

2

**NAVAL POSTGRADUATE SCHOOL
Monterey, California**

AD-A274 962



**S DTIC
ELECTE
JAN 27 1994
A**



THESIS

**TWO OPTIMIZATION MODELS FOR
SMALL-SCALE ROUTING OF MILITARY UNITS
IN A ROAD NETWORK**

by

Johann Thoma

September, 1993

Thesis Advisor:

Richard E. Rosenthal

Approved for public release; distribution is unlimited.

94-02742



94 1 26 184

REPORT DOCUMENTATION PAGE

Form Approved OMB Np. 0704

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 1993, September	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE TWO OPTIMIZATION MODELS FOR SMALL-SCALE ROUTING OF MILITARY UNITS IN A ROAD NETWORK			5. FUNDING NUMBERS	
6. AUTHOR(S) Johann Thoma				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (maximum 200 words) This thesis describes two integer programming models that are developed to support movement planners in optimally routing military units through a road network with minimal delays. The formulations are based on a multicommodity time-expanded network flow problem. The first model (called <i>model close column</i>) is designed for use when all vehicles of a military unit are required to move together over one route. The second model (called <i>model separate column</i>) allows a military unit to be split into groups of vehicles which may use different routes and start at irregular times. Variants of the models allow all units to move at just one speed, or allow units to move at one of two speeds, fast or slow. Both models are designed with the German Army in mind, but could be used with minor modifications by other Armies. The models are implemented via the General Algebraic Modeling System (GAMS). A hypothetical peacetime scenario based upon completion of a German brigade live exercise is used to test both models. In the test data, five military units are specified such that their movements inside an area of 24 by 12 km would conflict without careful planning. Depending on the model and whether movements are planned with different speeds, these five units are routed with solution times from 14 sec to 131 sec on a 486/33 MHz personal computer with the ZOOM solver. The advantage of movements at different speeds is clearly demonstrated. Evaluation of all results shows that they are accurate and that planners can be effectively supported by the models.				
14. SUBJECT TERMS Multicommodity, Time-Expanded, Dynamic, Transportation Network, Optimal Routing.			15. NUMBER OF PAGES 82	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

NSN 7540-01-280-3500

Standard Form 298 (Rev. 2-89)

Prescribed by ANSI Std. Z39-18

Approved for public release; distribution is unlimited.

TWO OPTIMIZATION MODELS FOR
SMALL-SCALE ROUTING OF MILITARY UNITS
IN A ROAD NETWORK

by

Johann Thoma
Captain, Federal Republic of Germany Army
Diplom-Ingenieur Univ, University of Munich, 1980

Submitted in partial fulfillment
of the requirements for the degree of

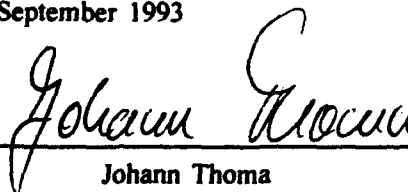
MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

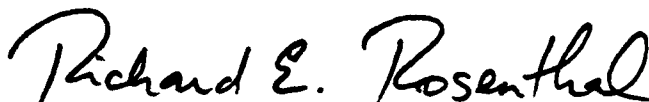
NAVAL POSTGRADUATE SCHOOL

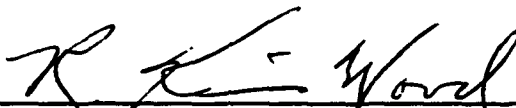
September 1993


Author:


Johann Thoma

Approved by:


Richard E. Rosenthal, Thesis Advisor


R. Kevin Wood, Second Reader


Peter Purdue, Chairman
Department of Operations Research

Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

ABSTRACT

This thesis describes two integer programming models that are developed to support movement planners in optimally routing military units through a road network with minimal delays. The formulations are based on a multicommodity time-expanded network flow problem. The first model (called *model close column*) is designed for use when all vehicles of a military unit are required to move together over one route. The second model (called *model separate column*) allows a military unit to be split into groups of vehicles which may use different routes and start at irregular times. Variants of the models allow all units to move at just one speed, or allow units to move at one of two speeds, fast or slow. Both models are designed with the German Army in mind, but could be used with minor modifications by other Armies. The models are implemented via the General Algebraic Modeling System (GAMS). A hypothetical peacetime scenario based upon completion of a German brigade live exercise is used to test both models. In the test data, five military units are specified such that their movements inside an area of 24 by 12 km would conflict without careful planning. Depending on the model and whether movements are planned with different speeds, these five units are routed with solution times from 14 sec to 131 sec on a 486/33 MHz personal computer with the ZOOM solver. The advantage of movements at different speeds is clearly demonstrated. Evaluation of all results shows that they are accurate and that planners can be effectively supported by the models.

TABLE OF CONTENTS

I. INTRODUCTION	1
A. PLANNING MOTOR MOVEMENTS	1
1. Types of Military Motor Movements	2
2. Techniques of Movement Planning	3
B. PROBLEM STATEMENT AND SCENARIO	6
C. DESCRIPTION OF THE MODEL	7
D. PREVIOUS STUDIES	8
II. MULTICOMMODITY TIME-EXPANDED NETWORK	10
A. SINGLE AND MULTICOMMODITY NETWORKS	10
B. TIME-EXPANDED NETWORK	11
III. MODEL DEVELOPMENT AND FORMULATION	14
A. GENERAL ROUTING CONSTRAINTS	14
B. MODEL ASSUMPTIONS	16
1. Assumptions on Units	16
2. Assumptions on Road Network	17
C. MODEL FORMULATION	17

1. Formulation for Model Close Column (MCC)	18
2. Formulation for Model Separate Column (MSC)	25
IV. MODEL ENHANCEMENT	30
A. MOVEMENTS WITH DIFFERENT SPEEDS	30
B. ENHANCED MODEL FORMULATION	31
1. Formulation for Model Close Column Two Speed (MCCTS) ...	31
2. Formulation for Model Separate Column Two Speed (MSCTS) .	34
V. PRE-OPTIMIZATION	35
A. USER INPUT	35
B. DATA DERIVATION	36
1. Derivation of $SHROUTE_{u,ij}$	37
2. Derivation of LT_u	38
3. Derivation of $C_{u,ij}$	39
VI IMPLEMENTATION AND RESULTS	41
A. PROBLEM GENERATION	41
B. RESULTS FOR MCC AND MCCTS	45
1. MCC (One Speed Class)	45
2. MCCTS (Two Speed Classes)	48
C. RESULTS FOR MSC AND MSCTS	50

1. MSC (One Speed Class)	51
2. MSCTS (Two Speed Classes)	53
VII CONCLUSIONS AND FUTURE DIRECTIONS	56
A. CONCLUSIONS	56
B. FUTURE RESEARCH DIRECTIONS	58
1. Improvements in the Models	58
2. Arc-Path Formulation	58
a. Arc-Path Formulation for Close Column Movements	59
b. Results with the Arc-Path Formulation for Close Column Movements	64
APPENDIX. ERROR IN SROUTEX IN GAMSLIB	66
LIST OF REFERENCES	68
INITIAL DISTRIBUTION LIST	69

LIST OF TABLES

Table 1: INFORMATION ON UNITS FOR PLANNING EXAMPLE	42
Table 2: INFORMATION ON ROAD NETWORK FOR PLANNING EXAMPLE	43
Table 3: STATISTICS FOR MCC	45
Table 4: MOVEMENT TABLE FOR MCC	47
Table 5: STATISTICS FOR MCCTS	48
Table 6: MOVEMENT TABLE FOR MCCTS	50
Table 7: STATISTICS FOR MSC	51
Table 8: MOVEMENT TABLE FOR MSC	53
Table 9: STATISTICS FOR MSCTS	54
Table 10: MOVEMENT TABLE FOR MSCTS	54
Table 11: STATISTICS FOR ARC-PATH FORMULATION FOR CLOSE COLUMN MOVEMENTS	65

LIST OF FIGURES

Figure 1: Static and Time-Expanded Network	13
Figure 2: Road Network for Planning Example	44
Figure 3: Routing for MCC	46
Figure 4: Routing for MCCTS	49
Figure 5: Routing for MSC	52
Figure 6: Routing for MSCTS	55

ACKNOWLEDGMENT

I extend my sincere appreciation to those who have significantly contributed to this thesis.

Professor Richard E. Rosenthal, who guided my earliest efforts in GAMS. He is always looking three steps ahead at where we ought to go next. I thank him particularly for his continuing assistance and thorough review.

Professor R. Kevin Wood, whose ability to discern the heart of a problem has been inspiring. His patience and many suggestions during the writing of the thesis have been invaluable.

Lastly, but most importantly, I thank my children, Jennifer and Bryan, and my wife Karin for their patience and understanding during this effort. Their support allowed me to spend the necessary time to explore new ideas and to understand problems.

EXECUTIVE SUMMARY

Planning motor movements for multiple military units that must travel at the same time through a common road network requires extensive and detailed study in order to obtain a feasible schedule with minimal delay for the units. When several units are involved, it may be impossible to avoid delays completely, but nonoptimal routing may create unnecessary delays and may lead to some roads being heavily clogged while other roads go unused. Determining optimal (minimal) delays is also important when planners want to consider various advantages and disadvantages of different types of column formations, or want to allow some units (such as truck units) to move faster than other units (such as tank units). Currently, most planning assumes a common unit speed because to allow or consider more diversity would impose a heavy computational burden on planners.

Planning aids like march formulas or time-space diagrams are usually used to manually calculate and plot the individual movements of units along a route versus time. Different units will not interfere as long as their plotted routes do not intersect on the diagram. The use of the space-time diagram, however, requires planners to choose between possible routes in advance. The shortest route is usually taken, but this choice may not be optimal. A small deviation from the shortest route may allow a unit to arrive earlier at its destination than if it waited for the shortest route to be clear of other units. The planning staff often misses such opportunities for improved efficiency

because it does not have enough time to try different routes and schedules for comparison. Consequently, a computer-based optimization model is necessary to develop routing plans more quickly than current methods allow.

Two models are presented for two kinds of movement formations. The first model (called *model close column*) is designed for use when all vehicles of a military unit move together over one route. The second model (called *model separate column*) allows a military unit to be split into groups of vehicles or *packages* which may use different routes and may start at irregular times. Neither model allows units or packages to intermingle with other units or packages and prevents units and packages from being overstretched. All movements can be planned at a single speed or at two different speeds. The objective in both models is to minimize the sum of delays over all military units subject to time and road restrictions. Fixing times by which units must reach their destination or must have left their staging areas is allowed as well as road limitations for heavy military vehicles.

The amount of data which must be input by the user is minimal to ensure easy model usability. The necessary input consists of information on the units and a description of the road system's intersections and the road segments joining them. The user is not required to understand how the models are formulated mathematically. All data which are necessary to represent the movements over time are derived from the basic input.

The models are implemented as prototypes on a personal computer. Both models are designed for the German Army but could be used with minor modifications by other

Armies. For each military unit, model solutions provide a recommended route along with the following details: the time for a column to pass an intersection at constant speed, the arrival times at intermediate intersections, the arrival time of the last vehicle at the destination, the amount of delay, and whether the shortest possible route can be scheduled. An intersection utilization report is also provided. It shows the usage of intersections by the units over time and helps to identify intersections where traffic control may be necessary to ensure smooth passage of the vehicles. Running both models for the same planning situation allows the user to compare the difference in minimal time delay for all units between the close column formation and the formation allowing dispatch of vehicles in groups.

All models are tested using data representing units concluding a German brigade live exercise. The solution times are 14 sec to 131 sec on a 486/33 MHz personal computer depending on the details of the scenario tested.

I. INTRODUCTION

Modern war is war of movement. The speed and facility with which a ground commander may effect the distribution or concentration of personnel, equipment, and supplies may decide the outcome of an operation. It is important that organization and training for efficiency in motor movements be stressed by all branches at all levels. Nonoptimal routing may lead to delays which make successful completion of a mission impossible. Therefore, this thesis considers the optimal routing of military units through a road network. The goal is to plan motor movements with minimal delays. Two integer programming models designed for the German Army, but applicable to other Armies, are presented to support movement planners in optimally routing units in two different contexts. Both models are implemented on a personal computer.

A. PLANNING MOTOR MOVEMENTS

For the purpose of background and for limiting the focus of this thesis, information is provided on how motor movements are planned in the German Army. Definitions of terms relating to motor movements and explanation of movement planning aids are included. References are made to field manuals of the United States Army because many planning techniques are similar in both services.

1. Types of Military Motor Movements

Functionally, military motor movements are divided into two general classifications, tactical and administrative. A *tactical movement* is conducted with primary emphasis on movement in combat-ready formations, based on the assumption of early ground contact with the enemy, either en route or shortly after arrival at the destination (FM 71-100, 1990, p. 3-9). *Administrative movements* are conducted when no enemy interference, except possibly by air, is anticipated. Planning of tactical movements is governed by factors such as combat effectiveness of units, military objectives, terrain, the tactical situation, probability of contact with enemy and effectiveness of enemy long-range weapons and aircraft. Including such factors into a model is not the goal of this thesis. Thus, we limit our investigations to administrative movements.

There are many considerations and objectives in planning administrative motor movements.

