Mathematical Models and Algorithms for Logistics Operations: The Case of Hazardous Materials

Konstantinos N. Androutsopoulos
Transportation Systems and Logistics Laboratory
Department of Management Science and Technology
Athens University of Economics and Business
translog@aueb.gr

Abstract

Hazardous Materials transportation risk mitigation may be achieved by reducing the accident probability and alleviating the associated potential consequences. The consideration of risk as a criterion for selecting Hazardous Materials distribution routes contributes substantially to the reduction of accident probability and the severity of the accident consequences. Furthermore the optimum deployment of the emergency response units for covering the hazardous materials routes contributes significantly to the reduction of the accident consequences. Vehicle routing and emergency response facility location are two of the most frequently faced logistical decisions in Hazardous Materials distribution management and emergency response planning respectively. In particular, Hazardous materials routing can be defined as a bi-objective vehicle routing problem with time windows since risk minimization accompanies the cost minimization in the objective function. Furthermore, the emergency response unit location problem aims at locating a set of response units in order to optimize the performance of the system. The heavy computational burden for solving the emerging mathematical models especially in cases of large-scale distribution networks, led to the development of two new heuristic algorithms. The algorithm for the Hazardous Materials routing problem has been applied to several benchmark problems providing results superior to other competing algorithms. Due to the lack of benchmark problems, the algorithm for the emergency response units location problem has been applied to several case studies providing near optimal solutions.

Introduction

Hazardous Materials Transportation play a significant role in modern economies. It has been estimated that one out of five trucks in the United States of America carries Hazardous Materials (Erkut and Verter, 1998) while the Hazardous Materials Transportation covers 5-15 % of the global land freight transport (Kuncyte et al, 1998). A major feature of the hazardous materials transportation relates to the potential severity of Hazardous Materials accidents. The en-route hazardous materials accidents may produce consequences (fire, explosion, chemical spills, infection etc.) that could endanger human lives and/or cause property damages and environmental pollution. For example a gasoline truck explosion in Afghanistan in 1982 caused 700 deaths. Apart from causing fatalities, the hazardous materials accidents require complex and effective emergency response and civil protection operations. A train accident in Missisagua of Ontario in 1979 required the evacuation of 200000 residents due to the radical release of toxic chemicals (Erkut and Verter, 1995). Although Hazardous Materials accidents are extremely rare events, the expected economic, environmental, and human losses from such accidents are high (List et al, 1991; Zografos & Davis, 1989). This fact raises dramatically the importance of securing a high level of safety for Hazardous Materials Transportation from both the perspective of carriers/shippers and the public authorities.

Transportation risk is expressed by the expected consequences of hazardous materials accidents (measured by the product of the accident probability and the measure of the associated consequences e.g. population exposure). In this context, the hazardous materials transportation risk mitigation involves the provision of preventive and repressive safety measures for decreasing the probability of hazardous materials accidents and reducing the relevant potential consequences. . Rationalising the planning process of hazardous materials transportation with the risk mitigation objective involves two critical logistical problems: (i) the specification of Hazardous Materials Distribution routes of minimum cost and risk, and (ii) the optimum deployment of emergency response units for covering efficiently the associated hazardous materials routes.

The determination of Hazardous materials distribution routes can be formulated as a bi-objective vehicle routing problem with time windows by adding the risk minimization criterion to the cost minimization in the objective function. Furthermore, the emergency response unit location problem aims at locating a set of first response emergency service units in order to minimize the total response time while complying with the following constraints: i) covering the hazardous materials routes within a specified time, and ii) equal workload distribution among the units.. The complexity of these two problems causes a heavy computational burden for their solution especially in cases of large-scale networks. Therefore, heuristic algorithms were proposed for the solution of the aforementioned problems (Desrosiers et al, 1995; Marianov & ReVelle, 1995). The objective of this paper is to present the aforementioned logistical problems and provide two new heuristic algorithms for their solution. Both algorithms have been applied to several benchmark problems in order to test their computational efficiency.

The rest of this paper consists of three sections. Section two defines the Hazardous Materials Distribution problem and describes the new algorithmic approach in solving the emerging vehicle routing problem. Section three describes the problem of locating the emergency response units for Hazardous Materials incidents and provides a detailed presentation of the proposed heuristic algorithm for solving it. Finally, section four presents concluding remarks on this work.

Hazardous Materials Distribution

The objective of this section is to define the problem of determining efficient distribution routes, present the mathematical formulation of the Hazardous Materials Distribution problem and describe a new heuristic algorithm for solving it.

Definition of the Problem

The objective of the Hazardous Materials Distribution problem is to determine the most economic and safe routes for shipping Hazardous Materials to prespecified intermediate stops using a given fleet of vehicles. The following assumptions are taken into account:

- each vehicle starts and terminates its route at the depot
- each stop (order) must be serviced within a prespecified time window
- the fleet of trucks is assumed non-homogeneous
- the shipments refer to either one material (e.g. gas) or to a mix of packed materials that can be carried on the same truck (thus excluding materials that require trucks carrying tanks with separate compartments)

It is evident that the Hazardous Materials Distribution problem belongs to the vehicle routing and scheduling problems category.

A basic characteristic of the Hazardous Materials vehicle routing problem relates to the introduction of the transportation risk into the criteria for route selection. The hazardous materials transportation risk experienced within roadway segment (s) is expressed by the expected consequences of a hazardous materials accident presented by formulae (1):

$$R_s = p_s C_s \tag{1}$$

where

p_s is the Hazardous Materials accident probability,

C_s is the population that is contained in distance l from segment s.

Furthermore the risk over the whole route is defined by formula (2):

$$R_p = \sum_{s \in p} p_s C_s \tag{2}$$

Given a mathematical graph $G(N, A)$ where N is the set of nodes, A is the set of links, each node in the set $N \setminus \{0\}$ denotes a customer, $\{0\}$ denotes the depot, and every link (i, j) denotes the transportation link (i.e. the shortest path) from customer i to customer j, the Hazardous Materials Vehicle Routing Problem could be formulated as a mixed integer programming model defined by (3)-(16).

$$Minimise \qquad z_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{v=1}^{NV} x_{ij}^{v} c_{ij} \tag{3}$$

$$z_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{v=1}^{NV} x_{ij}^{v} R_{ij} \tag{4}$$

subject to:

$$\sum_{i=1}^{n} \sum_{v=1}^{NV} x_{ij}^{v} = 1 \qquad j = 1, \ldots, n \tag{5}$$

$$\sum_{j=1}^{n} \sum_{v=1}^{NV} x_{ij}^{v} = 1 \qquad i = 1, \ldots, n \tag{6}$$

$$\sum_{i=0}^{n} x_{ip}^{v} + \sum_{j=0}^{n} x_{pj}^{v} = 0 \quad (7)$$

$$v = 1,..,NV,\ p = 0,..,n$$

$$\sum_{i=0}^{n} d_i (\sum_{j=1}^{n} x_{ij}^{v}) \le K_v \qquad v = 1,...,NV \quad (8)$$

$$x_{ij}^{v} = 1 \quad \Rightarrow \quad T_i + s_i + t_{ij} \le T_j \quad (9)$$

$$i = 1,...,n, \quad j = 1,...,n$$

$$x_{j0}^{v} = 1 \quad \Rightarrow \quad T_i + s_i + t_{i0} \le T_A^{v} \quad (10)$$

$$i = 1,...,n,$$

$$x_{0i}^{v} = 1 \quad \Rightarrow \quad T_D^{v} + t_{0j} \ge T_j \quad (11)$$

$$j = 1,...,n,$$

$$a_i \le T_i \le b_i \qquad i = 1,...,n \quad (12)$$

$$a_0 \le T_A^{v} \le b_0 \qquad v = 1,...,NV \quad (13)$$

$$a_0 \le T_D^{v} \le b_0 \qquad v = 1,...,NV \quad (14)$$

$$x_{ij}^{v} \in \{0,1\} \quad (15)$$

$$T_i \in \Re \quad (16)$$

where, x_{ij}^{v} is a binary variable that takes value 1 if vehicle v services the customer located at node i and heads towards node j and 0 otherwise, T_i is the moment in time that service at node i begins, T_A^{v}, T_D^{v} are two variables that express the arrival (A) and departure (D) time respectively of vehicle v at the depot, and s_i is the duration of service at customer i. The main attributes of links of G are: (a) the travel time (t_{ij}) of the shortest path from i to j, (b) R_{ij} which denotes the transportation risk generated on the shortest path from i to j. Each point of demand $i \in N \setminus \{0\}$ is associated with two attributes: (a) demand (d_i), and (b) a time window of service $[a_i, b_i]$.

