providers who design a cost-effective network for their own profit), or **public (**established and managed by public authorities, local or national government agencies, with long-term investment return objectives. In the latter case, hubs are meant to boost economic growth and attract sustainable international investment in local industry. The public authorities take user benefits and needs into consideration in designing a cost-effective hub network. The aim is to lower investment and operational costs and help investors to seize local and international market opportunities to drive national socio-economic growth. **Public-private** or semi-public/semi-private hubs are managed by multiple stakeholders resulting from public-private partnerships where responsibilities and profits are shared. Under thus scheme, the public sector will establish robust infrastructures and adopt suitable public policies while the private sector will finance the projects, design effective operational strategies and manage the day-to-day operations. Regardless of geographical location and market size, hubs may be **international logistics centers** situated near ports/airports. These will handle domestic distribution of of imported goods through regional hubs as well as international flows from international suppliers to international manufacturers. The r**egional logistics centers** located far from urban areas manage and concentrate flows of goods either imported from international logistics centers or produced locally, and distribute them nationwide using long-distance transportation means such as trucks carriers/trains/aircraft. The u**rban distribution centers** are located in the vicinity of urban areas to collect and manage flows of goods coming from senders or RDC to be distributed in the city center. This includes the well-known "logistics of the last mile" problem. Note that the above classification depends on country size. For example: hub network design in the USA is extremely different from that of the Netherlands. Thus, in the case of small countries, one may only find international hubs for both regional and urban distribution. Furthermore, the assumption of hierarchical network configuration only holds from the network designer' point of view. Looked at from an operational point of view, hubs are either **logistical hubs** with standard services such as international/ national transport, distribution, warehousing, inventory management etc... or **industrial hubs** offering high added-value services, such as order assembly, co-packing, post-manufacturing etc....

- **Hub interconnection**. Where there are multiple hubs, for reasons of flow consolidation, the problem of hub interconnection is posed. Interconnection may be ***complete*** or ***incomplete***. In the latter case, hub interconnections may be *predetermined* (and distant hub connections are eliminated, for example) or *to be determined*. Note that such determination impacts the definition of the transport network environment for the hubs (see section 4) but taking decisions at hub level simplifies the network definition problem.

II.4 Logistics Network Structure

Logistics network structures are defined by a number of end nodes (system customers), intermediary nodes (hubs) and a set of arcs that connect the nodes and represent transportation opportunities between two nodes. Network mapping, therefore, is an essential step in the definition of a logistics network.

- **Definition of network nodes**. The first problem lies in *determining the spokes*. A customer or group of customers in a given area may form an end-node enabling the pooling of transported goods, but their transport needs may differ. The enlargement of the catchment area of a node will increase its relative weight in the network organization but will multiply the drayage problems. The seriousness of these problems depends on the quality of local infrastructure. One may add that the nature of *spoke* transportation demand is related to the nature of the local transportation network offering. All these factors impact the definition of demand. As mentioned above, the position of *hubs* may either be known or to be defined (see section 3). In the latter case, the economic analysis partly depends on the rationale used in designing the transportation network (section 5).

There is a causality relationship between transportation infrastructure development and flow dynamics. Indeed, on one hand the establishment of a local transportation network may be justified by the existence of potential and regular traffic (transportation demand) related to terminal nodes. On the other hand, traffic growth may be driven by the availability of proper transportation infrastructure. Note that this causal link can only be observed if the time horizon comprises multiple periods. Furthermore, network overall performance depends on the number of nodes allocated to each hub. The network effect will impact overall network performance, negatively or positively: if additional terminal nodes increase costs and traffic congestion, this will reduce overall performance, but if this extension improves use of transport resources / handling, then this improves overall performance. Thus, it is important to assess the impact of an increase in the number of terminal nodes on overall network performance prior to making a final network design decision. Note that this network effect dimension is not dealt with adequately in literature.

- **Definition of network arcs**. The ***choice*** of arcs may stem from a *decision* to be taken (in the framework of an economic model) or be *predetermined*. One must first define the ***arcs that connect two end nodes***. This may be a *direct routing* between two system customers, without any hub involved. The arc may enable *pick up tours* to serve a number of customers in the same area as a first step to a hub and/or delivery tours. One will then need to define both the ***arcs that connect end nodes and hubs and hub interconnection arcs.*** An end node or hub may be connected to a single hub (*single allocation)* or be connected to several hubs (*multiple allocation*). The ***stability of allocation*** is the last feature of the problem: it may be defined once and for all (*static allocation)* or depend on demand trends (*dynamic allocation)* (see section 5).