- Military units may have to arrive at their destinations by certain times to be available for following military operations.
- Military units must finish their movements as early as possible to achieve or maintain the highest state of operational readiness.
- The available road system should be used efficiently. This prevents clogging some roads while leaving other roads unused.
- Short routes should be used to conserve fuel because fuel is often scarce during military operations and, at least in Germany, expensive.
- Intersections which are heavily used by units must be identified. This helps determine whether military police should be deployed at those intersections to ensure that units pass smoothly and as scheduled.

Not all objectives can usually be satisfied at the same time. They may conflict when several unit movements must be conducted on a limited road network simultaneously. In such cases delays cannot be avoided. Planning for minimal delays for efficient motor movements can be done, however. Therefore, in this thesis, our goal for optimal routing of military units is to minimize the total delay for all units to conduct their movements and to reach their destinations by specified times. Minimizing total delay leads to early arrivals at destinations, prevents clogged roads and will use shortest routes whenever possible. Thus, this goal implies a combination of all the planning objectives. The models could be formulated to achieve a more even mixture of delays. This is not considered because equal or nearly equal time delays for units involved are not a military objective. But, clearly, there might be other ways of defining an overall objective for optimal routing, e.g., to minimize the maximum of the time delays for the units. For the purpose of model tractability on a PC, we limit the scope of the thesis to small-scale administrative motor movements.

2. Techniques of Movement Planning

The definitions of unit, element and column must be given. The term *unit* refers to a military organization of battalion or company size. An *element* is the smallest subdivision of a military unit that can be tactically manoeuvred independently. The term *column* is used for all the vehicles involved in a single move over the same route under one common command. A column can consist of several units.

There are three fundamental types of column formations: *Close column*, *open column* and *infiltration*. Close and open column are regular formations; they

differ only in vehicle spacing. Infiltration allows vehicles to be dispatched, individually or in small groups (packages), at irregular intervals to reduce density or to execute a movement along a route that is heavily used by civilian traffic. It is the least preferable type of column formation because it requires more detailed briefing and maintenance arrangements for the groups to be able to operate separately. Nevertheless, there are quite a few instances in which the advantages and disadvantages of a close column formation and infiltration are counterbalanced. In such cases, it may be crucial to know which column formation results in smaller delays. (FM 55-30, 1980, pp. 5-3 - 5-5)

In order to plan effectively, the road movement planner must know the number of vehicles in a column and how long it will take a column to transit a road segment between two intersections. The terms used in this thesis to determine the time required to complete the move are described in the following.

- *Length of a column* is the length of roadway which it occupies, measured from front to rear in kilometers or miles.
- *Gap* is the space between successive vehicles in a column as measured from the rear of one vehicle to the front of the following vehicle (vehicle gap). If a column consists of several units, a gap can also be used to describe the space between successive units (column gap). In addition, the term can describe the space between successive columns.
- *Road distance* is the distance from point to point by road, expressed in kilometers or miles.
- *Pass time* is the time for a column, or group of vehicles, to pass a given point at constant speed.
- *Time distance* is the time to move from one point to another at a given speed. (FM 55-30, 1980, pp. F-1 - F-3)

Motor movements are scheduled from a start point to a release point. The start point (SP) must be a clearly defined point on the road network. It has to be easily recognizable on the map and on the ground and be readily accessible to facilitate forming the column and to eliminate the possibility of confusion. Thus, a road intersection is often chosen. A release point (RP) is a well defined point on the road network from which the elements composing a column continue their movements towards their own appropriate destination. Now, each one of these elements is under the authority of its respective commander. The RP, like the SP, should be easily recognizable and must be on the column's route to ensure a smooth and continued movement. Again, a road intersection is often chosen. (FM 55-30, 1980, p. 5-7)

Currently, both in the United States and in Germany, personnel planning an administrative movement use simple planning aids such as march formulas, road movement graphs, and road movement tables. These aids are described below.

- *March formulas* represent the basic arithmetic of march planning. Often tables are available which show for given speed, vehicle gap and number of vehicles the time distance, the length of a column and the pass time.
- A *road movement graph* is a time-space diagram on graph paper with the vertical axis showing road distance of a route and intermediate points like intersections and with the horizontal axis for time. The lower left-hand corner of the graph represents the start point at the earliest time movement is contemplated. Progress of the column's first vehicle between intermediate points is indicated by plotting time and road distance on the graph which is a straight line with slope equals speed (speed is always considered constant between intermediate points to ensure a smooth movement, but can change at those points). An additional parallel line is drawn for the last vehicle of the column. The vertical difference between these two lines corresponds to the length of the column, the horizontal line is just the pass time. It becomes immediately obvious how long it takes a column to clear a section of the road and different movements will not interfere as long as the lines do not intersect on the graph. (FM 55-100, 1980, pp. F-4 - F-8)

- A *road movement table* is a list showing the units composing a column and the time and space schedule for motor movements. The schedule is usually obtained by applying the march formulas on a road movement graph. The table is used as a convenient way of combining essential details.

B. PROBLEM STATEMENT AND SCENARIO

The focus of this thesis is to develop a model to support the planning of small-scale administrative motor movements. Although the basic methodology of the thesis is applicable to the movement of military units in many countries, we consider the optimal routing of German military units upon completion of a brigade live exercise. Several battalions and attached companies/batteries are spread over an area 15 by 25 km. Planning the administrative motor movement back to the garrisons is difficult, particularly when the routes of some units cross each other. The brigade must make sure that the movements start in a coordinated manner without any interference between units. If the units have different start and release points, it is not necessarily true that each unit will reach its RP fastest by using the shortest route between SP and RP. A small deviation from the shortest route may allow an earlier arrival at RP than waiting until the shortest route is clear of other units. The use of the road movement graph, however, requires the planner to make a choice between possible routes in advance to get the vertical axis. This choice may not be the optimal route and the planning staff usually does not have much time to try different routes and schedules for comparison. Consequently, an optimization process is necessary to develop routing plans more quickly than current methods allow. Our major concern is a prototype. Therefore,

questions on how to manage input files so that they are easy to use are beyond the scope of the thesis.

C. DESCRIPTION OF THE MODEL

Two models are developed for different types of column formations in a small-scale administrative motor movement. The first model supports planning a close column formation. The second model considers infiltration, the dispatch of vehicles in small groups. Both routing optimization models are integer programs. In initial formulations both models are designed for the usual planning practice which routes all units at the same speed. Then, in separate formulations, both models are enhanced to allow planners to route units at two different speeds. A column, or a group of vehicles, cannot be split into smaller elements and cannot intermingle with other military vehicles. This restriction is desired by German military planners to simplify the coordination of unit movements. A prototype is implemented on both a 486 personal computer and a 386 personal computer with a math coprocessor. The program is formulated via the General Algebraic Modeling System (GAMS) (Brooke, Kendrick and Meeraus, 1992) programming language and solved with the ZOOM solver (Singhal and Marsten, 1990). Specifically, the models

- schedule the arrival for units at RP by certain times such that the overall delay for all units is minimized,
- are generic so that the user can easily modify input parameters for different task organizations and road systems,
- list routes and schedules in a movement table for quick and easy analysis along with information on how heavily intersections are used,

- allow a comparison between delays in close column movements and dispatch of vehicles in small groups.

D. PREVIOUS STUDIES

In standard routing and scheduling problems, for each vehicle or driver, a sequence of locations to be visited is provided, together with the times at which activities at these locations are to be carried out (Bodin, and others, 1983, p. 71). These models are not at all like our models. Dynamic network optimization models describe network-structured, decision-making problems over time. They are of interest for our problem because of their numerous applications to traffic systems. For a survey of dynamic networks see, for instance, (Aronson, 1989). We especially mention two papers which have made substantial contributions to this thesis.

A space-time network that represents traffic flow over time for a capacitated road transportation system having one-way and two-streets was developed by Zawack and Thompson (1987). This automobile traffic flow model incorporates traffic lights and considers congestion effects explicitly, which is beyond the scope of this thesis. Their formulation of a multiple-source single-destination network for a single commodity was solved with the network flow code NETFLO. It was suggestive but not directly applicable to our vehicle scheduling problem. We must treat different columns as separate commodities, each with its own source and destination to maintain the identity of each column.

Moving military units through a road network was investigated using prototypic linear and integer mathematical programming models by Lee (1991). Based on GAMS implementations, optimal paths for three units were found to minimize the average arrival time at RP. However, the integer program assumed that all columns had the same length. The linear model allowed columns to split up in any group sizes and intermingle with other units. Lee's models were implemented on a mainframe computer, not a PC.

II. MULTICOMMODITY TIME-EXPANDED NETWORK

Different columns must be treated as separate commodities to guarantee that all moving columns can be identified unambiguously in a model. This chapter describes a multicommodity network and how a road network is transferred into a time-expanded network to represent the fact that a certain amount of time is necessary to transit roads.

A. SINGLE AND MULTICOMMODITY NETWORKS

A directed network is defined as a set N of nodes and a set A of ordered pairs of distinct nodes (i,j) called arcs. If (i,j) is an arc, we say that (i,j) is outgoing from node i and incoming to node j . We also say that arc (i,j) is incident to i and to j , and that i is the start node and j is the end node of the arc. A directed path in the network from node s to node t is an alternating sequence of nodes i_1, i_2, \dots, i_p , such that $i_1 = s$, $i_p = t$ and $(i_k, i_{k+1}) \in A$ for $k=0, \dots, p-1$, and p is some positive integer. A simple directed path is a path in which all nodes are distinct. A route is referred to as a simple directed path. For the road network, intersections, SPs and RPs are represented by nodes. In the road network, we define a *road segment* as a section of a road between two intersections. A one-way road segment refers to a road segment that is usable in one direction only. Thus, a one-way road segment is represented by a single arc in the directed network. A two-way road segment is represented by two arcs in anti-parallel. We don't have to model two lanes of a one-way road segment because two or more units will never be

scheduled to move side by side. This representation of the road network is referred to in the following as the *static network*. We call arcs in the static network *physical arcs*.

Consider now a single commodity which must be shipped through the road network. Defining supplies and demands, and using standard flow balance constraints, it is possible to model the shipment of this item in the network (Cormen, Leiserson and Rivest, 1991, p. 582). The decision variables determine the unknown amount of the commodity to be transported on each arc. For our purpose, separate columns must be distinguished by different commodities. Consequently, a multicommodity network model is considered for our routing problem.

In the multicommodity network each commodity has its own supplies and demands (Cormen, Leiserson and Rivest, 1991, p. 587). The supply for a given column at its SP and the demand for that column at its RP is the number of vehicles for that column. Node capacity constraints are added to the formulation to limit the throughput of an intersection. A more computationally intensive alternative is to replace nodes at intersections by a small subnetwork of eight nodes and as many as 16 arcs, which take all possible movements at an intersection into account explicitly (Potts and Oliver, 1972, p. 62). This detailed model is not used for the purpose of tractability. Additional constraints across anti-parallel arcs will be necessary and will be explained later.

B. TIME-EXPANDED NETWORK

The problem of routing military units over time is not a static or steady-state traffic flow application. The model is, therefore, defined on a *time-expanded network*

model (Ford and Fulkerson, 1962, p. 145). Let the time for a vehicle to move from the start node to the end node of an arc (i,j) be the travel time $t(i,j)$ which is a positive integer representing *time periods*. Then the time-expanded network is constructed from the static network with nodes representing both time and location. Planning over discrete time periods requires integral time distances and pass times for units. Consequently, real time distances and pass times can only be approximated. Shorter time periods improve the accuracy of the models but increase the total number of time periods for the study horizon which leads to bigger models which are harder to solve. If T is the total number of discrete time periods, then starting at time period one expands each node i into T nodes $i_t, t=1,2,\dots,T$. Each physical arc is expanded into several arcs of the form $(i_t, j_{t+t(i,j)})$, called *movement arcs*. A second type of arc that appears in a time-expanded network is called a *hold-over arc* (Aronson, 1989, p. 8). A hold-over arc (i_t, i_{t+1}) represents flow staying at node i for one time period, and is thus delayed and prevented from flowing on some movement arc (Zawack and Thompson, 1987, p. 155). Figure 1 shows the static and time-expanded network representation of a simple road network. For our routing problem, we must construct a time-expanded network for each commodity.

