forecasting has not progressed to the stage at which the behavior of individuals can be rationalized. Forecasting and modeling is still a crude process. It has been said that a forecast model is a muddled set of assumptions on an abstract piece of behavior.

The link model calibrated on the 49 links from the seven Maritime Province airports can be considered to produce marginally acceptable results. The statistical parameters associated with the forecasting model were significant at the 95 percent level.

The model, although not recommended for use in a detailed planning function, can be considered an acceptable departure point for the development of general aviation forecasting techniques for the Canadian Air Transport environment. The data supplied by Statistics Canada should be made available to other researchers so that development in this area can continue. The procedures for estimating commercial aviation activity are reasonably well advanced, and similar planning tools must become available for general aviation to enable the total air-transport mode to be evaluated on an ongoing basis.

ACKNOWLEDGMENT

We wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council and the cooperation of Statistics Canada in supplying data.

REFERENCES

1. Contracts. Canadian Aviation, March 1979.
2. Canadian Civil Aircraft Register. Transport Canada, Ottawa, 1978.

Publication of this paper sponsored by Committee on Aviation Demand Forecasting.

Air Traffic Control Network-Planning Model Based on Second-Order Markov Chains

NEIL W. POLHEMUS

A method designed to assess the impact of increased air traffic demand on flow rates in a network of en route air traffic control sectors is described. Given projected arrival and departure rates at airports within a given region, a second-order Markov-chain model is employed that has transition probabilities estimated from historical data. The technique is designed to serve as a planning tool and is demonstrated by using data from the New York Air Route Traffic Control Center.

The primary purpose of air traffic control (ATC) systems is to ensure the safe and efficient movement of air traffic. Given projected increases in traffic levels, it is important that a method be developed to predict the impact of additional demand on the system. In particular, the need to restructure existing sector boundaries depends on the distribution of flow in the current system.

As an example of the structure of ATC networks, the New York Air Route Traffic Control Center (ARTCC) consists of 32 sectors that cover the entire states of New Jersey and Delaware and parts of New York, Pennsylvania, Connecticut, and Maryland. The center controls en route traffic by dividing the low- and high-altitude airspace into sectors, each of which is handled by an individual controller who has an assigned communications frequency. Figure 1 shows the orientation of the low-altitude sectors. The high-altitude sectors are configured similarly and control traffic at or above 24 000 ft.

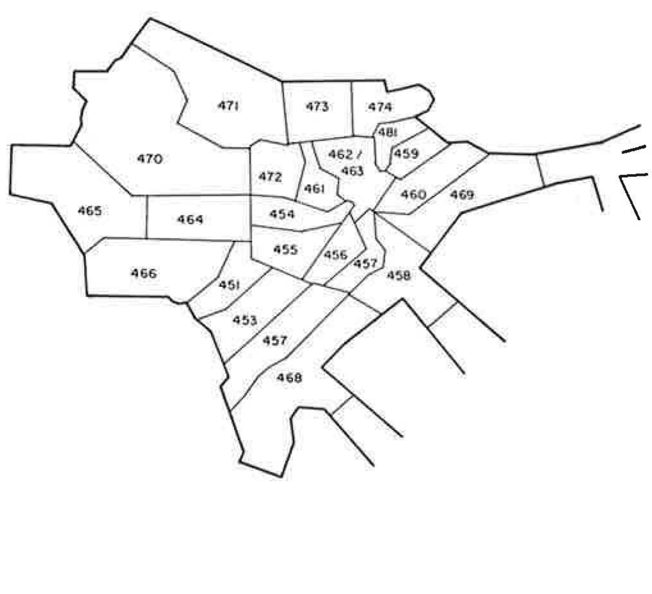
This paper describes a method designed to assess the impact of specified demand patterns on flow in the system. The approach is based on describing the sequences of sectors traversed by aircraft as second-order Markov chains. Although it is an approximation, the model provides a reasonable characterization of general system flow patterns with a simple-enough structure to allow for adequate parameter estimates. The need for a second-order Markov chain for terminal areas rather than a first-order chain as proposed earlier (1) is due to a lack of unidirectionality in the flow through many of the en route sectors.

The paper begins with a general formulation of the ATC system as a directed network and then considers characterizations of traffic generation and sector sequences. The use of the method in predicting system flows is discussed. Throughout, the techniques described are applied to the New York ARTCC.

NETWORK STRUCTURE

To represent an ATC system, let the sectors be rep-

Figure 1. New York ARTCC low-altitude sector control boundaries.



resented by the set of nodes $\{N_j, j = 1,2,..., m\}$. Feasible movement among the sectors is characterized by a set of arcs (A) between adjacent sectors. If we adopt the notation used by Ford and Fulkerson (2), Potts and Oliver (3), and others, the system is defined as the network G = [N;A].

For a network containing m sectors, A could consist of as many as m(m - 1) arcs. However, for most ATC systems only a very small subset of the possible arcs ever exists, since most pairs of sectors are not physically adjacent. To specify which arcs are present in a network, a node-node incidence matrix D of dimension (m x m) may be defined with element

$$d_{ij} = \begin{cases} 1 & \text{if flow is possible from node i to node j} \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

The node-node incidence matrix for the 32-sector New York en route system (Figure 2) consists of 152 arcs.

