Using Fuzzy Logic for Data Aggregation in Vehicular Networks

Irina Tal and Gabriel-Miro Muntean
Performance Engineering Laboratory, Network Innovations Centre,
School of Electronic Engineering, Dublin City University, Ireland
irina.tal2@mail.dcu.ie and munteang@eeng.dcu.ie

***Abstract*—Information provided in real-time via Vehicular Ad-hoc Networks is of great value in any kind of traffic systems. Bandwidth issues arise in this type of networks due to the potential large number of nodes. Data aggregation addresses these issues avoiding the dissemination of similar messages in the network. The lack of flexibility in the similarity criteria, security issues and the need of standardization were mentioned among the challenges of data aggregation schemes. Fuzzy Logic, very efficient in real-time systems, has been lately employed in data aggregation schemes. This paper analyzes various solutions for using Fuzzy Logic in data aggregation schemes and their mode of addressing the underlined challenges. The analysis conducted concludes that making use of Fuzzy Logic in data aggregation schemes is suitable to solving some of their issues and has great benefits in the development process of traffic systems that relies on these schemes.**

Keywords-component; aggregation, fuzzy, vehicular, networks, ad-hoc

I. INTRODUCTION

Vehicular Ad-hoc Networks (VANETs) are self-organizing networks formed from vehicles, capable of providing real time information about road conditions, vehicles speed, hazard events (e.g. accidents) which is of high value in any kind of traffic systems. Figure 1 presents a schematic network model for such a scenario. The data collected by the vehicles is further transmitted to the traffic system server via road-side units (RSU).

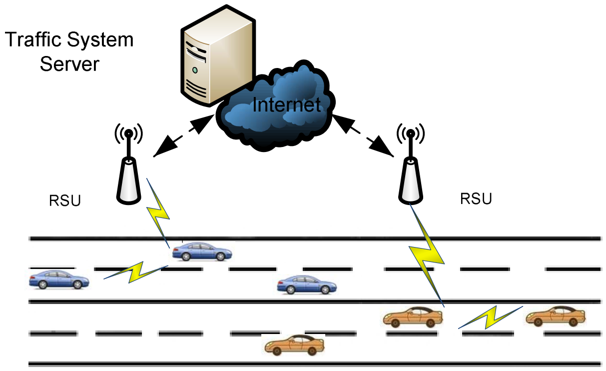

Figure 1. Network Model of a VANET-based Traffic System

One of the major challenges in VANET, a network with a potential large numbers of vehicle-nodes, is the efficient usage of the available bandwidth [1]. Data aggregation addresses this issue in the context of data collection, avoiding the dissemination of the similar information in the network. The existing data aggregation schemes have several limitations.

One common problem of data aggregation schemes is the lack of flexibility in the criteria of similarity between data. The information is correlated based on fixed or structured segmentation of the road assuming that two information items fit into the same spatial segment.

Another issue that arises in data aggregation is the security. The integrity of aggregated data is harder to be verified which opens new opportunities for attackers.

Moreover, most of the data aggregation schemes proposed so far are application-oriented and are not able to cope with different types of data or even less with simultaneous applications. In this context, European Telecommunications Standard Institutes (ETSI) underlined the need of data aggregation standardization [2].

Some of the latest data aggregation solutions employ Fuzzy Logic, in general very efficient in real-time systems [3]. This paper analyzes the existing modalities of making use of Fuzzy Logic in data aggregation schemes and their impact in solving some open issues in data aggregation. Based on this analysis, new modalities of employing Fuzzy Logic in data aggregation schemes are suggested.

This paper is structured as follows. Section II presents the concept of data aggregation in vehicular networks. Section III includes first an overview of Fuzzy Logic and its applications and then presents an analysis of the modalities of Fuzzy Logic employment in data aggregation schemes proposed in the literature so far. Conclusions are drawn at the end of the paper.

II. DATA AGGREGATION IN VEHICULAR NETWORKS

Data aggregation is used to combine correlated information from different nodes before redistributing the information into the network. The aim is to decrease the number of disseminated messages which in systems with a continuously update of information (e.g. traffic systems that keeps informs the drivers with the traffic conditions) will result in a dramatically decrease of the overhead.

