Sensor-Based Activity Recognition Using Extended Belief Rule-Based Inference Methodology

Calzada A. *IEEE Member*, Liu J., *IEEE Member,* Nugent C.D., *IEEE Member*, Wang H., *IEEE Member and* Martinez L., *IEEE Member*

***Abstract*— The recently developed extended belief rule-based inference methodology (RIMER+) recognizes the need of modeling different types of information and uncertainty that usually coexist in real environments. A home setting with sensors located in different rooms and on different appliances can be considered as a particularly relevant example of such an environment, which brings a range of challenges for sensor-based activity recognition.**

Although RIMER+ has been designed as a generic decision model that could be applied in a wide range of situations, this paper discusses how this methodology can be adapted to recognize human activities using binary sensors within smart environments. The evaluation of RIMER+ against other state-of-the-art classifiers in terms of accuracy, efficiency and applicability was found to be significantly relevant, specially in situations of input data incompleteness, and it demonstrates the potential of this methodology and underpins the basis to develop further research on the topic.

I. INTRODUCTION

Automatic recognition of human activities is a promising field that would aid developing solutions for applications in different domains such as healthcare, context-aware computing, security or ambient-assist living, to name but a few [1, 2]. Recent advances in several aspects of sensing technologies, such as miniaturization or decreasing costs, have progressed activity recognition related research substantially forward. Nevertheless, despite these advances, activity recognition is still regarded as a complex process where different types of information, data formats and elements of uncertainty are usually involved.

Traditionally, the methods utilized to approach activity recognition problems have been divided into two main groups: Knowledge-Driven Approaches (KDA) and Data-Driven Approaches (DDA) [2]. While KDAs make use of rich domain-specific expert knowledge to model activities, DDAs are based on learning from the information retrieved by sensors and labeling each activity. One key idea that this research considers is that the activities, whether they are modeled using a KDA or learnt and labeled using a DDA, need to be compiled in some type of Knowledge-Base (KB) in order to recognize or predict future activities. Depending on the methodology utilized, this prediction can be performed by means of induction, deduction or some other method that aggregates or infers the knowledge stored in the KB.

Invest Northern Ireland is acknowledged for partly supporting this project under the Competence Centre Programme Grant RD0513853 - Connected Health Innovation Centre.

A. Calzada, J. Liu, C.D. Nugent and H. Wang are with the School of Computing and Mathematics, University of Ulster, Belfast, BT37 0QB, UK (email: {a.calzada, j.liu, cd.nugent, h.wang}@ulster.ac.uk).

L. Martinez is with the Department of Computer Science, University of Jaen, Spain (email: luis.martinez@ujaen.es).

In order to take advantage of the main benefits provided by DDAs and KDAs [2] and to avoid some of their common disadvantages, this research utilizes the recently developed extended belief rule-based inference methodology (RIMER+) [3], which can be regarded as a combination of DDA and KDA. RIMER+'s KB is based on Extended Belief Rule-Bases (E-BRB), which are able to capture (i) sample data and expert knowledge in a homogeneous way; (ii) nonlinear and causal relationships; and (iii) several types of uncertainty related to expert knowledge and data. E-BRBs can be utilized to complement real data from sensors with expert knowledge in order to provide a more elaborate model of the activity recognition problem domain. This can be viewed as a great advance, given that one of the main inconveniences of DDA methods is their limited capacity to represent expert knowledge. The E-BRB can be generated from data using the novel rule generation scheme or provided by experts [3]. To produce a prediction result based on some input of the system, RIMER+ uses the Evidential Reasoning (ER) algorithm [4] to infer the information included in its E-BRBs.

The remainder of this paper is organized as follows: Section II briefly outlines the RIMER+ approach; Section III discusses how RIMER+ was adapted to specifically work with binary sensor data and Section IV details the case study and presents the results obtained. Finally, Section V concludes this paper.

