The History of Lift Traffic Control[1]

Dr Gina Barney

Gina Barney Associates, Gina Barney Associates, PO Box 7, Sedbergh, LA10 5GE, UK.
www.liftconsulting.org

Keywords: Lift, traffic, control, collective control, sectoring, hall call allocation.

Abstract. The advent of FAPB (Fully Automatic Push Button) made the human operator or dispatcher redundant. Then the way lifts responded to passenger demands was in the imagination of "programmers" using relay logic and then programmers using digital computers. This paper looks at the history of the early relay based controllers and draws attention to their remarkable sophistication. These include: nearest car, fixed sectoring and dynamic sectoring. The ultimate traffic control, now used extensively and often inappropriately, is Hall Call Allocation. First described by G D Closs in 1970 (extending Leo Port's 1961 work), analysed by Sergio dos Santos in 1974 and implemented by Joris Schroeder in 1990.

1 THE REQUIREMENTS FOR THE TRAFFIC CONTROL OF LIFTS

The traffic control requirement is to co-ordinate a group of lifts to best serve passengers with the minimum of equipment.

2 SINGLE LIFT TRAFFIC CONTROL

2.1 Single Call Automatic Control

The simplest form of automatic lift control is single call automatic control. Single pushbuttons are provided on the landings and a button for each floor in the car. Car calls are given absolute preference over landing calls. If the lift is in use, a new landing call can only be registered, when the lift is no longer in use. This type of control is only suitable for short travel passenger lifts serving up to four floors, with a light traffic demand and is suitable for goods lifts.

2.2 Collective Control

The most common form of automatic control used for a single lift is collective control. This is a generic designation for those types of control where all landing and car calls made by pressing pushbuttons are registered and answered in strict floor sequence. The lift automatically stops at landings for which calls have been registered, following the floor order rather than the order in which the pushbuttons were pressed. Collective control can either be of the single button, or of the two pushbutton types.

2.3 Non-directional collective

Non-directive collective control provides a single pushbutton at each landing. This pushbutton is pressed by passengers to register a landing call irrespective of the desired direction of travel. Thus, a lift travelling upwards, for example, and detecting a landing call in its path stops to answer the call, although it may happen that the person waiting at the landing wishes to go down. This type of control is only acceptable for short travel lifts.

2.4 Down collective (up-distributive, down-collective)

Single pushbutton call registration systems may be adequate in buildings where there is traffic between the ground floor and the upper floors only and no interfloor traffic is expected, e.g.: car parks, public high rise housing, flats. Retaining the single pushbutton on the landing, a suitable control system is the down collective control (sometimes called up-distributive, down-collective)

---

where all landing calls above the ground are understood to be down calls. A lift moving upwards only stops in response to car calls. A lift traveling downwards, answers car and landing calls in floor sequence.

2.5 Full collective (directional collective)

The two pushbutton full collective control provides each landing (except terminal landings) with one UP and one DOWN pushbutton and passengers press the pushbutton for the intended direction of travel. The lift stops to answer both landing calls and car calls in the direction of travel, and in floor sequence. This control system is suitable for single lifts or duplexes (two lifts) serving a few floors with some interfloor traffic. Typical examples are small office buildings, small hotels and blocks of flats. Directional collective control applied to a single lift car is also known as simplex control. The system can be applied to two or three interconnected lifts to work as a team, where a fully configured group control is not appropriate. Two lifts are termed a duplex and three lifts a triplex. Full directional collective control is the simplest form of group control.

3 GROUP TRAFFIC CONTROL

The purpose of group control is allocate landings calls in an optimal way to minimise: passenger waiting and journey times; system response time; energy consumption; maximise the handling capacity and reduce 'bunching'. These aims are sometimes in conflict.

4 LEGACY TRAFFIC CONTROL SYSTEMS

There were four basic (generic) types of traffic controller developed by the proprietary and independent manufacturers.

4.1 Nearest car

The simplest type of group control is the directional collective control described above. It is suitable for a group of two, or three lifts, each operating on the directional collective principles and serving seven or so floor levels. The controlled assignment of one lift only to a landing call can be achieved by the "nearest car" control algorithm.