The capacity of a movement arc is the upper bound on the number of vehicles that may enter the corresponding physical arc per time period (Weigel and Cremeaus, 1973, p. 77). In a column formation, all vehicles travel at a uniform speed with fixed vehicle gaps, so each column has a fixed positive pass time. Thus, for our model, the capacity of any movement arc is just one, i.e., at most one column can enter any arc at

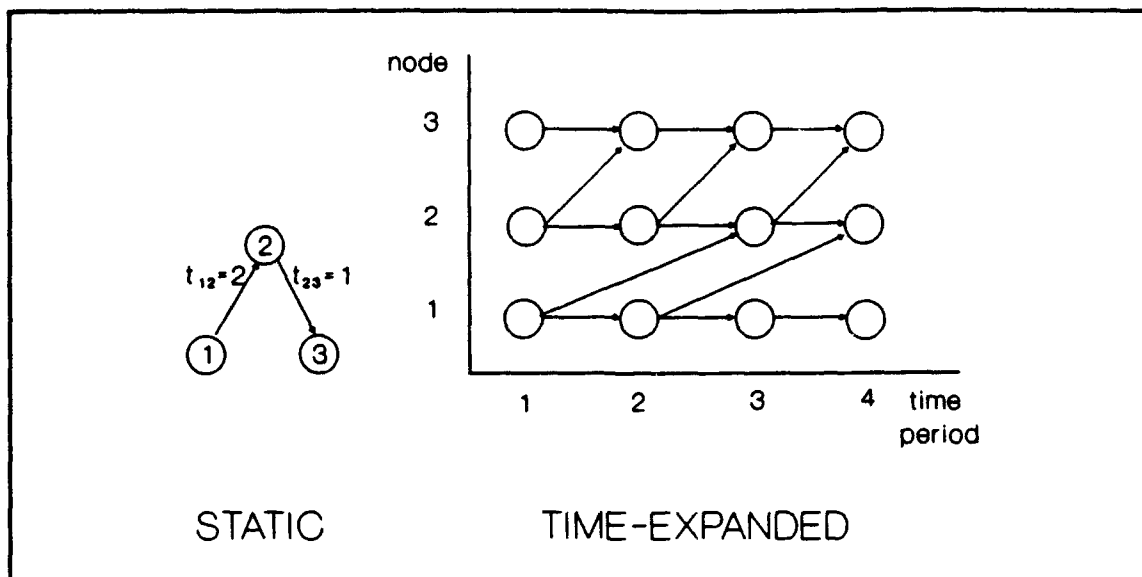


Figure 1: Static and Time-Expanded Network

a given time period. Hold-over arcs in the time-expanded network are assumed to have an unlimited capacity.

Time-expanded networks can get big really fast. Consider a static network with n nodes and e arcs and assume that, in a worst case, all arcs have a $t(ij)=1$. Then, in the time-expanded network for one commodity, starting at time period one, an upper bound on the number of arcs is $(n+e)*(T-1)$. An upper bound on the total number of nodes is $n*T$. The size of the expanded network increases with longer study horizons unless a longer time period is chosen. There are, however, quite a few movement arcs which will never be used and can be eliminated to keep the model tractable. Consequently, this effort will be of primary interest.

III. MODEL DEVELOPMENT AND FORMULATION

In this chapter we describe the general routing constraints which must be considered in the German Army. The assumptions underlying our models are then stated and finally, mathematical models are formulated and explained for both model close column and model separate column. Both models are only valid if all units move at the same speed. We will describe in Chapter IV how to account for different speeds.

A. GENERAL ROUTING CONSTRAINTS

General constraints for small-scale military movements that planners in Germany must cope with are described to illustrate some of the complexities of the routing task in our scenario. Basic planning standards are established here that must be represented in both models.

- Military units usually start movements at a *staging area*, a general locality that is established for the concentration of units between successive missions or for refueling. A staging area is near the road network and cannot serve as a SP (which must be a point on the road network). Thus, military planners who choose a SP for a unit must make sure that this unit can move from the staging area to the SP without any interference with other units. Consequently, a minor road from a staging area to SP is assigned which is not in consideration for the movements of other units.
- Each column moves between SP and RP at a fixed average speed (with a range of 30 km/h to 50 km/h) with fixed vehicle gaps (50 m or 100 m). Thus, pass time and time distances are fixed in the calculations of military planners.
- A column is not allowed to overstretch or to intermingle with other columns to facilitate column control and maintain uniformity of column movement.

- While a column is passing an intersection, no other column should arrive during this time to avoid vehicles (30, 60 or even more) clogging the road segment while they wait for clearance of the intersection.
- A German brigade does not have a homogeneous fleet of vehicles. Truck units normally move at a higher speed than tank units. However, fast units may not pass slower ones because this leads to intermingling of different units. In addition, most of the road segments used for small-scale motor movements in Germany only have one lane in each direction. Consequently, civilian traffic (at a speed of about 100 km/h) prevents military units from passing other units. This restriction requires careful calculations by the planner if he allows different speeds for different units which conduct movements at the same time. Thus, for the simplicity of the calculations, all columns are often required to move at the same speed, that of the slow tank units.
- Planners usually want at most one column moving through an intersection at a given time. This prevents an intersection from becoming totally blocked for civilian traffic during the pass time of one column. Of course, this restrictive way of planning does not apply to intersections that are bridged and have enough lanes so that the number of entering columns is limited only by the capacity of the outgoing road segments. Consequently, two kinds of intersections are modelled. However, intersections which serve as SP or RP must always have a capacity of one column at a time. This is desired because it supports the smooth beginning and ending of a movement, and avoids confusing situations for unexperienced drivers in unfamiliar terrain.
- It is forbidden to plan for more than one unit to start on a one-way road segment at the same time. Of course, two or more successive units which do not overlap can be distributed over a one-way road segment.
- Weight restrictions may exclude particular road segments for tank movements.
- Time restrictions that are known *a priori* may involve additional constraints. For instance, a unit may have to leave its staging area by a certain time or a tank unit may have to reach a railroad station to be loaded by a certain time.
- Some units may have movement priorities because of subsequent tasks. Thus, no or only a small delay at RP is desirable for these units.
- On narrow road segments, planners usually do not want columns to move in opposite directions at the same time, to avoid disturbing civilian traffic more than necessary. In the following we refer to columns in such situations as *encountering* columns.

B. MODEL ASSUMPTIONS

As with any model, there are some aspects of the problem which must be assumed in order to make the resulting formulation tractable.

1. Assumptions on Units

Units are classified according to weight limitations for road segments. Four categories are considered: tank units, units with armored infantry fighting vehicles or armored artillery vehicles (in the following referred to as units with Bradley Fighting Vehicles or BFVs), units with heavy equipment transporters (HETs) and truck units. Before the initial time period $t=1$, there is no movement on any road segments. Similarly, for the last time period $t=T$, all the vehicles must have reached their RP. Thus, we do not have any scheduled movements of units after that time.

Units are only allowed to wait at their staging areas. As soon as a column starts at SP it must keep moving (with constant speed) without any intermediate breaks or delays. This assumption is not really restrictive. It has often been implicitly made by planners because passing an intermediate staging area during a small-scale movement undermines the objective of bringing units rapidly to their destinations. If the time distance for a column is more than 2-3 hours, intermediate staging areas are usually visited for refueling or redistribution of elements. Such large time distances are not the case for our planning scenario.

2. Assumptions on Road Network

The administrative motor movement is to be conducted in a rural region as opposed to an inner city. Therefore traffic congestion and rush hours are not applicable to this thesis. The total time spent by a vehicle on a road segment is equal to the (uncongested) travel time and stays fixed during the study horizon. It takes no time for a vehicle that enters an intersection to cross it. We assume that there are enough minor road segments between the staging areas and RP in this region. Thus, movements from staging areas to SP are not modelled.

The road network is assumed unchanging during the study horizon. Thus, road segments will not be closed or opened during the study horizon because of information which is known *a priori*. This assumption allows us to compute the shortest route from SP to RP for each column. This serves as a reference point to calculate the delays for different routes.

C. MODEL FORMULATION

Two different models will be considered. The *model close column* (MCC) is developed to plan a close column formation. For infiltration we present the *model separate column* (MSC). Both models are designed for the, current, commonly used planning practice which routes all units at the same speed. The total delays to conduct the movements are critical.

The following way is chosen to present the models. First we describe a model in general terms. Second, the mathematical formulation is given. Then a detailed

description of some of the parameters and all the equations concludes the model presentation.

1. Formulation for Model Close Column (MCC)

In the following, for simplicity of description, a column consist of one unit (of battalion or company size). MCC schedules all vehicles of a military unit in a single move over the same route. Units cannot be split into smaller groups. All units move at the same speed. The decision variables in this model will be binary indicating whether or not a unit starts moving on a physical arc in the static network at a particular time. Thus, this model is a pure integer programming model (IP). The objective is to minimize the sum of delays for all columns. In addition to the basic flow balance, node and capacity constraints, constraints are necessary to restrict cases of encountering units.

Indices

i, j	Nodes of the network.
u, u'	Units that have to be scheduled.
t, t'	Time periods in minutes.
d	Node to represent a dummy sink for all units after reaching RP.

Parameters and sets

SOURCE _u	SP of unit u.
DESTIN _u	RP of unit u.

VEH_u	Pass time of column u (including gaps) in time periods.
$ESTAT_u$	Earliest start time (period) for unit u to leave from SP.
$EARRT_u$	Earliest arrival time (period) for unit u at RP.
PEN_u	Penalty for not scheduling unit u .
$DIST_{u,i,j}$	Time distance in time periods for unit u to move from i to j .
$ENC_{u,t,i,j}$	Minimum pass time over all units except u which are able to start their movements at time t on physical arc (j,i) and might encounter unit u which moves on arc (i,j) .
$C_{u,t,i,j}$	Equals one if unit u is allowed or able to enter physical arc (i,j) at time period t , and is zero otherwise. This parameter is used to reduce the number of decision variables and appears in each equation.
SF	Scale factor.
NC	Set of nodes without capacity limit of one unit.
NE	Set of physical arcs (i,j) on which encountering is not allowed.

Decision variables

$X_{u,t,i,j}$	Equals one if unit u starts to move at the beginning of time period t from i to j , and is zero otherwise.
$ELASTIC_u$	Binary variable which is one if a unit cannot be scheduled and is zero otherwise.
Z_u	Positive variable equals the time delay for unit u .

Formulation

Minimize

$$\sum_u Z_u + \sum_u \sum_t \sum_i \sum_j SF * DIST_{u,i,j} X_{u,t,i,j} + \sum_u PEN_u * ELASTIC_u$$

Subject to:

$$(1a) \quad Z_u \geq \sum_t X_{u,t,DESTIN_u,d} (t - EARRT_u) \quad \forall u$$

$$(2a) \quad \sum_t \sum_j X_{u,t,SOURCE_u,j} + ELASTIC_u = 1 \quad \forall u$$

$$(3a) \quad \sum_j X_{u,t,i,j} - \sum_j X_{u,t-DIST_{u,j,i,j,i}} = 0 \quad \forall u, t, i, i \neq SOURCE_u, \\ i \neq DESTIN_u, i \neq d$$

$$(4a) \quad X_{u,t,DESTIN_u,d} - \sum_j X_{u,t-DIST_{u,j,DESTIN_u,j,DESTIN_u}} = 0 \quad \forall u, t$$

$$(5a) \sum_u \sum_{\substack{t' \\ t-VEH_u < t' \leq t}} \sum_j X_{u,t',i,j} \leq 1 \quad \forall t, i, i \neq d, i \in NC$$

$$(6a) \sum_u \sum_{\substack{t' \\ t-VEH_u < t' \leq t}} X_{u,t',i,j} \leq 1 \quad \forall t, i, j, i, j \neq d, i, j \in NC,$$

i, j adjacent

$$(7a) \sum_{\substack{t' \\ t-DIST_{u,i,j} + ENC_{u,t,i,j} \leq t' \leq t, j \in NC \\ t-DIST_{u,i,j} - VEH_u < t' \leq t, j \in NC}} X_{u,t',i,j} + \sum_{u' \neq u} X_{u',t,j,i} \leq 1 \quad \forall u, t, i, j,$$

i, j \in NE, i \neq d

The objective function calls for the minimization of the sum of difference in time between earliest possible arrival time and scheduled time for each unit which is called *delay*. Constraint (1a) accounts for these delays. The delays could have been defined directly in the objective function. This particular way was chosen because it resulted in fewer iterations and CPU time during testing. The second term in the objective equation assures that in case of a choice between several routes of different length which would lead to the same delay at RP, the shortest one is taken. The default value for SF is 0.01. The objective function is penalized by ELASTIC_u if unit u cannot be scheduled. This can only occur if the time horizon of the model is too short to schedule all units.

Equation (1a) computes the individual delays for each unit by calculating the difference between arrival time at RP and $EARRT_u$. The arrival time of the first vehicle of any column at RP equals the start time at RP to move to the dummy sink and thus is easily available for computation. Otherwise the arrival time must be calculated from all physical arcs incident to RP. $EARRT_u$ is given by input or derived from the shortest distance from $SOURCE_u$ to $DESTIN_u$.

Equation (2a) says that a unit must start from its SP at some time period, but allows a unit to wait until a route is available. This formulation does not require hold-over arcs at the SP. Note that the summation on t can actually be restricted on the interval $(ESTAT_u, LSTAT_u)$ where $LSTAT_u$ is the latest possible start time and is easily derived from the input data. If a unit cannot be scheduled, the elastic variable becomes one and the unit is not sent out through the time-expanded network. In case of at least one unit that cannot be scheduled, the model branches to a subroutine that changes some parts of the derived data crucial for scheduling units. This is explained in detail in Chapter V. The model then restarts to optimally route all units. In the program, $ELASTIC_u$ is defined as a nonnegative variable. This uses the fact that the variable cannot obtain a fractional value given that other variables are integer and reduces the number of integer variables.

Equation (3a) ensures that the flow into any node other than SP or RP must equal the flow out of the node without allowing any units to wait at such nodes. These are the usual conservation of flow constraints.