II. RIMER+

One of the main features of RIMER+ is its E-BRB, which extends the KB used in the belief Rule-Base Inference Methodology (RIMER) [7], using belief degree distributions embedded in the consequent terms of its rules and also in each antecedent term. Take for example the following EBR:

IF Temperature is {(Hot, 0.7), (Warm, 0.1), (Cold, 0)} **THEN** Heating is {(ON, 0.8), (OFF, 0.2)} (1)

This kind of extended belief rule is generic in the sense that it is able to not only capture fuzziness (linguistic terms), uncertainty (beliefs), incompleteness (partially known belief or ignorance) and nonlinear relationships (IF-THEN rules) in an integrated way, however, also provides a flexible way to incorporate hybrid input information and an efficient rule generation scheme to build E-BRBs directly from sample data (refer to [3] for further details).

E-BRBs also provide the flexibility to incorporate context information in the KB. This context information could be vague, uncertain and/or incomplete and quantitative or qualitative in nature. In addition, they provide the means to tune the importance of different rules and antecedents, using rule weights (noted as Θ_k for the k^{th} rule) and antecedent relative weights (noted as δ_{ik}, for the i^{th} antecedent of the k^{th} rule), respectively. Once the E-BRB is generated from sample data and complemented with expert knowledge related to an environment, it can be used to recognize future activities, setting the values of sensors as inputs for the E-BRB. This recognition method of the RIMER+ approach is based on two main processes:

1) Rule Activation: evaluate which rules need to be activated, computing their activation weights (w_k) using the similarity of their antecedents against the given inputs.

2) Rule Inference: the ER approach [4] is applied to combine the activated rules and generate the final output.

Given that an E-BRB may be generated from sample data, the quality of data might be a big issue to be concerned when generating a reliable E-BRB. In this regard, a new Dynamic Rule Activation (DRA) algorithm has been proposed in [6, 13] as a method to select the most relevant information to be aggregated within the RIMER+ decision model. This is undertaken by considering that data incompleteness and inconsistency may be viewed as paired situations, given that the former appears due to lack of information while the latter can be considered as an excess of heterogeneous information [6]. This upgraded method is denoted as R+DRA, and is detailed in Fig. 1. This approach enhances the performance of RIMER+ (without DRA), especially in multi-class datasets. For details relating to the procedures and algorithms refer to [6]. It is important to note that the processes listed in Fig. 1 can be modified depending on each particular scenario. In this regard, the following Section details how the Individual Matching Degree and Rule Activation were upgraded to work with binary sensor data.

III. ADAPTATION FOR BINARY-SENSOR DATA

To recognize activities in a home setting, this research uses data from sensors that provide two possible values (0 or 1). Given that RIMER+ provides the flexibility to modify the similarity and aggregation functions (to calculate the individual matching degree and rule activation weights, as depicted in Fig. 1), this study proposes to replace these two functions with a popular similarity measure named hamming distance [8], particularly designed to work with binary (and qualitative) data:

$$H(\boldsymbol{\alpha}, \boldsymbol{A_k}) = \sum_{i=0}^{T_k} \delta_{ik} * d_H(\alpha_i, A_{ik}) \tag{2}$$

$$\text{where: } d_H(\alpha_i, A_{ik}) = \begin{cases} 0, & if\ |\alpha_i - A_{ik}| = 0 \\ 1, & otherwise \end{cases}$$

where $\boldsymbol{\alpha}$ is the input vector (α_i is the input for the i^{th} antecedent attribute) and $\boldsymbol{A_k}$ is the antecedent vector for the k^{th} rule (where A_{ik} is the i^{th} antecedent of the k^{th} rule). Note that in the case studies presented in this research, each α_i and A_{ik} may only take the values 0 or 1, since the activity recognition environment is based on binary sensors. If the similarity function H (Eq. (2)) returns zero this means that the input vector perfectly matches the antecedents of the k^{th} rule (i.e. the current sensor values totally match the description of one activity). The higher value H returns, the more dissimilar the input vector is to the k^{th} rule.