The nearest car traffic control system is expected to space the lifts effectively around the building, in order to provide even service. The group traffic control feature contained in this simple algorithm is the allocation of each landing call to the lift that is considered to be the best placed to answer this particular call and no other. The search for the "nearest car" is continuously performed using quite sophisticated rules, until the call is cancelled after being serviced.

4.2 Fixed sectoring – common sector system

A fixed sectoring common sector control system can be devised for dealing with off peak traffic and can be complemented with special features to cater for heavy unbalanced traffic. The system divides a building zone into a number of static demand sectors equal to the number of lifts. A sector includes both the up and down landing calls at the floors within its limits. A lift is allocated to a sector if it is present in that sector and the sector is not committed to another lift. Fully loaded lifts are not considered for assignment. An assigned lift operates on the directional collective principle within the limits of its range of activity. The de-assignment of a lift from its sector takes place when the lift leaves the sector. A lift picks up calls ahead when travelling in either direction, even if it is not assigned to the sector.

The system, by distributing the lifts equally around the building, presents a good performance for uppeak and balanced interfloor traffic. It lacks a proper procedure to cater for sudden heavy demands at a particular floor.

4.3 Fixed sectoring – priority timed system

A fixed sectoring systems can also allocate the lifts on a priority timed basis. The landings in the building zone served by the group of lifts are grouped into independent up and down sectors. Each sector is timed as soon as a landing call is registered within its limits. The timing is measured in predefined periods of time, designated the priority levels. The system is unique among the classical traffic control systems as it considers time when making an assignment. The other algorithms only consider position. The assignment of lifts to the sectors takes into account the number and positions of the available lifts and the sector priority levels. The control system provides a good up peak performance and good down peak performance, especially under very heavy traffic conditions. The interfloor traffic performance is fair.

4.4 Dynamic sectoring system

The dynamic sectoring group supervisory control system provides a basic algorithm that groups landing calls into dynamic sectors. The position and direction of each lift defines the dynamic sector. Each lift answers the landing calls in the sector "ahead" of it. In parallel with the basic traffic algorithm, another dynamic sectoring algorithm is provided to insert free lifts ahead of lifts serving a large number of floors or a large number of calls registered in their dynamic sector. The dynamic sectoring system provides a very good performance for uppeak and interfloor traffic conditions, but a poor performance for down peak.

5 MODERN CONTROL TECHNIQUES

5.1 Fundamental Limitations

Although computer based traffic control systems can allocate lifts more efficiently than the relay based traffic control systems, there is a limit to what can be done. The main limit is the finite handling capacity resource of the underlying equipment to handle the traffic demands. This relies firstly on good equipment, which is properly set up and secondly on advanced control systems. Once the major inefficiencies have been removed such as: single button calling; stopping full cars; faulty detection of car loads; inefficient door operations; etc., then it is only possible to "trade" one parameter against another. This means that one passenger's shorter waiting time is another passenger's increased waiting time. The effect on the second passenger could be so small that it is unnoticed, but the effect on the first passenger could be significant.

The opportunity exists with a computer to program complex tasks to assist the landing (hall) call allocation process, which are impossible to achieve with fixed program systems. This might be considered to lead to truly optimal traffic control. An Estimated Time of Arrival (ETA) based traffic control system is an example, which allocates lifts to landing calls, based upon computed car journey times, ie: how long a lift takes to arrive. Early systems of this type, developed in the 1970s, substituted relay or solid state fixed logic by a truly programmable computer. This technique was an obvious one to use once programming facilities were available. The ETA technique remains the underlying basis of many computer based systems on the market today. A variation of ETA is estimated time to destination (ETD). This system not only estimates the time to arrive and pick up the intending passenger(s), but also the time to take them to their destination.

5.2 Stochastic Traffic Control Systems

Observations of classically controlled lift systems have indicated that the response times to answer landing calls follow a curved shape similar to the Exponential Distribution curve of Figure 1 (a). This distribution curve has a large number of calls answered in zero time or during the first time band. However, there is a long tail to the distribution with some calls waiting very long periods of time. Thus the underlying premise of algorithm design should be to bring the tail closer to the average and to sacrifice the "instant" collection of some calls by moving the exponential away from the origin to a Gaussian shape similar to the Rayleigh Distribution curve of Figure 1 (c).