Constraint (4a) indicates that each unit u that reaches its RP has to move to a dummy node in the static network. Flows from each RP to a dummy node are proposed by Zawack and Thompson. This simplifies the calculation of the arrival times at RP for constraint (1a). But, more important for our models, the representation of a movement from RP to a dummy node becomes necessary in some special cases. To explain this, first note that the decision variable $X_{u,t,i,j}$ schedules the start time for any movement. Now, consider a unit u that starts on a physical arc incoming to its RP and assume that there is no requirement to schedule a following flow from RP to the dummy node (lack of constraint (4a)). Then, the decision variable $X_{u,t,i,j}$ cannot model a flow of this unit u at its RP. Consequently, the pass time of unit u at its RP cannot be modelled. But, there will be nodes in the static network which are an RP for at least one unit u and transshipment node for at least one other unit u' (and have a capacity limit of one column). While unit u is arriving at such a node, unit u' could pass that node at the same time. Thus, to represent the capacity limit in constraint (5a), it is necessary to model the pass time of unit u at its RP if this point is a node as described. This can easily be done by using a decision variable $X_{u,t,DESTINu,d}$. Therefore, the introduction of a dummy node is very helpful in our model.

Constraint (5a) limits the throughput of node i to at most one column leaving the node at any time. It blocks the intersection during the pass time of a column for other columns. The node throughput constraint forces all outgoing arcs to have an implicit arc capacity of one unit. In this case arc capacities do not have to be represented. Nodes without a capacity limit are not considered at all. This is possible

because adjacent nodes with capacity limit always make sure that at most one unit can pass these nodes. Thus, at most one unit at a time can be scheduled to start on a physical arc outgoing from a node without a capacity limit and incoming to node with capacity limit. Consequently, the arc capacity cannot be violated. This, of course, is not valid for adjacent nodes without a capacity limit.

Equation (6a) represents arc capacities for physical arcs whose start and end nodes do not have a capacity limit. These nodes are not restricted by equation (5a). Therefore, more than one unit at a time could start on a physical arc connecting them. Consequently, in this special situation, arc capacities must be included explicitly.

Constraint (7a) states that units are not allowed to encounter on certain road segments. A unit moving on the physical arc (i,j) within certain time periods blocks the anti-parallel arc (j,i) for movements of other units. However, nodes with a node capacity exclude certain encountering possibilities. Consider a node j with capacity limit. If some unit u' is leaving node j on physical arc (j,i) at time t, then constraint (5a) prevents any other units from entering node j during the pass time of unit u'. Thus, in particular, no other unit u on the anti-parallel arc (i,j) can enter node j during this time. Consequently, a unit u which starts on arc (i,j) during the time interval $(t-DIST_{u,ij}+1, t-DIST_{u,ij}+VEH_u)$ cannot enter node j if unit u' starts leaving node j at time t. Thus, there is no need to model this time interval again. For a given unit u, we can compute which units u' might encounter u and the minimum of their column lengths. This is $ENC_{u,t,ij}$. This parameter indicates that unit u cannot start on arc (i,j) during the time interval $(t-DIST_{u,ij}+1, t-DIST_{u,ij}+ENC_{u,t,ij})$ if unit u' starts on arc (j,i) at time t. It

does not matter which unit u' might be scheduled by the model because of the use of the minimum possible column length for the parameter. For nodes without a capacity limit the parameter $ENC_{u,t,i,j}$ is not applicable. The time interval to be examined for encountering must be extended by the unit's pass time to compensate for the lack of the node capacity constraint (5a).

2. Formulation for Model Separate Column (MSC)

In this integer programming model the vehicles of a unit are scheduled in groups (packages), each of them with a pass time of one time period. Thus, the total number of packages for a unit equals its column length in MCC. The packages can start at irregular time periods or at successive time periods. They can use different routes. All packages must move at the same speed. A split into smaller subgroups or intermingling with other packages is not allowed. The objective of MSC is to minimize the sum of delays for all vehicle groups. The delays are scaled so that we are able to compare MCC and MSC with respect to the resulting delays. In general, MSC must give results that are at least as efficient as those from MCC because it allows more flexibility for routing. Thus, the optimal objective function value from MCC can be used later as incumbent value for solving MSC. This reduces the solution time for MSC. The constraints are similar to those in MCC.

All definitions of indices, parameters and decision variables in MCC apply to MSC. The only difference is that they now refer to groups of vehicles instead of units. Therefore, we will not repeat the previous definitions, but include only the few additional items needed.

Parameters and sets

$M_{u,t}$ Upper bound on the number of packages of unit u which can be scheduled during some time interval whose range is indexed by t .

Decision variables

$ELAST_u$ Integer variable which is the number of packages of unit u that cannot be scheduled and is zero otherwise. ($ELAST_u$ can be greater than one. Therefore, $ELASTIC_u$ is not applicable in MSC).

Y_u Positive variable equals the average time delay for unit u .

Formulation

Minimize

$$\sum_u Y_u + \sum_u \sum_t \sum_i \sum_j SF * DIST_{u,i,j} X_{u,t,i,j} / VEH_u + \sum_u PEN_u * ELAST_u$$

Subject to:

$$(1b) Y_u \geq \left(\sum_t X_{u,t,DESTIN_u,d} (t - EARRT_u) - \sum_{t=1}^{VEH_u-1} t \right) / VEH_u \quad \forall u$$

$$(2b) \quad \sum_t \sum_j X_{u,t,SOURCE_u,j} + ELAST_u = VEH_u \quad \forall u$$

$$(3b) \quad \sum_j X_{u,t,i,j} - \sum_j X_{u,t-DIST_{u,j,i},j,i} = 0 \quad \forall u,t,i, \quad i \neq SOURCE_u, \quad i \neq d, \\ i \neq DESTIN_u$$

$$(4b) \quad X_{u,t,DESTIN_u,d} - \sum_j X_{u,t-DIST_{u,j,DESTIN_u},j,DESTIN_u} = 0 \quad \forall u,t$$

$$(5b) \quad \sum_u \sum_j X_{u,t,i,j} \leq 1 \quad \forall t,i, \quad i \neq d, \quad i \in NC$$

$$(6b) \quad \sum_u X_{u,t,i,j} \leq 1 \quad \forall t,i,j, \quad i,j \neq d, \quad i,j \in NC, \quad i,j \text{ adjacent}$$

$$(7b) \quad \sum_{\substack{t' \\ t-DIST_{u,i,j} < t'/t, \quad j \in NC \\ t-DIST_{u,i,j} \leq t'/t, \quad j \in NC}} X_{u,t',i,j} + M_{u,t'} * \left(\sum_{u' \neq u} X_{u',t,j,i} - 1 \right) \leq 0 \quad \forall u,t,i,j, \\ i,j \in NE, \quad i \neq d$$

The objective function minimizes the sum of the average time delay for the packages over all units as defined in equation (1b). It takes the distance which is travelled on the average by the packages into account. The objective function is penalized if a package cannot be scheduled. The default value for SF is 0.01.

Constraint (1b) indicates the average time delay of the packages for each unit. The computation considers that successive packages of a unit have *intrinsic delays* with respect to the first package to start. Therefore, the total delay is accordingly decreased before the average is calculated. The average is necessary. Otherwise we are not able to compare the results from MCC and MSC. Remember constraint (1a) which defines the delay for each unit in MCC. If we neglect the intrinsic delays, the delay of a unit in MCC can be interpreted as the delay of each vehicle or each element of the unit. Thus, Z_u of MCC is equivalent to an average delay which is the same for each vehicle of unit u . Consequently, Y_u from Equation (1b) in MSC can be directly compared to Z_u , and both objective function values can be compared.

Equations (2b) to (4b) are the flow balance constraints. Supply and demand are represented by the total number of all vehicles converted to time periods. This value equals the pass time of a unit in MCC. If at least one group of vehicles cannot be scheduled, the model restarts after some changes in its derived data to allow all units being scheduled. In the program, $ELAST_u$ is defined as a nonnegative variable because it cannot obtain a fractional value.

Equation (5b) limits the throughput of nodes which have a capacity limit. At most one package is allowed to leave such a node i at time t . Again, arc capacity does not have to be considered.

Constraint (6b) represents arc capacities (at most one group of vehicles can start on a physical arc (i,j) at a given time) for the cases in which start and end nodes of a physical arc do not have a capacity limit, and thus are not restricted by constraint (5b).

In equation (7b), units on arc (i,j) and (j,i) are prevented from encountering if specified by NOENC. The equation is only restrictive for possible movements of packages of unit u on physical arc (i,j) if a package of some other unit u' leaves at time t . As soon as no package of u' leaves at t , then $M_{u,t}$ is the upper bound on this constraint. This parameter is calculated as the minimum of the number of decision variables $X_{u,t',i,j}$ which are added in the sum with index t' and the number of packages comprising unit u . This formulation is weak and can be replaced with the stronger formulation (7b')

$$(7b') \quad \sum_{\substack{t' \\ t-DIST_{u,i,j} < t' \leq t, j \in NC \\ t-DIST_{u,i,j} \leq t' \leq t, j \in NC}} X_{u,t',i,j} + M_{u,t} * (X_{u',t,j,i} - 1) \leq 0 \quad \forall u, u', t, i, j, \\ i, j \in NE, u \neq u', i \neq d$$

IV. MODEL ENHANCEMENT

The previous chapter described models MCC and MSC. Both models support the common practice of routing all units at the same speed. However, as stated, truck units can conduct a motor movement at higher speed than tank units. Currently, this is seldom taken advantage of by planners. Therefore, we concentrate in this chapter on enhancing both models to support movement planning with different speeds.

A. MOVEMENTS WITH DIFFERENT SPEEDS

A model for optimizing routing of units should allow units to move at different speeds. This is desired by military planners to account for varying efficiency and speed of units conducting movements. MCC and MSC are not correct for simultaneous movements of units at different speeds. To explain this, consider a slow unit which starts from node i on physical arc (i,j) and a fast unit which departs a little later. The fast unit can mathematically pass the slow unit on this arc and move on from node j before the slow unit. This is not allowed in practice, but MCC and MSC cannot prevent it.

In certain cases, however, MCC and MSC prevent illegal passes from occurring. The node capacity constraint will exclude some passing conflicts because it prevents units from overlapping at nodes. Consequently, situations in which a fast unit would overlap with a slow unit (because the last vehicle of the fast column has not passed the

first vehicle of the slow column yet) cannot happen. And, passing will be less likely on physical arcs whose start and end nodes have a capacity limit, if the difference in travel time is small.

The problem of illegal passes that occur in MCC and MSC can sometimes be resolved by minor modification of the input data, for instance, by holding up a fast unit for a few time periods. This manual adjustment of the data requires experience with the model itself and can consume a large amount of time. Therefore, a model enhancement is needed to automatically resolve passing conflicts.

B. ENHANCED MODEL FORMULATION

We assume that all units move at one of two speeds, slow or fast, which is specified in the data. The passing problem is attacked in two steps. First, the models partition the units into two speed classes. Then, each unit of a speed class is related to all units belonging to the other speed class. Constraints are added to prevent illegal passes. The number of additional constraints will increase significantly for road segments available to slow and fast units. In the following, only definitions and constraints are presented which must be added to the existing formulation.

1. Formulation for Model Close Column Two Speed (MCCTS)

MCCTS is the enhancement of MCC. It allows units to move at two different speeds.

Additional indices

s Speed class for units, "SLOW" or "FAST".

Additional parameters and sets

NUM_i Number of units in each speed class.

DIFF_{ij} Difference in travel time (in time periods) on physical arc (ij) between "FAST" and "SLOW" units.

SLOW Set which contains units that are "SLOW".

Formulation

If NUM_{SLOW} ≤ NUM_{FAST}

$$(8a) \quad \sum_{\substack{t' \\ t-DIFF_{i,j} < t'st-VEH_u, i,j \in NC \\ t-DIFF_{i,j} < t'st, i \in j \in NC \\ t-DIFF_{i,j}-VEH_u < t'st-VEH_u, i \in j \in NC \\ t-DIFF_{i,j}-VEH_u < t'st, i,j \in NC}} X_{u,t',i,j} + \sum_{u' \in SLOW} X_{u',t,i,j} \leq 1 \quad \forall u,t,i,j, \\ u \in SLOW, i \neq d$$

If NUM_{SLOW} > NUM_{FAST}

$$(9a) \quad \sum_{\substack{t' \\ t < t'st+DIFF_{i,j}-VEH_u, i,j \in NC \\ t < t'st+DIFF_{i,j}-VEH_u, i \in j \in NC \\ t < t'st+DIFF_{i,j}, i \in j \in NC \\ t < t'st+DIFF_{i,j}, i,j \in NC}} X_{u,t',i,j} + \sum_{u' \in SLOW} \sum_{\substack{t'=t, j \in SEC \\ t-UNITVEH'_u < \\ t'st, j \in SEC}} X_{u',t',i,j} \leq 1 \quad \forall u,t,i,j, u \in SLOW, i \neq d$$

In constraint (8a), all fast units which might start on arc (i,j) at time t are related to each slow unit which might have started at some time $t' \leq t$ and might be passed. The lower limit of the summation on t' varies for different values of $DIFF_{ij}$. There is a change in the upper and/or lower limit of the summation on t' for physical arcs incident to nodes without a capacity limit.

Equation (9a) ensures that slow units leaving node i on arc (i,j) at t prevent fast units from starting on (i,j) at certain times t' . This equation holds up fast units for the number of time periods that is necessary to exclude illegal passes.

Not all physical arcs are feasible at all times for both slow and fast units. Thus, the number of constraints actually generated is less than that implied by the four-dimensional domain of the definition of the equations (8a) and (9a).

