Strategic Service Network Design for DHL Hong Kong

Waiman Cheung; Lawrence C Leung; Y M Wong
Interfaces; Jul/Aug 2001; 31, 4; ABI/INFORM Global
pg. 1

We developed a two-phase network planning methodology to
design a service network for DHL(HK) based on two measures
of service performance: service coverage and service reliability.
The methodology consists of first using an optimization model
to determine a least-cost distribution network, which forms the
basis of a simulation model for analyzing the network's operational
characteristics. It takes into consideration such design issues
as inbound and outbound shipment, equivalent shipment,
cutoff time, demand fluctuation, and random behaviors. DHL
decision makers can now examine trade-offs between coverage
and reliability by adjusting workforce, changing cutoff time, or
redesigning the service network.

DHL, an international air express
company, serves over 220 countries
around the world. Its worldwide opera-
tions network consists of over 40,000
employees, 1,900 international service sta-
tions, 30 regional gateways, and 28 interna-
tional sorting hubs. DHL's fleet of 177
aircraft and 11,000 vehicles transports mil-
ions of customers' packages daily. In addi-
tion to the hub-and-spoke routings of
the company's own aircraft, DHL also
uses commercial flights extensively for di-
rect point-to-point transport. Established
in San Francisco in 1969, DHL's major
markets are in Europe and the Asia Pacific
(AP) Region. Within the AP Region, DHL
is a leading carrier in terms of total vol-
ume, with Japan, Hong Kong, and Singa-
pore as its three largest subregions.

In 1998, the Hong Kong International
airport was relocated. As one of the largest air-express couriers in the Asia-Pacific region, DHL had to design a new service network for its Hong Kong operations. The network plan had to provide for strategic and timely installation of depots and service centers to meet future changing customer needs and provide competitive service. DHL(HK) has about 1,200 staff and 160 vehicles. Its main operations center is in a district three kilometers from the old international airport, and it has other smaller operating facilities within Hong Kong. Hong Kong's economic development has been strongly influenced by rapid growth in southern China. Business and social activities between mainland China and Hong Kong have intensified since July 1997 when Hong Kong was reunited with China, with southern China acting as a manufacturing base for Hong Kong businesses because of its cheaper land and labor. There is a growing need for express delivery of documents and packages between the two regions.

The New International Airport

The relocation of the international airport to Chek Lap Kok (CLK) on Lantau Island (to the far west of Hong Kong Island) significantly affects the effectiveness of DHL(HK)'s current service network. The new airport is much farther away from the city than the old airport, thus increasing commuting time and operating costs. Also, for safety reasons, the highway linkages between the city and the new airport restrict the type of truck that can be used. The customs office uses electronic-data-interchange (EDI) technology to shorten the customs-clearance process for those shippers who can provide EDI manifests before flight arrival. DHL has developed a new operating system, called the Air Cargo Clearance System, to be used at the new airport. Moreover, by setting up an express cargo terminal inside the new airport, DHL would be able to carry out some shipment registration processes while waiting for customs clearance. In many ways, the new airport has brought both challenges and opportunities to DHL(HK).

Competition Among Air Express Companies

DHL(HK) operates in a very competitive air-express market. Overnight service within the AP region and next-day service to Europe and North America are the industry standards. For selected sectors, such as banking and finance, same-day international service to the US is also available. These service standards are much higher than those offered by airfreight forwarders. In addition, air-express companies offer money-back guarantees in case of failure to keep delivery promises.

The factors customers consider in choosing an air-express company are:
—Short transit times and a large global network,
—Late cutoff time for package pickups without incurring delay,
—Convenient communications and order placement, and
—Competitive price.

Increasingly, to remain competitive, companies must reduce delivery times. To respond quickly to customer requests, couriers must stay close to the customers. They must also have the necessary logistical support to transport shipments to planned flights quickly. To continue to
meet or exceed industry service standards, DHL(HK) must provide superior service at a competitive price. It needs to balance service level with the fixed cost of installing facilities and the variable cost of operating the service network.

**The Service Network**

DHL’s service network consists of demand zones, satellite depots, service centers, and the airport. Demand zones are service areas determined according to the level of customer demand and geographical characteristics. Zones are smaller within busy commercial areas where orders are concentrated than in most outlying zones where demand is scattered.