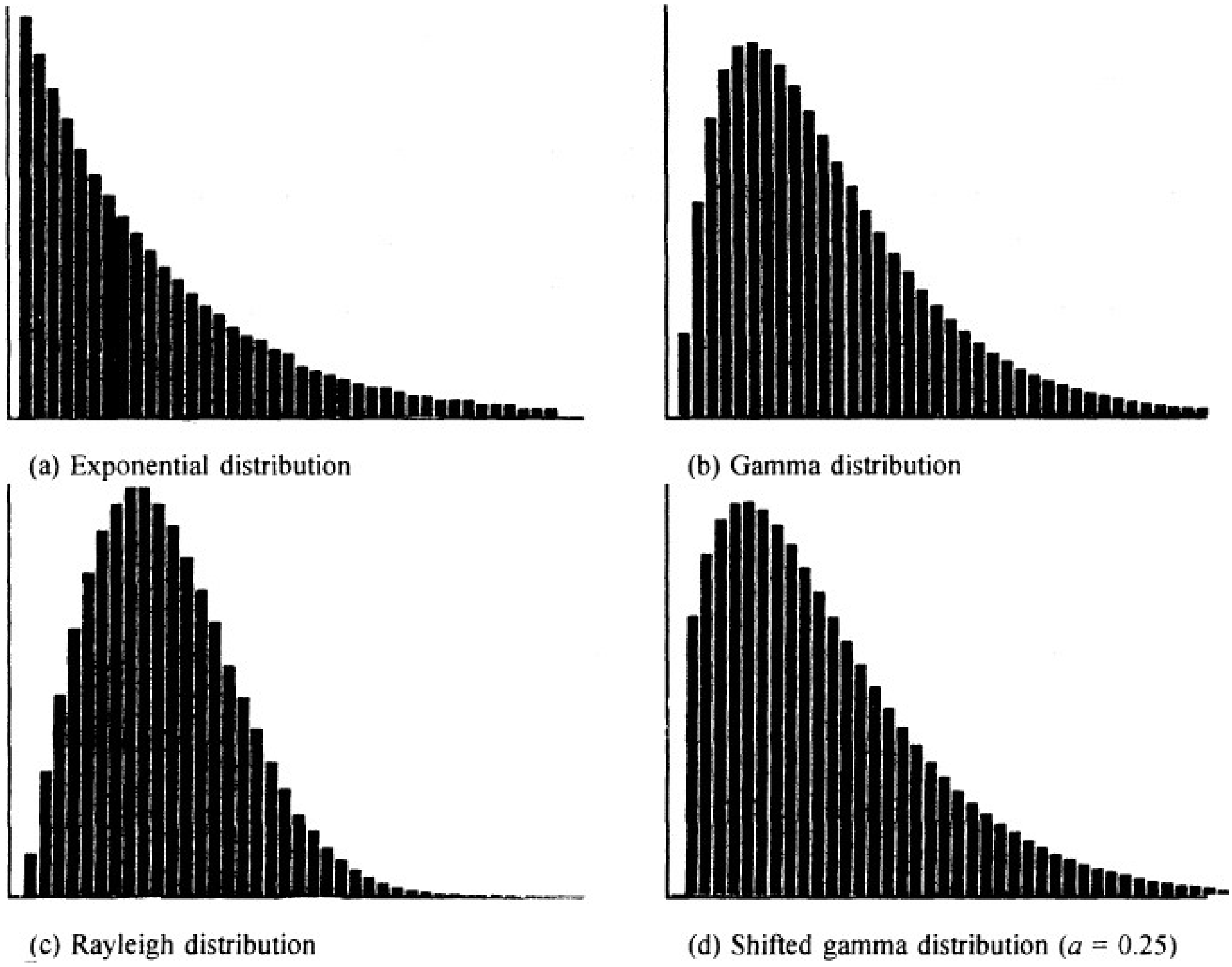


Figure 1 Statistical distributions (after Halpern, 1995)

Thus the stochastic control algorithm aims to provide an even service to all floors, where every landing call is given a fair consideration. This means that the landing call that has been waiting the longest should be given higher priority. The effect is to give a more consistent service to passengers; by trading the instant response calls to reduce long wait calls.