The objective function z_1 expresses the total travel time of the distribution process while objective function z_2 expresses the respective total transportation risk. Constraints (5)-(7) secure that all nodes are serviced only once by a unique vehicle. Constraint (8) implies that the total demand serviced by each route should not exceed the capacity of the vehicle. Constraints (12), (13) impose upper and lower bounds on the times that the vehicles may start servicing the demand at each node. Constraint (14) implies that a vehicle may not depart from the depot sooner than a_0 nor later than b_0. Furthermore, constraints (9), (10) and (11) assure that if a vehicle moves from node i to node j the time it needs to get at j is at least the service time at i plus the travel time from i to j.

It is evident from the mathematical formulation (3)-(16) that the hazardous materials distribution problem is modelled as a bi-objective vehicle routing problem with time windows. The solution of this problem is achieved through a new algorithmic approach presented at the next section.

Heuristic Algorithm for the Hazardous Materials Vehicle Routing Problem with Time Windows

The solution of the bi-objective vehicle routing problem with time windows should lead to the determination of the set of non-dominated (or non-inferior) solutions S. The Weighting Method was used as the solution approach for the problem at hand. The Weighting Method determines a set of non-dominated solutions I ($\subseteq S$) by solving the single objective vehicle routing problem with time windows where the objective function expresses the weighted sum of z_1 and z_2. The emerging integer-programming model is expressed through (5)-(18).

$$Min\{w_1 \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{v=1}^{NV} x_{ij}^{v} c_{ij} + w_2 \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{v=1}^{NV} x_{ij}^{v} R_{ij}\} \quad (17)$$

$$\sum_{i=1}^{2} w_i = 1 \text{ and } w_i \in [0,1] \quad (18)$$

The application of the Weighting method transforms the bi-objective problem into a set of single objective problems. This transformation justifies the use of a fast and efficient heuristic algorithm for approximating the set of non-dominated solutions of the problem at hand. However, exploring the stock of heuristic algorithms for the vehicle routing problem with time windows it can be verified that no such algorithm exists since the route building heuristics are fast but they do not produce near optimal solutions while more sophisticated algorithms like Improvement heuristics, Tabu search, and Genetic algorithms require a lot of computational time in order to determine solutions of high quality (Gendreau et al, 1997; Kontoravdis & Bard, 1995; Potvin & Rousseau, 1993, 1995; Potvin & Bengio, 1996; Russel, 1985). Alternatively, a new route building algorithm for the single objective vehicle routing problem with time windows has been developed that outperforms existing competitive algorithms. In what follows we present the new route building algorithm and a description of its use in solving the bi-objective vehicle routing problem with time windows. The heuristic algorithm proposed for solving the set of single objective vehicle routing problems is an insertion algorithm i.e. it builds the routes step by step by inserting in the already existing routes a new demand point at each iteration. Although the proposed algorithm has many common features with the respective insertion algorithm developed by Solomon (Solomon, 1987), it differs from it in the selection of the next demand point for insertion. More specifically, while the Solomon's algorithm allows the insertion of the unrouted demand points the proposed heuristic algorithm allows the insertion of both routed and unrouted demand points enabling in this way the reinsertion of the demand points at some more globally beneficial position. The basic steps of the proposed algorithmic approach are the following (Zografos and Androutsopoulos, 2004):

(1) Insertion of a new customer in a new route. In this step the construction of a new route is initiated. Three alternative ways are used for the initialisation of a route: (a) insertion of the unrouted customer that lies farthest from the depot, (b) insertion of the customer with the shortest lower time bound of its time window, or (c) insertion of the unrouted customer with shortest upper bound of its time window.

(2) If all customers are routed then the process terminates. Otherwise for every customer (both routed and unrouted) identify its best insertion position to the already existing routes. The best insertion position is defined on the basis of the minimization of the metric c_1:

$$c_1(i,u,j)=d_1(d_{i,u}+d_{u,j}-d_{ij})+ \\ d_2(R_{i,u}+R_{u,j}-R_{ij})+d_3(b_{j_u}-b_j) \qquad (19) \\ d_1,d_2,d_3\in[0,1]$$

where d_{ij} is the travel time from i to j, and b_i is the service start time at customer i, R_{ij} is the risk on link (i,j) and d_1, d_2, d_3 are weighting factors.

(3) Selection of the new customer to be inserted. The criterion for selecting the next customer to be routed is the minimization of metric c_2 .

$$c_2(u)=a_1t(u)+a_2R(u)+a_3S(u) \qquad (20)$$

where $t(u), R(u)$ are the total travel time and risk respectively after inserting u taking in to account that each unrouted customer adds the quantity $a_1(t(0,u)+t(u,0))+a_2(R(0,u)+R(u,0))$ to it. $S(u)$ denotes the total schedule time of the distribution process.

(4) Insertion of the selected customer, update the service start times of each routed customer and return to step 2.

It is necessary to emphasise at this point that this procedure is run for several values of $a_i\in[0,1]$ such that $\sum_i a_i=1$. These runs lead to the proposed set of solutions. Furthermore it is essential to highlight some crucial technical details of the implementation of the proposed algorithm. According to the state of the art and state of practice, the values of the Hazardous Materials accident probability range between 10^{-8}-10^{-6} (List et al, 1991). As a direct consequence the respective risk values are of the same magnitude. However this fact raises an important impedance in capturing non-dominated solutions since weights $a_i\in[0,1]$ are taking values from 0 till 1 with step 0.1 it is inevitable that any of these values would favour the travel time objective. In order to overcome this technical obstacle in applying the algorithm, a scaling technique has been used. The objective of this technique is to transform the values of travel time and risk into a common range of values. The scaling has been accomplished by dividing both attributes of a link (i,j) by specific scaling factors i.e. t_{sf}, R_{sf} respectively such that:

$$t_{sf}=\frac{\sum_{(i,j)}t_{ij}}{n(n-1)} \qquad (21)$$

$$R_{sf}=\frac{\sum_{(i,j)}R_{ij}}{n(n-1)} \qquad (22)$$

Based on this scaling technique, the attributes t_{ij}, R_{ij} appearing in the formulas of c_1, c_2 are replaced by the following attributes:

$$t'_{ij} = \frac{t_{ij}}{t_{sf}} \qquad (23)$$

$$R'_{ij} = \frac{R_{ij}}{R_{sf}} \qquad (24)$$

More details on the proposed algorithm can be found in (Zografos & Androutsopoulos, 2004).

Computational Performance of the Proposed Vehicle Routing and Scheduling Algorithm

The computational performance of the algorithm was tested in both single and bi-objective vehicle routing and scheduling problems. More specifically in the single objective case, the weights a_2, d_2 were set equal to 0 permanently and no scaling technique was applied. The proposed algorithm was applied on the set of Solomon's benchmark problems that consists of several representative single objective vehicle routing and scheduling problems of 100 customers and a single depot. The set of Solomon's benchmark problems consist of six categories namely R1,R2,C1,C2,RC1, and RC2. In categories R1, R2 the positions of the customers have been generated randomly from a uniform distribution. On the other hand the positions of the customers of problems within C1 and C2 have been generated in clusters while the respective positions in problems within RC1, RC2 are semi-clustered (Solomon, 1987). The assessment of the performance of the proposed algorithm on the Solomon's benchmark problems is based on the following criteria: (i) the average number of routes produced over the problems of each category, (ii) the average travel time over the problems of each category, and (iii) the average schedule time over the problems of each category.

The results of this application are presented in Table 1. Tables 2-4 presents the respective results of other competitive route building algorithms (Solomon, 1987). The comparison of the results of the alternative algorithms with the results of the proposed heuristic is quite encouraging for the performance of the new algorithm. The proposed algorithm outperforms the competitive route building algorithms i.e. Sweep, Solomon's I_1 and I_2 at problem sets R1, R2, RC1, and RC2.