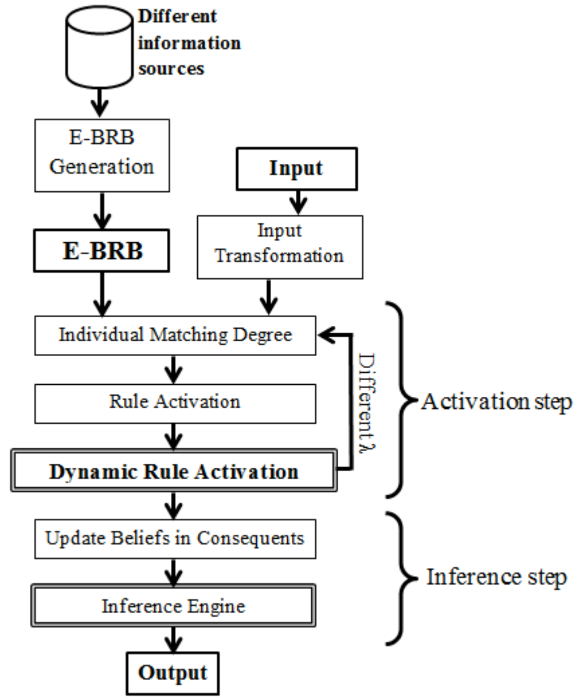

Fig. 1 The R+DRA decision-making process (from top to bottom).

In order to allow the decision model to accommodate noise in the input data, the rule activation weights (w_k) are calculated as follows:

$$w_k = \begin{cases} 1 * \theta_k, & if\ H(\boldsymbol{\alpha}, \boldsymbol{A_k}) = 0 \\ 0.2 * \theta_k, & if\ 0 < H(\boldsymbol{\alpha}, \boldsymbol{A_k}) \leq 1 \\ 0.1 * \theta_k, & if\ 1 < H(\boldsymbol{\alpha}, \boldsymbol{A_k}) \leq 2 \\ 0, & if\ H(\boldsymbol{\alpha}, \boldsymbol{A_k}) > 2 \end{cases} \tag{3}$$

Eq. (3) models the idea that some noise in the input vector is permitted (e.g., some sensors may have some technical failure or been deactivated, subsequently returning wrong values). Although rules containing a certain amount of noise (in 1 or 2 sensors) are not completely discarded, their activation weight is substantially lower than if no noise is found ($H(\boldsymbol{\alpha}, \boldsymbol{A_k}) = 0$). For the binary sensor data, the proposed H function suits better than the Euclidean distance used in [3]. The case studies presented in the following Section compare these two options and demonstrate how the use of H considerably enhances the results.

IV. CASE STUDY

This Section details the tests run to evaluate the performance of the method proposed in the previous Section.

A. Activity Recognition Dataset

The case study presented in this paper is based on a popular activity recognition dataset, presented in [1, 5]. The dataset will be used to compare the performance of RIMER+ against other popular classifiers. It contains 245 observations generated using 14 binary sensors, placed in 14 locations within a home setting: microwave, hall-toilet door, hall-bathroom door, cups' cupboard, fridge, plates' cupboard, front door, dishwasher, toilet flush, freezer, pans' cupboard, washing machine, groceries' cupboard and hall-bedroom door. Using this information, the 7 activities to be recognized were defined as: *going to bed*, *using toilet*, *preparing breakfast*, *preparing dinner*, *getting a drink*, *taking a shower* and *leaving the house*. It is important to note that the 7 classes are not evenly distributed; hence the dataset is unbalanced with just one class (*using toilet*) being assigned to 46.5% of the observations while others (like *preparing dinner*) are only represented in 4% of the observations.

When an E-BRB is built using the described dataset, it would have as many extended belief rules as training observations. Each one of those rules would have 14 antecedents (one per sensor) and a consequent, as follows:

IF Microwave is {(*ON*, 0), (*OFF*, 1)} **AND** … **AND** Hall-bedroom Door is {(*ON*, 0), (*OFF*, 1)} **THEN** Activity is {(*going to bed*, 0), (*using toilet*, 1), …, (*leaving the house*, 0)} (4)

It is important to note that all the attributes were defined as a set of qualitative values. The sensors included two possible discrete values: 1 and 0, which would translate as ON and OFF, respectively.