A stochastic[1] based traffic control system, named CGC was developed by Lim in 1983 and published (Barney and Dos Santos) in 1985 and implemented by at least one lift company (Godwin, 1986). It uses the principle that a landing (hall) call has to have waited a certain length of time before being considered for allocation (stops zero passenger wait times) and prioritises any call waiting over a high threshold time. The low and high thresholds are not fixed, but change to reflect demand by monitoring the average system response times.

What Lim proposed was subsequently analysed by Halpern (1992, 1993, 1995). Halpern showed that a classical traffic control system behaved as a Poisson process, but that computer based systems follow a shifted Gamma process, see Figure 1 (d). He also confirmed the premise of a finite (handling capacity) resource.

5.3 Hall Call Allocation[2]

5.3.1 Minimal Cost Functions

Calls are often allocated to a suitable lift using the concept of minimum cost, ie: a cost function[3]. This concept operates by performing a trial allocation to all available cars and allocating the call to

[1] The term "stochastic based", meaning "aim at a mark, guess".
[2] The term destination control, which is sometimes used is misleading. A lift traffic control system can only allocate a passenger's hall call to a suitable car, ie: Hall Call Allocation. The system cannot control the passenger's destination: that option belongs solely to the passenger.

the car presenting the lowest cost. There are criteria for selecting a suitable cost function. These can, for example, be based on either Quantity of Service, or Quality of Service, or both. In general terms, the Quantity of Service is a measure of the lift capacity consumed to serve a specific set of calls, indicated by the total of the journey times of all the cars. This could be minimised by keeping passengers waiting in a lobby until there were enough passengers to make a trip worthwhile. Airlines apply this principle. The Quality of Service is indicated by the average value of either the passenger waiting time or the passenger journey time (waiting time plus in-car travel time).

The minimisation of waiting time implies putting passengers into the first lift that arrives. This would result in no change from the usual procedure. The minimisation of the total car travel time implies using the smallest system capacity, which is equivalent to using the smallest possible number of cars. The result of this policy would be very large passenger waiting times, a result which would not be acceptable. This criterion alone is thus not suitable as a cost function. The minimisation of average passenger journey and waiting times are more acceptable objectives. Both times are interrelated and the minimisation of one might be achieved at the expense of the other. An accurate calculation of passenger journey time can only be achieved if passenger destinations are known at landing call registration time.

5.3.2 A new signalling system

The idea of destination buttons on the landing was first proposed by Leo Port (1961, 1968), but he only had relay logic in which to implement it and could not provide dynamic allocation, only fixed allocation. Installed in two buildings in Australia it functioned in one for some 20 years or more. A dynamic (ACA) system was first described by Closs in 1970, detailed by Barney & dos Santos in 1977 and partially implemented by a major lift company in 1990 (Schroeder, 1990c), when computer technology had caught up with the ideas. Now installed in many buildings, it has gained acceptance across the world as efficient. Most manufacturers have now applied the technique – some very badly.

Hall call allocation gives the opportunity to track every passenger through from registration to destination. This has great advantages during uppeak as passengers can be grouped to common destinations, as there are larger numbers of them. The individual waiting time may increase, the travel time may decrease, but there would be an overall reduction in journey time. During down peak there is no advantage as the destination floor is known. During reasonable levels of balanced interfloor traffic there is little advantage as most landing (hall) calls and car calls are not co-incident and car loading maybe one or two persons. However, during an uppeak with some down travelling traffic, or a down peak with some up travelling traffic, there are benefits. This leads to a conclusion that an optimum cost (money) system would have a full call registration station at the lobby and other principal floors and two button stations at all other floors. The control algorithm can go into "simple" mode, when dealing with the two button stations by knowing the direction and guessing the destination.