	NEW-I$_2$		
	NV	TT	ST
R1	13.25	1367	2503
R2	3.09	1261	2407
C1	10.44	1150	10299
C2	3.25	708	9756
RC1	13.25	1557	2673
RC2	3.6	1517	2673

TABLE 1 The results of the application of the heuristic algorithm New I_2 on the R1, R2, C1, C2, RC1, RC2 problems

	Sweep Algorithm		
	NV	TT	ST
R1	14.6	1449.7	2817.4
R2	3.2	1448.6	2590.1
C1	10.0	940.8	10133.8
C2	3.0	711.9	9755.2
RC1	14.9	1804.5	3094.5
RC2	4.0	1735.7	3007.9

TABLE 2 The results of the application of the heuristic algorithm Sweep on the R1, R2, C1, C2, RC1, RC2 problems

	Solomon's I$_1$		
	NV	TT	ST
R1	13.6	1436.7	2695.5
R2	3.3	1402.4	2578.1
C1	10.0	951.9	10104.2
C2	3.1	692.7	9921.4
RC1	13.5	1596.5	2775.0
RC2	3.9	1682.1	2955.4

TABLE 3 The results of the application of the heuristic algorithm I1 on the R1, R2, C1, C2, RC1, RC2 problems

	Solomon's I$_2$		
	NV	TT	ST
R1	14.5	1638.7	2888.1
R2	3.3	1470.7	2645.8
C1	10.1	1049.8	10174.3
C2	3.4	921.5	10151.4
RC1	14.2	1874.4	3029.5
RC2	4.1	1797.6	3128.4

TABLE 4 The results of the application of the heuristic algorithm I2 on the R1, R2, C1, C2, RC1, RC2 problems

Due to the lack of benchmark problems for the bi-objective vehicle routing and scheduling problems the proposed algorithm was applied in a set of problems designed by the authors. The relevant distribution scenario is defined by one depot, a set of eight customers with demands that range from 1-5 tons and an available homogeneous fleet of vehicles with capacity equal to 20 tons. The positions of the customers on the Euclidean plane were randomly selected within the square defined by the points (0,0), (0,100), (100,100), and (100,0). The service times at the customers ranged from 15-25 minutes. The travel time between each pair of customers was assumed equal to the Euclidian distance. On the other hand the risk between each pair of points in the plane (customer or depot) was measured by using the formula (25).

$$R_{ij} \coloneqq n_{ij} * d_{ij} * Pop_{ij} \qquad (25)$$

where n_{ij} is the frequency of Hazardous Materials accident per km, d_{ij} is the distance between i and j, and Pop_{ij} is the population within 2 km from the spot of a Hazardous Materials accident.Alternative problems were produced through changing the time windows and the risk values: (1) tight time windows and non-uniform risk values along the distribution route, (2) loose time windows and non uniform risk values, (3) tight time windows and uniform risk values, and (4) loose time windows and uniform risk values. The duration of a typical tight time window did not exceed 10% of the planning horizon of the problem while the respective value for the loose time window was 25%. In addition, the terms uniform and non-uniform risk values relate to the variance of the risk values between each pair of customers.

The scope of the evaluation of the algorithm was to test its capability to identify non-dominated solutions to all four scenarios. A subset of the set of non-dominated solutions was determined by solving 100 single objective problems for each scenario by using the Mathematical Programming Platform AIMMS (Bisschop & Entriken, 1997). The objective of the evaluation of the proposed algorithm is to secure that the generated solutions are not dominated by the efficient solutions identified by the AIMMS application. Tables 3 and 4 summarize the results of both solution methods (Algorithm vs AIMMS application respectively) under the four routing and scheduling scenaria. The efficient solutions produced by the AIMMS applications are denoted by S_{ij} where i denotes the identification number of the specific scenario and j denotes the identification number of the solutions within the set of solutions identified for scenario i. An analogous symbolism is used for the solutions produced by the proposed algorithm. Each of the two tables consists of five columns. Looking from left to right the first column contains the identification symbol of each solution, the second column contains the scenario identification number for each solution, the third column contains the travel time and the fifth column contains the total risk for each solution. As it can be verified the solutions that are generated by the algorithm are either equal to the efficient solutions defined by the AIMMS application or at least not dominated by them. More information on the evaluation of the proposed heuristic approach can be found in (Zografos & Androutsopoulos, 2004).

Solution ID	**Scenario ID**	**Number of Routes**	**Total Route Time (min)**	**Total Risk (x10^{-})^{6}**
S_{11}	1	3	651	54771,4
S_{12}	1	3	565	65150
S_{13}	1	3	482	83623
S_{14}	1	3	476	93228
S_{21}	2	2	597	22727,2
S_{22}	2	2	512	26188,4
S_{23}	2	2	426	36567,1
S_{31}	3	2	405	10116,3
S_{41}	4	3	588	14069,2
S_{42}	4	3	488	14812,6
S_{43}	4	3	476	18307,8

TABLE 3 Features of the non-dominated solutions produced by the AIMMS application to scenaria 1-4

Solution ID	**Scenario ID**	**Number of Routes**	**Total Route Time (min)**	**Total Risk (x10^{-})^{6}**
?$_{11}$	1	4	689	50924
?$_{12}$	1	3	620	65099
?$_{13}$	1	3	565	65145
?$_{14}$	1	3	482	83616
?$_{15}$	1	3	651	54766
?$_{16}$	1	3	649	5868
?$_{17}$	1	3	540	79668
?$_{21}$	2	2	426,1	36562
?$_{22}$	2	2	405	61970
?$_{23}$	2	2	512	26183
?$_{31}$	3	2	405	10110
?$_{41}$	4	3	495	15253
?$_{42}$	4	3	482	18749

TABLE 4 Features of the solutions produced by the application of the proposed algorithm to scenaria 1-4.

The efficiency of the proposed algorithm led the authors to the development of a computational tool for supporting decisions related to: (i) the identification of minimum cost routes, (ii) the identification of minimum risk routes, and (iii) the identification of alternative non-inferior minimum cost/risk routes, and (iv) specification of the schedule of each vehicle.

Hazardous Materials Emergency Response Unit Location Problem

Definition of the Problem

In general, hazardous materials accidents are treated by specially trained teams through performing specific repressive measures depending on the type and extend of the associated consequences. Usually these teams and the required equipment are located on specified sites while their involvement in a hazardous materials accident is activated only upon notification from the organization that verifies the accident (e.g. fire department). The first response actions involve the verification of the attributes of the accident and the notification of the actors commissioned with the management, coordination, an performance of the relevant emergency response process. It is evident that the timely arrival of the emergency response unit at the scene of the accident speeds up the activation of the emergency response process and increases the chances of avoiding or impeding the impacts of a hazardous materials accident. The first response emergency service units are not necessarily devoted to the hazardous materials accidents but they cover additional types of emergency calls too. However, locating the first response emergency service units in order to cover the hazardous materials routes within the area under study, aims to enhance the effectiveness of the emergency response system.

The emergency response calls for the fire department can be traced either on a roadway segment or on a residential area (e.g. a house). In order to facilitate the mathematical formulation of the location problem, each potential source of calls is associated with the nearest roadway segment. Based on this assumption we may consider that calls can be generated on the roadway network solely. Furthermore, the roadway network is transformed in to a mathematical graph by applying the following basic rules:

(i) Roadway Segmentation. Every roadway link is divided into segments of equal length (e.g. 800m)

(ii) Node Definition (N). Each segment (as resulted from the segmentation process above) is represented in the mathematical graph by a node.

(iii) Links definition (A). Two nodes (i,j) defined in (ii) are joined with a link in A only if i and j are consecutive (i.e. directly connected) roadway segments

There are also other special considerations that rule the transformation of the roadway network to a mathematical graph. However these details are out of the scope of this paper.

