Best Time and Content for Delay Notification

Markus Schaal
Computergestützte InformationsSysteme (CIS), Technische Universität Berlin,
Sekr. E-N 7, Einsteinufer 17, D-10587 Berlin, Germany
e-mail: schaal@cs.tu-berlin.de

Hans-Joachim Lenz
Institut für Produktion, Wirtschaftsinformatik und Operations Research, Freie Universität Berlin,
Garystr. 21, D-14195 Berlin, Germany
e-mail: hjlenz@wiwiss.fu-berlin.de

Abstract

We consider an active information system, which aims to notify a traveler timely about a likely delay. Besides providing the right content, it should also wait for the best time for notification. This is especially true for mobile information services, where every notification may irritate the recipient.

We formulate and solve such a problem involving best choices both for content and time of notification by modeling it as an influence diagram. The chosen example stems from the domain of personal travel assistance.

Note that the action of the traveler is nothing else than a reaction to the notification from the information system.

1 Introduction

We aim to give a formal representation for a common but apparently not yet formalized decision problem. An information system (IS) providing timely notifications to a recipient faces the following task. Not only is it neccessary to decide on the best notification content at a given moment, but the IS must also decide whether to send the notification right away or to wait until later in order to decide on the notification content with improved precision. This is especially true for mobile or active decision support. Related work includes Mixed-Initiative-Assistance (cf. Ferguson et al. [5] and Horvitz [7]) and Just-In-Time-Information-Retrieval (JITIR) agents (cf. Rhodes [9]). A research project on information logistics at the FhG-ISST[1] specifically addresses the aspect of timely notification.

We consider the domain of personal travel assistance. Currently, the traveler consults an information system for the best route before he starts his journey and follows this route until he reaches his destination. However, delays may change the optimal route during the trip. This is especially true when parts of the routes are served by different transport providers, and connecting vehicles may not wait for a delayed passenger. Therefore, travelers must be informed about delays and alternative routes during their trip. The problem here is that delays are not certain in advance and early warnings could be misleading.

The following example will be studied in this article.

A traveler may reach his final destination on two distinct routes. The first route (by train) is considered to be faster than the second route (by taxi), but the latter is usually available without waiting, while the train may arrive late, causing the traveler to wait inefficiently.

The traveler has not yet decided which route to take, but will choose either of them with equal probability unless otherwise notified in time by the IS.

We will model the decision of the IS about *time* and *content* of the notification to be sent to the traveler. More specifically, we will decide whether to provide the best route now or to wait to provide a better route later.

The assumption here is that there is just one notification of the traveler. In other real-world problems, this constraint must be relaxed. In any case, extra notifications should be avoided.

2 Problem Formulation

Reasoning about the best time and content is a non-trivial task. The a-posteriori analysis may prove the a-priori decision to be ultimately wrong. In our example, the time of train departure influences the benefit of taking the train. A

[1] Fraunhofer Gesellschaft - Institut für Software- und Systemtechnik (Germany).

notification supporting the traveler's decision for the right mode of transport should be sent before the train leaves the station.

Uncertainty will be modeled by probability theory, thus we employ random variables. The following points of time must be considered.

- t^*_{clock}, the current time or clock reading
- t^*_d, the real departure time of the train
- t^*_s, the scheduled time for the train departure
- t^*_m, the time of traveler notification

We intend to use time-dependent probabilities that do not refer to absolute times. Instead, these probabilities depend on the temporal distance from the scheduled time for the train departure. Therefore all time variables will be defined as follows.