~O~O~O~

INTERLUDE

As so few people understand Hall Call Allocation and its derivative Adaptive Call Allocation, including most manufacturers it is worth an interlude to explain the basics.

I – 1The simple cost function

During an uppeak, the obvious cost function to implement with call allocation is journey time. This is because a waiting time allocation criterion would do no more than allocate every new call to the first available lift at the main terminal which possessed space capacity, in the same way as the collective-distributive algorithm. If journey time is the cost function, calls terminating at the same

[3] "Cost function" is optimal control theory terminology and its equivalent inverse, the "performance index", is sometimes quoted. Its converse is a "penalty function".

floor tend to be allocated to the same lift, hence reducing the number of stops per trip and the round trip time. The system handling capacity is increased and the main terminal floor more frequently served. However, a waiting passenger may not be allocated to board the first available lift, and this may produce increased waiting times. The overall effect is that better journey times are produced, in comparison to conventional algorithms, for the whole range of traffic intensities, but can result in longer waiting times. It is better to sacrifice some passenger waiting time and use passenger average journey time as the cost function. The maths is as follows.

Consider that a new call is to be allocated to a system of *L* lifts, each lift (*I*) with *N(I)* calls to answer and *JT(I)* accumulated journey time for the *N(I)* calls.

Assume that *NJT(K)* is the new accumulated journey time for *N(K) + 1* calls, when the new call is allocated to lift *K*. The average journey time for the complete set of calls is:

$$AJT = \frac{NJT(K)+\sum_{I=1,I\neq K}^{L} JT(I)}{1+\sum_{I=1}^{L} N(I)} \quad (1)$$

This can be written as:

$$AJT = \frac{NJT(K)-JT(K)}{1+\sum_{I=1}^{L} N(I)} + \frac{\sum_{I=1}^{L} JT(I)}{1+\sum_{I=1}^{L} N(I)} \quad (2)$$

As the two summations in Equation (2) do not depend on the allocation *K*, the minimisation of *AWT* only requires the minimisation of the term *NJT(K) - JT(K)*. This simplifies the evaluation of the cost function, as only this incremental cost is to be evaluated instead of the whole expression for *AWT*. The quantities *NJT(K)* and *JT(K)* are evaluated by simulation.

It should be noted that the incremental cost *NJT(K) – JT(K)* is made up of several terms. It includes the waiting and journey times for the new call and the increase in the waiting and journey times of calls already allocated to lift *K*, the extra passenger transfer time resulting from the new call, and any extra stops to pick up and discharge the new passenger.

I – 2 Average Journey Time with Maximum Waiting Time Constraint

A third type of cost function, proposed by Closs (1970), uses average journey time with a maximum waiting time constraint. It operates by costing each allocation against an average journey time cost function, but penalising any solution for which the waiting time of the new call exceeds a predefined value of maximum wait (*MWT*). The Adaptive Call Allocation algorithm operates as follows:

(1) Evaluate cost of allocation of the new landing (hall) call to lift 1:

COST(1) = NJT(1) - JT(1) (3)

(2) Compare the new call waiting time *NCWT(1)* with the predefined value *MWT*. If it is smaller than *MWT*, then *COST(1)* is not altered, but if it is greater a penalty is added to the cost:

COST(1) = COST(1) + penalty (4)

The penalty is made up of a fixed value added to a term proportional to the excess of waiting time above *MWT*. For example:

penalty = 300 + 10 *(NCWT(1) - MWT)* (5)

(3) Repeat the procedure from (1) for all lifts.

The effect of using a penalty is to force the elimination of the allocation to lifts with an existing high number of allocations from receiving another allocation, making it easier to select a more lightly loaded lift.

I – 3 Reduction in Number of Stops

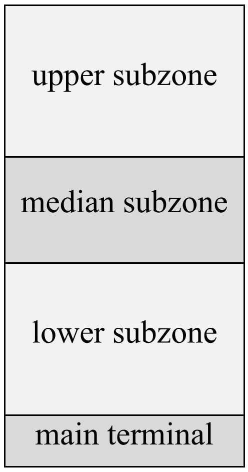
The "positive" concept of using a cost function as a performance index can be transposed into a "negative" concept of penalty functions in order to promote higher efficiency. An example of a penalty function is the rejection of an allocation which introduces an additional stop.