Given a mathematical graph G(N,A) and assuming $p_i \quad (i \in N)$ is the frequency of emergency calls from node i the proposed mathematical model is expressed by constraints (26)-(34). The proposed mathematical formulation resembles the mathematical model that has been developed and implemented for the optimum deployment of Emergency Restoration Units of Electric Utility Companies (Zografos et al, 1998), and the districting problem accompanying the location of a set of Emergency Response Units for roadway incidents (Zografos et al, 1993, 1994; Zografos & Androutsopoulos, 2002). The location problem for hazardous materials emergency response units has been previously addressed by Saccomanno & Allen (1988). However, their approach did not take into account the equal workload constraints introduced in (Zografos et al, 1993, 1994) while the coverage of the units referred solely to risk minimization ignoring the overall performance of the system.

$$Minimize \quad \sum_i \sum_j \left(\sum_{k=1}^{K} l_i^k\right) t_{ij} Y_{ij} \tag{26}$$

subject to:

$$Y_{ij} \le x_j \qquad \forall i, j \tag{27}$$

$$\sum_j x_j = p \tag{28}$$

$$\sum_j Y_{ij} = 1 \qquad \forall i \tag{29}$$

$$\frac{\sum_i p_i}{p}(1+e) \ge \sum_i Y_{ij} p_i \ge \frac{\sum_i p_i}{p}(1-e) \qquad \forall j \tag{30}$$

$$\sum_{j \in J} Y_{ij} t_{ij} \le S_1, i \in I \tag{31}$$

$$\sum_{j \in J} Y_{ij} t_{ij} \le S_2, i \in H \subseteq I \tag{32}$$

$$Y_{ij} \ge 0, \quad \forall i, j \tag{33}$$

$$x_j \in \{0,1\}, \quad \forall j \tag{34}$$

where e is the acceptable deviation from average workload, K is the number of alternative types of emergency calls, H is the set of nodes that hazardous materials are passing through (Hazardous Materials nodes), t_{ij} is the travel time between points i and j, S_1 is the maximum travel time between the nodes within a district in G from their respective base, while S_2 is the maximum travel time between the Hazardous Materials nodes within a district and their respective emergency response unit base, l_i^k is the rate of call of type k at node i, x_j is a binary variable that takes value 1 if a response unit is located at node i and 0 otherwise, Y_{ij} is a binary variable that takes value 1 if source of calls j is covered by facility location i and 0 otherwise. It should be clarified at this point that the workload at each node i is defined by formulae (35).

$$p_i = \sum_{k=1}^{K} l_i^k S^k \qquad (35)$$

where S^k is the average service time of an emergency call of type k.

The objective function (26) expresses the expected response time of the system. Constraint (27) implies that the source of calls j can be serviced by node i if and only if a facility is located at node i. Constraint (28) implies that the total number of emergency response units is p. Constraint (29) implies that each source of emergency response calls is covered by a unique emergency response unit while constraint (30) implies that the workload assigned to each facility should not deviate more than (e x100)% from the average workload. Constraint (31) implies that the travel time between a response unit base and the nodes within its area of responsibility should not exceed the standard value S_1 while constraint (32) implies that the travel time between the response unit base and the Hazardous Materials nodes within its area of responsibility should not exceed the standard value S_2.

The emergency response unit location problem defined by (26)-(34) belongs to the category of p-median problems, which are hard to solve combinatorial problems. For this reason a heuristic algorithm has been developed for solving it. The proposed heuristic is presented in the next section.

Heuristic Algorithm for the Emergency Response Units Location Problem

The location problem defined by (26)-(34) consists of two basic interrelated subproblems: (i) the identification of the p bases of the emergency response units and (ii) the determination of the area of responsibility of each emergency response unit. The proposed heuristic algorithm is based on this intuition. Every iteration of the algorithm basically consists of two phases: (i) at the first phase a set of p bases is selected, (ii) at the second phase the districts are formed around each base aiming to minimize the total response time and secure the distance constraints (31) and (32). The basic variables involved in the proposed algorithm are: (i) x^* that denotes the best solution, (ii) "bestvalue" that denotes the objective function value of the optimal solution x^*, (iii) x that denotes the solution identified at the end of each iteration of the algorithm, and (iv) k is the counter of iterations. Furthermore, N denotes the set of nodes and I denotes a subset of N.

At the first step of the algorithm, k is set equal to 1, x^* is assumed empty and "bestvalue" is set equal to a very large positive number M. At the second step, the set I is set equal to N and for each node in I the quantity $V(i)$ named position function at i is calculated. The position function expresses the weighted sum of the travel times between the node i and its nearest nodes, constituting the set N_i, under the constraint that total workload of the set of the neighboring nodes does not exceed the average workload. If a neighboring Hazardous Materials node of i is placed at a travel distance that exceeds the standard value S_2 then a penalty quantity is added to the value $V(i)$. At the third step we set k as the first base of the solution and subtract the set N_k from set I. At the fourth step, $V(i)$ is calculated for each node I in the updated set I. In addition, the ratio $U(i) = \frac{V(i)}{d(i)}$ is calculated where $d(i)$ is the total distance of node i from the nodes that have already been assigned an emergency response unit base. The node i^* with the minimum value of $U(i)$ is selected as the next base. When the number of bases reaches (p) then the process moves to step five or otherwise the fourth step is repeated. At the fifth step the district of each base is formed determining in this way the solution of the iteration (x_i^k, Y_{ij}^k). At the sixth step, the solution is improved by exchanging as appropriate nodes between two districts. At the seventh step, the objective function value "fvalue" of the improved solution $(x_i^k, Y_{ij}^{'k})$ is calculated. If the fvalue is less than the currently optimal value bestvalue then the bestvalue is set equal to the fvalue and the $x^* = x$. Finally at the eighth step, k is increased by 1. If k is above n then the procedure stops, or otherwise the control returns to step 2. Thus the procedure defined by steps 2-7 is repeated (n-p) times starting at each iteration with a different first base.

A basic procedure of the algorithm is the construction of the area of jurisdiction around each base. The procedure consists of n-p steps. At each step r a node i is assigned to one of the bases i_k according to the minimization of metric h_r defined by (36):

$$h_r(i,i_k) = t(i,i_k)W(i_k,r) * penalty1(i) * penalty2(i) \qquad (36)$$

where W_{ri_k} is the temporary total workload of base i_k, $penalty1(i)$ is a penalty quantity that is applied only if the travel time of node i from each assigned base exceeds the standard value S_1, while $penalty2(i)$ is a penalty quantity that is active only if i is a Hazardous Materials node and its travel time from its assigned base exceeds S_2.

Assessment of the Computational Performance of the Algorithm

Due to the lack of benchmark problems in the state of the art and practice, the proposed algorithm was tested on a set of emergency response unit location problems defined on a network consisted of 200 nodes and two routes of Hazardous Materials. The position of the nodes were randomly selected within a square of the Euclidean plane. The travel times between the nodes were set equal to the respective Euclidean distances. Four types of calls have been considered while the rates of the calls per type at each node were assigned random values within the interval (0,1). The workload at each node was calculated by considering the following service times (per type): 10,15, 15, and 20 minutes.

The number of available response units (p) ranged from 10 to 2 and for each value of (p) two alternative values of S_2 were set: (i) the minimum value of S_2, (ii) a relaxed value of S_2 i.e. 20% plus the value of minimum S_2. The parameter S_1 was assigned a relative high value so that it does not affect the solution of the problem (i.e. 80 minutes). The optimum solutions of these problems were identified by mathematical programming application developed in AIMMS 3.0. Table 5 presents the results of this application. The first and second columns of the table present the number of available response units and standard value S_2 respectively. The third column presents the proportional deviation of the workload from the average workload. The fourth and fifth column present the objective function value of the solution produced by the algorithm and the optimum value respectively. The sixth column of the table presents the proportional deviation of the objective function value of the algorithmic solution from the optimum. The seventh and eighth columns present the computational time of the algorithm and the mathematical programming application respectively. Furthermore, the last row of the table provides the average value of the proportional deviation of the workload and the percentage deviation of the objective function values of the algorithmic from the optimum solutions.