II.5 Transport Organization

The transportation of goods between two network nodes (*spokes* and/or *hubs*) involves means of transport and a packaging as well as scheduling.

Means of transportation. Several ***types*** of vehicles (in the wide sense) may be used: trucks, trains, aircraft, ships…. And the network may rely on a ***single*** or ***multiple*** type(s) of vehicles: when the goods are always transported between shipper and addressee using the same type of vehicle, transport is *unimodal,* if not it is called *multimodal*. At a node corresponding to a switch of transport mode (rail/road, for example) specific infrastructure (such as a lifting device) may be required. A fleet of vehicles of a certain type may be required to perform a range of transport services, and this fleet may be ***homogenous***, i. e. sharing the same features (chiefly their mass or volume capacity) that may or not be predetermined; the ***number*** of vehicles in a homogenous fleet (or sub-set of it) may be known or to be determined.

Transport packaging. The goods to be transported may be in *bulk* (packs…), grouped *identical containers* (crates, pallets, containers…) or in *heterogeneous containers* of a limited number of families of identical containers. The type of packaging has an impact on handling time and cost as well as on vehicle occupancy.

Transport scheduling. The aim is to establish optimal pickup tours of goods between two network nodes (*spokes* and/or hubs), which implies pickup timing compatible with vehicle availability (see description of means of transport), demand for transport and node capacity. The ***tours*** may be *simple* (vehicle arrives empty at departure node and loads the goods and then fully unloads at recipient node) or *complex* (vehicle passes through multiple nodes and may load and/or unload at each one of them). Complex tours, applied to a set of hub customers, enable to accelerate pickup (or delivery) frequency and to reduce the size of batches collected (or delivered).

Note that this transportation organization requires (more or less precise) prior knowledge of demand.

II.6 Physical Data

As in all industrial management problems, knowing the data used to describe a real situation in a model strongly impacts the choice of model approach as well as the relevance and robustness of recommendations stemming from the models. The characterization of these data (transport requests, transportation time, nodes processing time, etc.) involves two dimensions, which features may be combined freely:

- **Degree of accuracy of knowledge on data**. The data may be ***deterministic, stochastic or fuzzy*** (the latter case is difficult to deal with and is only rarely used). Note that taking into account the service level is only justified in a random approach; the indicators normally used correspond to averages of random variables (time spent in a system…) that cannot be determined in a deterministic approach.
- **Stability of data over time**. The data may be deemed stable for at least a certain time (steady state) and qualified as ***static*** or deemed to be variable over time (trend, seasonality…) and qualified as ***dynamic***. Static problems involve a **single planning period** where data is measured based on an average evaluation whereas dynamic problems consider **multiple planning periods** enabling the consideration of data trends such as flow dynamics and feedback decisions.

The model approaches often also use valuation systems (see section 7) to enable an economic optimization approach (see section 8).

II.7 Economic Valuation System

Papers dealing with hubs do not systematically enter into an economic analysis. Where this is available, the economic valuation system is limited to logistics network hub setting up and/or operating costs. Four dimensions are relevant to describe this expenditure.

- **Nature of costs**. This includes standard management accounting ***recurring expenses*** (incurred in the delivery of logistics services) with varying degrees of detail (transport, handling, inventory…). In strategic decision-making models costs include ***investment expenditure*** (studies, new facilities, acquisition of handling and transportation equipment…) to deliver or adjust resources dedicated to logistics services. Note that a discounting device must be included in any economic evaluation combining recurring and investment expenditure; the discount rate used and the economic horizon for the study have a major impact on decisions.
- **Cost allocation** depends on the definition of the actor(s) who will bear the logistics costs described above regardless of the business organization governing the hubs (single actor, multi-actors). The selection and measure of those costs depend on who the costs will be allocated to. We distinguish the **user costs** and **operator costs** points of view. The latter involves costs that operators (either single actor of multi-actors) bear in order to set up and manage hub networks; those costs include investment expenditure and recurring expenses (service costs) whereas the users' costs include use of the hub network and are mainly recurring expenses (service price).
- **Method of cost calculation**. This problem concerns recurring costs. Operational research specialists frequently use traditional accounting approaches that rely on the ***standard cost*** method measured in unit of work (tonne per mile…). This approach is generally used though of doubtful relevance wherever the physical referential used for calculation differs from that resulting from implementation of decisions. In addition, these systems can introduce distortions on economic recovery, as some activities are not triggered by volume work units. For example: according to the point of view of fleet owner operator, transportation cost does not only vary according to transported quantity and travelled distance a generally argued but also don other drivers such as shipment rate,

