ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 2012
Seria: TRANSPORT z.74 Nr kol. 1861

Bartłomiej PŁACZEK

UNCERTAINTY AND INFORMATION GRANULATION IN ROAD TRAFFIC CONTROL

Summary. The contribution discusses the impact of information granulation on uncertainty of decisions in traffic control. Road traffic measurements have a predetermined precision, thus it is convenient to express their results in terms of information granules. A formal model of adaptive traffic control procedure is defined using set theory and relational representation of uncertainty. According to the proposed model, the traffic control decisions are based on a performance comparison for candidate control strategies. In this context, an uncertainty determination algorithm is introduced, which can be used for finding an optimal size of the traffic information granules.

NIEPEWNOŚĆ I GRANULACJA INFORMACJI W STEROWANIU RUCHEM DROGOWYM

Streszczenie. W artykule przedstawiono problem granulacji informacji w systemach sterowania ruchem drogowym oraz omówiono wpływ granulacji informacji na niepewność podejmowanych decyzji sterujących. Pomiary ruchu drogowego są realizowane z określoną precyzją, czego skutkiem jest ziarnisty charakter informacji opisującej parametry ruchu. Zdefiniowano formalny model adaptacyjnej procedury sterowania ruchem drogowym, wykorzystując do tego celu podstawowe definicje teorii zbiorów oraz relacyjny opis niepewności. Zgodnie z przyjętym modelem decyzje sterujące są podejmowane na podstawie porównania efektywności dla różnych wariantów sterowania. Zaproponowano algorytm określania niepewności decyzji sterujących, który może zostać wykorzystany do znajdowania optymalnego rozmiaru ziarna informacji opisującej warunki ruchu drogowego.

1. INTRODUCTION

The main task of the road traffic control is to select an optimal control strategy among several alternatives on the basis of currently available traffic information (e.g. optimal selection of travel route or traffic signal timing) [6]. For many cases the scope of real-time data available in modern road traffic monitoring systems (e.g. vision-based monitoring systems [7], vehicular sensor networks [5] and floating car data systems [10]) exceeds the requirements of particular traffic control implementations. Moreover, the transfer of all data records from traffic sensors to the control units is not advisable due to the high communication and processing costs.

The contribution of this paper is the proposal of a novel method of information granulation for traffic control applications, which provides a reduction in data amounts that have to be delivered from the traffic monitoring system. The underlying idea is to detect the necessity of data transfers on the basis of uncertainty determination of the traffic control decisions. This approach allows the precision of traffic measurements to be dynamically adapted to changing traffic conditions. As a consequence, the volume and complexity of processed traffic data can be significantly reduced.

The rest of the paper is organised as follows: Section 2 introduces basic definitions related to the granular information processing in road traffic control procedures. Section 3 describes the basic steps of decision making in the traffic control procedure. In Section 4, the impact of information granulation on the decision uncertainty is discussed and an uncertainty determination algorithm is proposed. Section 5 contains a numerical example on information granulation in signal traffic control at an intersection. Finally, in Section 6, conclusions are given and some future research directions are outlined.

2. INFORMATION GRANULATION

Granular computing has emerged as a new computational paradigm for complex information processing [1]. In this paradigm, information granules are the basic information units that group data according to certain criteria of similarity, indiscernibility, proximity or integrity [4]. The information granules are considered as atomic entities – equivalents of bits in digital information processing. In the granular computing theory several formal methods are used for information granules description, e.g.: numeric intervals, classical sets, fuzzy sets, rough sets, shadowed sets, and random numbers [9]. This section introduces definitions of basic concepts that are related to the granular information processing in road traffic control procedures.