To represent arrivals to and departures from the system, it is convenient to construct a supersource (s) and a supersink (t) for the point of entry and exit, respectively, for all traffic in the network. Arcs are then constructed from s to the sectors in node N and from N to t. The actual source of traffic, however, is one of the airports in the region or an en route sector in another center. To represent the actual sources, we may construct the set of sources $\{s_i, i = 1, 2, ..., m_s\}$ and, in a similar manner, a set of sinks for the termini $\{t, k = 1, 2, ..., m_t\}$. This formulation is illustrated in Figure 3.

In the New York center, traffic was observed departing from and arriving at 12 separate airports in the region. During a 2-h sample, there were (a) 253 aircraft departures from airports within the region covered by the New York en route sectors, (b) 238 aircraft arrivals at airports within the region, and (c) additional en route traffic that had both source and terminus outside the region. For this system, $m_s = m_t = 13$, and there is one source and one terminus for each airport and an additional

Figure 2. Node-node incidence matrix for New York en route network.

SECTOR	451	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83
451	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
453	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
454	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
455	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
456	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
457	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
458	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1
459	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
460	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
461	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0	0	0	0	0
462	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
463	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0
464	1	0	1	1	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0
465	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
466	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
467	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
468	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
469	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
470	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
471	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0
472	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	0
473	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	0	0
474	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0
475	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	1	1	1	0	0	0
476	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
477	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1
478	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
479	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0
480	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	1	0	0
481	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1
482	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
483	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Figure 3. Schematic diagram of en route network traffic flow.

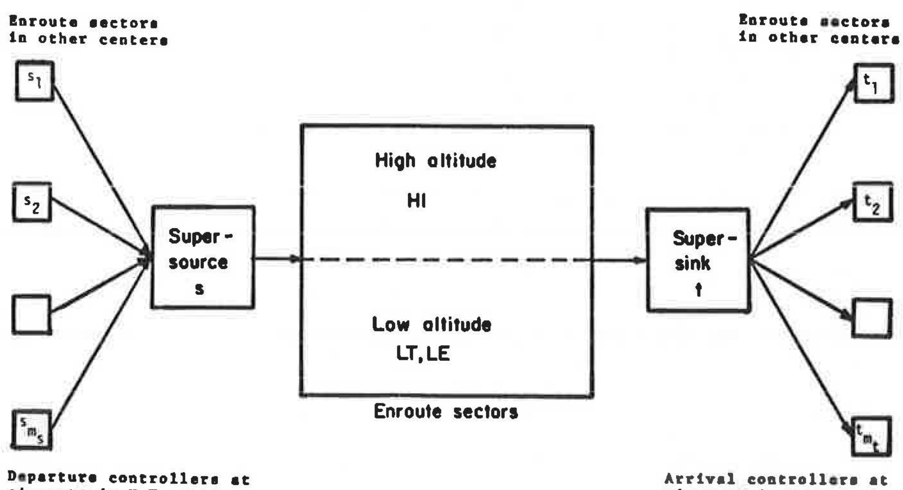


Table 1. Traffic to and from airports within New York region.

Airport	Code	Traffic Through En Route Network[a] Departures	Arrivals
Newark	EWR	53	40
John F. Kennedy	JFK	52	64
LaGuardia	LGA	58	56
Philadelphia	PHL	36	40
Atlantic City	ACY	9	5
Wilmington	ILG	5	9
Wilkes-Barre	AVP	3	2
Binghamton	BGM	3	5
Harrisburg	HAR	12	0
Allentown	ABE	8	3
Elmira	ELM	7	3
Westchester	HPN	7	11
Total		253	238

[a]Aircraft per 2-h sample.

source and terminus for en route sectors outside the New York center. Table 1 is a summary of traffic to and from airports within the region.

To characterize flow through the network, we then need to determine the following:

1. The manner in which aircraft are generated at the various sources,
2. The arcs over which they enter the en route network,
3. The sequence of sectors through which they proceed, and
4. Their final termini.

Given a finite set of data, achievement of such a characterization in a meaningful and consistent fashion raises various problems. In particular, one must ensure that the flow-conservation equations are satisfied yet allow for manipulation of system input at a sufficiently macroscopic level to provide a usable tool for the decision maker. The technique described below is designed to generate meaningful predictions of system flows in a manner suitable for planning purposes.

TRAFFIC SOURCES

To characterize sources of traffic in the network, let λ_i be the rate of traffic generated at source s_i. If

$$p_j = \text{prob [aircraft enters network over arc } (s,N_j)] \tag{2}$$

where prob represents probability and if

$$p_{j|i} = \text{prob [aircraft enters network over arc } (s,N_j) \text{ given generation at source i]} \tag{3}$$

then

$$f(s,N_j) = \sum_{i=1}^{m_s} p_{j|i}\,\lambda_i \tag{4}$$

For specified flow rates λ_i, estimation of the entry-arc flow rates $f(s,N_j)$ requires estimation of the conditional entry probabilities $p_{j|i}$.