A survey of the data aggregation solutions in VANETs reveals that there are two high level approaches for making use of data aggregation schemes in VANETs. These are described in Figure 1. In both approaches, data is collected locally (e.g. from local sensors) or received from other vehicles. Data is then analyzed to see whether there is any correlation between atomic data items. If it is decided that there is data correlation, then data is fused. Data, aggregated or not depending on the correlation, is further disseminated in the network. The difference between the two approaches is when the data is stored locally. In one of the approaches the data is stored after the fusion while in the other approach, based on the argument that local storage is cheap [1], the data is stored immediately after collection. However, both approaches have the same steps: **decision**, **fusion**, **storing** and **dissemination**, but in a different order.

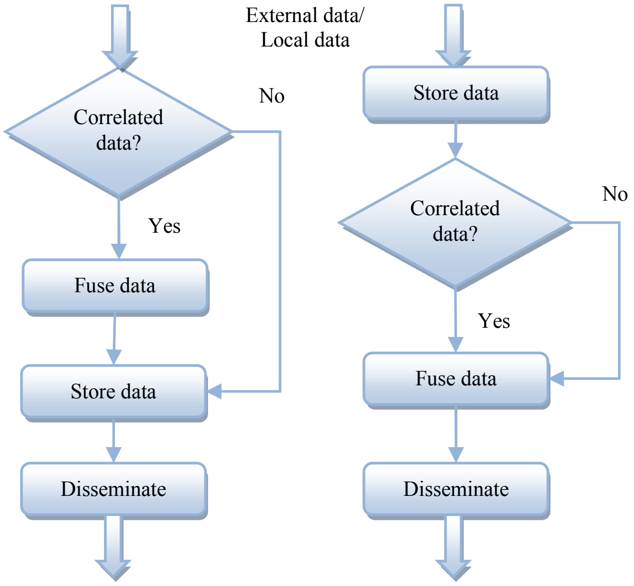

Figure 2. High-Level Approaches for Implementing Data Aggregation Schemes

Based on this observation, Dietzel et al. proposed an architecture model for modeling data aggregation schemes in vehicular networks [1]. The architecture model has four functional components, one for each outlined step: **decision component**, **fusion component**, **world model** – the storage component also referred to as knowledge source [7] –, and **dissemination component**. Each data aggregation scheme can be modeled in these four functional components, but the communication between components it is particular to each solution. This architecture model represents a step in the declared purpose of the authors ([1], [4]) of defining a generic data aggregation scheme that could be further considered for the needed standardization of data aggregation schemes. In consequence, this architecture model has been chosen in this work as support for the analysis of Fuzzy Logic employment in data aggregation.

III. DATA AGGREGATION AND FUZZY LOGIC

A. *Fuzzy Logic – overview*

Fuzzy Logic, first introduced in 1965 by Prof. Zadeh, was defined as an "*attempt to mimic human control logic*". Fuzzy Logic, an extension to the binary logic, uses continuous values in the [0, 1] interval and allows for the usage of linguistic variables (i.e. low, medium, high). This makes it suitable to modeling and solving a wide-range of real-world problems. Solving problems in Fuzzy Logic implies a Fuzzy inference system or a Fuzzy Logic controller, concepts introduced by Mamdani in 1975 [5]. Other Fuzzy Logic controllers were introduced in time (e.g. Sugeno's Fuzzy controller [6]). The structure of a Fuzzy Logic controller is presented in Figure 3 and its components are further described in the next paragraphs.

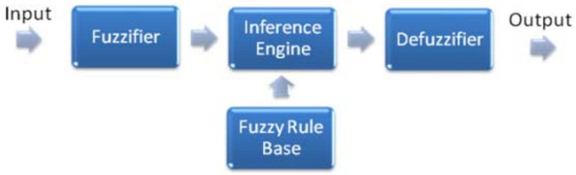

Figure 3. The Structure of a Fuzzy Logic Controller

The ***Fuzzifier*** component takes the crisp values of input parameters and gives as output their corresponding fuzzy degree of membership based on the defined membership functions.

The ***Inference Engine*** component does the mapping of the input fuzzified values on the output fuzzy set described by the output membership function. The mapping is done based on the rules contained in ***Fuzzy Rule Base***. The rules are most commonly of type "IF-THEN" as presented in equation (1).

$$\begin{aligned}&\textbf{\textit{IF}}\quad \textbf{\textit{input1}} = \textit{fuzzy\_term\_input1}\\&\qquad \textbf{\textit{AND/OR input2}} = \textit{fuzzy\_term\_input2}\\&\qquad \ldots\\&\qquad \textbf{\textit{AND/OR inputN}} = \textit{fuzzy\_term\_inputN}\\&\textbf{\textit{THEN output}} = \textit{fuzzy\_term\_output} \qquad (1)\end{aligned}$$

The ***Defuzzifier*** component takes the fuzzy set given as output by the inference process and transforms it in a crisp value.