Physical arcs with start and end nodes which both have a capacity limit are a special case that allows further significant reduction of the number of constraints. The reduction is based on the node throughput constraint (5a). A value of $DIFF_{ij}$ that is less than $VEH_u + VEH_v$ prevents a passing on road (i,j) at any time since there is not enough time for a fast unit during its travel time $t(ij)$ on the physical arc (i,j) to pass a slow unit. In such a case, unit u' , either in constraint (8a) or in constraint (9a), does not have to be considered. When this happens for all units u' , the constraint is not necessary at all for a given physical arc.

2. Formulation for Model Separate Column Two Speed (MSCTS)

MSCTS is the enhancement of MSC. It allows packages of different units to move at two different speeds. Packages which belong to the same unit move at the same speed. Indices and parameters defined for MCCTS are also applicable to MSCTS.

Formulation

$$(8b) \quad \sum_{t'} X_{u,t',i,j} + M_{u,t'} * (X_{u',t,i,j} - 1) \leq 0 \quad \forall u, u', t, i, j,$$

$$\begin{array}{l} t-DIFF_{i,j} < t' < t, i, j \in NC \\ t-DIFF_{i,j} < t' \leq t, i \in, j \in NC \\ t-DIFF_{i,j} \leq t' < t, i \in, j \in NC \\ t-DIFF_{i,j} \leq t' \leq t, i, j \in NC \end{array}$$

$$u \in SLOW, u' \notin SLOW, i \neq d$$

Constraint (8b) ensures that a slow package must start a minimum number of time periods ahead of a fast package to avoid being passed. It only restricts possible starts of packages of a slow unit u from node i on physical arc (i,j) , if a package of some fast unit u' leaves from i on arc (i,j) at time t . Conversion of constraints (8a) and (9a) for MCCTS would result in weak constraints like constraint (7b) for encountering military traffic in MSC.

V. PRE-OPTIMIZATION

Pre-optimization is a phase between formulation and solution of the model. It refers to data derivations and elementary operations that can be performed to improve or simplify the formulation. Our main concern is to reduce the number of decision variables to keep both models and their enhancements tractable on a PC. We start with a description of necessary and optional input data which all other data in the models are derived from.

A. USER INPUT

One of the goals in implementing the models is to make it easy for the user to enter and modify input parameters. To ensure easy model usability, minimal input should be required from the user in order to derive the data required by the optimization. The following data about the road network which is to be represented and about the characteristics of units whose routing is to be optimized must be entered.

- The total number of intersections and road segments in the road network and the total number of units.
- Description of the road segments containing road number, start and end point, and the road distance in kilometers. Information on weight limitation, one way or two way military traffic is optional.
- Information about the units, such as the number of vehicles, SP, RP, earliest possible start time, unit type, march organization (speed, vehicle gap and column gap). Additional information on time restrictions such as earliest possible arrival time or latest start time or information on priorities is optional.

- Duration of a time period in minutes.
- Number of time periods within which all movements should be completed. If the user is very pessimistic in this guess and specifies a number which is much too big, the model will compute a tighter upper bound on the time horizon. A guess too small causes some units to be not scheduled. In this case, the user must increase the total number of time periods.

Usually there are road segments which an experienced planner will never allow one or more units to use. For instance, some road segments may be usable by small tank units, but not for bigger tank units, a judgment based on current traffic flow information. The user is given a way to eliminate road segments for each unit individually which reduces the number of decision variables.

Models with incorrect data should not be solved. Thus, in addition to the checks performed by GAMS, several check statements are built in to terminate the model if some logical or numerical condition is violated. More sophisticated checks may be necessary. They should be part of a data management program which elicits all user input, formats it and moves it to the GAMS model, so that little or no user experience with GAMS code is necessary. This, however, is beyond the scope of this thesis.

B. DATA DERIVATION

Many parameters are derived from the information provided by the user. The data derivation is implemented within the GAMS code. A lot of derivations are simple and obvious. In the following, only main derivation procedures which require some

execution time or are important for reducing the number of decision variables are presented.

1. Derivation of $SHROUTE_{u,i,j}$

The parameters $EARRT_u$ and $C_{u,t,i,j}$ can only be built if the model has information on travel time between nodes. For the planning problem a shortest path from each node to each other node must be found.

The GAMS system includes a large library of 100 models, collectively called GAMS LIB (Brooke, Kendrick and Meeraus, 1992, pp. 177-181). The model SROUTEX (sequence number 93) finds the shortest route between all pairs by the Floyd-Warshall algorithm (Cormen, Leiserson and Rivest, 1991, pp. 558-562). To perform the derivation, another index jp for nodes is required which is not part of the formulation. Mathematically, the Floyd-Warshall algorithm works as follows:

for all units u ,

for all nodes i ,

while at least one improved shortest distance is found,

for all j ,

$$SHROUTE_{u,i,j} = \min_{jp} (SHROUTE_{u,i,jp} + DIST_{u,jp,j}).$$

The GAMS LIB implementation of Floyd-Warshall contains an error which is explained and corrected in the Appendix.

For real planning, deviations from the shortest route from SP to RP are permissible, but only within reasonable limits. Therefore, the model deletes,

individually for each unit, all the arcs in the static network whose inclusion in a route would lead to a route length which must be greater than or equal to a specified number greater than one times the shortest possible route. By default this number has a value of 1.33. So physical arcs are only considered by the models if their usage can still allow a route which is at most 33% longer than the shortest route. This arbitrary choice was found to work well during the tests, but can be easily changed by the user. The gain for the model is eliminating unnecessary arcs in the static network, and thus avoiding their replication in the time-expanded network.

2. Derivation of LT_u

Another way to make the model smaller is to introduce a unit-specific time horizon LT_u . This is the time period by the beginning of which the last vehicle of unit u must have arrived at RP. LT_u limits the total number of ways for each unit to conduct movements. This restriction is necessary since the total number of time periods for the model as entered by the user will usually be related to the unit with the latest estimated arrival time at RP. But, this does not necessarily mean that other units need the same number of time periods to complete their movements. The parameter LT_u , as used in the models, depends on $EARRT_u$, the shortest route and the number of vehicles in the unit as follows:

$$LT_u = EARRT_u + VEH_u + \text{Floor}(LF * (VEH_u + SHROUTE_{u,SOURCE_u,DESTIN_u})),$$

where LF is a factor which limits the time horizon. If LF equals zero, this forces the model to try scheduling on the shortest route without any time delay. With a value of

one, LF allows a maximum delay of $VEH_u + SHROUTE_{u,SOURCE_u,DESTIN_u}$ in MCCTS. In MSCTS, an LF of one allows a maximum delay for the last package of vehicles of $2 * VEH_u + SHROUTE_{u,SOURCE_u,DESTIN_u} - 1$. If a unit cannot be scheduled because LT_u is too small, the model automatically restarts with an increased value for LF . If LT_u is larger than an optional input that specifies when a unit u must have passed RP , then LT_u is decreased accordingly.

The author can give no guidelines in choosing a reasonable value for LF at this time. Big values which, for example, lead to $LT_u =$ time horizon of the model for all units u , may increase solution times dramatically. A value too small may result in two model runs (the first one unsuccessful), but the total execution and solve time may be less than that of a single run of a model where LF is too large.

3. Derivation of $C_{u,t,i,j}$

Arcs and nodes in the time-expanded networks for all commodities which will never be used should be eliminated to keep the model as small as possible. It was already shown that the model formulation does not need any hold-over arcs. The parameter representing unit-time-arc compatibility, $C_{u,t,i,j}$, is constructed to eliminate all impossible movement arcs. The values of $C_{u,t,i,j}$ could be manually entered by the user, as in (Lee, 1991, pp. 31-33), but for large numbers of units and physical arcs an automatic method is really necessary.

The derivation of $C_{u,t,i,j}$ is divided in three parts. First, only feasible movement arcs are considered. In MCCTS, the following algorithm is used:

for all units u ,

if u can start at node i at a time period not earlier than $ESTAT_u +$

$SHROUTE_{n,SOURCE_u,i}$ (no way to get to node i earlier),

and if u can leave node i on movement arc (i,j) at a time period not later

than $LT_u - VEH_u - SHROUTE_{u,j,DESTIN_u} - DIST_{u,i,j}$ (later start in i on (i,j)

does not allow to complete movement by LT_j),

$$C_{u,i,j} = 1.$$

In MSCTS, the parameter VEH_u is replaced by one, the length of a package of vehicles. Second, the derivation checks for time restrictions for units specified by the user. If a unit must have left SP by a certain time, then all movement arcs $(SOURCE_u,j)$ after that time are deleted. Movement arcs $(i,DESTIN_u)$ are excluded if they allow the unit to reach RP earlier than an earliest arrival time specified by input. In that case, the "distance-covered" part of the objective function becomes especially important to keep units at SP instead of circling around the road network in an absurd way until a move to $DESTIN_u$ is feasible. In the final part of the derivation, priorities of units for early completion of movements are considered and movement arcs $(i,DESTIN_u)$ for high priority units are removed if their use would result in an unacceptable delay.

VI IMPLEMENTATION AND RESULTS

A planning problem of routing units was designed as an example to show some performance aspects of the models under a variety of conditions. The GAMS implementation was tested on an AMDAHL 5995-700A at the Naval Postgraduate School and on a 486/33 MHz personal computer, with the solver ZOOM and a computer memory of 4 Megabytes. Results for different planning situations will be discussed to evaluate effectiveness of the proposed schedule and road utilization and to validate the model.

A. PROBLEM GENERATION

The hypothetical scenario studied in this chapter is based upon completion of a German brigade live exercise. It is assumed that the movements of only five units must be optimized. The exercise may contain other units, but they do not interfere and their movements can be planned manually. The five units, their strengths and deployments, are arbitrarily chosen. Information on each unit, A through E, and the road network is given in Tables 1 and 2. Unit E is designed to move at a different speed than the other units to invoke passing situations. A high priority for routing is entered for unit D. Figure 2 shows the road network with 10 intersections and 16 road segments that was arbitrarily chosen from a German map, and some basic data on the units (SP, RP, type). Basically, two west-east movements (units B and D) cross with north-south (units C and

Table 1: INFORMATION ON UNITS FOR PLANNING EXAMPLE

Unit	Nveh	SP	Estat	RP	Calc	Type	Prio	Gap
A	21	6	19.15	1	5	2	0	0
B	65	9	19.27	7	5	2	0	6
C	45	4	19.15	8	5	1	0	6
D	4	9	19.00	7	5	3	1	0
E	45	1	19.15	8	5(11)	0	0	4

Nveh:	Number of vehicles in u	Type:	Type of unit
SP:	Start point	0:	Truck
Estat:	Earliest time to start	1:	BFV
RP:	Release point	2:	Tank
Calc:	Code for speed and vehicle gap	3:	HET
	5: 35 km/h at 50 m	Prio:	Priority for unit
	11: 50 km/h at 50 m	0:	No priority
Gap:	All gaps between elements	1:	High priority
	of unit in minutes	2:	Top priority

E) and south-north (unit A) movements. Units obviously conflict on some road segments and at some intersections so careful planning is required.

All intersections are entered with a throughput capacity. The roads (3,7) and (7,8), drawn with bold lines, allow encountering units at any time. Encountering military traffic on the other road segments is undesirable, but excluded explicitly only for those road segments where the planner is particularly concerned based on his judgment. These road segments are displayed in the map. A model horizon of 50 time periods is specified. With a duration of two minutes per time period, this should be sufficient to complete all movements. The static network contains 10 nodes and 32 directed arcs. If all possible nodes and arcs were generated in all time-expanded

Table 2: INFORMATION ON ROAD NETWORK FOR PLANNING EXAMPLE

Road	Start point	End point	Distance in km	Oneway	Wclass	Encount
1	2	1	5.0	0		1
2	1	3	7.5	0		1
3	2	5	7.5	0		
4	2	4	7.0	0	3	
5	4	3	4.0	0		1
6	10	8	13.5	0		
7	3	7	6.5	0		
8	4	6	3.5	0		1
9	5	4	5.0	0	2	
10	5	9	6.0	0	2	
11	5	10	8.0	0		
12	10	6	8.5	0	1	1
13	6	7	4.0	0		
14	9	10	3.5	0		
15	10	7	10.5	0	1	
16	7	8	14.0	0		

<p>Oneway: 0 One-way road 1 Two-way road</p> <p>Encount: 0 Encounter allowed 1 No encounter</p>	<p>Wclass: Information on weight limitation of road</p> <p>0: No restriction 1: No HET 2: No HET, no tanks 3: Only trucks</p>
---	--