- Clock $t_{clock} = t^*_{clock} - t^*_s$, the (relative) current time or clock reading
- Delay $t_d = t^*_d - t^*_s$, the (relative) departure time of the train
- Notification time $t_m = t^*_m - t^*_s$, the (relative) time of traveler notification

Beside continuous time variables the following discrete variables are needed (possible values are given in curly brackets):

K(t) $\{$*delay*,*unknown*$\}$, the knowledge of the IS at time t on the delay of the train. For *K(t)=delay*, the train is known to be delayed, otherwise the IS has no knowledge about whether the train is delayed or not;

Weather $\{$*sunny*,*rainy*$\}$, an external influence on the departure time. *Weather* serves as a representative example for other external influences as well;

Timely $\{$*yes*,*no*$\}$, denoting the timely or late arrival of the notification with the traveler. (only timely notifications are effective)

Content $\{$*take train*,*take taxi*,*none*$\}$, the content of the notification to be sent to the traveler;

Action $\{$*takes train*,*takes taxi*$\}$, the action taken by the traveler.

The problem considered in this article can be stated now as follows. For each point in time $t \in dom(t_{clock})$, the state space S is the cross product $dom($*K(t)*$) \times dom($*Weather*$)$ and the decision space D is the cross product $dom(t_m) \times dom($*Content*$)$. For a given state $s \in S$, the best decision $d \in D$ is the one with optimal expected reward. The influences between variables and reward are described below.

2.1 Reward based on Action and Delay

Both train delay and the action to be taken by the traveler influence the reward of the notification. The reward could, for instance, represent the likelihood of the traveler's arrival in time. In order to exclude trivial cases, we require the existence of two distinct train delays t^1_d and t^2_d, such that different actions have to be taken by the traveler in order to maximize the reward, i.e.:
Reward(*takes train*,t^1_d) $<$ *Reward*(*takes taxi*,t^1_d) $\wedge$
Reward(*takes train*,t^2_d) $>$ *Reward*(*takes taxi*,t^2_d)
This means, that there is no traveler action that dominates the other one for all possible delays t_d.

2.2 Delay influences Knowledge

The delay t_d is related to the current knowledge *K(t)* about this delay. The IS changes its knowledge due to external information processes which are not under consideration here. The following assumptions are made:

1. *K(t)* is *unknown* until some point in time t_f (flipping time), where it flips to *delay*. It never flips back.
2. If *K(t)* flips to value *delay*, then the train must be delayed, i.e. *K(t)=delay* is free of error.
3. If the train is delayed ($t_d > 0$), then flipping time t_f and delay t_d are independent, i.e. the delay does not influence the time the IS may learn about this delay.

$F(t) = P(t_f \leq t \mid t_d > 0)$ is the distribution function of the flipping time t_f (a random variable) given that $t_d > 0$. $F(t)$ is the probability of *K(t)=delay* given that the train is delayed ($t_d > 0$).

2.3 Action based on Timeliness and Content

As argued in the introduction, timeliness of notification may be influenced by many factors. Here we are primarily interested in the aspect of effectiveness, i.e. whether or not the action can be chosen after notification. There is a critical time after which notification becomes ineffective. In our setting, this would be the time when the traveler chooses between two alternative actions which exclusively lead either to *takes train* or *takes taxi*.

3 The Model

We model the decision problem with an influence diagram. Influence diagrams are directed graphs with three types of nodes (cf. Shachter [11] and Pearl [8]). Chance nodes (shown as ovals) represent uncertain quantities, decision nodes (shown as rectangles) represent possible decisions and value nodes (shown as diamonds) represent rewards and costs for decisions and outcomes of uncertain quantities. Directed links leading to chance nodes denote conditional dependency, directed links leading to value nodes denote functional dependency and directed links leading to decision nodes are informational, i.e. the respective quantity is known before the decision has to be made. For demonstration purposes, discrete points in time were used instead of continuous time.