Although Fuzzy Logic is a new technology it already has a large impact, being used on a wide scale and in many domains, such as engineering, medicine, science and business. Applications of Fuzzy Logic include: control applications, information systems, pattern recognition systems (e.g. image processing, machine vision), and decision support systems. Control applications are used in Robotics, Automotive and Consumer Electronics. For example the subway systems controlled by Fuzzy Logic controllers are widely spread in Japan. The first such a system, the Sendai subway developed by Hitachi [8], had a great success, improving transportation efficiency and, consequently encouraged development of more such subway systems.

In the Automotive industry there are two famous Fuzzy applications deployed in vehicles: efficient and stable control of car-engines and cruise-control [8]. The first application is deployed in Nissan vehicles and the second in Nissan and Subaru. This demonstrates that the basic infrastructure needed for deploying Fuzzy Logic controllers in vehicles which encourages the designing of Fuzzy-based solutions in vehicular networks is already present. However, in recent vehicular networks-related research there are still very few Fuzzy Logic-based solutions proposed. Some of these solutions were applied in data aggregation and will be exhaustively discussed next.

B. Fuzzy Logic in Data Aggregation

This section proposes to discuss some Fuzzy Logic-based data aggregation schemes in VANETs. The employment of Fuzzy Logic in data aggregation schemes is analyzed at the level of each of the four functional components of the architecture model for data aggregation schemes described in Section II: **decision, fusion, world model** and **dissemination**.

Fuzzy-based **decision components** were designed in [10] and [11]. Both approaches emphasize the limitations of the existing non-fuzzy decision components. The information is correlated based on fixed or structured segmentation of the road assuming that two information items fit into the same spatial segment. The data aggregation schemes based on these decision components are not flexible and they are static as they depend either on fixed road segments ([12], [13]) or predefined structures as trees ([15]) or clusters [14]. The Fuzzy-based solutions ensure flexibility and extensibility in the set of criteria used for correlating the information for aggregation. This resulted in structure-free and dynamic aggregation approaches.

In [10], the Fuzzy solution proposed takes into consideration two parameters: space – the approximate location of the provenience of information, and time – the persistence of information. Their Fuzzy terms are: LOW, MEDIUM and HIGH. The output parameter is the correlation between the pieces of information to be aggregated and has two possible values: YES and NO. The inference process/fuzzy reasoning is based on “IF-THEN” rules. The solution is open to be extended by considering other input parameters that might be dependent on the application type (e.g. speed, distance between vehicles etc.).

The Fuzzy-based decision component designed in [11] follows the same principle, but it takes a step forward in showing the generality, flexibility and extensibility that can be reached using a Fuzzy-based approach. The input parameters are named influences: *Influence 1*, … , *Influence n*. Unfortunately, only one parameter, speed difference between the vehicles, is given as a specific example in the exemplification of the rules and inference process.

Although none of the previous schemes specifies clearly, the steps indicated in the solutions are following in fact the structure of a Fuzzy Logic controller. As such, a generalized design (Figure 4) could be proposed for the Fuzzy-based decision components of data aggregation schemes: a Fuzzy Logic controller having as inputs the parameters that influence the correlation and as output {0, 1} values that stand for {Yes = correlated data, No= no correlated data}.

The flexibility and extensibility provided by the Fuzzy-based approaches are of great importance in the development of a system for data aggregation. As such, in the early stages of the development, such as the architectural description or design of the architectural components, there is no need of knowing all or even none of the parameters that influence the correlation between the pieces of information supposed to be aggregated. Moreover, even if in late stages of the development (e.g. after the implementation is finished) it will be concluded that new parameters that influence correlation must be considered, there will not be a considerable impact over the development of the system. The architecture of the system and most probably the design of components will not be affected. Only the implementation will be extended with new functionality. The *Fuzzifier* component will be extended with new membership functions for the newly considered parameters and new rules will be added in the *Rule Base*.