Customer requests for pickups are forwarded to couriers responsible for their zone. Couriers are assigned to specific satellite depots, each of which covers pickups in several zones. The depots consolidate packages and deliver them to the corresponding service center. Each service center, which also functions as a depot, is responsible for several depots. Service centers handle all major processing, such as labeling, X-ray screening, reweighing, sorting, documentation, and formality follow-up. The service centers further consolidate shipments into air containers or bags and transport them to the airport for transfer onto the appropriate aircraft.

DHL(HK) must manage this process from pickup to delivery at the airport effectively. The service network is at the heart of this process. The critical decisions in the design of the service network concern the installation of the depots and service centers:
- The locations of depots and what demand zones they cover,
- The locations of service centers and what depots they cover,
- The capacities of these facilities, and
- The installation schedule for these facilities.

**Objectives of the Study**

DHL(HK) asked the authors (Cheung and Leung) to perform a study to improve their services. The objective of the research study was to design a service network that would be the most desirable economically and operationally for DHL over a multi-year period. In designing the network, we had to consider long-term goals and short-term operational goals. The design had to include the strategic and timely installation of depots and service centers based on a judicious examination of service performance. The study included designing the overall framework, formulating models, collecting and preparing data, interpreting results, setting operating rules and policies, and making recommendations to the top management. The principal strategic recommendations concerned
- Installation decisions for depots and service centers, and
- A strategic cutoff time for pickups that balances capturing more customers with providing reliable service.

**Inbound and Outbound Shipments**

DHL(HK) handles both inbound and outbound shipments. For DHL(HK), the inbound-shipment-flow and the outbound-shipment-flow patterns are basically the same. The volume shipped is slightly less for inbound, while the shipment patterns

---

July–August 2001

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
are roughly identical. The two flows do not compete for resources because they occur at different times of day. Unlike inbound shipments, which DHL collects at the airport according to a predetermined schedule, outbound shipments require more processing and must reach the airport by a specific time. Therefore, outbound shipments have more stringent time requirements. To simultaneously consider inbound and outbound shipments in planning the network would unnecessarily complicate the design task. Also, it is very likely that optimal conditions for outbound shipments are close to optimal for inbound shipments as well. In our study, we considered only outbound traffic.

**Multiyear Demand**

The businesses located in Hong Kong and the business volume between Hong Kong and mainland China are expected to increase gradually in the years to come. These changes will affect the quantity of shipment demand in the various zones and the characteristics of the shipments. Today's optimal service network may not be adequate in the future; however, DHL(HK) cannot abandon, relocate, or scale down facilities without penalty. Therefore, we needed to base the network design on a long-term demand forecast. The marketing department of DHL(HK) has the task of preparing the long-term demand profile, examining demand fluctuations, such as seasonality, daily variations, and peak loads; types of shipments (volume and weight); potential changes in market structure; regulatory issues; the macroeconomic profile; and the development of new industrial and residential zones.

**Equivalent Shipment**

The demand profile has to be refined to reflect the amount of resources consumed by shipments, which depends largely on their weight. In demand zones where the shipments are heavy, requirements for human and material-handling resources are proportionately higher. To capture this weight factor, we used a unit of measure called equivalent shipment, which we defined as one shipment whose weight is one kilogram. We consulted ground-service managers, service-center managers, industrial engineers, and couriers on their perceptions of the impact of a range of shipment weights on the amount of resources required. We collected these expert opinions and performed a simple regression. We concluded that we could correlate the weight and equivalent shipment-weight approximately using a power correlation function \( y = ax^b \) with \( a \) as 1 and \( b \) as 0.5. We suggested the following rule of thumb:

—The equivalent number of shipments is equal to the square root of the actual shipment weight in kilograms.

For example, we could count a shipment that weighs four kilograms as two equivalent shipments, implying that this shipment would consume twice the resources a one kilogram shipment consumed. We used this rule of thumb to transform all shipments with differing weights into corresponding equivalent demands.