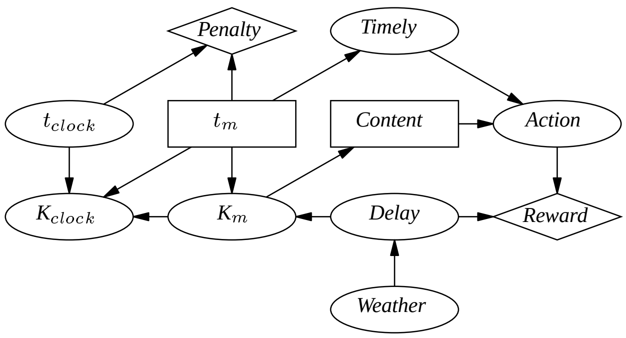

Figure 1. The influence diagram

The influence diagram for the example is shown in figure 1. First, a short description of the nodes is given, including the variables which have been introduced before (cf. Section 2 for details).

Chance Nodes:

Action $\{$*takes train*/*takes taxi*$\}$, the action of the traveler.

Weather $\{$*sunny*/*rainy*$\}$, the external influence of the weather.

Timely $\{$*yes*/*no*$\}$, timeliness of notification is modeled by this chance node. Only timely information may influence the *Action* of the passenger. Timeliness is guaranteed for early information and impossible for late information. For timely information, timeliness is reached with high probability.

K_{clock} $\{$*delay*/*unknown*$\}$, the IS knowledge at current (clock) time.

K_m $\{$*delay*/*unknown*$\}$, the (expected) IS knowledge at notification time.

t_{clock} $\{$*-20*/*-10*/*0*$\}$, the current clock time. The values may be interpreted as minutes.

Delay $\{$*in time*/*delayed*$\}$, the train delay. Delay t_d has been replaced by this discrete two-valued variable[2].

Decision Nodes:

t_m $\{$*-20*/*-10*/*0*$\}$, the time of notification.

Content $\{$*take train*/*take taxi*/*none*$\}$, the notification content.

Value Nodes:

Reward, the reward given on timely arrival.

Penalty, a value node to ensure temporal consistency (see below).

The existence of most links follows directly from the discussion in Section 2.1-2.3 and will be continued in Section 3.1-3.4 for a concrete problem instance. The following comments consider the remaining issues.

Since we consider two different points in time, we also need to represent the IS knowledge for both points in time. t_{clock} and t_m can be viewed as parameters on the conditional distributions of K_{clock} and K_m (cf. Section 3.2). K_m separates K_{clock} from *Delay*, since *K(t)* is free of error and independent of the delay.

Temporal consistency, i.e. $t_{clock} \leq t_m$ is ensured by the *Penalty*-node. The penalty for $t_{clock} > t_m$ is chosen to be greater than the absolute value of the maximum reward, i.e. for the maximum reward R_{max}, penalty is $P = -(R_{max} + \epsilon)$. For any choice of t_{clock} and t_m, the expected value of *Penalty* is either $-(R_{max}+\epsilon)$ or 0. Therefore, the expected value of *Reward*+*Penalty* is less than or equal to $-\epsilon$ for choosing t_m inconsistently and greater or equal to 0 for choosing t_m consistently. Therefore, the consistent choice of t_m will always be preferred versus the inconsistent choice of t_m.

The informational link between K_m and *Content* is crucial as the content does not need to be determined before t_m is reached and thus K_m is known (and equal to K_{clock}).

The decisions on notification time (t_m) and notification content (*Content*) are modeled as decision nodes, while the traveler's action (*Action*) is modeled as chance node. This is justified by the fact, that changes in the traveler's actions are only reactions to notifications stemming from the IS.

The joint distribution of all variables is given by the marginal distributions for clock t_{clock} and *Weather* together with conditional probability tables for all other variables with respect to their predecessors. Instances of the problem are calculated by entering evidence for t_{clock}, K_{clock}

[2][10] gives a quantitative model for the train delay t_d and provides a graphical model for stochastic reasoning there-about.

and *Weather* and propagating this information through the network until the new joint distribution is computed.