Taking into consideration these advantages brought over the non-fuzzy solutions, the generalized Fuzzy-based design for decision components could be imposed as a model. This could represent a step forward in the modeling and the standardization of data aggregation schemes.

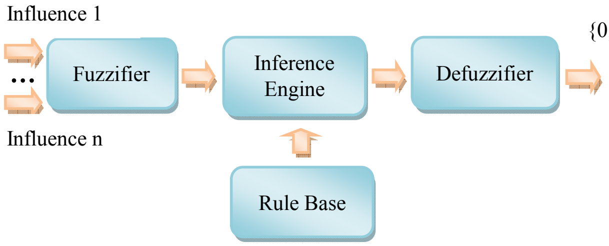

Figure 4. Generic Design for a Fuzzy-based Decision Component

To the best knowledge of the authors, there is only one solution in the literature that employs Fuzzy Logic in the **fusion component** of a data aggregation scheme. In [16], the authors enhanced the fusion component that resulted in a Fuzzy-based trust fusion component aiming to address the security issues imposed by data aggregation in vehicular networks. This approach solves the limitations of the other security mechanisms existing in data aggregation in VANETs that are summarized in the next paragraphs.

In [17], a security method using random checks that results in achieving a probabilistic verification of aggregates without a large communication overhead has been proposed. This method supposes the existence of a *tamper-proof service* in each vehicle that requests for integrity *proofs* from the randomly chosen original records. The main disadvantage of this approach consists in the dependency on the tamper-proof service that can be easily by-passed by attackers.

Another probabilistic verification-based security mechanism is proposed in [10]. This does not require the existence of a tamper-proof service on each vehicle, but relies in a certain level of trust in terms of node reputation which is not easily achieved in the self-organized networks such as VANETs [18].

A secure aggregation mechanism dependent on fixed segmentation road is presented by Raya et al. [19]. It has been shown that this type of aggregation schemes have scalability issues [20]. Moreover, the road segment-based group members have a common view of their environment which is achieved either by a multi-round scheme or by quantization. If first is applied, a serious overhead is included; if second is employed then a reduction in the value of the aggregated information is registered.

The approach that includes the Fuzzy-based trust fusion component removes the need for a tamper-proof service in each car. It is neither dependent on node reputation nor on any kind of structures (e.g. fixed road segments). Additionally it does not rely on agreeing to common views. The security mechanism is data-centric. A probabilistic scheme is employed to achieve a selective attestation of the aggregates. The selective attestation process results in clues leading to trust in the correctness of an aggregate. The clues are used as input parameters of a Fuzzy Logic controller that has as output parameter the Trust in the {0%-100%} range. The inference process is based on "IF-THEN" rules such as described in equation (2). The Fuzzy-based design allows for extensibility and flexibility in the considered clues depending on the type of the application.

IF **CLUE1** = *fuzzy_term_clue1*

AND/OR …

THEN **TRUST** = *fuzzy_term_trust* (2)

In [10] a Fuzzy-based data aggregation solution that implements the high-level approach where the storage is performed immediately after receiving the data is proposed. Based on the probable assumption that the communication channel could get overloaded restricting the amount of data that can be sent from the vehicle to the network, the authors proposed a Fuzzy-based selection scheme to be implemented in **world model.** This selects the most relevant data items for aggregation. The scheme is not detailed, only the possible parameters that could be considered in the selection are mentioned, such as severity or antiquity of information. However, the scheme is open to extension as other parameters could also be considered.

Some of the researchers argue that the selection scheme should be implemented in the **dissemination component,** before the aggregated data is sent to the next vehicles. This approach is suitable for data aggregation schemes that are implementing a high-level approach where the storage is done after the fusion. So far, no Fuzzy-based selection solutions were proposed for the dissemination component, but the previously described solution for the world model could be deployed in the dissemination component.

IV. CONCLUSIONS

This paper discusses how Fuzzy Logic can be used in data aggregation in VANETs. The analysis revealed that Fuzzy Logic can be successfully used in the design of all functional components of an aggregation scheme. The major benefits of its usage are the flexibility in the decision criteria and the extensibility of the resulted solution. This has a great positive impact on the development of systems that incorporate Fuzzy-based data aggregation schemes. Fuzzy Logic was successfully employed to address the security problem of data aggregation schemes as well. So far, the major contributions of Fuzzy Logic employment are in the decision making process that establishes whether the information is correlated or not. Based on these contributions, a generalized design model of the decision component was introduced that could be further considered in the standardization steps of data aggregation.