The call allocation algorithm causes calls requesting the same destination floors to be carried by the same lift. This has the effect of reducing the number of stops. However, in some cases the cost of allocating a new landing (hall) call to a lift already stopping at the calling landing (hall) or destination floor is marginally greater than the cost to allocate the call to another lift not stopping at either floor. Although the allocation is perfectly proper, it might be better not to allocate the new call to the lift with the lowest cost, as by not doing so capacity is reserved for future calls. To cater for this idea a penalty p% is introduced for each extra stop motivated by the new call. To prevent operation of this penalty under low traffic conditions, the penalty is made dependent on the incremental cost of the allocation and is proportional to car load.

$$\text{penalty} = \frac{p}{100} \times \text{incremental cost} \times \frac{\text{load}}{AC} \qquad (6)$$

where, AC is the actual car capacity and the load is measured as the average value of the number of passengers inside the lift, or queuing for service. The procedure improves performance for values of p up to 10%. For larger values of p the algorithm is self-defeating, as it produces less appropriate allocations.

I – 4 Dynamic Uppeak Sub-zoning



Uppeak sub-zoning is sometimes used by conventional group control systems to improve the uppeak handling capacity. Sub-zoning is very sensitive to where the zone partition is fixed and should ideally be adjusted for every traffic situation. As in practice a fixed partition is implemented, it cannot respond to the wide fluctuations found in arrival traffic patterns. Knowing the advantages of uppeak sub-zoning, and the adaptability of a computer implemented algorithm in coping with input traffic variations, a dynamic sub-zoning concept can be implemented in the ACA system. The building is divided into three sub-zones, as shown in the figure.

The lifts are divided into two subgroups, one for the lower sector and the other for the upper sector. No indication of this partition is given to the passengers. A newly registered landing (hall) call is allocated to a lift in the usual way, by evaluating the costs of the allocation of the call to every lift and choosing the allocation giving the lowest cost. However, during the evaluation of the cost, the allocation of a call registered for the lower subzone to a lift allocated to the upper subzone is penalised, and so is the allocation of a call with a destination in the upper subzone to a lift in the subgroup serving the lower subzone. The penalty, which is added to the cost of the allocation, is a function of the load of the two subgroups of lifts, and can be expressed as:

$$\text{penalty} = \left(1 + \frac{b}{100}\right) M \qquad (7)$$

where, M is a constant value and b measures the imbalance of lift loads between the upper and lower subgroups as a percentage of the highest subgroup lift load.

The fact that the loads of the two subgroups of lifts are taken into account contributes to equalise these loads. For example, the allocation of a call terminating at a floor in the lower subzone to a lift assigned to the upper subzone can be penalised by a quantity ranging from zero, if all the upper subzone lifts are idle, to $2M$, if the lower subzone lifts are idle.

A call registered to the median subzone can be allocated to either subgroup of lifts, with preference for the subgroup with the smallest load. The allocations to the lifts assigned to the heavier loaded sub-group are penalised by a quantity which equals the absolute value of b multiplied by M.

A correction mechanism allows this technique to deal with extremely unbalanced traffic destinations, as if excessive unbalance between the subgroup loads is detected, the subzone limits are automatically adjusted.

I – 5 Walking time

A further feature is necessary in the call allocation control algorithm. After registering the required destination floor and receiving a reply as to which lift will service the landing (hall) call, a passenger must walk to the lift. Thus, the allocation procedure must allow sufficient walking time for the passenger to reach the lift from the landing (hall) call station when allocating the landing (hall) call to a lift.

I – 6 Look ahead (*K*)

Although the mathematics suggest hall call allocations to up to *K* lifts (see equations (1) and (2)), in practice a "look ahead" (*K*) of from 2 to 4 only is practical. This also implies groups of six or more cars.

