Handling Uncertainties and Unreliability in Data

Pitu B. Mirchandani School of Computing, Informatics and Decision Systems Engineering Arizona State University

M. Gentili Department of Mathematics and Computer Science University of Salerno, Italy

May 12, 2010 – Tempe

Outline



1. Sensor Technologies and Sensors Measurements

2. Location of Sensors to Maximize Variance Travel Time Reduction

- The basic AVI Reader location models
- Considering Link Failure
 - Considering Sensor failure
- 3. Conclusions and Further Research

-

Need for Sensors



We need sensors to detect and monitor traffic for:

Providing travel time and congestion information to motorists

Providing data for traffic management so that traffic congestion is reduced

Assisting in managing incidents and evacuations

Providing data for trucking, transit, navigation, route guidance, transportation planning, ...

Sensor Technologies and Sensors











 $\left(6 \right)$

13



OD pair	Path ID	Trips	Links	•				
1-3	1	100	(1,2)	(2,3)				
1-9	2	10	(1,2)	(2,3)	(3,4)	(4,7)	(7,8)	(8,9)
7-11	3	10	(7,8)	(8,9)	(9,10)	(10, 11)		
5-13	4	10	(5,4)	(4,7)	(7,8)	(8,9)	(9,12)	(12, 13)

Criteria for the location:

- Number of readings
- Number of AVI readers
 - Length of the readings
- Number of Covered OD pairs
- Flow captured

Locate AVI Readers to Monitor Travel Times: Proposed Mathematical Models



travel time predictions

Locate AVI Readers to Monitor Travel Times: Model AVI3 – a simple case



Three routes in the network:

R1 with a volume of 20 vehicles that pass locations A, B, and C

Model AVI4 + Sensor

Failure

Conclusions

R2 with a volume of 50 vehicles that pass locations A, B, and D

R3 with a volume of 10 vehicles that pass locations A and D <u>If two sensors are to be located</u>, by locating them at <u>A and D</u>, we will measure (50 + 10) vehicles x 2 miles = 120 vehicle miles, which is the optimal solution.

Challenge!



 x_8, x_{13} measure blue route (route 2) and x_3, x_9 as well as (x_8, x_{13}) measure red route (route 1)

Challenge!



Now use upstream/downstream decision variables y_{ij} , z_{ij}

Locate AVI Reader to Monitor Travel Times: Model AVI3 – formulation (1/4)



Indices:

- *i* = 1, 2, ..., *l* represents the **set of routes**
- *j* = 1, 2, ..., J represents the **set of potential sites** for locating sensors in the network

Parameters:

- *d_{ii}* is Site *j*'s distance from the origin of Route *i*
- v_i is the traffic volume on Route i
- B is the number of new sensors that can be installed (budget limits)
- *w* is the number of existing sensors in the network

Locate AVI Reader to Monitor Travel Times: Model AVI3 – formulation (2/4)



Locate AVI Reader to Monitor Travel Times: Model AVI3 – formulation (3/4)



 $\sum_{j\in R_i} y_{ij} - \sum_{j\in R_i} z_{ij} = 0 \quad \forall i$

Locate AVI Reader to Monitor Travel Times: Model AVI3 – formulation (4/4)



Locate AVI Reader to Monitor Travel Times: Proposed Mathematical Models



locates B readers to maximize both the total number of readings and the total number of covered OD pairs Model AVI2 : associates different benefit to the OD pairs and maximizes the total benefit accrued by locating at most B readers Model AVI3: locates B readers to maximize total vehicle miles that are measured

Model AVI4:

locates B readers to maximize the reliability of the obtained travel time predictions

Locate AVI Reader to Monitor Travel Times: Proposed Mathematical Models



Model AVI4:

locates B readers to maximize the reliability of the obtained travel time predictions

The motivation for this objective function is for traffic managers to inform to commuters as reliable as possible estimates for travel times for route segments.

Locate AVI Reader to Monitor Travel Times: Model AVI4 – concept & findings (1/3)



We want to improve our ability to predict network travel times by sampling vehicle travel times with a prescribed number of sensors installed in the network

Due to the dynamic and stochastic nature of the transportation network, the prediction of travel times is represented by confidence intervals for different time periods

We need to determine on which of the routes to locate the sensors to minimize the variance of the predictive distribution of network travel times in order to improve confidence intervals

Locate AVI Reader to Monitor Travel Times: Model AVI4 – concept & findings (2/3)



Travel time on an arc is a random variable, that includes variations due to travelers, and variations due to traffic conditions.