To estimate these probabilities from a finite set of data, define a source-node entry count matrix C of dimension (m_s x m) with elements

$$c_{ij} = \text{number of aircraft generated at source } s_i \text{ that entered sector } N_j \tag{5}$$

Then the total number of aircraft generated at source s_i is given by

$$c_{i.} = \sum_{j=1}^{m} c_{ij} \quad i = 1, 2, \ldots, m_s \tag{6}$$

and the number of aircraft that enter the network over arc (s,N_j) is given by

$$c_{.j} = \sum_{i=1}^{m_s} c_{ij} \quad j = 1, 2, \ldots, m \tag{7}$$

The top 13 rows of Figure 4 show part of a matrix determined from the New York sample period in which 000 indicates entry from en route sectors in another center. The totals $c_{i.}$ and $C_{.j}$ are given in the last column and row of the figure.

If the selection of entry sector N_j for the arrivals from source i is independent, the probabilities $p_{j|i}$ are parameters of a multinomial distribution. The maximum-likelihood estimates are given by

$$\hat{p}_{j|i} = c_{ij}/c_{i.} \tag{8}$$

and the estimated entry-arc flow rates by

$$\hat{f}(s,N_j) = \sum_{i=1}^{m_s} \hat{p}_{j|i}\,\lambda_i \tag{9}$$

To test the assumption of independence in selection of entry sector, the selections of consecutive departures from the four major airports in the region were examined. By using a technique described by Anderson and Goodman (4), χ^2-test statistics indicated significant violation of the assumption only at LGA, at which successive departures tended to alternate between sectors 461 and 462.

Entry of aircraft to the network is completely determined by the set $\{p_{j|i};\ i = 1, 2, \ldots, m_s,\ j = 1, 2, \ldots, m\}$. The movement of aircraft after they enter the initial sector is the subject of the next section.

CHARACTERIZING SECTOR SEQUENCES

As aircraft move through an ATC system, they pass from sector to sector (from node to node) in sequences affected by their origin and destination. In a network of many sectors, the number of possible sequences is enormous, which makes the specification of the relative frequencies of all such sequences prohibitive. In order to reduce the complexity of the problem and still maintain the general patterns of network flow, an approach based on Markov chains will be presented.

To state the problem formally, consider a Markov chain with $M = m_s + m + m_t$ states, where the states represent the m_s-sources, m-en route sectors, and m_t-sinks, numbered in that order. Further, let $\{s_n(h),\ h = 0, 1, 2, \ldots\}$ be the sequence of sectors through which the nth aircraft passes, in which

$$s_n(0) = i \quad \text{if n th aircraft is departure from i th source} \tag{10}$$

$$s_n(h) = m_s + j \quad \text{if h th sector entered by n th aircraft is sector j} \quad 1 \leq h \leq m_n \tag{11}$$

$$s_n(h) = m_s + m + k \quad \text{if n th aircraft is arrival at k th sink} \quad h > m_n \tag{12}$$

where m_n is the number of network sectors in the sequence for the nth aircraft. Then $\{s_n(\cdot)\}$ is a realization from a Markov chain of unknown order.

In the above formulation, the nodes and sources

Figure 4. Counts of departures, arrivals, and transitions in New York network.

	EWR	JFK	LGA	PHL	ACY	ILG	AVP	BGM	HAR	ABE	ELM	HPN	000	451	453	···	TOTAL
EWR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		53
JFK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		52
LGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		58
PHL	0	0	0	0	0	0	0	0	0	0	0	0	0	22	1		36
ACY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		9
ILG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2		5
AVP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		3
BGM	0	0	0	0	0	0	0	0	0	0	0	0	C	0	0		3
HAR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		12
ABE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		8
ELM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		7
HPN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		7
000	0	0	0	0	0	0	0	0	0	0	0	0	0	6	28		373
451	0	0	0	0	0	1	1	0	0	0	0	0	24	0	3		55
453	0	0	0	5	1	1	0	0	0	0	0	0	6	2	0		55
454	1	0	18	7	0	0	0	0	0	0	0	0	3	20	0		55
455	30	0	3	0	0	0	0	0	0	0	0	0	4	3	0		42
456	0	30	0	0	0	0	0	0	0	0	0	0	3	2	0		42
457	0	6	16	0	0	0	0	0	0	0	0	0	3	0	0		26
458	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0		50
459	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0		32
460	0	28	0	0	0	0	0	0	0	0	0	0	15	0	0		47
461	4	0	0	0	0	0	0	0	0	0	0	0	1	0	0		53
462	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		40
463	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0		52
464	0	0	0	0	0	0	0	0	0	2	0	0	4	1	0		72
465	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0		31
466	0	0	0	20	0	1	0	0	0	0	0	0	18	3	0		56
467	0	0	0	8	3	5	0	0	0	0	0	0	18	0	9		56
468	0	0	0	0	1	1	0	0	0	0	0	0	24	0	1		40
469	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0		28
470	0	0	0	0	0	0	1	0	0	0	2	0	18	0	0		42
471	0	0	0	0	0	0	0	5	0	0	1	0	11	0	0		38
472	5	0	0	0	0	0	0	0	0	1	0	0	3	0	0		65
473	0	0	1	0	0	0	0	0	0	0	0	6	13	0	0		65
474	0	0	0	0	0	0	0	0	0	0	0	3	23	0	0		43
475	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0		65
476	0	0	0	0	0	0	0	0	0	0	0	0	3	0	12		36
477	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0		52
478	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0		36
479	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0		48
480	0	0	0	0	0	0	0	0	0	0	0	0	41	0	0		75
481	0	0	18	0	0	0	0	0	0	0	0	2	5	0	0		56
482	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0		18
483	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0		29
TOTAL	40	64	56	40	5	9	2	5	0	3	3	11	373	59	56	···	2126

are transient states, whereas the sinks are absorbing. Further, the nodes form a communicating class that is accessible from the sources, but the sources (nonreturn states) are not accessible from any states in the chain. The characterization of the sequences will thus involve state transition matrices of very special form. Although all sequences begin in one of the source states, the probability of ever returning to those states is zero. In describing the sequences, we state first the initial distribution of $s_n(0)$ and then discuss the state-transition probabilities.