**Cutoff Time and Decentralization**

Determining a service cutoff time and the level of decentralization (or centralization) of services are related policy issues for DHL(HK). The cutoff time is the time
of day before which a customer’s order is 
guaranteed to be delivered to the airport 
on the same day. It is a critical competitive 
factor among air-express couriers. An un-
necessarily early cutoff time would mean 
a loss of potential customers, whereas an 
unreasonably late cutoff time would leave 
DHL little time for shipment processing

Before we implemented this methodology, managers’ decisions were based on 
intuition.

and transportation, thus increasing the 
risk of not reaching the airport in time. 
One way to achieve late cutoff while meet-
ing service promises is to decentralize the 

service network. In the extreme case of de-

centralization, this would mean a depot 
and a service center in each demand zone. 
Clearly, this would shorten the response 
time for pickups and delivery times. How-

ever, the facility installation cost would be 
tremendous, and since there would be no 
consolidation, the total transportation cost 
would be very high and resource utiliza-
tion would be very poor. On the other 
hand, a centralized service network would 
mean having a super-service center that 
doubled as a superdepot. All couriers 
would originate from this superfacility for 
pickups and would return to it for pro-
cessing and consolidation. This would 
provide economies of scale, more efficient 
consolidation, better utilization of re-
sources, and better control of shipments, 
but the response time for pickups would 
be longer on average and meeting flight 
schedules for pickups from some outlying 
zones would be difficult. In determining 

the optimal cutoff time, we had to care-
fully weigh the trade-off among level of 
customer coverage, risk of missing flights, 
and costs of decentralization.

Capacities and Costs of Depots and 
Service Centers

For each potential depot and each po-
tential service center, we had to identify 
what capacity sizes could be installed and 
the costs. Because of tremendous variation 
in real estate costs, our analysis had to be 
specific to location. The elements we in-
cluded in cost assessments were physical 
plants, material-handling systems (trucks, 
vans, conveyance systems), human re-

sources, and information-processing capa-
bilities. Other capacity-related factors we 

had to determine were the utilization rates 
of the existing facilities and acceptable 
utilization rates, and overtime costs for fa-
cilities and human resources.

Fluctuations and Random Behaviors

DHL(HK)’s operating environment con-
sists of many dynamic elements that can 
fluctuate drastically and unpredictably. 
Typically, customer demand varies ac-
cording to the time of day, the day of the 
week, and the month. Customer requests 
peak around the daily cutoff time. They 
are higher on Fridays, on days before holi-
days, and during major holiday seasons, 
such as Christmas. Facilities with enough 
capacity to handle average demand may 
not be adequate to cover peak hours and 
days. Consequently, we needed to con-
sider facility utilization. Furthermore, the 
randomness in transportation times means 
that some shipment routings that are feasi-
ble under normal traffic conditions may 
not be feasible during peak hours when 
traffic is congested. To examine the service
performance of the network, we needed voluminous operational data. We wanted to characterize hourly customer demand by volume and to describe the randomness in such elements as pickup times, transport times, and processing times.

**The Planning Methodology**

Clearly, the design of a long-term service network cannot be separated from short-term operational considerations. Within the context of operations research methodology, design of a service network can be approached using facility-location models or variants of such models. The approach is a normative one to determine least-cost locations that meet certain demand and logistics requirements. Because of the problem size, such models can use only aggregate information, such as average annual cost, average utilization, and average annual demand. But a service network based only on an optimization model is unlikely to be adequate. An optimization model is not suitable for handling such operational factors as demand uncertainty, randomness in travel and processing times, and dynamic consolidation. The model prescribes the overall least-cost network design but lacks specific operational details. An air-express courier cannot use a network design that does not satisfy important daily and weekly operational specifications.

Analysts can use simulation to model uncertainty and fluctuations in large dynamic systems, allowing decision makers to examine the behaviors of complex systems operating in uncertain environments. However, before we could simulate operational details (product flow, pickup and delivery schedules, disaggregated product-mix demand patterns, and probabilistic behavior of travel times and processing times), we needed a network design on which to base the simulation.