~O~O~O~

5.3.3 Conclusions on Modern Traffic Control Techniques

There are a number of other techniques, which can be applied to the conventional two button and hall call signalling systems. These include: expert systems (Qun et al., 2001); fuzzy logic (Ho and Robertson, 1994); dynamic programming (Chan and So, 1996); genetic algorithms (Siikonen et al., 2001; Miravete, 1999); knowledge based systems (Prowse et al., 1992); neural networks (Barney and Imrak, 2001) and optimal control (Closs, 1970). Many of the advanced control techniques employ complex mathematics and involved programming, which makes the practical implementation of the traffic controllers difficult. Also the proper understanding and correct adjustment on site by installation and service persons is doubtful and there is also an increased risk of system unreliability. Powell (2001) states "… *the added complexity involved in creating these (neural) networks and putting them into production could not be justified on the (slightly) expected gains in dispatching performance … over less complicated techniques*".

The use of any of the techniques during a dominant traffic flow, such as uppeak or down peak, is unlikely to improve traffic handling over a minimum cost algorithm. The provision of additional destination information, as with call allocation, is unnecessary during light traffic conditions, ie: balanced interfloor, and becomes most effective for heavy traffic situations, particularly uppeak. Then passengers for common destinations can be assembled to travel together. The technique improves the handling capacity for uppeak, but does not assist down peak or interfloor traffic handling (Barney 2000a, 2000b).

Once a computer is employed to implement the control strategy, the final algorithm is limited only by the imagination and ability of the program designer. For example, the search for a "bumpless" transfer of control strategy can be dealt with by having one algorithm able to adapt to changing traffic conditions. Also the Hall Call Allocation algorithm becomes the Adaptive Call Allocation by detecting when to switch from a waiting time to a journey time cost function. The stochastic algorithm CGC could easily be married to the Hall Call Allocation to restrict the allocation of landing calls to those that have been waiting for a threshold period of time. Learning algorithms can be added to "predict" outcomes and learn to improve the calculation processes such the estimated time to reach a landing (hall) call.

All these techniques allow the use of the underlying resource (handling capacity) more effectively for the benefit of all passengers. An added advantage is the systems become more consistent in their response to passenger demands.

6 COMPARISIONS

Readers are invited to examine Figure 2. The three main (pure) traffic demands are shown. Note how no one algorithm works for all three.

LEGEND COL = Collective, DSF = Dynamic sectoring with sub-zoning, FSO = Fixed up/down sectoring, FS4 = Fixed sectoring, priority timed, HCA = Hall Call Allocation, ACA = Adaptive Call Allocation