The application of the algorithm to this set of location problems led to the following major conclusions:

- The maximum proportional deviation of the solutions produced by the algorithm from the optimal solutions ranges is 9.13% while the average deviation amounts to 7.02%.
- The average proportional deviation of the workload of the districts from the average workload was 6.7%.
- The computational time for determining the exact solutions through the AIMMS application was extremely high for some of the location problems (e.g. the problem presented at the sixth row of Table 5 that took more than a day to determine the exact solution), as compared to the computational time of the proposed heuristic that does not exceed 1.5 minutes.

p	**S_2**	**e (%)**	**Exact-f**	**Heur-f**	**%**	**H-time (sec)**	**E-time (sec)**
10	14	9,88	158524	169789	7,11	68	2055
10	17	8,41	144070	164242	14,00	68	6585
9	14	8,43	177914	184242	3,56	70	761
9	17	9,3	155583	180367	15,93	80	1292
8	17	9,7	177398	199681	12,56	80	7237
8	20	7,45	167993	180325	7,34	80	84000
7	17	9,95	205398	244685	19,13	80	9615
7	20	3,98	181000	212667	17,50	80	20000
6	20	8,44	196476	214791	9,32	80	8703
6	24	7,37	191427	204212	6,68	80	190
5	23	5,02	230186	238877	3,78	79	6233
5	28	5,17	211107	225187	6,67	81	4450

4	26	2,78	243953	244650	0,29	81	1619
4	31	2,78	237837	244650	2,86	76	2378
3	34	2,85	298275	336677	12,87	80	455
3	41	0,3	283082	287155	1,44	80	8015
2	48	0,74	346992	351926	1,40	75	809
2	58	0,45	345230	358695	3,90	77	1000
Average Workload		6,70	Average Deviation		7,02		

TABLE 5. The results of applying the location algorithm to networks of 200 nodes.

Concluding Remarks

Risk mitigation constitutes a major issue in Hazardous Materials Transportation planning problem. This paper is focused in two logistical decisions related to the distribution of the Hazardous Materials and the optimum deployment of the Hazardous Materials emergency response resources. Both types of decision tend to decrease the probability of a hazardous material accident and reduce the potential consequences of an accident. The analysis of the Hazardous Materials Distribution problem resulted in a bi-objective vehicle routing and scheduling problem. The solution of the problem was achieved through the application of the weighting method that transforms the bi-objective problem into a set of single objective vehicle routing and scheduling problems. A new route building heuristic algorithm has been presented for solving the aforementioned set of single objective problems. The new algorithm has been applied to the Solomon's benchmark problems providing encouraging results. Furthermore, the proposed algorithm has been applied on a set of Hazardous Materials Distribution problems providing non-inferior routes of minimum cost and risk.

The optimum deployment of the Hazardous Materials emergency response resources is expressed by a discrete location model. The heuristic algorithm that has been developed for solving the location model, belongs to the category of myopic algorithms. The proposed algorithm has been applied to numerous Hazardous Materials emergency response units location problems providing efficient solutions.

References

Bisschop J. & Entriken F. (1997). AIMMS: The Modeling System. Paragon Decision Techmnology.

Desrosiers J., Dumas Y., Solomon M., Soumis F. (1995). Time Constrained Routing and Scheduling. In Network Routing, Handbooks in Operations Research and Management Science.

Erkut E., Verter V. (1998). Modelling of Transport Risk for Hazardous Materials. *Operations Research, Vol. 45, No 5*, 625-642.

Gendreau M., Laporte G., Potvin J. (1997). Vehicle Routing: Modern Heuristcs. In Local Search in Combinatorial Oprimization, edited by E. Aarts and J. Lenstra, John Wiley &Sons Inc.

Golden B., Assad A. (1986). Vehicle Routing with Time Windows. *American Journal of Mathematical and Management Sciences*.

Kontoravdis G., Bard J. A GRASP for the Vehicle Routing Problem with Time Windows. ORSA Journal on Computing, vol. 7, no 1, 1995.

Kuncyte, R., Crainic, T.G., Laberg-Nadeau, C., Read, J.A. (1998). Le Systeme Internationale deTransport des Merchandises Dangerouses et son Harmonisation. In Association Quebecoise du Transport et des Routes, *Konies Transports 27 (3)*, 10-20.

List, G.F., P.B.Mirchandani, M. Turnquist , Zografos K. (1991). Modeling and Analysis for Hazardous Materials Transportation: Risk Analysis, Routing/Scheduling and Facility Location. *Transportation Science, Vol. 25, No.2*, 100-114.

Marianov V., ReVelle C. (1995). Siting Emergency Services. In Facility Location: A Survey of Applications and Methods, edited by Z. Drezner, Springer Series in Operations Research.

Potvin J., Rousseau J. (1995). An Exchange Heuristic for Routeing Problems with time windows. *Journal of the Computational Research Society, vol. 46*, 1433-1446.

Potvin J., Bengio S. (1996). The Vehicle Routing Problem With Time Windows, Part II: Genetic Search. *INFORMS Journal on Computing, vol. 8, No 2*, 165-172.

Potvin, J., Rouseau J. (1993). A Parallel Route Building Algorithm for the Vehicle Routing and Scheduling Problem with time Windows. *European Journal of Operational Research, vol. 66*, No. 3, 331-340.

Russell R. (1995). Hybrid Heuristics for the Vehicle Routing Problem with Time Windows. *Transportation Science, Vol. 29, No 2*, 156-166.

Saccomanno F. Allen, B. (1988). Locating emergency response capability for dangerous goods incidents on a road network. *Transportation Research Record*, 1193, 1-9.

Solomon M. (1987) Algorithms for the Vehicle Routing and Scheduling Problems with the Time Window Constraints. *Operations Research, Vol. 35,* 154-165.

Zografos K.G., Androutsopoulos K.N. (2004). A Heuristic Algorithm for Solving Hazardous Materials Distribution Problems. *European Journal of Operational Research 152 (2),* 507-519.

Zografos, K.G., Androutsopoulos, K.N., Vasilakis G.M. (2002). A Real Time Decision Support System for Roadway Network Incident Response Logistics. Transportation Research Part C, Vol. 10, 1-18

Zografos K.G., C. Douligeris, P. Tsoumpas. (1998) An Integrated Framework for managing Emergency-Response Logistics: The Case of the Electric Utility Companies. *IEEE Transactions on Engineering Management*, *Vol. 45, No. 2*, 115-126.

Zografos, K.G., C. Douligeris, Chaoxi, L. (1994). Simulation Model for Evaluating the Performance of Emergency Response Fleet. *Transportation Research Record, 1452,* National Research Council, Washington, D.C., 27-34.

Zografos, K.G., Nathanail, T., Michalopoulos, P. (1993). Analytical Framework for Minimizing Freeway Incident Response, *Journal of Transportation Engineering, Vol. 119* (4), 535-549.

Zografos, K.G., and Davis, C.F. (1989). A Multiobjective Programming Approach for Routing Hazardous Materials, *ASCE Transportation Engineering Journal*, Vol. 115 (6), 661-673.

Re-Examination of Human-Computer Interaction Methodology in the Domain of Interactive TV Applications

Konstantinos Chorianopoulos
Marie Curie Fellow, Bauhaus University of Weimar, Department of Architecture
k.chorianopoulos@archit.uni-weimar.de

Abstract

The user-centric mentality is usually materialized into the usability engineering techniques. In Human-Computer Interaction (HCI), the generic user-centric approach includes the following activities: 1) user studies and evaluation, 2) user analysis and modelling, 3) prototyping, and 4) implementation. These activities are highly iterative with the exception of user studies and implementation, which are performed less frequently during the product lifecycle. There is an extensive body of previous research in user-centric methods and techniques for personal computers and internet applications, but there is limited research about interactive TV applications. This work investigates the question: 'How does the generic notion of 'user-centric' translate into specific actions for interactive TV (ITV) user interface (UI) design, development and evaluation?'

Introduction

The user-centric mentality is usually materialized into the usability engineering techniques. Accordingly, previous evaluations of popular ITV applications, such as the Electronic Program Guide (EPG) and digital video navigation, have employed task efficiency concepts. Since ITV applications gratify entertainment needs and leisure activities in a domestic environment, there is a need to re-examine the suitability of the traditional usability engineering conceptualizations.

It is argued that the UI in ITV applications should move beyond the productivity paradigm implied by the contemporary notion of usability. Media studies literature was investigated, in order to augment the user-centric approach with a toolset of concepts, methods and techniques that are suitable for interactive TV. Indeed, the media studies field is very relevant to the analysis of TV and it has been highlighted as a significant area for further research in HCI (Macdonald 2004).