Our network-planning framework uses an optimization model and a simulation model (Figure 1). The optimization model is an economic network-planning model that determines the macro optimal network configuration in an aggregate fashion. The simulation model evaluates the daily operational performance of the recommended network configuration. Analysts have used optimization models and simulation models jointly to analyze hospital layout [Butler et al. 1992], freight operations [Moore, Warmke, and Gobran 1991], ambulatory health care [Kropp and Carlson 1997], manufacturing [Leung, Maheshwari, and Miller 1993], and defense logistics planning [Nolan and Sovergin 1972]. These are situations in which design and operations are closely related. While the air-express industry is unique, it is similar in that a courier’s service network is also closely tied to its service operations. We found no published work concerning an air-express network within this design-operations context.

**The Macro-Network-Planning Model**

The multiperiod model, a mixed 0–1 linear program, takes on the following framework (Appendix):

Minimize

\[
\text{present value of (fixed cost + variable cost)},
\]

Subject to

flow distribution = demand, individual flow \(\leq\) capacity, flow time \(\leq\) time window, only one installation at a location, conjunctive or mutually exclusive installations.
The objective of the model is to minimize the sum of present-value costs of transportation and facility installation. The variable transportation costs depend on the assignment decisions of shipments from zones to depots and from depots to service centers. The installation costs depend on installation decisions, the choice of capacity level, and the schedule of the installations.

We need constraints to ensure that the assignments of shipments from zones to depots to centers will meet forecast demand. Also, we need capacity constraints to ensure that decisions to assign shipments to particular facilities correspond to facility-installation decisions and appropriate capacity decisions. Further, the total flow time for any zone-depot-center-airport assignment cannot exceed the maximum time window (elapsed time between cutoff and due time at the airport). We also needed logical constraints to ensure that only one installation per site was allowed and that certain correlated time-dynamic installation logic is not violated.
The Micro Operation-Simulation Model

Based on the results of the macro planning model, the simulation model simulates daily operational activities to evaluate the service performance of the network. We constructed multiple scenarios in terms of peak loads and average loads and collected operational statistics including the workloads of couriers, capacities and utilization of facilities, arrival patterns of vehicles to the facilities, pickup requests not honored, and shipments that miss flights.

The Simulation Environment

We used the simulation model to examine the dynamics of courier pickups, deliveries to depots and service centers, and deliveries to the airport.

The locations of the depots and service centers, as well as shipment assignments from zones to depots to centers, are in accordance with the results of the planning model. DHL(HK) uses three types of vehicles: vans to transport shipments from customers in the zones to depots, trucks to carry shipments from depots to service centers, and lorries to carry shipments from the service center to the airport. We simulated the consolidation of shipments from vans to trucks, and from trucks to lorries. The entire operation is applied to two major product types, documents and packages, collectively representing almost 90 percent of the shipments.

The workforce includes couriers and shipment-processing workers. A typical working day starts at 8:00 am, when the couriers leave these depots in vans to deliver and pick up shipments in zones. The usual cutoff time is 5:15 pm, but it can be adjusted depending on the cutoff policy. Each week has five-and-a-half workdays. We incorporated in the model probabilistic behaviors exhibited in three categories of events—shipment arrivals and characteristics, travel time, and processing time.

The Simulation Model

The simulation model is divided into four modules (Figure 2): (1) simulation initialization, (2) shipment pickups in the zones, (3) shipment consolidations in the depots, and (4) processing in the service centers and transport to airport. Essentially, we modeled both zones and service centers as one-line multiple-server queueing resources. The capacity of a resource is related to either the number of couriers in a zone or the number of processing workers in a service center. At the end of each day, documents and packages that remain in the queue of each zone are considered to be lost-sale shipments. Documents and packages that remain in the queue of each service center are undelivered shipments and are to be delivered the following day. Since the depots do no pro-

![Figure 2: The simulation model is divided into four modules, (1) initialization, (2) shipment pickups in zones, (3) shipment consolidation in the depots, and (4) processing in the service centers and transport to the airport.](image-url)
cessing except consolidating shipments, we modeled depots as simple storage with a queue and incorporated a time delay for unloading and loading vehicles. We created three types of transporters (representing vans, trucks, and lorries) to move entities from zone to depot, from depot to service center, and from service center to the airport. In the simulation model, we used one week as a cycle. It is a duration in which the system resets itself, and it typically reflects the real-life behavior of DHL.