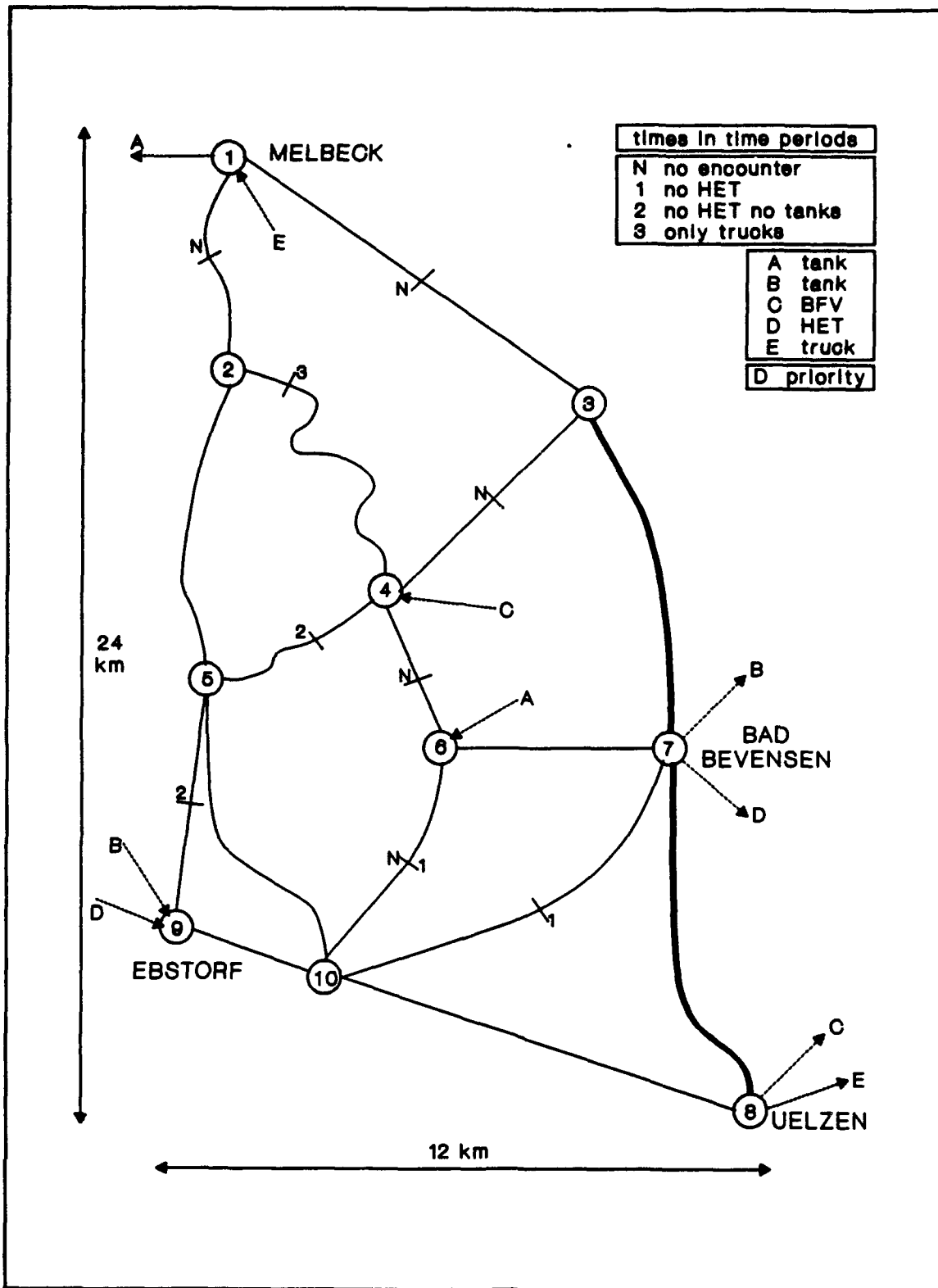


Figure 2: Road Network for Planning Example

networks (one for each unit), the complete program would contain up to 10290 decision variables (movement arcs and hold-over arcs).

B. RESULTS FOR MCC AND MCCTS

The close column formation is preferable to the dispatch of vehicles in groups and, consequently, more often used in practice. Therefore, we start with the results for MCC and MCCTS. In both implementations, an extensive use of multidimensional dynamic sets in GAMS is avoided to minimize memory requirements. This causes total run time to increase a bit, but it is believed that a small increase in run time is preferable to possibly running out of computer memory.

1. MCC (One Speed Class)

All units move in individual columns at a speed of 35 km/h, with a vehicle gap of 50 m. Tables 3, 4 and Figure 3 display the statistics and the optimal plan generated by MCC with LF set to 0.5. The scheduled times are represented in time periods. Unit D which is to be scheduled with high priority has no delay. Two units

Table 3: STATISTICS FOR MCC

	AMDAHL 5995-700A	486/33 MHz
GAMS time	6.4 sec	19 sec
Equations	433	433
0-1 variables	474	474
Non-zero elements	3420	3420
ZOOM time	3.7 sec	33 sec

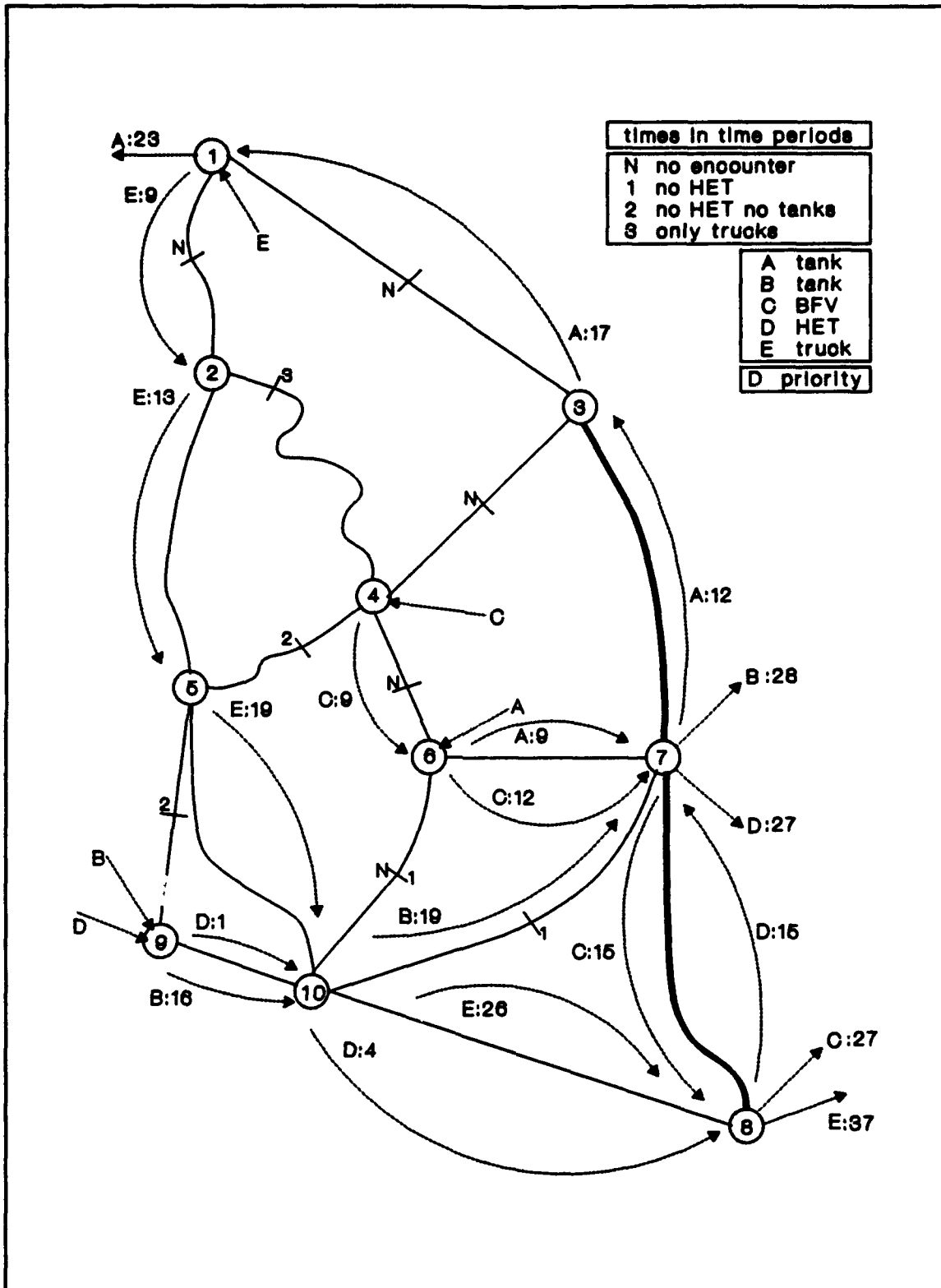


Figure 3: Routing for MCC

Table 4: MOVEMENT TABLE FOR MCC

Units	A	B	C	D	E
Earliest departure time	9	15	9	1	9
Earliest arrival time	21	27	27	27	32
Pass time of column	2	7	6	1	5
Time for shortest route	12	12	18	26	23
Scheduled departure at SP	9	16	9	1	9
Scheduled arrival at RP	23	28	27	27	37
Arrival of last vehicle	24	34	32	27	41
Delay for unit	2	1	0	0	5
Schedule of shortest route	no	yes	yes	yes	no
Total time delay for all units: 8					
All times in time periods (1 time period = 2 minutes)					

are scheduled for routes other than their shortest ones. The shortest route for unit A includes road (6,4) usable only for one unit at a given time. Instead of waiting until unit C passes intersection 6 so that road (4,6) is clear for A, a detour via intersection 7 delays the earliest possible arrival for A by only two time periods. Unit E, if scheduled for its shortest route via intersections 3 and 7, would interfere with unit A either at intersection 3 or on the road (1,3). Thus, it is beneficial in this situation to schedule two units for routes other than the shortest ones. The intersection utilization report that is provided by the model shows for the optimal plan that, for example, intersection 7 is used by units during the time periods 12 to 13, 15 to 20, 27 to 34. Units C and D encounter on road (7,8) which is allowed. Making this encounter not acceptable by appropriate input would cause a serious delay for at least one of the units.

2. MCCTS (Two Speed Classes)

In this scenario, the march speed for unit E is increased to 50 km/h, while all other parameters described above are held constant. Tables 5, 6 and Figure 4 summarize the statistics, the generated movement table and optimal schedule. The GAMS time for the model increases which is to be expected since, in addition to the previous run, constraint (9a) must be generated and additional computations are implemented to avoid generating unnecessary constraints (as described for this constraint). Units A, B, D keep their schedules. If the units C and E had the same schedule as in the previous run, they would interfere, now, at intersection 8 (the fast unit E would arrive while unit C is still passing). This conflict is resolved by scheduling unit E on its shortest route without a departure delay. Unit C waits long enough so that unit E will just have cleared intersection 7 when unit C arrives. An earlier start of unit C would lead to a passing conflict on road (7,8). The last vehicle of unit E clears intersection 7 at the end of time period 20. Thus, the earliest possible time for C to leave intersection 7 is time period 21. Intersection 8 is not used by any

Table 5: STATISTICS FOR MCCTS

	AMDAHL 5995-700A	486/33 MHz
GAMS time	10.0 sec	29 sec
Equations	374	374
0-1 variables	424	424
Non-zero elements	2990	2990
ZOOM time	3.8	48 sec

Table 6: MOVEMENT TABLE FOR MCCTS

Units	A	B	C	D	E
Earliest departure time	9	15	9	1	9
Earliest arrival time	21	27	27	27	25
Pass time of column	2	7	6	1	4
Time for shortest route	12	12	18	26	16
Scheduled departure at SP	9	16	15	1	9
Scheduled arrival at RP	23	28	33	27	25
Arrival of last vehicle	24	34	38	27	28
Delay for unit	2	1	6	0	0
Schedule of shortest route	no	yes	yes	yes	yes
Total time delay for all units: 9					
All times in time periods (1 time period = 2 minutes)					

vehicles between time periods 28 and 32. This time interval corresponds to the difference in travel time between a slow and a fast unit on road (7,8). Intersection 7 is frequently used with continuous utilization during the time periods 17 to 34. On road (3,7), two units move in both directions at the same time.

C. RESULTS FOR MSC AND MSCTS

The results from MCC and MCCTS were used to calculate a minimum value for LF, the factor which determines the upper bound on the unit-specific time horizons, LT_u , such that LT_u is still big enough for all units u to ensure that the optimal solution will be obtained. We found that $LF=0.2$ satisfied the condition. This value was used in the following runs, together with the objective function values of the previous runs as incumbent values.

1. MSC (One Speed Class)

Each group of vehicles is scheduled at a speed of 35 km/h with a vehicle gap of 50 m. The statistics and the optimal plan generated by the model are presented in Tables 7, 8 and Figure 5. The units B, C and D are scheduled for their shortest routes. The packages of B follow closely behind each other like a close column formation, but the last package of unit C is separated from the other packages of C. The vehicles of units A, B and D have the same schedule as in MCC. However, two packages of unit E are now able to use their shortest routes without any conflict at intersections 3 and 7. They hold up the last vehicle group of C at intersection 7, but the total delay for all units decreases from 8.0 to 6.3. The other packages of unit E must use the same route that was scheduled for MCC. The intersection utilization report reveals that the usage of intersection 7 increases by two time periods compared to MCC: This intersection is used during the time periods 12 to 13, 15 to 22 and 27 to 34, and each unit uses this intersection at some time. Employment of military police at this intersection to ensure that units can pass it smoothly may be worth considering.