3.1 *Reward* based on *Action* and *Delay*

The reward is given in Table 1 below. For *Action=takes taxi* the reward is independent of the delay. 50 is the reward for taking the taxi. We can think of the reward as giving the likelihood of timely arrival at some destination (in percent). Then the values can be interpreted in the following manner:

- By taxi, the traveler will reach her destination in time with a likelihood of 50%.
- By train, the traveler will reach her destination in time with a likelihood of 100%, if the train is in time, and with 0%, if the train is delayed.

Action	*takes train*		*takes taxi*	
Delay	*in time*	*delayed*	*in time*	*delayed*
Reward	100	0	50	50

Table 1. *Reward*

3.2 *Delay* influences Knowledge *K(t)*

The relationship between the delay (*Delay*) and the imperfect knowledge *K(t)* is shown in Table 2 for K_m and in Table 3 for K_{clock}. In Table 2, the distribution of K_m depending on the random variables *Delay* and t_m is shown, while in Table 3 the distribution of K_{clock} is shown depending on random variables K_m, t_m and t_{clock}.

If *Delay=in time* and t_m =*-20*, then K_m=*delay* with a probability of 0% (cf. Table 2). If the train will be delayed (*Delay=delayed*), then K_m =*delay* with probabilities of 20%, 50% and 90% at t_m equal to *-20*, *-10* and *0* respectively.

For K_m=*unknown*, K_{clock} is *unknown* due to *K(t)* being free of error. The last column of Table 3 represents this fact in a shorthand representation. For $t_{clock} = t_m$, also K_{clock} =K_m. Since $t_{clock} > t_m$ is formally not forbidden, we assume equivalence of K_m and K_{clock} in these unspecified cases. For $t_{clock} < t_m$, the conditional probabilities can be inferred from the distribution of the flipping time. The distribution function for the flipping time is implicitly given in Table 2 by $F(t) = P(K_m\text{=}delay|Delay\text{=}delayed \wedge t = t_m)$. The following values are known:

$F(\text{-}20) = 0.2$
$F(\text{-}10) = 0.5$
$F(\text{-}0) = 0.9$

Hence the conditional probabilities can be calculated for Table 3 as well. As an example, we consider the probability of K_{clock}=*delay* for K_m=*delay*, t_m =*-10* and t_{clock} =*-20*. The probability is given by:

$$P(K_{clock}=delay|K_m=delay \wedge t_m=\text{-}10 \wedge t_{clock}=\text{-}20) =$$
$$P(K(\text{-}20)=delay|K(\text{-}10)=delay) =$$
$$P(K(\text{-}20)=delay \wedge K(\text{-}10)=delay) \div P(K(\text{-}10)=delay) =$$
$$F(\text{-}20) \div F(\text{-}10) = 0.2 \div 0.5 = 0.4$$

The value is underlined in the Table 3. The qualification on *Delay=delayed* can be omitted here, as *K(t)=delay* is sufficient for *Delay=delayed*.

	Delay	*in time*			*delayed*		
	t_m	*-20*	*-10*	*0*	*-20*	*-10*	*0*
K_m	*delay*	0	0	0	0.20	0.5	0.9
	unknown	1	1	1	0.80	0.5	0.10

Table 2. Knowledge K_m

	K_m	*delay*					
	t_m	*-20*			*-10*		
	t_{clock}	*-20*	*-10*	*0*	*-20*	*-10*	*0*
K_{clock}	*delay*	1	1	1	<u>0.40</u>	1	1
	unknown	0	0	0	0.60	0	0

	K_m	*delay*			*unknown*
	t_m	*0*			*
	t_{clock}	*-20*	*-10*	*0*	*
K_{clock}	*delay*	0.22	0.56	1	0
	unknown	0.78	0.44	0	1

Table 3. Knowledge K_{clock}

Temporal consistency ($t_{clock} \leq t_m$) is enforced for the decision on t_m by the *Penalty*-node (Table 4). The penalty for $t_{clock} > t_m$ is chosen to be -101 thus being absolutely greater than the maximum reward which is 100.