At any time, component dues to traffic conditions, has mean μ_i , which remains constant over short periods of time; μ_i has a mean η_i and variance γ_i (assume Normal and known from data)

With the observed travel times, by appealing to the Bayesian statistical analysis (prior \rightarrow posterior), the mean travel time distribution on the segment is updated

Locate AVI Reader to Monitor Travel Times: Model AVI4 – concept & findings (3/3)



For any pair of routes R_1 and R_2 with volumes V_1 and V_2 and prior knowledge of the means and variances of μ_i on both routes, the following sampling rule can be proven:

Sample the route R with higher value of f(R) = volume + β(R) × prior travel-time variance where β(R) route-specific parameter, to get a higher decrease in the variance of the predicted mean network wide travel time.

The general network sensor location problem that minimizes the posterior variance of network travel time distribution can be formulated as an integer program based on the above rule (due to Bayesian statistics), with constraints similar to AVI₃



Subject to: other constraints similar as Model AVI3.

For those interested, objective of Model AVI4





Study Area Network

Install **B** = 5 Sensors



B = 5	reduced	of total	by
	7,578.79 (3,361.08)	60.25% (26.72%)	33.53%

Inoroace

Install **B** = 10 Sensors



Install **B** = 15 Sensors

Link failure:

Some links in the network have a known probability of disruption due to external events.

The model seeks for the optimal location of B sensors to minimize the expected posterior variance over all possible link failure scenarios.

Sensor failure:

We assume each sensor has a failure probability of q.

The resulting model minimizes the expected posterior variance in the prediction of travel time when B sensors are to be located.

Sensor Technologies

Conclusions

Extensions of models (2/2)

Model AVI3 + Sensor Failure : Addressed by Li and Ouyang (2010)

We consider: Model AVI3 + Link Failure Model AVI4 + Link Failure Model AVI4 + Sensor Failure

Model AVI4 + Link failure

Route $R_1 = \{a_1, a_3, a_5\}$ Route $R_2 = \{a_1, a_4, a_5\}$ Route $R_3 = \{a_2, a_3, a_5\}$ Route $R_4 = \{a_2, a_4, a_5\}$ Route $R_5 = \{a_1, a_3, a_6\}$ Route $R_6 = \{a_1, a_4, a_6\}$ Route $R_7 = \{a_2, a_3, a_6\}$ Route $R_8 = \{a_2, a_4, a_6\}$

Model AVI4 + Link failure

Route $R_{4} = \{a_{4}, a_{3}, a_{5}\}$ Route $R_{2} = \{a_{1}, a_{4}, a_{5}\}$ Route $R_{3} = \{a_{2}, a_{3}, a_{5}\}$ Route $R_{4} = \{a_{2}, a_{4}, a_{5}\}$ Route $R_5 = \{a_4, a_3, a_6\}$ Route $R_6 = \{a_1, a_4, a_6\}$ Route $R_7 = \{a_2, a_3, a_6\}$ Route $R_8 = \{a_2, a_4, a_6\}$

Model AVI4 + Link failure

Model AVI4 + Link failure: Mathematical Formulation

Indices:

- *i* = 1, 2, ..., *l* represents the **set of routes**
- *j* or *k* = 1, 2, ..., J represents the **set of potential sites** for locating sensors in the network

• s= 1,2,... S represents the **set of possible scenarios** for link failure

Parameters:

• $\Delta \gamma_{is}^{jk}$ decrement in variance of mean route travel times when Route *i* is observed from Site *j* to Site *k* during the scenario *s*

- p_s is the probability for scenario s
- B is the number sensors that can be installed (budget limits)

 $x_{j} = \begin{cases} 1 & \text{if a sensor is located at Site } j \\ 0 & \text{otherwise} \end{cases}$

 $z_{iks} = \begin{cases} 1 & \text{if Site } k \text{ is the most downstream site in Route } i \text{ with a sensor installed} \\ & \text{on scenarios} \\ 0 & \text{otherwise} \end{cases}$

 $y_{ijs} = \begin{cases} 1 & \text{if Site } j \text{ is the most upstream site in Route } i \text{ with a sensor installed} \\ & \text{on scenario } s \\ 0 & \text{otherwise} \end{cases}$

Model AVI4 + Link failure: Mathematical Formulation

 $\max \frac{1}{2} \sum_{i} \sum_{k} \sum_{i} \sum_{k} \Delta \gamma_{is}^{jk} (y_{ijs} + z_{iks}) p_s$ s=1 i=1 $j\in R_i$ $k\in R_i$