Table 7: STATISTICS FOR MSC

	AMDAHL 5995-700A	486/33 MHz
GAMS time	4.7 sec	14 sec
Equations	309	309
0-1 variables	286	286
Non-zero elements	1228	1228
ZOOM time	1.6 sec	14 sec

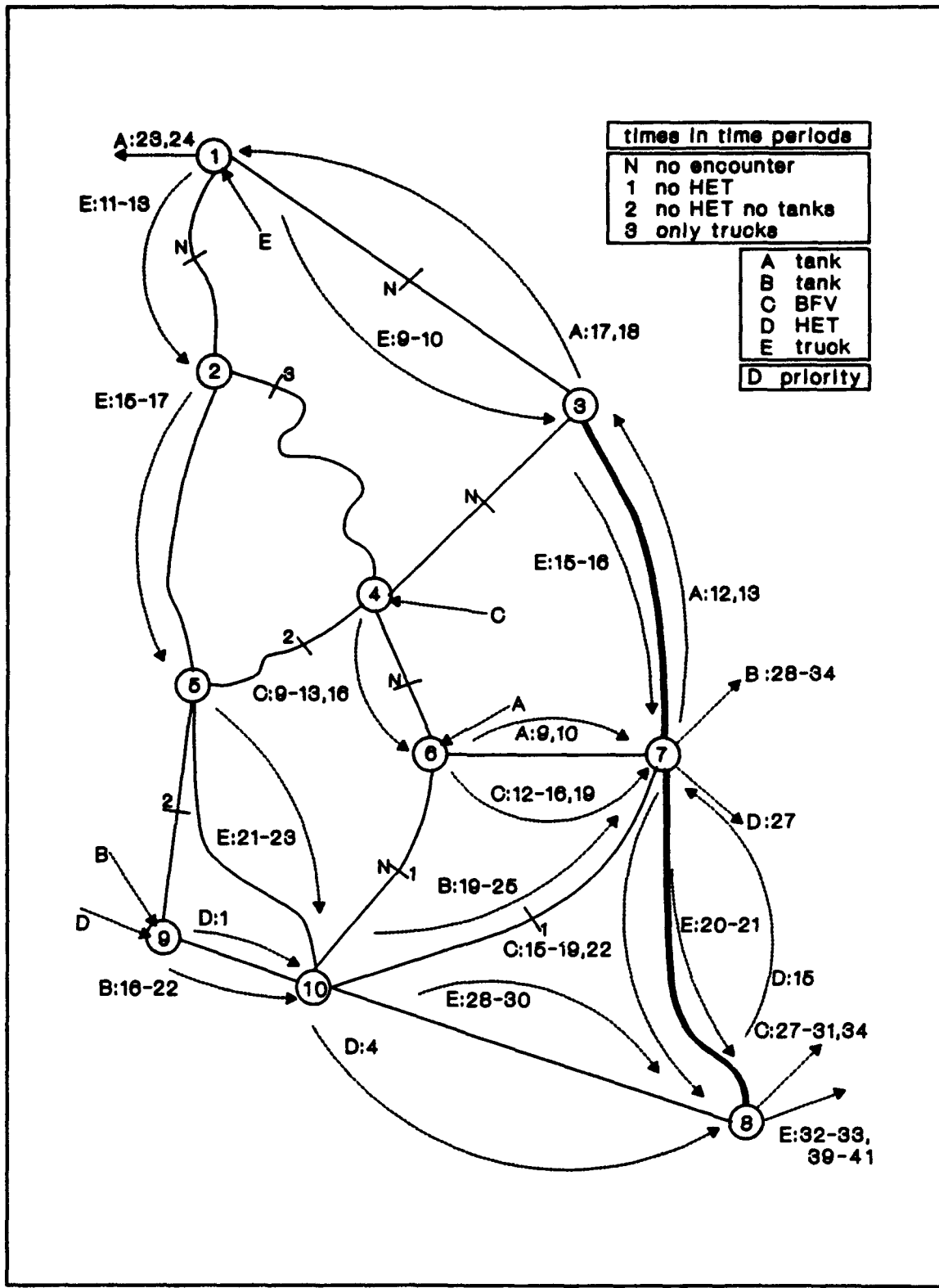


Figure 5: Routing for MSC

Table 8: MOVEMENT TABLE FOR MSC

Units	A	B	C	D	E
Earliest departure time	9	15	9	1	9
Earliest arrival time	21	27	27	27	32
Total length of packages	2	7	6	1	5
Time for shortest route	12	12	18	26	23
Scheduled departure at SP	9	16	9	1	9
Scheduled arrival at RP	23	28	27	27	32
Arrival of last vehicle	24	34	34	27	41
Delay for unit	2.0	1.0	0.3	0.0	3.0
Schedule of shortest route for all groups	no	yes	yes	yes	no
Total time delay for all units: 6.3					
All times in time periods (1 time period = 2 minutes)					

2. MSCTS (Two Speed Classes)

The packages of unit E move at an increased speed of 50 km/h, while all other parameters from MSC are held constant. Tables 9, 10 and Figure 6 show the statistics and the optimal plan. A comparison with MCCTS reveals that the packages of units D and E are scheduled like the vehicles in MCCTS. Four packages of unit C use the same route as scheduled in MCCTS. The other two packages depart from SP at the earliest possible time and move via intersections 5 and 10. Instead of waiting until intersection 7 is cleared by the fast packages of E, they bypass road (7,8) and manage to arrive at their RP at time periods 31 and 32, just before the other packages of C. This reduces the average delay for unit E because, in MCCTS, the first vehicle of E reaches RP at time 33. These two packages of C do not use road (6,10) to avoid encountering two packages of unit B. Splitting unit B is the only way to get

Table 9: STATISTICS FOR MSCTS

	AMDAHL 5995-700A	486/33 MHz
GAMS time	5.1 sec	15 sec
Equations	304	304
0-1 variables	257	257
Non-zero elements	1175	1175
ZOOM time	11.8 sec	131 sec

intersection 10 clear of packages of B at time periods 20 and 21, so that the packages of C can pass. Note that the packages of unit A arrive at RP as in MCCTS, but separate scheduling allows one package to move on its shortest route. The total delay for all units decreases from 9.0 to 7.7.

Table 10: MOVEMENT TABLE FOR MSCTS

Units	A	B	C	D	E
Earliest departure time	9	15	9	1	9
Earliest arrival time	21	27	27	27	25
Total length of packages	2	7	6	1	4
Time for shortest route	12	12	18	26	16
Scheduled departure at SP	10	15	9	1	9
Scheduled arrival at RP	23	28	31	27	25
Arrival of last vehicle	24	35	36	27	28
Delay for unit	2.0	1.7	4.0	0.0	0.0
Schedule of shortest route for all groups	no	no	no	yes	yes
Total time delay for all units: 7.7					
All times in time periods (1 time period = 2 minutes)					

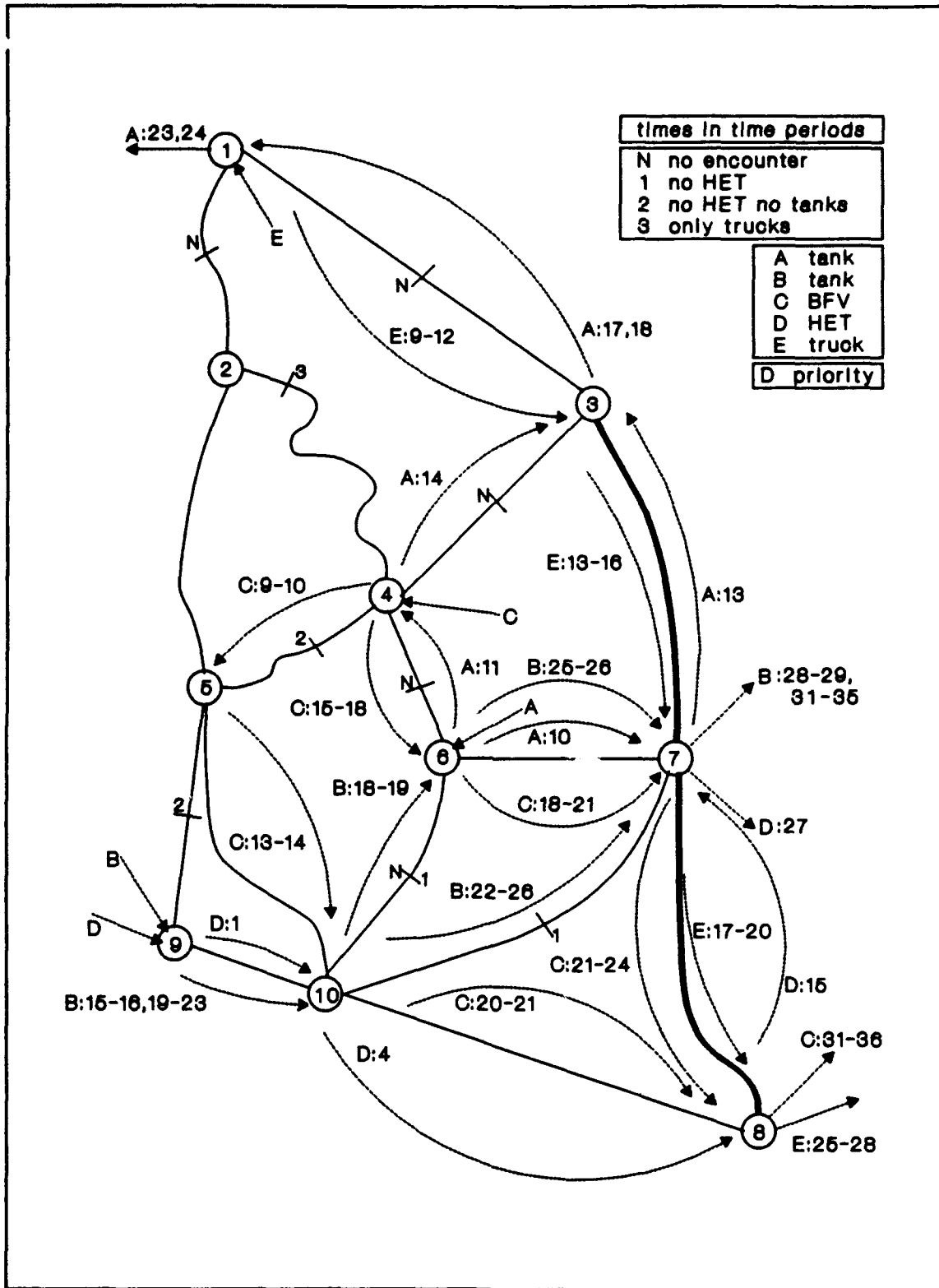


Figure 6: Routing for MSCTS

VII CONCLUSIONS AND FUTURE DIRECTIONS

Two different models for small-scale administrative motor movements have been described and compared in a realistic example. This chapter summarizes the presentation of the models and then lists two future directions of research.

A. CONCLUSIONS

Two optimization models, together with two enhancements, have been created to support movement planners in optimally routing military units in a road network. The objective is to minimize the total time delays of all units involved. MCC (model close column) schedules all vehicles of a military unit in a single move over the same route. All units move at the same speed. MCCTS (MCC two-speed) allows units to move at one of two different speeds. More flexibility in planning is allowed in MSC (model separate column) by splitting units into packages of vehicles which move at the same speed. MSCTS (MSC two-speed) routes packages of different units at one of two different speeds. The single speed models have been designed because routing all units at the same speed is the current norm. The two speed models account for the ability of truck units to move at higher speed than tank units. In the planning example, truck unit E could finish its movement 13 time periods earlier for both kinds of column formations if it was allowed to move at higher speed. At the same time, the total delay for all units involved increased by a small amount (1.0 time period for close column

formation, 1.4 time periods for scheduling units in separate groups of vehicles). Nevertheless, this clearly demonstrates the advantage of the two speed models versus the single speed models. Allowing certain units to move faster ought to result in earlier arrival times. The delay for a specific unit or the total delay of all units might increase, but the gain in time by earlier arrivals will outweigh a possible increase in delays. The planning example was constructed to invoke passing conflicts deliberately and, in this case, these conflicts could only be resolved by slightly increasing the total delay. A comparison between the two kinds of column formations in the planning example reveals that the dispatch of vehicles in groups resulted, as expected, in fewer delays, both in the single speed and the two speed models.

None of the models take variable transit times resulting from traffic congestion into account. During peacetime, it is standing procedure to plan movements at times and via routes which avoid rush hours and traffic congestion. Thus, the assumption of constant transit times and the other minor assumptions described in Chapter III (no changes in road network and no *a priori* scheduled movements during the study horizon) are not too limiting. *Node-arc formulations* of all the problems were solved using an exact algorithm to identify optimal solutions. Analysis in Chapter VI illustrated that the results are accurate and reasonable, and that planners can be effectively supported by the models.

B. FUTURE RESEARCH DIRECTIONS

Future research should take at least two directions. The first direction would improve the existing models. The second direction would take a different modelling approach to the unit routing problem.

1. Improvements in the Models

The initial success of both models leads to several areas for further investigation. The following improvements should be examined.

- The model should allow combinations of two or more units in a single column with specified gaps between the units. This issue arises when several units have a different SP, but a common RP.
- Additional constraints might be necessary to ensure that a slow unit does not get passed by a fast unit which uses the same road as the slow unit beyond RP, but is scheduled right behind the slow unit at their common RP. Defining RP "further down the road" would not help if the remaining road distance were 50 or more kilometers because the models would become intractable.
- The user should be given a choice of particular routes from SP to RP for each unit. In some cases, this input might be more precise than excluding road segments for movements which is currently implemented.
- The model could allow speed of units to become a decision variable. An appropriate choice of different speeds might reduce overall time delays. This would be quite difficult, however, because pass time, transit time on any physical arc and length of shortest path would no longer be constant in the mathematical formulation.