The following tests analyze one important aspect: the performance of the classifiers in situations of incompleteness of input information. To assess this, several sensors will be manually deactivated to simulate a situation where they did not activate and hence will have their output permanently set to zero.

B. Tests definition

A series of tests were run to compare the performance of RIMER+ (using Euclidean and Hamming - Eq. (2) - similarity distances) against other popular classifiers used as DDA approaches for activity recognition. Among them, the Naïve Bayes classifier (NB) has been considered to retrieve promising results despite its simplicity [2, 14, 15], and Support Vector Machines (SVMs) also have been recognized to work consistently well in sensor-based activity recognition [2, 16]. Hence, these two classifiers are included as reference methods for comparative purposes in this study. In addition, the Nearest Neighbor (NN) [9], and Decision Table (DT) [10] classifiers were also included as reference methods.

To assess their performance in different situations, three commonly-used types of tests were run to evaluate the classification accuracy of each method: (i) 10-fold Cross-Validation (CV10); (ii) using 66% of samples for training and the remaining for testing (66%T); and (iii) using half of samples for training and the other half for testing (50%T).

Moreover, as previously mentioned, additional series of tests were performed posing the situation where a number of sensors were faulty for any reason (e.g. low battery, sensor failure). In this regard, the tests were emulated for the cases where 4, 5 and 6 sensors always provided an output of 0 to simulate them being permanently inactive.

C. Results and Discussion

For the first series of tests, the accuracy of the selected methods was analyzed. The accuracy is measured as the percentage of observations correctly predicted over the total number of observations. As Table I illustrates, the proposed method (RIMER+ enhanced with DRA using the Hamming distance as described in Section III – noted as R+DRA**H**) outperforms most methods in every case. Only SVM and NB are able to match the performance of the proposed method, however, in just one test with each one of them. Therefore, the hamming-based R+DRA method presented in this research can be considered as the most consistent in its positive performance.

TABLE I. ACTIVITY RECOGNITION DATASET RESULTS

Method	**Accuracy (%)**			
	CV10	***66%T***	***50%T***	***Mean***
R+DRA	93.06	91.46	91.87	92.13
SVM	**96.73**	93.90	90.16	93.59
NN	96.33	95.12	95.08	95.51
NB	96.33	95.12	**96.72**	96.05
DT	95.51	95.12	93.44	94.69
R+DRA**H**	**96.73**	**96.34**	**96.72**	**96.59**

The second series of tests considered a situation where 4 sensors from the activity recognition environment had some type of failure. These 4 deactivated sensors were: microwave, dishwasher, washing machine and pans' cupboard. Therefore, the aim of the following series of tests was to evaluate the performance of the methods in this situation of lack of information (input incompleteness), where the scenario is affected by an increasing amount of uncertainty. In addition to 4 sensors deactivated in the previous tests, another series of tests that also deactivated the cup's cupboard sensor were run. The results of these tests are summarized in Table II:

TABLE II. ACTIVITY RECOGNITION DATASET RESULTS WITH 4 AND 5 SENSORS DEACTIVATED

Method	**Accuracy (%)**					
	4 sensors deactivated			**5 sensors deactivated**		
	CV10	***66%T***	***50%T***	***CV10***	***66%T***	***50%T***
R+DRA	95.51	92.68	92.68	95.92	92.68	93.44
SVM	96.73	93.90	90.16	96.73	93.90	90.16
NN	**97.55**	93.90	95.08	97.14	92.68	94.26
NB	96.73	**95.12**	95.08	93.87	92.68	94.26
DT	94.28	**95.12**	93.44	93.87	**95.12**	93.44
R+DRA**H**	**97.55**	**95.12**	**96.74**	**97.55**	93.90	**95.94**

These series of tests reflect a similar situation than the previous ones: the proposed method is the best in accuracy in almost every situation, while other methods are just able to match its performance, in one test at most (like the NN in CV10). The fact that the accuracy of R+DRA-based methods sometimes increases under these levels of inconsistency is not strange. This is because when fewer sensors are utilized in the system, the EBRs contain fewer antecedents and therefore need fewer conditions to be activated. Accordingly, the E-BRB is less restrictive, and more information is activated for an input, presumably leading to more accurate outcomes.