**The Preliminary Service Network**

DHL(HK) has established 33 demand zones in Hong Kong, 15 of which are candidates for depots and nine of which are potential sites for service centers. Using inputs from the demand profile, fixed and variable cost estimates, travel and processing times, capacity alternatives for facilities, and a cutoff time of 5:15 pm, we solved the macro planning model using the PC-based MPSIII [Ketron 1992]. There are 49,500 continuous variables, 1,050 binary variables, and 820 constraints. Solution takes about seven hours of computation time on a Pentium-166 machine.

In the resulting service network, there are five depots and four service centers (Figure 3).

To protect the interests of DHL(HK), we will not show the corresponding costs of this configuration nor the recommended future expansion plan. The distribution network is neither centralized nor decentralized and remains somewhere in-between. This network configuration is only a preliminary result; we had not yet scrutinized the network’s service.
performance.

**Service Performance of a Network: Coverage and Reliability**

To evaluate the service performance of a service network, we defined two related performance measures:
- Service coverage, the percentage of requests that arrive before cutoff time, and
- Service reliability, the percentage of pickups that make the same-day flight.

These two criteria are central to an air-express firm's service performance. The former addresses its coverage of requests and the latter its success in keeping promises of same-day delivery. The actual pickups may go beyond the cutoff time because same-day delivery applies to all requests as long as they arrive before or on the cutoff time.

In the macro model, service coverage and service reliability are both 100 percent, since all shipment requests are met and all assignments are within the maximum time window. But the parameters of the macro model are merely deterministic averages. The actual daily fluctuations and randomness of activities will likely result in varying levels of coverage and reliability. We used the simulation model to make a realistic assessment of the network's coverage and reliability.

Based on the preliminary network configuration, we simulated the daily activities. We implemented the simulation experiment using a SIMAN-based simulation software package, ARENA [System Modelling Corporation 1995; Pegden 1995], in a Pentium-166 machine with 128 MB RAM. A simulation run of one week takes approximately 14 minutes. With the simulation results, we plotted the levels of service coverage and service reliability with respect to a range of cutoff times (Figure 4).

For example, if the cutoff time is 5:15 pm, coverage is at 92 percent and reliabil-

![Service Reliability vs. Cutoff Time]

**Figure 4:** Based on the simulation results, the chart shows the trade-off between service reliability and coverage. For example, extending cutoff time will result in a higher level of service coverage but a lower level of reliability.
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
network using a more stringent time window. That is, we could solve the least-cost optimization model again but with a later cutoff time. Because the time constraint is tighter, the resulted network would have a higher cost but would also be more time efficient (or at least as good) since less time is allowed for services. But this more stringent time window is not implemented during actual operation. Instead, we would use the old time window. Consequently, service reliability would improve because of the more stringent network that would provide acceptable reliability and coverage with respect to the actual cutoff (use in the simulation). Again, whether the incremental cost of constructing a more time-efficient network justifies the improvement in reliability will be a decision top management has to make. There are essentially two cutoff times, one for the network model (planning cutoff) and one for the simulation (operational cutoff). The former determines the preliminary network, while the latter provides a realistic assessment of the effects of cutoff times in actual operations. In designing the network, we had to explore carefully the relationship between planning and operational cutoff times.

**Recommendations**

Based on the current preliminary network, we investigated several scenarios to improve service coverage (without affecting service reliability). We iteratively used the optimization model and the simulation model to analyze different levels of human resources and different combinations of planning cutoff time and operational cutoff time. Several variants of the preliminary design showed strong potential. In addition, we determined several alternative network designs that are not least-cost designs but perform satisfactorily in terms of service coverage and reliability. We presented these network designs and their individual merits to DHL(HK)'s top management. Together with the installation decisions and the cutoff time, we described each candidate design by three attributes, cost, coverage, and reliability. The DHL decision makers will select the eventual design.

**Iterative Use of Optimization and Simulation Models**

In iterating between the macro model and the simulation model, there is no guarantee that a convergence will always take place. Convergence means that we can progressively zero in on a least-cost network that satisfies all the operational objectives. To do this, we emphasize diagnosing the outputs of both models judiciously and the art of revising the inputs. Arbitrary modification of input parameters would likely cause divergence or oscillations.