t_m	*-20*			*-10*			*0*
t_{clock}	*-20*	*-10*	*0*	*-20*	*-10*	*0*	*
Cost	0	-101	-101	0	0	-101	0

Table 4. *Penalty*

3.3 *Action* based on *Timely* and *Content*

Late notification results in failure to inform the traveler in time. This is represented by *Timely* (Table 5). Earliest notification (t_m =*-20*) results in *Timely=yes*, while late notification (t_m =*0*) results in *Timely=no*. An intermediate notification time (t_m =*-10*) will be timely with a probability of 90%. The *Action* (Table 6) does not depend on the *Content* for *Timely=no* or *Content=none*. In these cases *Action* takes both values with equal probability.

	t_m	*-20*	*-10*	*0*
Timely	*yes*	1	0.9	0
	no	0	0.10	1

Table 5. *Timely*

	Timely	*yes*			*no*
	Content	*take train*	*take taxi*	*none*	*
Action	*takes train*	1	0	0.5	0.5
	takes taxi	0	1	0.5	0.5

Table 6. *Action*

3.4 External factor *Weather*

The influence of the external factor *Weather* on the *Delay* is shown in Table 7.

	Weather	*sunny*	*rainy*
Delay	*in time*	0.7	0.30
	delayed	0.30	0.7

Table 7. *Delay*

3.5 Marginal Distributions

For completeness, the marginal distribution of *Weather* is given in Table 8 (left). t_{clock} is given with an evidence (*-10*) in Table 8 (right). Since this is no distribution, t_{clock} may also bear other evidences.

Weather	*sunny*	0.5
	rainy	0.5

t_{clock}	*-20*	0
	-10	1
	0	0

Table 8. *Weather* **and clock** t_{clock}

4 What-if Scenarios

Some scenarios will be presented in the following. The expected value of *Reward+Penalty* (also referred to as utility) will be denoted by E.

4.1 Scenario I(a): *Weather=sunny* / t_{clock}*=-20*

Evidence for *Weather* (*sunny*) is entered, clock time t_{clock} is *-20* and K_{clock} is *unknown*. The resulting utilities for different decisions on notification time t_m and *Content* are shown in Table 9 below.

t_m	E
-20	74.47
-10	77.55
0	62.23

Content	E
take train	74.47
take taxi	62.23
none	62.23

Table 9. Results for Scenario I(a)

The optimal notification time t_m is *-10*. Therefore, notification can be deferred. The optimal content (*Content*) cannot be determined, until notification time t_m is fixed.

4.2 Scenario I(b): *Weather=sunny* / t_{clock}*=-10*

Now, clock time t_{clock} is *-10* and K_{clock} is still *unknown*. Again, *-10* is the optimal notification time (Table 10 (left)). Notification has to be made immediately. For immediate notification, the resulting utilities for different decisions on *Content* can be seen in Table 10 (right).

t_m	E
-20	-26.53
-10	80.74
0	66.18

Content	E
take train	80.74
take taxi	51.62
none	66.18

Table 10. Results for Scenario I(b)

Obviously, the notification content *take train* should be delivered.

4.3 Scenario II: *Weather=rainy* / t_{clock}*=-20*

Evidence for *Weather* (*rainy*) is entered, clock time t_{clock} is *-20* and K_{clock} is *unknown*. The resulting utilities for different decisions on the notification time t_m are shown in Table 11 (left). In this case, *-20* is the optimal notification time, although the utility for t_m *=-10* is also near-optimal. The utilities for notification time t_m and *Content* are shown in Table 11 (right).

t_m	E
-20	50
-10	49.24
0	42.44

Content	E
take train	42.44
take taxi	50
none	42.44

Table 11. Results for Scenario II

Obviously, the notification content *take taxi* should be delivered.