The rest of the paper is structured as follows: Section 2 presents the motivation for re-examining the traditional usability engineering approach. Section 3 explores the dimensions of the quality of an ITV UI. In Sections 4 to 8, the usability engineering techniques for UI design, development and evaluation are re-examined, under the light of the new definitions for UI quality in ITV applications. Finally, Section 9 provides a brief future research agenda.

Previous research

Previous approaches to popular ITV applications (e.g. EPG) have employed concepts, such as task efficiency and effectiveness. Indeed, the usability of the EPG is very similar to the usability of productivity software, because it involves more information processing than enjoyment of ITV content. Several aspects of EPG navigation can be modelled after the traditional HCI tasks and goals. Nevertheless, there are some aspects of the EPG design, and many other types of ITV applications that would benefit by a consideration of the affective dimension of the UI.

Most notable among the recent findings for ITV applications is the realization that users' subjective satisfaction is at odds with the established metrics of efficiency. A usability test of three video skipping UIs revealed that user satisfaction was higher for the UI that required more time, more clicks and had the highest error rate. In other words, the most efficient UI was not the most favoured one. This result is opposite to the assumptions of the efficient usability paradigm, which conceives the efficient as more usable and thus preferable.

The satisfaction questionnaires exposed that the users regarded their preferred UI as more fun and relaxing compared to the most efficient one (Drucker et al. 2002). Accordingly, Chorianopoulos and Spinellis (2004b) let their subjects use a video skipping application without specifying any task, besides the suggestion to 'watch TV for a period of time.' Moreover, they employed the hedonic quality construct (Hassenzahl et al. 2001) and validated that users preferred a video skipping UI, although it was coupled with a dynamic advertisement insertion feature, which increased the total number of advertisements they were exposed to.

Table 1 Methodological issues in ITV UIs compared to traditional domains (PC, Web)

HCI concepts/Domain	**PC, Web**	**Interactive TV**
User, context	PC user, work	Viewer, domestic (leisure time)
Goal	Effectiveness, efficiency, productivity	Entertainment, relaxation
Activities	Task execution	Free exploration
Techniques	Task completion, errors, efficiency	Affective state, emotions

In summary, the majority of previous studies have considered only the efficient aspect of the ITV UI. Since ITV applications gratify entertainment aspirations, in a leisure context of use, for a wide diversity of users, there is a need to extend the HCI toolset, so that it also considers the affective quality of an ITV UI (Table 1).

User interface quality

It has been argued that people spend most of their leisure time trying to moderate their moods. Daniel Goleman (1995, p.57) said that 'managing our emotions is something of a full-time job: much of what we do —especially in our free time— is an attempt to manage mood. Everything from reading a novel or watching television to the activities and the companions we choose can be a way to make ourselves feel better. The art of soothing ourselves is a fundamental life skill.' Television entertainment is a multidimensional construct that cannot be measured as such, but consists of several parameters that could be measured (Vorderer 2001). For example, Reeves and Nass (1996) assert that a mediated experience elicits an emotional response, which is partly valence (pleasure) and partly arousal. There are also additional elaborate models of the uses and gratifications when watching TV (Lee and Lee 1995, Rubin 1983). Therefore, the UI of an ITV application could be conceptualized as an additional —to channel changing and program selection— means to moderate the mood of the TV viewer.

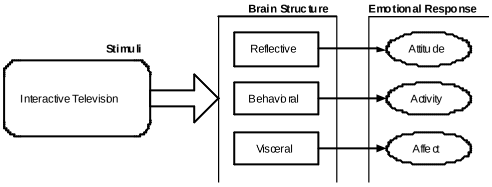

Figure 1 The ITV entertainment experience elicits three types of emotional responses (attitude, activity, affect), which correspond to the three-level model of affect —adapted from Norman et al. (2004)

According to Norman et al. (2004) there are three distinct levels of brain mechanism: 1) the visceral level, which is the pre-wired part of the brain and acts automatically to external stimuli, 2) the behavioural level, which contains the brain processes that control everyday behaviour and 3) the reflective level, which is the contemplative part of the brain. Each level could be associated to a different class of constructs, which could then be employed to evaluate the differences between the emotional responses to alternative UI designs (Figure 1). For example, an ITV application may elicit enjoyment (e.g. pleasure, or arousal) at the visceral level. Then, the user may continue using the ITV application for a long time and become emotionally absorbed (e.g. involvement and engagement). Finally, the user may decide that she likes the specific ITV application, which leads to the formation of an attitude (e.g. program liking).

The emphasis on an affective conceptualization for ITV applications does not entail a complete abandonment of the efficient usability paradigm. For example, an ITV news application, used in the morning before leaving the home for work, should be very efficient in terms of fast information retrieval and navigation. The same application, used in the evening after returning home from a long day at work, should be more automated and encourage relaxed use (Steve Draper, personal communication). In general, the UI evaluation should be regarded to have both an affective and an efficiency dimension. In the ITV case, the leisure context of use and the need for gratifying entertainment goals might push the balance towards the affective dimension of the UI. UI designers should explicitly set the goals of each UI depending on the nature of the ITV application and then they should employ the appropriate assortment of efficient usability and affective quality methods for the evaluation.

Besides media studies, there might be alternative paradigms for conceptualizing the affective dimension of an ITV UI. ITV applications are supposed to offer more than just an improved version of the traditional TV experience. One potential benefit of ITV applications would be the creation of optimal experiences through flow (Csikszentmihalyi 1991), which requires the establishment of a match between the viewer skills and the challenge posed by the ITV application. There are also a few additional paradigms that should be investigated in the context of ITV applications. For example, the HCI research is gradually diversifying its focus, in areas such as: 1) influencing the user through persuasion (Fogg 2002), which offers concepts related to the trust in ITV applications that have an advertising and commerce characteristics and 2) video-games and fun (Draper 1999), which offers concepts related to the game-play dimension of ITV usability.

In brief, depending on the application domain (e.g. entertainment, learning, e-commerce, game-play, information) the design and evaluation of ITV applications should employ the most suitable UI conceptualization, instead of employing implicitly on the traditional usability engineering mentality.

Conceptual models

Previous research about conceptual models for ITV is only limited to the EPG and to the transfer of experience from the Web and the PC. Instead, the development of conceptual models for ITV should be based on the identification of: 1) the familiar patterns) and 2) the emerging usage patterns. Moreover, designers should distinguish between two levels of conceptual models: 1) operating system and 2) application. In this way, many types of ITV applications could be modelled and approached in a consistent way. For example, ITV news and music could be organized in virtual TV channels.

The Virtual Channel is a conceptual model for UI design; it augments the familiar access method to broadcast programming (i.e. the notion of a channel), to an integrated model for accessing interactive audiovisual content from diverse sources, which supports the emerging usage patterns (e.g. digital video recording). First, the organization of digital media content into a small number of spatio-temporally personalized virtual channels simplifies choosing from a vast array of available broadcasts, stored programs and Internet resources. Second, presenting media programming from Virtual Channels gives more control to the user, who can actively shape the content flow. The virtual channel model suggests only a minimal shift from current media use, while it focuses further research on the design of a content-specific UI (e.g., music TV).

A virtual TV channel is a TV channel that is not a fixed video signal shared by all viewers in the same way, but a dynamic synthesis of discrete video, graphics, and data controlled by a computer program, which runs on each digital set-top box (Chorianopoulos 2003). The traditional television experience consists of video and overlaid graphics-text created at the media source (the TV broadcast station or the TV production studio); thus it is fixed for all TV viewers. The Virtual Channel model shifts the decision-making about TV programming from the media source to the set-top box (STB). The television experience is now created and controlled at the STB from a combination of locally stored material, real-time broadcast transmissions, and Internet resources (Figure 2).