Traffic information granule is a subset of universe of discourse $G \subseteq R$, representing a group of traffic states that can be recognised as identical for a given task of information processing [8]. The universe of discourse *R* is a multidimensional space of traffic parameters. In other words, the points in universe *R* correspond to all possible traffic states. This definition enables an arbitrary choice of formal method for representation of the information granules. The choice of formal method involves definition of theoretical framework for information processing (set theory, algebra of interval numbers, algebra of fuzzy numbers etc.) Information granulation is defined as a division of the universe of discourse *R* into a set of granules $\{G_1, \ldots, G_m\}$, which fulfils the following condition:

$$\bigcup_{i=1}^{m} G_i = R. \tag{1}$$

It should be noted, that the division mentioned above does not need to be crisp and the particular granules can overlap. The information granulation can be also expressed as a mapping:

$$M : R \to P(R), \tag{2}$$

where $P(R)$ denotes the power set of *R*.

The type of mapping *M* have to correspond to the formal framework used for granules representation. E.g., if the information granules are represented as fuzzy sets, then *M* is a fuzzy relation; in the case of rough sets application *M* is an indiscernibility relation.

In discrete universe of discourse *RD* the size of traffic information granule *G* is determined by the following sum [11]:

$$s(G) = \sum_{RD} g(r_i), \quad (3)$$

where $g(r_i)$ determines the membership of traffic state r_i in granule G.

If the information granule is represented by a classical set, then the binary membership definition is used:

$$g(r_i) = \begin{cases} 1, & r_i \in G \\ 0, & r_i \notin G \end{cases}. \quad (4)$$

3. TRAFFIC CONTROL PROCEDURE

This section discusses the basic steps of decision making process in a traffic control application.
A formal description of a generalised traffic control procedure is introduced using set theory [12] and relational representation of uncertainty [2]. Traffic information granules are defined as relations that include data collected from traffic measurements. Let us consider the road traffic control procedure as an iterative process that is based on control decisions making in successive time intervals. Each control decision determines the selection of control strategy, which will be implemented in a given moment of time. A set of consecutive decisions creates a traffic control programme for the defined time period T. The objective of traffic control procedure is to find an optimal programme, which minimises the objective function. The task o traffic control procedure can be interpreted as a problem of finding function $d(t)$:

$$d : T \to W, \quad (5)$$

where $W = \{w_i\}$ is a set of all applicable control strategies.

The function d takes integer values and it is referred to as control programme. Symbol $d(t)$ denotes the control decision, i.e. the number of control strategy selected for the time moment t.

The objective of traffic control is to find an optimal programme d^*, which satisfies the following condition:

$$e(d^*, f) \to \max, \quad (6)$$

where e denotes the performance of traffic control and f is a function describing parameters of traffic flow for the time period T:

$$f : T \to R, \quad (7)$$

where R is a multidimensional space of traffic parameters (universe of discourse).

The performance of road traffic control is determined (predicted) on the basis of two elements: the traffic model at hand and the available information on current traffic parameters (function f). The information necessary for this task is achieved by traffic monitoring and measurements. The traffic model includes knowledge about the control plant, which can be used to predict the performance of traffic control.

We will assume that the control decisions are made at some time steps, thus for the following discussion the time period T will be defined as a set: $T = \{t_0, t_0+1, ..., t_n\}$. In the context of control theory [2], the road traffic need to be considered as a dynamical and nondeterministic control plant. Therefore, the statement that function d^* describes the optimal control programme is usually uncertain. Main reason of this uncertainty is the lack of full knowledge, which is necessary to unambiguously determine the form of functions e and f. At time moment t_0, the function f can be just evaluated on a prediction basis. At the subsequent time steps, new traffic data are delivered from the traffic sensors and in consequence the function f can be evaluated more precisely. The complete information on function f is available at time step t_n. However, the performance function e has an approximate character at all times as it is impossible to precisely anticipate the influence of control decisions on the traffic control performance.