The initial distribution of $s_n(0)$ has parameter set

$$\underline{\theta} = \{\lambda_1/\lambda_s, \lambda_2/\lambda_s, \ldots, \lambda_{m_s}/\lambda_s\} \tag{13}$$

where

$$\mathrm{prob}[s_n(0) = i] = \lambda_i/\lambda_s \tag{14}$$

$$\sum_{i=1}^{m_s} \lambda_i = \lambda_s \tag{15}$$

Thus the relative generation rates at the sources determine the probability distribution for $s_n(0)$ in a natural way.

To determine the movement of aircraft through the network, suppose that the sector sequences $\{s_n(\cdot)\}$ can be regarded as realizations of a Markov chain of order q. Then the distribution of $s_n(h)$ depends on the history of the sequence only through $s_n(h - 1)$, $s_n(h - 2)$, ..., and $s_n(h - q)$. To be more explicit, let

$$p_k = \mathrm{prob}[s_n(h) = k] \tag{16}$$

$$p_{jk} = \mathrm{prob}[s_n(h) = k | s_n(h - 1) = j] \tag{17}$$

$$p_{ijk} = \mathrm{prob}[s_n(h) = k | s_n(h - 1) = j, s_n(h - 2) = i] \tag{18}$$

Then, if the sequences are zero-order Markov chains,

$$p_k = p_{jk} = p_{ijk} \tag{19}$$

For first-order Markov chains,

$$p_k \neq p_{jk} = p_{ijk} \tag{20}$$

For second-order chains,

$$p_k \neq p_{jk} \neq p_{ijk} \tag{21}$$

The extension to higher orders is direct.

In studying sequences of sectors, it is therefore necessary to determine both the order of the chain and all relevant transition probabilities. This is most easily handled by defining a series of transition matrices $P^{(1)}$, $P^{(2)}$, ..., where $P^{(q)}$, the q-step transition matrix, has element

$$p_{jk}^{(q)} = \mathrm{prob}[s_n(h) = k | s_n(h - q) = j] \tag{22}$$

For a zero-order Markov chain,

$$P^{(q)} = [p_{jk}] \quad p_{jk} = p_k \tag{23}$$

For first-order Markov chains,

$$P^{(q)} = [P^{(1)}]^q \tag{24}$$

Of particular interest are both the limiting matrix,

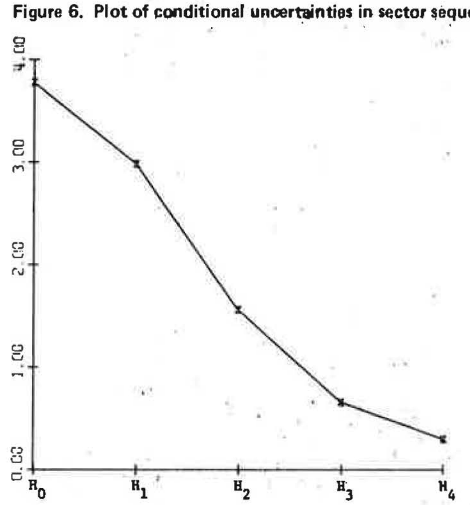
$$P^{\infty} = \lim_{q \to \infty} P^{(q)} \tag{25}$$

which can be used to determine the distribution of

Figure 5. Transition-count matrix for sector 451.

	000	453	455	464	466	479	ILG	AVP	TOT
000	1	0	1	0	2	0	0	1	5
453	0	0	0	2	0	0	0	0	2
454	15	2	0	0	0	0	0	0	17
455	3	0	0	0	0	0	0	0	3
456	1	0	0	0	0	0	0	0	1
464	0	1	0	0	0	0	0	0	1
466	1	0	0	0	0	0	1	0	2
PHL	0	0	0	6	5	10	0	0	21
TOT	21	3	1	8	7	10	1	1	52

Figure 6. Plot of conditional uncertainties in sector sequences.



exits from the system, given the entry sector, and the total-flow matrix

$$F^{(q)} = \sum_{r=1}^{q} \underline{\lambda} P^{(r)} \tag{26}$$

which measures the impact of given entries on sectors throughout the network for a given flow vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{m_s}, 0, 0, \ldots, 0)$ of dimension $(1 \times M)$.