Model AVI3

 $\sum x_j \le B$

 $y_{ijs} \leq x_j$

 $z_{iks} \leq x_i$

Model AVI4

Model AVI4 + Link Failure Model AVI4 + Sensor Failure

Sensor Technologies and Sensors

Readers

Location of AVI

Measurements

Conclusions

 $\forall j \quad \forall i \quad \forall s$

 $\forall k \quad \forall i \quad \forall s$

Locate sensors on different routes in the network so that the expected variance travel time reduction on the routes is maximized

the expected total reduction in the variance of mean route travel times

Let us consider the random variable S=number of failures

 $Pr(S=0) = q^{0} (1-q)^{4}$ $Pr(S=1) = q^{1} (1-q)^{3}$ $Pr(S=2) = q^{2} (1-q)^{2}$ $Pr(S=3) = q^{3} (1-q)^{1}$ $Pr(S=4) = q^{4} (1-q)^{0}$

Sensor Technologies and Sensors

Readers

Location of AVI

Model AVI3

Model AVI4

Conclusions

Model AVI4 + Link Failure

Model AVI4 + Sensor Failure

Measurements

Sensor Technologies and Sensors

Readers

Location of AVI

Model AVI3

Model AVI4

Conclusions

Model AVI4 + Link

Failure

Model AVI4 + Sensor Failure

Measurements

We could compute the expected value also in a different way by observing that it depends only on the active upstream and downstream sensors.

Sensor on a_1 is the 1st active upstream sensor (we say it is active at *level 0*) Sensor on a_3 is the 2nd active upstream sensor (we say it is active at *level 1*) Sensor on a_4 is the 3rd active upstream sensor (we say it is active at *level 2*) Sensor on a_5 is the 4th active upstream sensor (we say it is active at *level 3*)

 \rightarrow (1-q)² q² q⁰

Model AVI4 + Sensor failure: Mathematical Formulation

Indices:

• *i* = 1, 2, ..., *l* represents the **set of routes**

• *j* or *k* = 1, 2, ..., *J* represents the **set of potential sites** for locating sensors in the network

• *h* or *t* =0,1,2, ... |R_i| represents the **set of possible levels** for sensors (upstream/downstream) on route R_i

Parameters:

• $\Delta \gamma_i^{jk}$ reduction in variance of mean route travel times when route R_i is observed from Site *j* to Site *k*

- q is the failure probability for each located sensor
- B is the number of sensors that can be installed (budget limits)

 $\begin{cases} 1 & \text{if Site } j \text{ is the most upstream site of level } h \text{ in Route } i \text{ with a sensor installed and} \\ \text{Site } k \text{ is the most downstream site of level } t \text{ in Route } i \text{ with a sensor installed} \end{cases}$

0 otherwise

 $\delta_{iht}^{jk} =$

Model AVI4 + Sensor failure: Mathematical Formulation

 $\sum_{i=1}^{|R_i|} y_{ijh} = x_j \qquad \forall j \quad \forall i$

 $\sum z_{ikt} = x_i \qquad \forall k \quad \forall i$

 $\sum x_j \le B$

 $|R_i|$

t=0

Locate sensors on different routes in the network so that the Expected variance travel time reduction on the routes is maximized

Sensor Technologies and Sensors

Measurements

Conclusions

Model AVI4 + Sensor failure: Mathematical Formulation

Conclusions

$$\forall i \quad h = 0$$

$$\sum_{i} y_{ijh} \leq \sum_{j \in R_i} y_{ij(h-1)} \quad \forall i \quad h = 1, ..., | R$$

 $\sum_{j \in R_i} Z_{ijh} \le \sum_{j \in R_i} Y_{ijh}$

$$\forall i \quad h = 0, ..., |R_i|$$

For each route R_i:

- -all the upstream (downstream) levels start from o
- each located sensor is assigned exactly one upstream level
- and one downstream level
- the level assignment forms a consecutive sequence

$$\delta_{iht}^{jk} \leq \frac{1}{2} (y_{ijh} + z_{ikt})$$

$$\forall j,k \quad \forall i$$
$$\forall h,t = 0,1,..., |R_i|$$

If *j* is an upstream sensor Site of level *h* and *k* is a downstream sensor Site of level *t* on Route *i* then the corresponding variable delta is assigned value equal to 1

Conclusions

What we did:

-We have reviewed the problems of optimally locating AVI READERs on a traffic network to monitor travel time performances

- We presented two basic models to optimally locate AVI READERs either to maximize the total vehicle-miles monitored (MODEL AVI3) or the total travel time variance reduction (MODEL AVI4)

-We presented the mathematical formulations of both the models when - link failures are taken into account - sensor failures are taken into account

Conclusions

What's more to be done:

- Experiments with travel times predictions
 - Need data

-...

-...

- Need to calibrate the distributions

 Extensions of the models when the sensor failure probabilities are site specific

 Development of good heuristics to solve the problems on real case instances