transport modes, size and type of container etc... . Indeed, if we compare the transportation cost for goods on 20-feet equivalent units (TEUs) containers with weekly shipment and that for 40-feet equivalent units containers every two weeks over the same distance, we might conclude, from a standard costs point of view, that transportation cost will double in proportion to load. This traditional conclusion, however, is erroneous, as physical configurations and conditions actually vary. Thus, an economic assessment must be based on the cost drivers. The ***cost driver*** approach (*activity-based costing*) though far more relevant, is rarely used (ill-known of OR specialists and/or difficult to implement due to inherent high degree of granularity). In contrast to traditional approaches where the work unit serves to assign costs without reference to factors that generate costs, the cost driver approach relies on the physical relationship between the expenses and activities that generates it. The cost drivers are the resource-consuming activities triggered by decisions.

- **Consistency analysis**: the economic evaluation system is based on physical and economic description. The interdependency of these interrelated descriptions calls for a consistency analysis (iterative loop). Indeed, as underlying assumptions to costing design may not match the physical representation related to the optimal solution, some costs will be adjusted along with the underlying assumptions. As a result, the optimal solution will be systematically modified. Note that consistency analysis is rarely conducted in the literature, and we feel this is a major progress avenue.

II.8 Decision Evaluation Criteria

These criteria may be economic or physical and enforce or drive decisions. This reflects the traditional dichotomy between management efficiency and effectiveness.

Economic evaluation criteria. Generally an aggregation of costs (see section 7) that varies from paper to paper. Thus, if optimization is made from a hub operator's point of view, one would include assets into global logistics cost, as the economic aim would be to maximize cashflow.

Physical performance criteria. These criteria depend on spatial-temporal granularity of the model used; they are often difficult to include in strategic decision approaches. These criteria usually hinge on *transportation time* (between two *spokes*, etc.).

II.9 Application Field

Hubs are used in the transportation of goods, people and information. The nature of the transported entities (tangible or intangible) may restrict the scope of the conclusions of some papers. Accordingly, this aspect is included in the analytical map. This point shall be dealt with in further detail in a forthcoming work.

III. Conclusion

Our analytical grid is currently applied to a corpus of some 150 papers identified in various documentary databases. The eight first dimensions are broken down into sub-dimensions corresponding to the items in bold prints in the corresponding above sections; these may be further subdivided as per the items bold/italic print in the same sections. The resulting mapping of the papers serves to form a clear picture of the contingencies specific to the different approaches reviewed and to identify potentially valuable research avenues.

IV. Bibliography

Alizadeh, M., (2009). Facility Location in Supply Chain. *Facility Location: Concepts, Models, Algorithms and Case Studies 1*.

Alumur, S., Kara, B.Y., (2008). Network hub location problems: The state of the art. *European Journal of Operational Research* 190, 1–21.

Campbell, J.F., Ernst, A.T., Krishnamoorthy, M., (2002). Hub Location Problems, in: Drezner, Z., Hamacher, H.W. (Eds.), Facility Location. *Springer Berlin Heidelberg, Berlin, Heidelberg*, pp. 373–407.

Campbell, J.F., O'Kelly, M.E., (2012). Twenty-Five Years of Hub Location Research. *Transportation Science* 46, 153–169

Daskin, M.S., Snyder, L.V., Berger, R.T., (2005). Facility location in supply chain design, in: *Logistics Systems: Design and Optimization.* Springer, pp. 39–65.

Farahani, R.Z., Hekmatfar, M., Arabani, A.B., Nikbakhsh, E., (2013). Hub location problems: A review of models, classification, solution techniques, and applications. *Computers & Industrial Engineering* 64, 1096–1109.

Kara, B.Y., Taner, M.R., (2011). Hub Location Problems: The Location of Interacting Facilities. *Foundations of Location Analysis* 273.

Melo, M.T., Nickel, S., Saldanha-Da-Gama, F., (2009). Facility location and supply chain management–A review. *European journal of operational research* 196, 401–412.

O'Kelly, M.E., Bryan, D.L., (1998). Hub location with flow economies of scale. *Transportation Research Part B: Methodological* 32, 605–616.

O'Kelly, M.E., Miller, H.J., (1994). The hub network design problem: a review and synthesis. *Journal of Transport Geography* 2, 31–40.