5 Related Work

Reasoning with imperfect information has led to the theory of POMDP (Partially Observable Markov Decision Processes, cf. Hauskrecht [6]). Our model differs from previous work in this area:

- The temporal distance between successive states is not fixed, but given by the time variables t_m (notification time) and t_{clock} (clock reading).
- The transition probabilities do not only depend on the temporal distance between successive states, but also on the relative distance from the scheduled time t_s.

The representation of time in Bayesian networks has led to various distinct approaches. Berzuini [3] introduced a network of dates in order to reason about the probabilistic nature of event occurrence times for medical applications. Temporal random variables and continuous time is used in this work.

Dean and Kanazawa [4] propose random variables for duration as a means to represent semi-Markov processes in probabilistic networks. Tawfik and Neufeld [12] employ Temporal Bayesian Networks (TBN) for the representation of probabilities as functions of time. Arroyo-Figueroa and Sucar [2] model event occurrence times as nodes with respect to time intervals as developed by Allen [1]. This last approach is actually very similar to Berzuini's approach, but it is restricted to a finite number of intervals for the occurrence time of events.

We employ temporal variables in a manner similar to the one found in Berzuini, and Arroyo-Figueroa and Sucar. However, we employ relative points in time in order to represent the influence of temporal distances on the conditional probabilities.

6 Conclusion

The presented model enables decisions on immediate vs. deferred notification. The expected improvement of information precision is traded off against the expected loss of effectiveness.

Further investigations are aimed at incorporating two or more traveler (re-)actions and allowing for additional notifications.

Acknowledgment

This research was supported by the DFG (Deutsche Forschungsgemeinschaft), Berlin-Brandenburg Graduate School in Distributed Information Systems (DFG grant no. GRK 316).

References

[1] J. F. Allen. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26:832–843, 1983.

[2] G. Arroyo-Figueroa and L. E. Sucar. A Temporal Bayesian Network for Diagnosis and Prediction. In *Proceedings of the Fifteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–99)*, pages 13–20, San Francisco, CA, 1999. Morgan Kaufmann Publishers.

[3] C. Berzuini. Representing Time in Causal Probabilistic Networks. In M. Henrion, R. D. Shachter, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 5*, volume 10 of *Machine Intelligence and Pattern Recognition*, pages 15–28. North-Holland, Amsterdam, 1990.

[4] T. Dean, J. Kirman, and K. Kanazawa. Probabilistic Network Representations fo Continuous-Time Stochastic Processes for Applications in Planning and Control. In J. Hendler, editor, *Artificial Intelligence Planning Systems: Proceedings of the First International Conference (AIPS 92)*, pages 273–274, College Park, Maryland, USA, June 1992. Morgan Kaufmann.

[5] G. Ferguson, J. Allen, and B. Miller. Towards a Mixed Initiative Planning Assistant. In B. Drabble, editor, *Proceedings of the 3rd International Conference on Artificial Intelligence Planning Systems (AIPS-96)*, pages 70–77. AAAI Press, 1996.

[6] M. Hauskrecht. *Planning and Control in Stochastic Domains with Imperfect Information*. PhD thesis, MIT, 1997.

[7] E. Horvitz. Principles of Mixed-Initiative User Interfaces. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems*, volume 1 of *Characters and Agents*, pages 159–166, 1999.

[8] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1991. (Revised 2nd Edition).

[9] B. J. Rhodes. *Just-In-Time Information Retrieval*. PhD thesis, MIT Media Lab, 2000.

[10] M. Schaal. Probabilistic Plan Evaluation - A Prerequisite for a Model of Information Value Dynamics. Forschungsberichte des Fachbereichs Informatik 2000-16, Technische Universität Berlin, 2000.

[11] R. D. Shachter. Probabilistic Inference and Influence Diagrams. *Operations Research*, 36(4):589–604, 1987.

[12] A. Y. Tawfik and E. Neufeld. Temporal bayesian networks. In *Proceedings of First International Workshop on Temporal Representation and Reasoning (TIME)*, Pensacola, Florida, 1994. FLAIRS.