Many variations and scenarios go into implementing the macro-micro iterations to arrive at the most desirable network. Invariably, the number of iterations required depends on the quality of the input-revision process. We have established some basic algorithms and rules in revising the model inputs during the iterative process. Still, many intangibles in evaluating a network require expert opinions. We found valuable inputs from decision makers at DHL highly useful in streamlining the process of determining the most desirable network and several good alternatives.
Regarding the simulation exercise, we recommend doing simulations using demand profiles for a busy week, an average week, and a slow week. Decision makers can then judge the variations in utilization and service reliability. For each simulation run, we performed 10 replications. The number of replications should be statistically supported. This should be done for each year in the planning horizon open (using the demand profile for that year) or to a point at which the decision makers are comfortable with the current decisions. However, this can be computationally cumbersome, and its extent is also related to the judicious use of the macro-micro iterations.

**Conclusion**

Our methodology helped DHL decision makers to gain insight into many facets of their planning and operations. Before we implemented this methodology at DHL(HK), managers’ decisions about network design and network operation were based on intuition. Using our method as a framework, they can now analyze many network-related issues explicitly. Decision makers can now test many what-if scenarios and see the impact of certain strategic decisions. DHL(HK) top managers took our work very seriously and adopted most of our recommendations. They have continued to use our methodology to update their network design.

Since DHL(HK) bases its multiyear distribution plan on many elements in the future, it must continually update the distribution network with its expansion plan as it learns of changes in demand and transport structure. Essentially, DHL(HK) should perform the network-planning process periodically (for example, every six months) while using the micro model regularly to analyze its operational performance.

**Acknowledgment**

This research project is funded in part by the Hong Kong Research Grant Council (CUHK#4037/99H).

**APPENDIX**

The general mixed 0–1 LP formulation in matrix notation is

Minimize

\[ \sum_t \left( c_{ft} x_t + c_{vt} p_t \right) \]

Subject to

\[ A_{fp} \leq d_p \]

\[ p_t \leq (M_1 x_1 + \cdots + M_n x_n) \]

\[ e_{x_t} \leq 1, \quad G(x_1, \ldots, x_n) = 0, \]

\[ x_t \in (0,1), \quad p_t \geq 0, \quad \forall t = 1, \ldots, n \]

where

- \( x_t \) = decision vector for the facility and capacity installations in period \( t \)
- \( p_t \) = flow assignment vector in period \( t \)
- \( c_{ft} \) = vector for the fixed charges of the installations in period \( t \)
- \( c_{vt} \) = vector for variable cost per unit flow in period \( t \)
- \( A_{fp} \) = flow matrix in period \( t \)
- \( d_p \) = demand vector
- \( e_{x_t} \) = logic matrix (0–1) in period \( t \)
- \( M_t \) = capacity matrix of the installations in period \( t \)

and

- \( n \) = planning horizon.

All cost factors are in terms of present value. The flow constraints, \( A_{fp} \leq d_p \), make sure that assignment of shipment meets demand. The capacity constraints, \( p_t \leq (M_1 x_1 + \cdots + M_n x_n) \), ensure that the assignment of shipments to a facility is accompanied by the decision to install the facility along with the corresponding capacity decision; the shipment assigned must not exceed the accumulated capacity of the facility. Constraint sets \( e_{x_t} x_t \leq 1 \) ensure one installation per site, and \( G(x_1, \ldots, x_n) = 0 \) is a system of logic constraints representing additional correlated...
time-dynamic installation logic. All flow decisions $p_i$ must not exceed the maximum flow time.

References


Gary Wong, Planning and Development Manager, DHL International (HK) Ltd., DHL House, 13 Mok Cheong Street, Tsimshatsui, Kowloon, Hong Kong, writes: “As a Planning and Development Manager of DHL (HK), I found their work very useful. We have used it and will continue to use it as an important resource for our future distribution network planning. As a matter of fact, I was very impressed that the study was both scientific and practical, which included a simulation model that reflected the actual courier activities (such as fluctuation and random behavior). I also recall the important idea of “shipment weight”, which was used to distinguish the different level of resources needed to handle light and heavy shipments. From a bigger picture, the models allowed us to have a better understanding of the relationships among service coverage, reliability, and costs.”