In order to take into account the information granulation and to implement the relational description of uncertainty, let us replace functions f and e with the following relations: $F_t \subseteq T \times R$ and $E \subseteq (R \times W)^{(n+1)} \times \mathbf{R}$ that include the information available at time t ($\mathbf{R}$ denotes the set of real numbers). Relation F_t represents the traffic information granule. The above definition includes family of relations $(F_t)_{t\in T}$, because at each time moment $t\in T$ the information on road traffic parameters is updated using the sensor data. The successively collected information granules are used to make the control decisions during the execution of traffic control procedure. At a given time moment $t\in T$, the $d(t_i)$ are already made decisions for $t_i \leq t$, and for $t_i > t$ they are decisions to be made in the future. Thus, the control programme at time step t can be expressed by the relation $D_t \subseteq T \times W$.

On the basis of the information available at time step $t\in T$ in relations E and F_t, we can predict the traffic control performance as follows:

$$\hat{E}_t = \{x \in \mathbf{R} : (r_0, w_0, r_0, w_0, \ldots, r_n, w_n, x) \in E \wedge \forall i \in \{0,1,\ldots,n\}((t_i.r_i) \in F_t \wedge (t_i, w_i) \in D_t)\} . \quad (8)$$

4. UNCERTAINTY OF CONTROL DECISIONS

The objective function of traffic control optimisation problem can be any combination of the following performance measures [3, 6]: average delay per vehicle, maximum individual delay, percentage of cars that are stopped, average number of stops, throughput of intersections, travel time, etc. Prediction of objective function values for all applicable strategies is required to decide the optimal control strategy. This prediction is always uncertain due to the nondeterministic nature of traffic flow. The additional source of uncertainty is the granulation of information describing the current traffic state. These two types of uncertainty affect control decisions. However, the uncertainty related to the traffic information granulation can be reduced by delivering new data from traffic sensors.

By $\hat{E}_t^{d(t)}$ we will denote the values set of objective function predicted at a time moment $t\in T$ for the case when the control decision $d(t)$ is made. It means that $(t, w) \in D_t \Rightarrow w = d(t)$. We will say that the decision $d'(t)$ is more effective than decision $d(t)$ if the following inequality of probabilities holds:

$$P(e' > e) > P(e' < e), \quad (9)$$

where: $e' \in \hat{E}_t^{d'(t)}$, $e \in \hat{E}_t^{d(t)}$.

To simplify notation, let $L[d'(t), d(t)] = d'(t)$ denote that the condition (10) is satisfied i.e. decision $d'(t)$ is more effective than $d(t)$. Uncertainty of this conclusion will be determined using the following formula:

$$UNC_{L[d'(t),d(t)]} = 1 - P(e' > e) + P(e' < e). \tag{10}$$

The conclusion $L[d'(t), d(t)] = d'(t)$ is certain if $P(e' > e) = 1$. Uncertainty is zero in that case, because $P(e' < e) + P(e' > e) \leq 1$ is always true. The uncertainty value increases when the probability $P(e' > e)$ falls, however this value is always lower than one since the condition (10) has to be satisfied.

A control decision $d^*(t)$ is an optimal one at a time moment t if:

$$\forall d(t) \in W\left(d(t) \neq d^*(t) \Rightarrow L[d^*(t), d(t)] = d^*(t)\right), \tag{11}$$

and uncertainty associated with this decision is given by:

$$UNC_{d^*(t)} = \max_{d(t) \in W - d^*(t)} \{UNC_{L[d^*(t),d(t)]}\}. \tag{12}$$

The control decision is based on evaluation of the objective function (9) for all applicable strategies. If the objective function can be evaluated precisely, then the uncertainty level of control decision is low. It means that the uncertainty depends directly on cardinality of set $\hat{E}_t$. Using the equation (9) it can be found that the cardinality of set $\hat{E}_t$ is proportional to the cardinality of relation F_t. Thus, in most cases the uncertainty of control decision can be reduced by decreasing the size of traffic information granule (cardinality of relation F_t). This reduction is achieved through updating the traffic data in F_t. The updating operation involves traffic data collection and prediction. The data collection is necessary to determine current traffic parameters. Moreover, the collected data are further used as a starting point of the traffic prediction and allow the prediction to be made with more confidence. The main insight is that an uncertainty threshold can be used to decide if new traffic data have to be delivered at a given time step of the traffic control procedure. Namely, the data collection is necessary only when the level of uncertainty, evaluated using equation (13), exceeds a predetermined uncertainty threshold. This technique will enable the considerable reduction of number of traffic measurements that are necessary to capture sufficient information for the traffic control optimisation.