Table II also illustrates that, when 5 sensors were deactivated, although the proposed method is the best performing in the CV10 and 50% train tests, Decision Table outperforms it in the 66% train. In order to contrast these results, one final series of tests deactivating one more sensor (hall-bathroom door) were run. Note that in these final tests, 6 sensors are deactivated (43% of the total), hence the uncertainty of the activity recognition environment can be considered as substantial. As Table III illustrates, methods like NN, NB and DT are affected by the lack of information and increasing uncertainty. Nevertheless, R+DRA methods are barely affected by this factor. Especially in the 50% train tests, where fewer observations are used for training, R+DRA still maintains its accuracy over 95% of the testing observations, which is almost 2% better than the second best-performing method in the case where 6 sensors were deactivated. This capability of R+DRA to maintain a high accuracy despite increasing levels of incomplete data and uncertainty can be of great use when approaching real scenarios, like real-time activity recognition.

TABLE III. ACTIVITY RECOGNITION DATASET RESULTS WITH 6 SENSORS DEACTIVATED

Method	Accuracy (%)		
	CV10	*66% T*	*50% T*
R+DRA	95.10	91.46	92.68
SVM	95.10	89.02	92.62
NN	95.51	89.02	91.80
NB	93.46	91.46	93.44
DT	92.25	**92.68**	91.80
R+DRA**H**	**96.73**	**92.68**	**95.12**

Finally, in order to contrast the results, a series of one-tailed, unpaired (Type2) T-tests were run to measure the statistical significance of the accuracies obtained from the tested methods against the R+DRA**H** ones. Table IV includes the *p*-values retrieved from these tests as well as the average time complexity (in seconds) of the tests run.

TABLE IV. TIME COMPLEXITY AND SIGNIFICANCE TESTS OF THE STUDIES PERFORMED

Method	Summary comparison against R+DRAH				
	T test-results (*p*-values)	Time (s)			Software used
		CV10	*66%T*	*50%T*	
R+DRA	0.0001	3.3	0.5	0.6	RIMER [12]
SVM	0.0051	3.8	1.3	1.0	Weka [11]
NN	0.0428	0.3	0.0	0.1	Weka [11]
NB	0.0191	0.1	0.0	0.0	Weka [11]
DT	0.0005	1.8	0.3	0.4	Weka [11]
R+DRA**H**	-	3.2	0.5	0.6	RIMER [12]

As Table IV illustrates, R+DRA**H** is significantly more accurate than every other method (at 95% confidence, since every *p*-value is less than 0.05). Regarding the time complexity, the software in which each method is implemented plays a crucial role, since it is not possible to verify which implementation is more efficient. While SVM, NN, NB and DT were run using Weka [11], the RIMER-based methods were tested using the RIMER Tool [12]. Despite this fact, it can be considered that the proposed method is highly efficient since it does not need any complex pre-processing method to generate rules [3] and it retrieves the decision results in a short period of time: 3.2secs for the 10-fold Cross Validation tests, the longest run in this study. This means that the proposed method needs just about 13 milliseconds to provide an output for each input vector (3200 milliseconds / 245 observations).

V. CONCLUSION

This paper presents RIMER+ as a suitable decision support approach that can be used to model the sensor data and expert knowledge needed to deliver reliable activity recognition. The E-BRBs included in RIMER+ contain mechanisms to represent various information-related uncertainties, such as vagueness, imprecision and incompleteness. The case studies presented in this paper illustrate that the proposed R+DRAH is able to overcome the accuracy retrieved by some of the most popular classifiers, preserving a considerably low time complexity. Moreover, when some of the sensors were deactivated simulating uncertainty in the sensor data, R+DRAH demonstrated to be a very robust method in situations of input incompleteness, performing significantly better than any other classifier.