Higher-step transition matrices determine the accessibility of sectors in the network. Since all flow originates at a source and ends at a sink, the only elements of the matrix that do not converge to zero as q becomes large are those that correspond to the source rows and sink columns. Let the limiting values of these elements be given by

$$e_{ik} = \lim_{q\to\infty} p_{i}^{(q)}{}_{,\,k\text{-}m_s\text{-}m} \quad i = 1, 2, \ldots, m_s \quad k = 1, 2, \ldots, m_t \tag{27}$$

Then the sink-attraction rates $\mu_1, \mu_2, \ldots, \mu_{m_t}$ are related to the source-generation rates by

$$\mu_k = \sum_{i=1}^{m_s} \lambda_i e_{ik} \quad k = 1, 2, \ldots, m_t \tag{28}$$

This link between entry and exit rates is an important consideration in attempting to estimate system flows, given projected levels of both arrivals and departures at airports in the region. It is discussed more fully in the next section.

The problem of estimating transition probabilities in Markov chains has been studied by several authors (4-6). Suppose that there are available records of c independent sector sequences, each of which is assumed to be a random observation from a qth-order Markov chain with M states. Let n_{jk} be the number of times an aircraft enters state k from state j, n_{ijk} be the number of times an aircraft enters state k from state j after it has entered state j from state i, and so forth. If we assume stationary transition probabilities, n_{jk} forms a set of statistics sufficient for the state-transition probabilities. For a second-order Markov chain, n_{ijk} is sufficient. The results can be generalized to higher-order chains.

The maximum-likelihood estimates of the transition probabilities depend on the order of the Markov chain. For a second-order chain, $s_n(0)$ and $s_n(1)$ are assumed to be nonrandom, whereas $s_n(k)$, $k \geq 2$, are assumed to be random variables. Then the maximum-likelihood estimates of the transition probabilities in Equations 16-18 are given by

$$\hat{p}_{ijk} = n_{ijk} / \sum_{l=1}^{M} n_{ijl} \tag{29}$$

$$\hat{p}_{jk} = \sum_{i=1}^{M} n_{ijk} / \sum_{i=1}^{M} \sum_{l=1}^{M} n_{ijl} \tag{30}$$

$$\hat{p}_{k} = \sum_{i=1}^{M} \sum_{j=1}^{M} n_{ijk} / \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{l=1}^{M} n_{ijl} \tag{31}$$

Note that, since $s_n(0)$ and $s_n(1)$ are assumed to be nonrandom, the estimates of the transition probabilities involve summations over n_{ijk} rather than the direct use of n_{jk} and n_k (the results are not equivalent).

For the New York en route network, transition-count matrices that use n_{ijk} were obtained for each of the 32 sectors. Figure 5 is the matrix obtained for one of the sectors. The sector shown was evidently handling traffic that departed from Philadelphia (PHL).

To determine the order of Markov chain appropriate for a given set of data, a likelihood-ratio test was derived by Anderson and Goodman (4). The technique, however, can be applied effectively to ATC sector sequences only if the number of sectors is small and the number of observed sector sequences is quite large. For other situations, a graphical technique based on information theory given by Chatfield (5) (which can be related to the likelihood-ratio test) is all that the data will support. The technique involves plotting the conditional uncertainties about the next sector that an aircraft will enter if we are only given knowledge of its current sector, of the previous sector, of the two previous sectors, and so forth. The reduction in conditional uncertainties as more and more of the past is known helps to indicate the order of Markov chain necessary to characterize the sequences.

Figure 6 is a plot of the estimated conditional uncertainties in the sector sequences made by using all observed quadruplets in the sample. From the New York data, $N_4 = 777$ quadruplets were tabulated. The following formulas were used to calculate the conditional uncertainties:

$$H_0 = \log 44 \quad \text{(44 states in chain)} \tag{32}$$

$$H_1 = \log N_4 - N_4^{-1} \sum n_{i...} \log n_{i...} \tag{33}$$

$$H_2 = N_4^{-1} \left(\sum_i n_{i...} \log n_{i...} - \sum_{i,j} n_{ij..} \log n_{ij..} \right) \tag{34}$$

Table 2. Specified source-generation rates and sink-attraction rates observed and computed by model, for airports.

Airport	Specified Source-Generation Rate[a]	Sink-Attraction Rate[a] Computed	Observed
EWR	26.5	20.6	20.0
JFK	26.0	38.5	32.0
LGA	29.0	29.8	28.0
PHL	18.0	20.8	20.0
ACY	4.5	2.8	2.5
ILG	2.5	5.3	4.5
AVP	1.5	1.1	1.0
BGM	1.5	3.3	2.5
HAR	6.0	0.0	0.0
ABE	4.0	1.6	1.5
ELM	3.5	1.6	1.5
HPN	3.5	5.3	5.5
000	186.5	182.3	186.5
Total	313.0	313.0	305.5

[a]Aircraft per hour.

Table 3. Observed and computed sector-flow rates computed by model, for sectors.