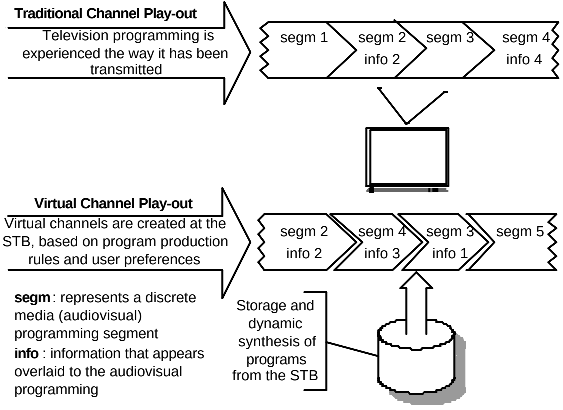

Figure 2 Generic model of a system with the virtual channel metaphor, in contrast to the traditional broadcasting scheme

The main implication for practice is that neither the vision of 500 channels, nor that of a single personalized channel is suitable for giving consumers access to the digital STB. Instead, it is proposed that a small number of dynamic virtual channels could offer enough choices for serendipity in media experiences, while simplifying access to vast and diversified sources of interactive multimedia content. The Virtual Channel model has been applied to personalized advertising (Chorianopoulos and Spinellis 2002) and to interactive music TV (Chorianopoulos and Spinellis 2004b).

Principles and guidelines

User interface principles enable the translation from the semantic level of the conceptual model to the syntactic components of the system's functionality (Wickens et al. 1998, p.462). There are generic and application specific principles that address the multitude of issues associated with UI design, such as screen design, user needs and performance, input device, etc (Nielsen 1994, p.91). For example, a generic principle is that 'a UI should provide visibility of the current status.' Most principles seem to be common sense, but it is argued that it is impossible for the designer to have all of them in mind, so principles are offered as checklists (Wickens et al. 1998, p.465). Nevertheless, HCI has grown up alongside the desktop PC and most of the current UI principles reflect a productivity context of use, instead of considering the domestic entertainment activities, such as ITV.

ITV usability fits within the broad continuum between the traditional TV and the desktop PC-Web user experience. Therefore, the ITV experience may be considered as the gamut of different combinations between the traditional TV and the interactive PC-Web UI principles. The UI principles were formulated after a critical review of previous research in: 1) mass communication, 2) advertising and 3) usability for the PC and the Web. The principles are presented as trade-offs, in order to allow ample space for the selection of the most appropriate UI design (Table 2). This list of UI principles for ITV has been applied to personalized advertising (Chorianopoulos and Spinellis 2002) and to interactive music TV (Chorianopoulos 2004).

Table 2 User interface principles for interactive TV applications

Design Factor	**Description**
Low Vs High Attention	A viewer may sit down and watch a TV program attentively, or leave the TV open as a radio
Group Vs Individual Watching	Distinguish between group and solitary use
Planned Vs Impulse Program Selection	Consider multiple levels of planning: from channel surfing to appointment viewing
Real Time Vs Time Shift	Both stored and broadcast content and applications should be available and complement each other
Entertainment Vs Information	Support a wide variety of user goals, from entertainment to information
Interactive Vs Passive	Support the predominately passive patterns of use, but offer the option for interactivity
Computer generated Vs fixed content	Replace static video elements (channel logo, ticker, info, sport statistics, etc) with computer generated content
Computer Vs Television Visual Design	Enhance, but do not replace, the main TV elements (characters, stories) with PC elements (objects, actions)

Before applying the principles to practical design problems, it should be made explicit that the design of a UI for an ITV application is very different from that for a PC application. The major difference is that the UI for ITV is content specific, while the UI for the PC is application specific. For example, in a word processor most of the UI principles are common with those of other productivity applications, while the UI for a music ITV application is very different from the UI for a news ITV application. From that perspective, UI design for ITV is very similar to UI design for the Web, in which each Web site has different content, structure and navigation.

The ITV UI principles were presented as a list of high-level and generic design factors, although UI principles are usually more specific and concern particular parts of the interaction, such as dialog box, menu, icon etc. Moreover, there are UI guidelines, which are quantitative reformulations of principles. For example, the generic principle 'respond fast to user commands' may be transformed to 'respond in less than 1sec to user commands' as a guideline for a specific system. Then, the high-level ITV UI trade-off for 'multiple levels of attention' may be transformed to a more specific UI principle, such as 'remove a dialog box, if the user does not interact with the TV system' or transformed to a guideline such as 'remove a dialog box, if the user does not interact with the TV system after 3-5 seconds.' Therefore, further research should refine the proposed set of UI design trade-offs into longer lists of more specific principles and guidelines for particular types of ITV applications.

Prototyping techniques

Previous studies with ITV applications have revealed that whatever is presented on the TV screen is assessed in comparison with the current television experience. Computer-like menus, toolbars, scrolling pages and form-navigation look irrelevant on a TV screen, even when used by experienced computer users. Before a high-fidelity prototype is presented to users, there is a clear need for television-values content development.

The main requirement for an ITV prototyping platform is that it imitates the look and feel of everyday TV by employing (Chorianopoulos and Spinellis 2004a): 1) remote control and TV screen connected to a portable PC, 2) audiovisual content (e.g. video clips) and 3) TV-language rules and aesthetics (e.g. transparent overlays, smooth transitions).

Figure 3 Experimental set-up for unobtrusive and seamless usability evaluation of ITV applications

The central element for the above experimental set-up was a portable PC. The ITV application was designed to run in full-screen and windowless mode and was set to give audiovisual output to a TV set. The PC's serial port was connected to an infrared sensor (http://www.evation.com/irman/) that receives the signals from the remote control. The whole set-up was unobtrusive and seamless to the television viewer (Figure 3). When the portable PC was hidden away, the only item of the set-up that was left hanging around was the small black box of the infrared sensor.

User interface toolkits

The development of an ITV application is usually reflective of an IT developer viewpoint, instead of that of a TV producer. The creative part of the development process has a subordinate-decorative role, because the requirements in technical knowledge for using the popular authoring tools demand a strong IT background. Despite the progress, most of the contemporary ITV authoring tools are still closer to the IT developer than the TV producer.

Since compelling ITV applications are most likely to be developed by TV producers, it makes sense to design ITV authoring tools that make ITV application development more accessible to TV producers. Chorianopoulos and Spinellis (2004a) have designed a high-level ITV Application Programming Interface (API), which facilitates the workflow of the TV producer and is based on the properties of the Virtual Channel model, described earlier. In the following paragraphs, the Virtual Channel Control Library (VCCLib) is briefly presented.

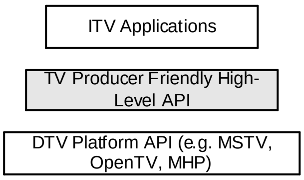

Figure 4 Using a high-level API to make ITV development accessible to TV producers

The VCCLib is a higher-level ITV API that takes interactive computer graphics further away from the specifics of the underlying implementation and closer to the traditional TV production values (Figure 4). A partial class diagram of the VCCLib is provided (Figure 5), so that it becomes feasible to employ it in ITV application development and for realizing in other contexts, with alternative implementation tools and platforms. The Virtual Channel Control is the central element of the VCCLib and provides methods and events for handling the flow of a virtual channel.

Event-driven computer programming might feel familiar for the majority of developers who have used object-oriented languages to build interactive applications. Nevertheless, event-driven programming of ITV applications might be different from that of productivity applications, in the sense of being more time-driven than user-action-driven. ITV applications are in need to organize the UI and the user experience temporally, instead of spatially, which has been the norm for computer application development so far. Therefore, an API for ITV applications should support the programming of time-driven UIs. The Timers Control is based on the Timer Control and enables the definition and handling of pre-scheduled (author) events. The VCAgent Control is a simple wrapper-class around the MS Agent Control.

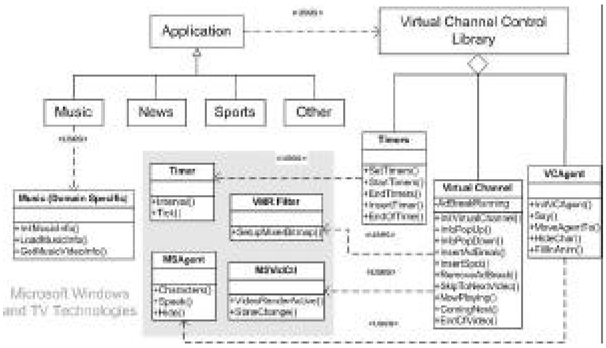

Figure 5 Class diagram for the Virtual Channel API with references to a MSTV implementation

An example interactive music TV application has been developed with the aid of the VCCLib

(Chorianopoulos and Spinellis 2004a). The application allows a television viewer to skip a music video clip —an action that comes at the cost of watching a targeted advertisement. In addition to clip skipping, video overlays were used to superimpose information over the music video. Music information provides trivia about the artist, biographical information and discography.