As it was shown, the performance of RIMER+ can be considerably enhanced when the hamming distance is used to measure the similarity between inputs and rules in a binary sensor-based scenario. Hence, it is worth noting that many other elements could be modified to further improve R+DRAH, e.g., adding expert knowledge to the E-BRB or using other inference engines of similarity functions.

ACKNOWLEDGMENT

Invest Northern Ireland is acknowledged for partly supporting this project under the Competence Centre Programme Grant RD0513853 - Connected Health Innovation Centre.

REFERENCES

[1] Van Kasteren T., Noulas A., Englebienne G. and Krose B., (2008), *"Accurate Activity Recognition in a Home Setting"*. International Joint Conference on Pervasive and Ubiquitous Computing, September 21-24, Seoul, South Korea.

[2] Chen L., Hoey J., Nugent C.D., Cook D.J. and Yu Z.W., (2012), *"Sensor-Based Activity Recognition"*. IEEE Transactions on Systems, Man, and Cybernetics – Part C, 42(6): 790-808.

[3] Liu J., Martinez L., Calzada A. and Wang H., (2013), *"A novel belief rule base representation, generation and its inference methodology"*. Knowledge-Based Systems, 53:129-141.

[4] Yang J.B. and Xu D.L., (2002), *"On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty"*. IEEE Transactions on Systems, Man and Cybernetics – Part A, 32(3): 289-304.

[5] Van Kasteren T., Englebienne G. and Krose B., (2010), *"Transferring knowledge of activity recognition across sensor networks"*. In *Pervasive Computing*, Springer Berlin Heidelberg: 283-300.

[6] Calzada A., Liu J., Wang H. and Kashyap A., (2013), *"A new dynamic rule activation method for extended belief rule-bases"*. Submitted to IEEE Transactions on Knowledge and Data Engineering: Under Review.

[7] Yang J.B., Liu J., Wang J., Sii H.S. and Wang H.W., (2006), *"Belief rule-base inference methodology using the evidential reasoning approach – RIMER"*. IEEE Transactions on Systems, Man and Cybernetics – Part A, 36(2): 266-285.

[8] Hamming R. W., (1950), *"Error detecting and error correcting codes"*. Bell System Technical Journal, 29(2): 147–160.

[9] Cherkassy V. and Mulier F., (2007), *"Learning from Data: Concepts, Theory, and Methods, Second Edition"*. John Wiley and Sons, Inc.

[10] Kohavi R., (1995), *"The power of Decision Tables"*. In 8th European Conference on Machine Learning: 174-189.

[11] Weka Project (2013), *"Environment for knowledge analysis. Weka 3: Data mining software in Java"*. Machine Learning Group, University of Waikato. [Available: http://www.cs.waikato.ac.nz/ml/index.html, last accessed on 1st, April, 2014].

[12] Calzada A., Liu J., Wang H., Martinez L. and Kashyap A., (2011), *"An intelligent decision support tool based on belief rule-based inference methodology"*. IEEE International Conference on Fuzzy Systems, June 27-30, Taipei, Taiwan: 2638-2643.

[13] Calzada A., Liu J., Wang H. and Kashyap A., (2013), *"Dynamic rule activation for extended belief rule bases"*, International Conference in Machine Learning and Cybernetics (ICMLC), July 14-17, Tianjing, China.

[14] Cook D. and Schmitter M., (2009), *"Assessing the quality of activities in a smart environment"*. Methods of Information in Medicine, 48(5): 480-485.

[15] Van Kasteren and Krose B., (2008), *"Bayesian activity recognition in residence for elders"*. In Proceedings of International Conference of Intelligent Environments: 209-212.

[16] Brdiczka O., Crowley J. and Reignier P., (2009), *"Learning situation models in a smart home"*. IEEE Transactions on Systems, Man and Cybernetics – Part B, 39(1):56-63.