5. NUMERICAL EXAMPLE

In this section a numerical example is discussed, which deals with decision making in road traffic control procedure using information granules represented by sets (relations). This example includes computations that are necessary for two control decisions in successive time steps: $t \in T$, $T = \{0, 1\}$. The control decision $d(t) \in W$ involves the control strategy selection from two available options: $W = \{0, 1\}$. Control strategies are some predefined programmes (timings) of traffic signals at an intersection. Each signalisation programme includes definitions of the cycle length, split and phase sequence. At the time steps 0 and 1 a signalisation programme is selected for implementation. The control step in this case correspond to the length of signalisation cycle.

The selection of control strategy (signalisation programme) is made on the basis of the currently available information, which describes traffic parameters and control performance.

In the presented example only one traffic parameter (the flow conditions) is used. It was assumed that this parameter takes three values: $R = \{0, 1, 2\}$ that denote respectively free flow, synchronised flow and congestion. Prediction of the traffic control performance is determined using the information included in relation $E = \{(r_0, w_0, r_1, w_1, x)_j, j = 1 \ldots 24\}$, which is presented in Table 1. Each row from Table 1 can be translated into an if-then rule, e.g. for the first row the following rule is obtained: if the flow is synchronised at time step 0 and control strategy 0 is selected and the flow is free at time step 1 and control strategy 0 is selected then the control performance is 3.

Table 1

Information on the performance of traffic control (relation *E*)

lp.	r_0	w_0	r_1	w_1	x
1	1	0	0	0	3
2	1	1	0	0	4
3	2	0	0	0	2
4	2	1	0	0	3
5	1	0	0	1	2
6	1	1	0	1	3
7	2	0	0	1	1
8	2	1	0	1	2
9	1	0	1	0	2
10	1	1	1	0	3
11	2	0	1	0	1
12	2	1	1	0	2
13	1	0	1	1	3
14	1	1	1	1	4
15	2	0	1	1	2
16	2	1	1	1	3
17	1	0	2	0	1
18	1	1	2	0	2
19	2	0	2	0	0
20	2	1	2	0	1
21	1	0	2	1	2
22	1	1	2	1	3
23	2	0	2	1	1
24	2	1	2	1	2

Traffic data describing the flow volume are partially collected from measurements and partially predicted. The traffic predictions are omitted in the first part of this example, thus at first only the measurement data will be used. Let us assume that the available traffic data allow the flow volume to be categorised into two states: a) free or synchronized, b) congested or synchronised. In other words, the granular traffic information is delivered in this case that includes two types of information granules describing the flow volume.

At the first time step ($t = 0$) of the traffic control procedure the flow volume is recognised as congested or synchronised. No additional prediction results are available. Therefore, for the second time step ($t = 1$) it was assumed that the flow can be free or synchronised or congested. Te available information on traffic state is given by the following granule:

$$F_0 = \{(0,1),(0,2),(1,0),(1,1),(1,2)\}. \tag{13}$$

Using the information in relation E (Table 1, rows 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23) the traffic control performance is predicted for strategy 0:

$$\hat{E}_0^0 = \{3,2,2,1,2,1,3,2,1,0,2,1\}, \tag{14}$$

and for strategy 1 (rows: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24):

$$\hat{E}_0^1 = \{4,3,3,2,3,2,4,3,2,1,3,2\}. \tag{15}$$

The following probabilities are then computed:

$$P(e_0 > e_1) = \tfrac{15}{144},\ P(e_0 < e_1) = \tfrac{95}{144}, \tag{16}$$

where $e_0 \in \hat{E}_0^0$ and $e_1 \in \hat{E}_0^1$.