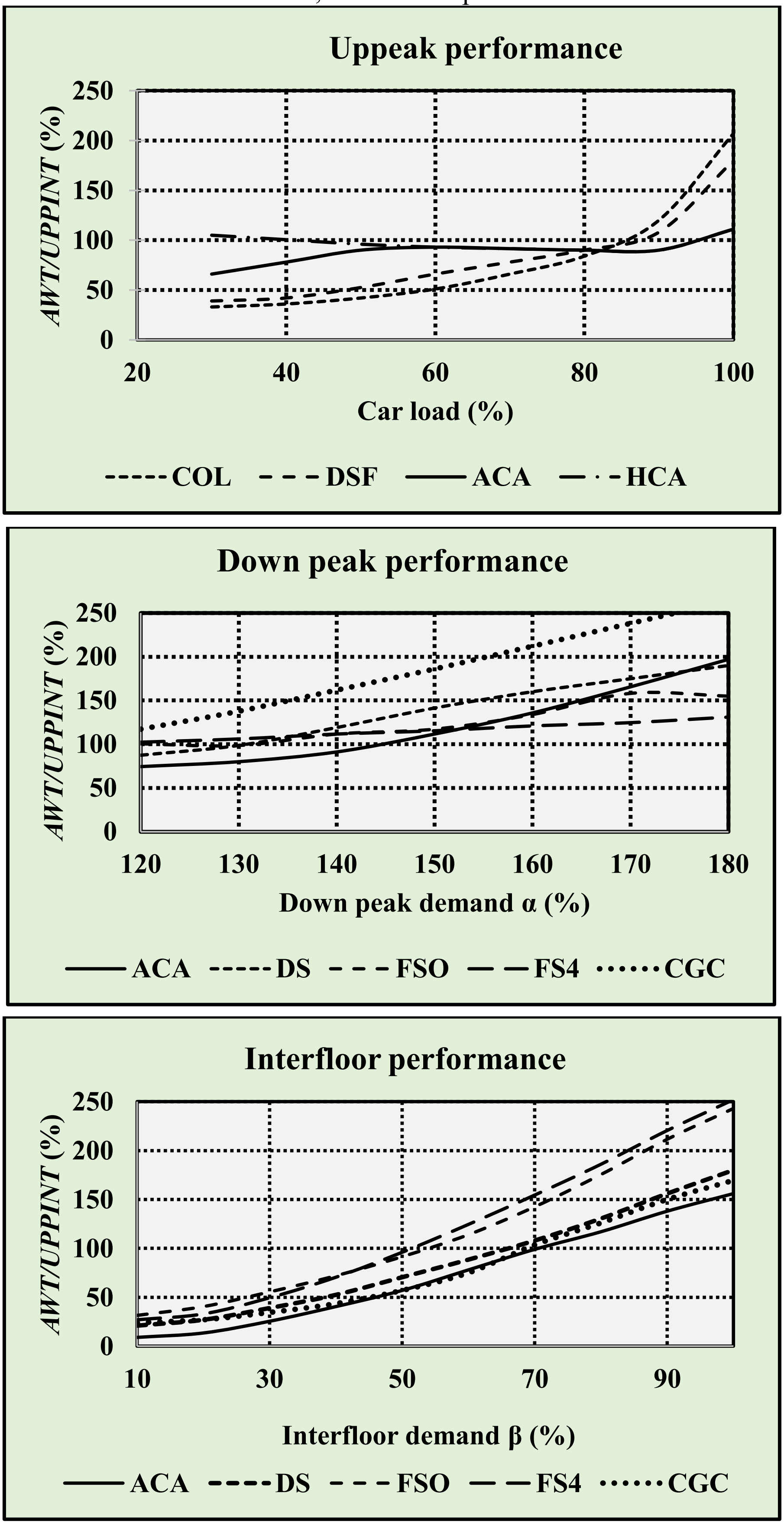


Figure 2 Comparison of traffic control algorithms for three traffic demands

ACKNOWLEDGEMENTS

This author acknowledges the work of her team at the University of Manchester. She started to study lift traffic control and design in January 1968, after she had graduated with an MSc in control theory (1962) and a PhD in control practice (1965). The principal members of her team included: David Closs, who researched simplex control (1968) and then analysed hall call allocation in 1970, Sergio dos Santos, who wrote an interactive lift simulation "app" known as LSD in 1972 and analysed several control algorithms in1974, Bill Swindles, who programmed up many of the algorithms in 1975. Other postgraduates included: Mahommed Moussalati (1974), Saideh Hirbod (1975), Nick Alexandris (1977), Jonathan Beebe (1980), Sinha Lim (1983), Lutfi Al-Sharif (1992), Erdem Imrak (2001). There were at least a score of undergraduates who also contributed.

WHERE CAN FURTHER INFORMATION BE FOUND?

As time passes the brain's neurons cease to connect and this author forgets. Fortunately anyone wishing the study lift traffic control in depth can do so as the author wrote it all down before she forgot. If anyone is seriously studying traffic control then they will already have two books Elevator Traffic Analysis Design & Control (1985) and the Elevator Traffic Handbook (2003). All the references referred in this paper are there. New researchers are directed to: Chapter 3 of Lift Traffic Analysis, Design and Control, Barney and dos Santos, 1/ed, 1977for legacy systems and a comprehensive description of Hall Call Allocation. Section 7.2 of Elevator (*sic*) Traffic Analysis, Design and Control, Barney and dos Santos,2/ed, 1985 for computer group control. And Pages 245–302 of Elevator (*sic*) Traffic Handbook, Barney, 2003.

BIOGRAPHICAL DETAILS

Dr Gina Barney, PhD, MSc, BSc, CEng, FIEE, HonFCIBSE is an independent vertical transportation consultant, working with the lift industry since 1968. She is an author, co-author and editor of over 120 papers and books.