Sector	Sector-Flow Rate[a] Computed	Observed	Sector	Sector-Flow Rate[a] Computed	Observed
451	29.5	27.5	469	13.9	14.0
453	28.1	27.5	470	22.7	21.0
454	29.3	27.5	471	24.3	19.0
455	20.1	21.0	472	34.3	32.5
456	21.9	21.0	473	32.0	32.5
457	12.8	13.0	474	21.3	21.5
458	23.5	25.0	475	33.5	32.5
459	15.2	16.0	476	19.4	18.0
460	25.1	23.5	477	24.8	26.0
461	25.5	26.5	478	18.1	18.0
462	20.8	20.0	479	26.0	24.0
463	25.6	26.0	480	38.0	37.5
464	36.9	36.0	481	26.7	28.0
465	13.4	15.5	482	7.8	9.0
466	31.0	28.0	483	17.1	14.5
467	28.4	28.0	Total	765.9	750.0
468	18.9	20.0			

[a]Aircraft per hour.

$$H_3 = N_4^{-1} \left(\sum_{i,j} n_{ij..} \log n_{ij..} - \sum_{i,j,k} n_{ijk.} \log n_{ijk.}\right) \quad (35)$$

$$H_4 = N_4^{-1} \left(\sum_{i,j,k} n_{ijk.} \log n_{ijk.} - \sum_{i,j,k,l} n_{ijkl} \log n_{ijkl}\right) \quad (36)$$

The sharp drop from H_0 to H_2 shows the importance of knowing the current sector when determining the next. The drop from H_2 to H_3 is almost as sharp, which indicates significant information in the previous sector. The drop from H_3 to H_4 may or may not be significant, but it does not appear to be so important as the earlier drops. No rigorous statistical tests were performed because of the large number of states in the chain and the consequently small number of counts for all observed pairs, triplets, and higher sequences during the 2-h sample period.

On the basis of the above analysis, it appears that second-order Markov chains are sufficient to describe the patterns observed in the sector sequences. The maximum-likelihood estimates of p_{ijk} can thus all be developed from the transition-count matrices by means of

$$\hat{p}_{ijk} = n_{ijk} / \sum_{l=1}^{c} n_{ijl} \quad (37)$$

where c is the number of columns in the matrix.

Traffic in the network is then completely described by arrival rates λ_i, conditional entry probabilities $p_{j|i}$, and transition probabilities p_{ijk}. The next section considers the use of such a formulation in predicting network flow patterns.

APPLICATION OF MODEL

To use the above method to predict sector flows, the arrival-rate parameters λ_i are specified and an arc-flow matrix $F^{(1)}$ of dimension (M x M) is formed from Equation 26 with $M = m_s + m + m_t$. The elements of the matrix are

$$f_{ij}^{(1)} = \begin{cases} p_{j-m_s|i}\,\lambda_i & \begin{cases} i = 1, 2, \ldots, m \\ j = m_s + 1, m_s + 2, \ldots, m_s + m \end{cases} \\ 0 & \text{otherwise} \end{cases} \quad (38)$$

After q transitions, the arc flows are given by

$$f_{jk}^{(q)} = \sum_{i=1}^{M} f_{ij}^{(q-1)} p_{ijk} \quad j = 1, 2, \ldots, M; \; k = 1, 2, \ldots, M \quad (39)$$

After many transitions,

$$\lim_{q \to \infty} f_{jk}^{(q)} = \begin{cases} \mu_{k-m_s-m} & j = k = m_s + m + 1, m_s + m + 2, \ldots, m_s + m + m_t \\ 0 & \text{otherwise} \end{cases} \quad (40)$$

In other words, all flow eventually reaches and remains in one of the sinks. Further, total sector flows are given by

$$f_j = \lim_{q \to \infty} \sum_{r=1}^{q} \sum_{i=1}^{M} f_{ij}^{(r)} \quad \text{for } m_s < j \leq m_s + m \quad (41)$$

Tables 2 and 3 show the observed sink-attraction and sector-flow rates computed by the above method with source-generation rates λ_i set equal to that estimated from the sample data. Good correspondence between the observed and computed rates resulted. After q = 10 iterations, 99.99 percent of the flow had reached a sink and there was little change in computed rates beyond that point. Any of various stopping criteria could be used to stop the iterative process.

To demonstrate the use of the model as a planning tool, the rate of traffic that departed from Newark (EWR) was increased by 50 percent, which yielded the computed flow rates shown in Tables 4 and 5. Increases in sector-flow rates of more than 10 percent occurred in sectors 454, 472, and 480. Although most of the additional traffic terminated in the en route sink, a certain proportion became

Table 4. Flow rates for EWR departures increased by 50 percent, for airports.

Airport	Specified Source-Generation Rate[a]	Computed Sink-Attraction Rate[a]	Percentage of Change[b]
EWR	39.75	20.6	0.2
JFK	26.0	38.6	0.3
LGA	29.0	29.9	0.5
PHL	18.0	21.4	2.7
ACY	4.5	2.9	4.2
ILG	2.5	5.3	0.1
AVP	1.5	1.1	0.0
BGM	1.5	3.6	8.2
HAR	6.0	0.0	0.0
ABE	4.0	1.6	4.6
ELM	3.5	1.6	0.0
HPN	3.5	5.6	4.7
000	186.6	194.1	6.4
Total	326.3	326.3	4.2

[a]Aircraft per hour. [b]Compared with computed rates in Table 2.

Table 5. Flow rates for EWR departures increased by 50 percent, for sectors.