ACKNOWLEDGMENT

This work was supported, in part, by Science Foundation Ireland grant 10/CE/I1855 to Lero - the Irish Software Engineering Research Centre (www.lero.ie). This material is based upon works supported by Dublin City University under the Daniel O'Hare Research Scholarship scheme.

REFERENCES

[1] S. Dietzel, F. Kargl, G. Heijenk and F. Schaub, "Modeling in-network aggregation in VANETs," in *Communications Magazine, IEEE* , vol.49, no.11, pp.142-148, 2011.

[2] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications," 2010.

[3] J. M. Mendel, "Fuzzy logic systems for engineering: a tutorial," *Proc. of the IEEE*, vol. 83, no. 3, pp. 345–377, 1995.

[4] S. Dietzel F. Kargl, G. Heijenk and F. Schaub, "On the Potential of Generic Modeling for VANET Data Aggregation Protocols," *IEEE Vehicular Networking Conference,* 2010, pp 78-85,

[5] E.H.Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, Vol. 7, No. 1, 1975, pp. 1-13.

[6] Sugeno, M., "Industrial applications of fuzzy control", Elsevier Science Pub. Co., 1985.

[7] C. Lochert, B. Scheuermann and M. Mauve, "A probabilistic method for cooperative hierarchical aggregation of data in VANETs," *Ad Hoc Networks* vol. 8, no. 5, pp. 518-530, July 2010.

[8] B. Kosko and S. Isaka, "Fuzzy Logic," in *Scientific American*, 1993.

[9] L.Reznik, "Fuzzy Controllers*",* Elsevier, 1997, ISBN: 9780750634298. pp. 9-10,

[10] P. Caballero-Gil, J. Molina-Gil and C. Caballero-Gil, "Data Aggregation Based on Fuzzy Logic for VANETs", in *Int. Conf. on Computational intelligence in security for information systems*, 2011.

[11] S. Dietzel, B. Bako, E. Schoch and F. Kargl, "A Fuzzy Logic based Approach for Structure-free Aggregation in Vehicular Ad-hoc Networks", *ACM International Workshop on VehiculAr InterNETworking,* 2009.

[12] L. Wischhof, A. Ebner, H. Rohling, M. Lott and R. Halfmann, "SOTIS – A Self-Organizing Traffic Information System," IEEE VTC, pp. 2442–46, 2003.

[13] B. Yu, J. Gong and C.-Z. Xu, "Catch-up: a data aggregation scheme for vanets", *VANET'08: ACM Int. Workshop on VehiculAr Inter-NETworking,* pp. 49-57, NewYork, NY, USA, ACM, 2008.

[14] K. Ibrahim and M.C. Weigle, "CASCADE: Cluster-Based Accurate Syntactic Compression of Aggregated Data in VANETs", *GLOBECOM Workshops*, pp. 1-10, 2008.

[15] M. Caliskan, D. Graupner and M. Mauve, "Decentralized Discovery of Free Parking Places", *ACM VANET '06,* New York, USA 2006, pp. 30–39.

[16] S. Dietzel, E. Schoch, B. Konings, M. Weber and F. Kargl, "Resilient Secure Aggregation for Vehicular Networks", *IEEE Network*, pp. 26-31, 2010.

[17] F. Picconi, N. Ravi, M. Gruteser, L. Iftode, "Probabilistic Validation of Aggregated Data in Vehicular Ad-Hoc Networks," *ACM Int. Workshop on Vehicular Ad Hoc Networks*, New York, USA, 2006, pp. 76–85.

[18] S. Buchegger, J. Mundinger and J. Le Boudec, "Reputation Systems for Self-Organized Networks: Lessons Learned," *IEEE Technology and Society Magazine*, vol. 27, no. 1, pp. 41–47, 2008.

[19] M. Raya, A. Aziz and J. Hubaux, "Efficient Secure Aggregation inVANETs," ACM Int. Wksp. Vehicular Ad Hoc Networks, New York, USA, 2006, pp. 67–75.

[20] B. Scheuermann, , "A Fundamental Scalability Criterion for Data Aggregation in VANETs," *ACM Int. Conf. on Mobile Computing and Networking*, New York, USA, 2009, pp. 285–96,