User studies and evaluation

The selection of a UI evaluation method depends on the type of the research problem. For example, an ethnographic study may provide in-depth insights about the uses of TV in everyday domestic life (O'Brien et al. 1999). Then, a survey may reveal relationships between the uses and the type of the family or the viewers' profile, and to give quantitative results (Freeman and Lessiter 2003).

Previous findings demonstrate that the consumers' perceptions are very elastic and prone to change with the passage of time (Petersen et al. 2002). Therefore, a longitudinal study may be used to study the evolution of important variables for longer periods of time (Kubey and Csikszentmihalyi 1990). Furthermore, focus groups and interviews have been used in ITV research. The latter methods are useful for requirements collection and for investigating the long-term effects of ITV applications, while usability tests are more appropriate and cost-effective during the development process.

The majority of UI evaluation studies have been conducted in the laboratory with experimental methods. Mass communication research employs large (compared to HCI experiments) samples of people, in order to study the effect of features that exist in TV content. On the other hand, HCI research focuses on informing product development and employs small numbers of subjects iteratively with discount usability engineering techniques (Nielsen 1994).

Maguire (2002) raised the research question: 'Should tasks be fixed or should users be allowed to use the service as freely as they wish?' It has been argued that the users should be allowed to use the service for a predefined but flexible duration of time (e.g. 15-30 minutes), without any particular task to complete (Chorianopoulos and Spinellis 2004). In this way, the traditional usability test reflects the tradition of the selective exposure paradigm (Zillmann and Bryant 1985), which has been also used to study the media effects of interactive products in contemporary mass communication research (Knobloch and Zillmann 2002). Since viewers select TV channels and watch TV programs to their discretion, without performing any particular task, then the evaluation of an ITV UI should facilitate free exploration and enjoyment of the ITV application.

The evaluation methods employ one or more data collection techniques, which may be qualitative, or quantitative. Many UI evaluation studies are qualitative (e.g. observation, thinking aloud, interview, focus group), but some of them (e.g. thinking aloud) may not be suitable for ITV (Maguire 2002). The quantitative methods provide explicit results for formulated hypotheses and concrete UI issues, while the qualitative methods are used to reveal UI issues that have not been identified by the designers. Ideally, the qualitative measurement techniques should be used to complement the quantitative ones (Eronen 2001). Chorianopoulos and Spinellis (2006) have identified and suggested constructs that are relevant to the affective quality dimension of an ITV UI and presented the respective measuring instruments. These instruments are easy to administer and are compatible with popular UI evaluation methods, such as the usability test.

Future research agenda

Further research should consider the emerging paradigms in the HCI discipline, such as persuasion, and play. Further application areas should consider the design of UIs for novel ITV delivery channels, such as mobile and peer-to-peer networks. Moreover, ITV should be considered in the context of an ecosystem of consumer devices, such as mobile phones, PCs, video-game consoles, home media servers, portable media players and information appliances. Peer-to-peer (P2P) technologies such as Bittorrent are currently employed by early adopters in order to gain prompt access to TV content. In the legitimate side of P2P delivery, BBC is employing a PC-based P2P client in order to give access to its vast archive of audiovisual content. Moreover, all major search engines are augmenting their systems to support video searches of broadcast and user contributed video. Then, the research question is: What is an ITV UI for P2P delivery of multimedia content? Similar research questions apply also to Internet delivery, cross media delivery and user-contributed ITV content.

In conclusion, the employment of the HCI theory and methods in the ITV field is a worthwhile area for research and practice. The implications of further research will have a significant impact on the HCI theory, media industry and consumers.

References

Chorianopoulos, K., 2004: *Virtual Television Channels: Conceptual model, user interface design and affective usability evaluation*. PhD Thesis. Athens University of Economics and Business.

Chorianopoulos, K. 2003: The digital set-top box as a virtual channel provider. In *Adjunct Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 666-667, ACM press,.

Chorianopoulos, K. and Spinellis, D. , 2003: A metaphor for personalized television programming. In Carbonelle, N. and Stephanidis, C., editors, in *Proceedings of the 7th ERCIM International Workshop on User Interfaces for All (2002), LNCS 2615*, pages 187–194. Springer-Verlag.

Chorianopoulos, K. and Spinellis, D., 2004a: User interface development for interactive television: Extending a commercial DTV platform to the virtual channel API. *Computers and Graphics*, 28(2):157–166, Elsevier.

Chorianopoulos, K. and Spinellis, D., 2004b: Affective usability evaluation for an interactive music television channel. *Computers in Entertainment*, 2(3):14, ACM Press.

Chorianopoulos, K. and Spinellis, D., 2006: User Interface Evaluation of Interactive TV: The Contribution of Media Studies. *Universal Access in the Information Society*, To Appear, Springer.

Csikszentmihalyi, M., 1991: *Flow: The Psychology of Optimal Experience*. Perennial.

Desmet, P.M., 2003: Measuring emotions: Development and application of an instrument to measure emotional responses to products. In Blythe, M., Monk, A., Overbeeke, K., and Wright, P., editors, *Funology: from usability to enjoyment*. Kluwer Academic Publishers.

Draper, S.W., 1999: Analysing fun as a candidate software requirement. *Personal and Ubiquitous Computing*, 3(3), Springer.

Drucker, S. M., Glatzer, A., Mar, S. D., and Wong, C., 2002: Smartskip: consumer level browsing and skipping of digital video content. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 219–226. ACM Press.

Eronen, L., 2001: Combining quantitative and qualitative data in user research on digital television. In *Proceedings of PC HCI 2001*. Typorama publications.

Fogg, B.,, 2002: *Persuasive technologies: Using computer power to change attitudes and behaviors. San Francisco*, CA: Morgan Kaufmann.

Freeman, J. and Lessiter, J., 2003: Using attitude based segmentation to better understand viewers' usability issues with digital and interactive TV. In *Proceedings of the 1st European Conference on Interactive Television:from Viewers to Actors?*, pages 19–27.

Goleman, D., 1995: *Emotional Intelligence*. Bantam Books.

Hassenzahl, M., Beu, A., and Burmester, M., 2001: Engineering joy. *IEEE Software*, 18(1):70–76, January/February, IEEE.

Knobloch, S. and Zillmann, D., 2002: Mood management via the digital jukebox. *Journal of Communication*, 52(2):351–366.

Kubey, R. and Csikszentmihalyi, M., 1990: *Television and the Quality of Life: How Viewing Shapes Everyday Experiences*. Lawrence Erlbaum.

Lee, B. and Lee, R.S., 1995: How and why people watch TV: Implications for the future of interactive television. *Journal of Advertising Research*, 35(6):9–18.

Livaditi, J., Vassilopoulou, K., Lougos, C., and Chorianopoulos, K., 2003: Needs and gratifications for interactive TV applications: Implications for designers. In *Proceedings of the HICSS 2003 conferece*. IEEE.

Macdonald, N., 2004: Can HCI shape the future of mass communications? *Interactions*, 11(2):44–47, ACM Press.

Maguire, M., 2002: Applying evaluation methods to future digital TV services. In Green, W. and Jordan, P., editors, in *Pleasure with products beyond usability*, pages 353–366. Taylor and Francis.

Nielsen, J., 1994: *Usability Engineering*. Morgan Kaufmann, San Francisco.

Norman, D.A., 2004: *Emotional Design: why we love (or hate) everyday things*. Basic Books.

O'Brien, J., Rodden, T., Rouncefield, M., and Hughes, J., 1999: At home with the technology: an eth-nographic study of a set-top-box trial. *ACM*

Transactions on Computer-Human Interaction (TOCHI), 6(3):282–308, ACM Press.

Petersen, M. G., Madsen, K. H., and Kjaer, A., 2002: The usability of everyday technology: emerging and fading opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(2):74–105.

Reeves, B. and Naas, C., 1996: *The media equation: How people treat computers, television and new medialike real people and places*. CLSI.

Vorderer, P., 2000: Interactive entertainment and beyond. In Zillmann, D. and Vorderer, P., editors, *Media entertainment: The psychology of its appeal*, pages 21–36. Lawrence Erlbaum Associates.