Since the probability $P(e_0 < e_1)$ is higher than $P(e_0 > e_1)$, we can conclude that the selection of strategy 1 is more favourable in this situation ($L(0, 1) = 1$). The uncertainty of this conclusion is determined using equation (11): $UNC_{L(0,\,1)} = 0{,}44$. On this basis the control decision is made $d^*(0) = 1$ and the control programme takes the following form:

$$D_0 = \{(0,1),(1,0),(1,1)\}, \tag{17}$$

At the second time step ($t = 1$) new data from the traffic measurements become available and the traffic flow is identified as free or synchronised:

$$F_1 = \{(0,1),(0,2),(1,0),(1,1)\}. \tag{18}$$

Using the above traffic information granule, the control performance can be predicted for the first strategy (Table 1, rows: 2, 4, 10, 12):

$$\hat{E}_1^0 = \{4,3,3,2\} \tag{19}$$

and for the second strategy (rows: 6, 8, 14, 16):

$$\hat{E}_1^1 = \{3,2,4,3\}. \tag{20}$$

Results of the performance prediction are identical for the two traffic control strategies. In consequence the probabilities are equal:

$$P(e_0 > e_1) = P(e_0 < e_1) = \tfrac{5}{16}, \tag{21}$$

where $e_0 \in \hat{E}_1^0$ and $e_1 \in \hat{E}_1^1$.

There is no basis for the control decision in the presented situation (the uncertainty of control decision equals 1). The available traffic information is insufficient to enable the selection of more effective control strategy. More precise traffic data have to be used in order to eliminate this problem. Additional data delivered from traffic sensors will allow us to reduce the size of the information granules describing flow volume.

In the next part of this example, the traffic control procedure will be reconsidered for the case of enhanced traffic information (reduced size of granules). To this end, we assume that the data obtained by traffic measurements enable recognition of the three traffic flow categories: free flow, synchronised flow and congestion. Additionally, a traffic prediction is introduced to determine the traffic state for the second time step of control procedure. According to the assumptions, the results of flow prediction allow the future traffic state to be determined as: a) free or synchronised, b) congested or synchronised.

At the first time step (t = 0) the result of traffic measurement indicates the congested flow and simultaneously the free or synchronised flow is predicted for the second time step. Thus, the following traffic information granule is available:

$$F_0 = \{(0,2),(1,0),(1,1)\}. \tag{22}$$

The control performance is predicted for the first control strategy using the information from Table 1 (rows 3, 7, 11, 15):

$$\hat{E}_0^0 = \{2,1,1,2\}, \tag{23}$$

and similarly for the second strategy (rows 4, 8, 12, 16):

$$\hat{E}_0^1 = \{3,2,2,3\}. \tag{24}$$

Let $e_0 \in \hat{E}_0^0$ and $e_1 \in \hat{E}_0^1$, then:

$$P(e_0 > e_1) = 0,\ P(e_0 < e_1) = \tfrac{12}{16}. \tag{25}$$

On this basis we recognise the second strategy as the more effective one, i.e.: $L(0, 1) = 1$. Uncertainty of this conclusion is calculated using equation (11): $UNC_{L(0, 1)}$ = 0,25. In consequence the control decision is $d^*(0) = 1$ and the control programme is determined as follows:

$$D_0 = \{(0,1),(1,0),(1,1)\}. \tag{26}$$

At the second time step of the procedure (t = 1) the free flow state is registered:

$$F_1 = \{(0,2),(1,0)\}, \tag{27}$$

and the control performance is predicted for the two alternative control strategies accordingly (Table 1, row 4 and 8):

$$\hat{E}_1^0 = \{3\},\ \hat{E}_1^1 = \{2\}. \tag{28}$$

Assuming that $e_0 \in \hat{E}_1^0$ and $e_1 \in \hat{E}_1^1$, the following probabilities are computed:

$$P(e_0 > e_1) = 1,\ P(e_0 < e_1) = 0. \tag{29}$$

The above results allow the control strategy to be unambiguously selected: $L(0, 1) = 0$. Uncertainty of the control decision $d^*(1) = 0$ is determined as $UNC_{L(0, 1)} = 0$. And finally, the control programme is obtained:

$$D_1 = \{(0,1),(1,0)\}. \tag{30}$$

The above numerical example shows that the reduced size of traffic information granules corresponds with a substantially lower uncertainty of the control decisions. The reduction of granules size was necessary in this example for the determination of traffic control programme. In general, the uncertainty of control decisions depends on the precision of the delivered traffic information. If the precision of the resulting information is insufficient, the optimal control strategy cannot be derived without ambiguity. As a result the control decision becomes uncertain and it is a signal informing that new traffic data are necessary in the system to provide the information granules of lower size (higher precision) and to reduce the uncertainty of decision.

6. CONCLUSIONS AND FUTURE WORK

In this paper a novel approach was proposed to effectively reduce the amount of data that have to be delivered from traffic measurements in order to collect the sufficient information for traffic control applications. The offered approach is based on an original concept of the uncertainty dependent information granulation. According to this concept the reduction necessity of information granules size can be detected by the evaluation of uncertainty of traffic control decisions. Data from traffic sensors are collected to reduce the size of the information granules describing traffic flow. The basic formulas for the uncertainty calculations were derived using a formal description of generalised traffic control procedure. The uncertainty evaluation was carried out by means of set theory. In this solution, the set-theoretic notion of relation was used to represent imprecision of the available traffic information. The approach introduced in this paper provides the foundation for further research on uncertainty dependent traffic data granulation. In future studies the feasibility of the proposed information granulation method will be verified experimentally using a simulation environment and realistic mobility models. Tests will be performed to determine the effect of traffic information granulation on performance of the traffic control.

Bibliography

1. Bargiela, A., Pedrycz, W.: Granular Computing. An introduction. Kluwer Academic Publishers, Boston, Dordrecht, London, 2002
2. Bubnicki, Z.: Modern control theory. Springer-verlag, Berlin, Heldeiberg, 2005.
3. Gazis, D. C.: Traffic Theory. Kluwer Academic Publishers, Boston, Dordrecht, London, 2002.
4. Hoeppner F., Klawann F: Systems of Information Granules. In: Pedrycz W. et al. (eds.) Handbook of Granular Computing. John Wiley & Sons, Ltd, 2008, pp. 187–203.
5. Lee, U., Gerla, M.: A survey of urban vehicular sensing platforms. Computer Networks, vol. 54(4), 2010, pp. 527–544
6. Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y.: Review of road traffic control strategies. Proceedings of the IEEE Vol. 91, No. 12, 2003, p. 2043.
7. Płaczek, B.: Vehicles Recognition Using Fuzzy Descriptors of Image Segments. [in:] Kurzyński M. et al. (eds.) Advances in Soft Computing. Computer Recognition Systems 3. Springer-Verlag, Berlin Heidelberg, 2009 , p. 79.
8. Płaczek, B.: Przetwarzanie informacji ziarnistej w systemach sterowania ruchem drogowym. (in Polish) Magazyn Autostrady, nr 10, 2010, s. 98.
9. Qiu T.R., Liu Q., Huang H.K.: A Granular Computing Approach to Knowledge Discovery in Relational Databases. Acta Automatica Sinica, Vol. 35, No. 8, 2009, p. 1071.
10. Rahmani, M., Koutsopoulos, H. N., Ranganathan, A.: Requirements and potential of GPS-based floating car data for traffic management: Stockholm case study. [in :] 13th Int. IEEE Conf. on Intelligent Transportation Systems ITSC 2010, IEEE, 2010, p.730.
11. Thiele H.: On algebraic foundations of information granulation. I. The crisp case. [in:] The Ninth IEEE International Conference on Fuzzy Systems, FUZZ IEEE 2000, Vol. 2, p.780.
12. Vaught, R.L.: Set Theory. An Introduction. Second edition. Brikhauser, Boston, 2001.