Sector	Computed Sector-Flow Rate[a]	Percentage of Change[b]	Sector	Computed Sector-Flow Rate[a]	Percentage of Change[b]
451	31.7	7.6	468	18.9	0.0
453	28.3	0.9	469	14.2	2.3
454	33.9	15.6	470	23.6	4.1
455	20.2	0.2	471	24.7	1.7
456	23.0	5.1	472	41.3	20.4
457	12.8	0.0	473	33.5	4.7
458	23.6	0.2	474	21.8	2.3
459	15.2	0.0	475	36.8	9.7
460	25.2	0.4	476	20.5	5.8
461	28.0	9.8	477	25.4	2.6
462	21.3	2.4	478	19.7	8.5
463	26.4	3.1	479	26.1	0.5
464	38.0	2.8	480	42.2	11.2
465	14.0	5.1	481	26.9	0.5
466	32.0	3.3	482	7.8	0.0
467	28.6	0.7	483	17.6	2.5

[a]Aircraft per hour. [b]Compared with computed rates in Table 3.

Table 6. Results of combined forward and backward analyses, for airports.

Source and Sink	Target Rate[a]		Model-Specified Rate[a]	
	Source Generation	Sink Attraction	Source Generation	Sink Attraction
EWR	26.5	20.0	13.11	19.53
JFK	26.0	32.0	12.22	12.28
LGA	29.0	28.0	14.04	12.77
PHL	18.0	20.0	9.62	9.39
ACY	4.5	2.5	2.43	1.06
ILG	2.5	4.5	1.42	1.81
AVP	1.5	1.0	0.65	0.45
BGM	1.5	2.5	0.79	0.83
HAR	6.0	0.0	3.24	0.00
ABE	4.0	1.5	1.54	0.69
ELM	3.5	1.5	1.79	0.68
HPN	3.5	5.5	1.71	2.80
000	179.0	186.5	95.69	95.00
Total	305.5	305.5	158.25	147.29

[a]Aircraft per hour.

arrivals at other airports in the region. Although this is consistent with the observed behavior of the system, it points out the interdependencies between source-generation rates and sink-attraction rates.

Although departure rates from airports can be easily manipulated, given the above formulation, arrival rates cannot. Given a single source for all entries from outside the region, it is not possible to set the arrival rate at each of the airports. However, if the role of sources and sinks is reversed and the network is run backward, the sink-attraction rates (μ_k) can be set as desired and the source-generation rates determined from the analysis.

To perform a backward analysis, the following adjustments are necessary. Conditional exit probabilities must be estimated by

$$\hat{p}_{j|k} = c_{jk}/c_{.k} \tag{42}$$

where c_{jk} is the number of aircraft attracted to sink t_k directly from sector N_j. Transition probabilities must be estimated by

$$\hat{p}_{ijk} = n_{ijk}/\sum_{i=1}^{c} n_{ijk} \tag{43}$$

where c is the number of rows in the transition-count matrix for sector j. The initial flow vector must be estimated by

Table 7. Results of combined forward and backward analyses, for sectors.

Sector	Sector-Flow Rate[a]		Sector	Sector-Flow Rate[a]	
	Computed	Observed		Computed	Observed
451	29.1	27.5	468	19.3	20.0
443	27.3	27.5	469	14.4	14.0
454	27.6	27.5	470	21.6	21.0
455	20.6	21.0	471	21.1	19.0
456	19.7	21.0	472	34.1	32.5
457	12.4	13.0	473	30.9	32.5
458	24.1	25.0	474	20.5	21.5
459	16.5	16.0	475	32.7	32.5
460	22.9	23.5	476	17.6	18.0
461	25.5	26.5	477	24.6	26.0
462	19.4	20.0	478	16.4	18.0
463	24.8	26.0	479	25.3	24.0
464	36.0	36.0	480	37.0	37.5
465	14.4	15.5	481	26.0	28.0
466	29.4	28.0	482	6.8	9.0
467	28.0	28.0	483	14.7	14.5

[a]Aircraft per hour.

$$f_{jk}^{(1)} = \begin{cases} p_{j-m_s/k} & \begin{cases} j = m_s+1, m_s+2, \ldots, m_s+m \\ k = 1, 2, \ldots, m_t \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

$$f_{ij}^{(q)} = \sum_{k=1}^{M} f_{jk}^{(q-1)} p_{ijk} \tag{45}$$

In practice, a decision maker who wishes to predict sector flows will most likely want to specify both arrival rates and departure rates. To do so, the forward and backward analyses may be combined. Suppose the desired (target) source-generation and sink-attraction rates are $\{\lambda_i^{(t)},\ i = 1, 2, \ldots, m_s\}$, and $\{\mu_k^{(t)},\ k = 1, 2, \ldots, m_t\}$. Then the total generations and attractions that result from the sum of the forward and backward analyses will equal their targeted values if the model analysis rates $\lambda_i^{(s)}$ and $\mu_k^{(s)}$ are set to satisfy

$$\mu_k^{(s)} + \sum_{i=1}^{m_s} \lambda_i^{(s)} e_{ik}^{(\text{forward})} = \mu_k^{(t)} \quad k = 1, 2, \ldots, m_t \tag{46}$$

$$\lambda_i^{(s)} + \sum_{k=1}^{m_s} \mu_k e_{ik}^{(\text{backward})} = \lambda_i^{(t)} \quad i = 1, 2, \ldots, m_s \tag{47}$$

in which e_{ik} is determined from Equation 27 for the forward analysis and in a similar manner for the backward analysis. Note that, since the above set of equations does not have a unique solution, only $m_s + m_t - 1$ rates can be specified separately; the other rate is determined by the fact that the sum of the source-generation rates must equal the sum of the sink-attraction rates.