2. Arc-Path Formulation

This thesis has concentrated on a node-arc formulation of the routing problem; other formulations are possible, for example, an *arc-path formulation*. In the

arc-path formulation a network is denoted by paths from the source to the destination of a commodity and by enumeration of the individual arcs which belong to each path. The decision variables specify how much quantity of a commodity is to be transshipped on which path.

Running the model with bigger planning problems than described in Chapter VI (up to 870 binary variables) results in a large increase in solution time. Applying the X-System (Brown and Olson, 1993) as a solver instead of ZOOM enhances the performance significantly. Nevertheless, an arc-path formulation may ameliorate the problem of increased solution times in general. Such a formulation is described here for possible future research on the routing problem. This reformulation is presented only for MCC and MCCTS. The MSC and MSCTS reformulation is similar.

a. Arc-Path Formulation for Close Column Movements

In the following, a *physical path* for unit u is an alternating sequence of physical arcs that join its SP and RP. All military units are scheduled in a single move in individual columns. The decision variables are binary indicating whether or not a unit starts moving on a physical path at a particular time. All intersections are assumed to have a capacity limit of one column. All parameters which have the same notation as in the node-arc formulation (except the subscripts) maintain their original purposes.

Indices

i, j Nodes of the network.

u, u'	Units which have to be scheduled.
t, t'	Time periods in minutes.
e, e'	Arcs in the static network.
p	Physical paths in the static network.
s	Speed class for units, "SLOW" or "FAST".

Parameters and sets

VEH_u	Pass time of column u (including gaps) in time periods.
$ESTAT_u$	Earliest start time (period) for unit u to leave at SP.
PEN_u	Penalty for not scheduling unit u .
$DIST_{u,e}$	Time distance in time periods for unit u to move on arc e .
$ENC_{u,t,e}$	Minimum pass time over all units except u which are able to start their movements at time t on arc e' and might encounter u which moves on the anti-parallel arc e .
$C_{u,t,p}$	Equals one if unit u is allowed or able to start on path p at time period t , and is zero otherwise. This parameter is used to reduce the number of decision variables and appears in each equation.
$TRAVEL_{u,p,e}$	Time periods which are necessary for unit u to reach the start node of arc e if path p is chosen. Thus, all travel times to intermediate nodes are known and the model has to schedule only start times at SP.
$LENGTH_{u,p}$	Time periods required for unit u to traverse path p .
SHP_u	Length of shortest path in time periods for unit u .

NUM _i	Number of units in each speed class.
DIFF _e	Difference in travel time on arc e between "FAST" and "SLOW" units.
SF	Scale factor.
PATH _p	Set of physical arcs which belong to path p.
LINK _{e,e'}	Set which contains the anti-parallel arcs e and e' in the static network.
SLOW	Set which contains units that are "SLOW".
NE	Set of physical arcs on which encountering is not allowed.

Decision variables

X _{u,t,p}	Equals one if unit u starts on path p at the beginning of time period t, and is zero otherwise.
ELASTIC _u	Integer variable which is one if a unit cannot be scheduled and is zero otherwise.
Z _u	Positive variable equals the time delay for unit u.

Formulation

Minimize

$$\sum_u Z_u + \sum_u \sum_t \sum_p SF * LENGTH_{u,p} X_{u,t,p} + \sum_u PEN_u * ELASTIC_u$$

Subject to:

$$(1c) \quad Z_u \geq \sum_t \sum_p X_{u,t,p} (t - \text{ESTAT}_u + \text{LENGTH}_{u,p} - \text{SHP}_u) \quad \forall u$$

$$(2c) \quad \sum_t \sum_p X_{u,t,p} + \text{ELASTIC}_u = 1 \quad \forall u$$

$$(3c) \quad \sum_u \sum_{\substack{p \\ e \in \text{PATH}_p \\ e \text{ outgoing } i}} \sum_{\substack{t' \\ t - \text{TRAVEL}_{u,p,e} - \text{VEH}_u < t' < t - \text{TRAVEL}_{u,p,e}}} X_{u,t',p} + \\ + \sum_u \sum_{\substack{p \\ p \text{ ends in } i}} \sum_{\substack{t' \\ t - \text{LENGTH}_{u,p} - \text{VEH}_u < t' < t - \text{LENGTH}_{u,p}}} X_{u,t',p} \leq 1 \quad \forall t, i$$

$$(4c) \quad \sum_{\substack{p \\ e \in \text{PATH}_p}} \sum_{\substack{t' \\ t - \text{TRAVEL}_{u,p,e} - \text{DIST}_{u,e} + \text{ENC}_{u,t,e} \\ st' < t - \text{TRAVEL}_{u,p,e}}} X_{u,t',p} + \sum_{u' \neq u} \sum_{\substack{p \\ e' \in \text{PATH}_p \\ e' \in \text{LINK}_{e,e'}}} X_{u', t - \text{TRAVEL}_{u',p,e'} \cdot p} \leq 1 \quad \forall u, t, e, e \in \text{NE}$$

If $\text{NUM}_{\text{SLOW}} \leq \text{NUM}_{\text{FAST}}$

$$(5c) \quad \sum_{\substack{p \\ e \in \text{PATH}_p}} \sum_{\substack{t' \\ t - \text{TRAVEL}_{u,p,e} - \text{DIFF}_e \\ st' < t - \text{TRAVEL}_{u,p,e} - \text{VEH}_u}} X_{u,t',p} + \sum_{u' \in \text{SLOW}} \sum_{\substack{p \\ e \in \text{PATH}_p}} X_{u', t - \text{TRAVEL}_{u',p,e} \cdot p} \leq 1 \\ \forall u, t, e, u \in \text{SLOW}$$

If NUM_{SLOW} > NUM_{FAST}

$$(6c) \sum_{\substack{P \\ e \in PATH_p}} \sum_{\substack{t' \\ t - TRAVEL_{u,p,e} \leq t' \leq \\ t - TRAVEL_{u,p,e} + DIFF_e - VEH_u}} X_{u,t',p} + \sum_{u' \in SLOW} \sum_{\substack{P \\ e \in PATH_p}} X_{u',t-TRAVEL_{u',p,e}} \cdot P \leq 1 \quad \forall u, t, e, u \in SLOW$$

The objective function calls for the minimization of the sum of difference in time between earliest possible arrival time and scheduled time for each unit which is called delay. These individual delays are defined in constraint (1c). The second term in the objective equation assures that in case of a choice between several physical paths of different length which would lead to the same delay at RP, the shortest one is taken. The default value for SF is 0.01. A penalty is paid in the objective if a unit u cannot be scheduled and $ELASTIC_u$ takes on the value 1.

Equation (2c) says that a unit must start on exactly one physical path at some time period, but allows a unit to wait until a physical path is available. This equation will also set the elastic variable to one in case unit u is not scheduled. Again, $ELASTIC_u$ is defined as a nonnegative variable in the program. In this formulation flow balance constraints are not necessary.

Equation (3c) limits the throughput of node i to at most one column leaving the node at a time and it blocks the intersection during the pass time of a column for others. A physical arc cannot be entered by two or more units at the same time. Physical paths which end at node i must be included in the equation to represent the pass time of units whose RP is at node i (same problem as in constraint (5a)).

Constraint (4c) states that units are not allowed to encounter on certain road segments. A unit moving on arc e within certain time periods blocks the anti-parallel arc e' for a movement of other units.

In constraint (5c), all fast units which might start on arc e at time t are related to each slow unit which might have started at some time $t' \leq t$ and might be passed. Equation (6c) ensures that slow units starting on arc e at t prevent fast units from starting at certain times $t' \geq t$.

b. Results with the Arc-Path Formulation for Close Column Movements

With slight changes, almost all preprocessing features of the node-arc implementation can be used. A major problem, of course, is the generation of physical paths that are feasible, reasonable and sufficiently close to exhaustive. For this preliminary investigation, physical paths were enumerated and input manually. Physical paths with a length greater than 33% of the shortest possible path length are automatically excluded by the model. The results in Table 11 for the previously generated planning problem show that with respect to solution time the arc-path formulation outperforms the node-arc formulation whose ZOOM time is displayed in parentheses. We cannot compare the GAMS time of Table 11 because it does not account for the time which will be necessary to generate physical paths automatically. Several trials with other, bigger problems (up to 250 binary variables) always resulted in smaller solution times for the arc-path formulation.

The new formulation can be adapted to MSC and MSCTS with slight changes. The optimal solution for the planning problem with two speed classes was

Table 11: STATISTICS FOR ARC-PATH FORMULATION FOR CLOSE COLUMN MOVEMENTS

AMDAHL 5995-700A	One speed class	Two speed classes
GAMS time	3.8 sec	4.3 sec
Number of generated physical paths	16	16
Equations	166	134
0-1 variables	122	122
Non-zero elements	2482	2065
ZOOM time arc-path formulation	0.7 sec	0.6 sec
node-arc formulation	(3.7 sec)	(3.8 sec)

found with the solver ZOOM after 2.6 sec. (versus 11.8 sec. with the node-arc formulation).

The node-arc formulation was not in vain. Our experiences with preprocessing the data and the results from testing different ways of generating constraints could be used almost unmodified in the arc-path formulation. But, a lot of work might be necessary to automatically generate the set of possible physical paths.

APPENDIX. ERROR IN SROUTEX IN GAMSLIB

The GAMSLIB implementation of the Floyd-Warshall algorithm contains an error that was detected during the test runs of our models. For each unit and root node i the original model SROUTEX calculates, after each iteration within the while-loop, the sum of the current estimates of shortest distances from the current root node to all other nodes j . If two successive iterations give the same value for this sum, then the algorithm assumes it cannot improve the current solution and moves on to the next root. SROUTEX uses *INF* as the starting value for all estimates of shortest distances. *INF* is defined in GAMS as a special symbol to represent an arbitrarily large number that solver systems regard as infinite. GAMS has also defined the results of all arithmetic operations using this special value, e.g., $1+INF$ evaluates to *INF* (Brooke, Kendrick and Meeraus, 1992, p. 70) or $INF+INF=INF$. Thus, if some current estimates remain *INF* after the first and second iteration for a given root node (for example, for a sparse network), the sum of the estimates remains unchanged at *INF* after two successive iterations even if some other estimates might have improved during these iterations. Consequently, this algorithm falsely declares optimality when, in fact, a shorter path might be found at a later iteration.

Several ways to fix the problem are possible. The current model prevents the error by inserting a check for "sum equals *INF*" in the optimality test and by assuring

that $\text{SHROUTE}_{u,ij}$ is only computed to nodes j with at least one incoming arc. The problem could also be resolved by using a starting value different than INF, e.g., 10000 or 100000.

LIST OF REFERENCES

- Army Motor Transport Units and Operations*, HQDOA, FM 55-30, March 1980.
- Division Operations*, HQDOA, FM 71-100, June 1990.
- Aronson, J.E., A Survey of Dynamic Network Flows, *Annals of Operations Research*, v. 20, pp. 1-66, August 1989.
- Bodin, L., and others, Routing and Scheduling of Vehicles and Crews, *Computers and Operations Research*, v. 10, no. 2, pp. 63-311, 1983.
- Brooke, A., Kendrick, D., and Meeraus, A., *GAMS: A User's Guide*, Release 2.25, The Scientific Press, 1992.
- Brown, G. G. and Olson, M. P., *Dynamic Factorization in Large-Scale Optimization*, draft 93.01.25, Operations Research, Naval Postgraduate School, Monterey California 93943, January 1993.
- Cormen, T.H., Leiserson, C.E., and Rivest, R.L., *Introduction to Algorithms*, The MIT Press, 1991.
- Ford, L.R., and Fulkerson, D.R., *Flows in Network*, Princetown University Press, 1962.
- Lee, D.K., *Optimal Routing of Military Convoys Through a Road Network*, Master's Thesis, Naval Postgraduate School, Monterey, California, March 1991.
- Potts, R.B., and Oliver, R.M., *Flows in Transportation Networks*, Academic Press, 1972.
- Singhal, J.A., and Marsten R.E., *ZOOM/XMP-386 Version 2.2*, XMP Optimization Software Inc., 1990
- Weigel, H.S., and Cremeaus, J.E., The Multicommodity Network Flow Model Revised To Include Vehicle Per Time Period And Node Constraints, *Naval Research Logistics Quarterly*, v. 19, pp. 77-89, March 1972.
- Zawack, D.J., and Thompson, G.L., A Dynamic Space-Time Network Flow Model for City Traffic Congestion, *Transportation Science*, v. 21, pp. 153-162, August 1987.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
Cameron Station
Alexandria, Virginia 22304-6145
2. Superintendent 2
Attn: Library, Code 052
Naval Postgraduate School
Monterey, California 93943-5002
3. Richard E. Rosenthal, PhD 2
Department of Operations Research
Naval Postgraduate School
Monterey, California 93943
4. R. Kevin Wood, PhD 2
Department of Operations Research
Naval Postgraduate School
Monterey, California 93943
5. Fachinformationszentrum der Bundeswehr 2
Friedrich-Ebert-Allee 34
D 53113 Bonn
Federal Republic of Germany
6. Panzerflugabwehrbatallion 12 2
- Kommandeur -
Carl-Schurz-Kaserne
D 74736 Hardheim
Federal Republic of Germany
7. Johann Thoma 2
Schloßstraße 2
D 74722 Buchen-Bödigheim
Federal Republic of Germany