Tables 6 and 7 summarize the results of applying the above procedure to the New York center at the observed source-generation and sink-attraction rates (the en route source rate has been adjusted down to make the sum of the source and sink rates equal). Again, close agreement with observed sector flows is evident. To predict sector flows under projected increased traffic rates in and out of the region, the decision maker need only select new target values and repeat the above procedure.

CONCLUSION

The characterization of ATC network flows by using second-order Markov chains provides a technique for predicting sector flows that, although it is relatively easy to apply, can readily indicate potential areas of excess traffic loading. Based on empirical data, the method preserves general patterns of network flow without specifying the actual geometry of each aircraft's flight. For ATC network planners, such a method for predicting traffic distribution could provide a useful tool for ensuring safe and efficient movement of air traffic.

ACKNOWLEDGMENT

I wish to acknowledge the assistance and support of the Federal Aviation Administration's Technical Center, Atlantic City, New Jersey, in pursuing this work. Portions of the study were supported by the U.S. Department of Transportation.

REFERENCES

1. N.W. Polhemus. Modeling Aircraft Flow in Air Traffic Control Systems. Proc., Transportation Research Forum, 1974, Vol. 15, pp. 526-539.
2. J.R. Ford and D.R. Fulkerson. Flows in Networks. Princeton Univ. Press, Princeton, NJ, 1962.
3. R.B. Potts and R.M. Oliver. Flows in Transportation Networks. Academic Press, New York City, 1972.
4. T.W. Anderson and L.A. Goodman. Statistical Inference About Markov Chains. Annals of Mathematical Statistics, Vol. 28, No. 1, 1957, pp. 89-110.
5. C. Chatfield. Statistical Inference Regarding Markov Chain Models. Applied Statistics, Vol. 22, No. 1, 1973, pp. 7-20.
6. G.T. Duncan and L.G. Lin. Inference for Markov Chains Having Stochastic Entry and Exit. Journal of the American Statistical Association, Vol. 67, No. 340, 1972, pp. 761-767.

Publication of this paper sponsored by Committee on Airfield and Airspace Capacity and Delay.

Analyzing Ticket-Choice Decisions of Air Travelers

SCOTT D. NASON

This paper examines the nature of the problem that faces air travelers confronted with choosing from among a variety of air fares, each associated with different service characteristics, and the problem of forecasting these decisions. A theoretical framework is developed that views the problem at the level of the individual traveler; the ticket-type choice is expressed in terms of the individual's socioeconomic characteristics, the characteristics of the trip in question, and the level of service associated with each available alternative. Logit models are suggested as the preferable functional form on the basis of theoretical and computational grounds, and the properties of logit models are briefly described. A pilot application of the method is presented for a two-alternative situation (full fare versus standby) by using a small sample of interview data collected from departing passengers at Boston's Logan Airport. A calibrated model is presented that demonstrates a statistically significant relationship between the ticket-type choice and the fare, fare differential, trip purpose, automobile ownership (as a proxy for income), and the passenger's perception of the delays that may be expected if flying standby. This application merely demonstrates a method and could easily be improved by using the airlines' onboard surveys for estimation.

Events during the last few years have substantially altered the air-travel-demand forecasting requirements of the individual airlines. Until recently, the number of different fares available was quite limited, and differences among the fare packages available from individual airlines were almost nonexistent. In this environment, the crucial requirements were for an aggregate estimate of the size of an individual city-pair market, which may or may not have been based on the level of service available in that market, and a carrier's share of the total, based on a measure of that carrier's frequency share (or a more-sophisticated model that took into account the timing of those flights).

With the advent of deregulation, pricing freedom has emerged as a major factor that influences air-travel-demand decisions. Discount fares have stimulated new travel. Just as important to airline marketing departments is the impact on the yield per passenger or per passenger mile, which is affected by the passenger's choice of ticket type, as well as the impact of discount fares on the passenger's carrier-choice decisions. Passengers have always made minor distinctions between carriers on the basis of food, cabin attendants, or advertisements, but more and more there is a tangible economic incentive to choose one carrier over another. Examples include the unlimited-mileage tickets available on Eastern and Allegheny; the straight price reductions offered in some markets by National, Braniff, Texas International, World, and Transamerica (among others); and half-price coupon offers from United and American.

This paper examines the nature of these new decisions that face air travelers and proposes a technique that should prove useful in analyzing the passenger's ticket-choice decisions. The ticket-type choice is viewed within the context of the entire trip-planning process. Each individual's decision is based on that person's characteristics and the characteristics of each available alternative--travel time, price, reservation, length-of-stay restrictions, etc. This type of problem has exact parallels in other decision-making processes, and the modeling of personal preferences, which is well developed elsewhere, is adapted to the problem at hand.

TRIP-PLANNING PROCESS

There are several decisions involved in planning a trip by air; these include (a) a decision to travel somewhere, (b) a choice of destination or destinations and departure and return times, (c) a decision to fly in preference to other modes of travel, and (d) a selection of the least-expensive and most-convenient flight and ticket combination. For many trips, some of these decisions may be trivial or made simultaneously with other decisions. The first three (or even all four) are likely to be made simultaneously and without much