Volume 8, No. 5, May-June 2017

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Survey of Multi Entity Bayesian Networks (MEBN) and its applications in probabilistic reasoning

Shrinivasan R Patnaik Patnaikuni
Department of Computer Science & Engineering
Walchand Institute of Technology
Solapur, India

Dr. Sachin R Gengaje
Department of Electronics Engineering
Walchand Institute of Technology
Solapur, India

Abstract: Bayesian networks have been at the core of the research pertaining to probabilistic reasoning. Several machine learning algorithms and techniques highly rely on Bayesian networks for their reasoning capabilities. Since several decades Bayesian networks proved to be the essential tool in the hands of researchers working in artificial intelligence domain. Yet Bayesian networks do have certain limitations [1], Bayesian networks need extensions to be more expressive and functional for probabilistic reasoning needs of various domains. Multi Entity Bayesian Networks (MEBN) is a theory combining expressivity of first-order logic principles and probabilistic reasoning of Bayesian networks. The paper thoroughly introduces the MEBN theory with an example modelling for a problem description and discusses various other works which included MEBN theory as a core of their research. The study in this paper concludes with the current highlights and challenges in MEBN and future developments.

Keywords: Bayesian Networks, MEBN, Probabilistic Reasoning.

1. INTRODUCTION

In this paper, we present a detailed study of MEBN (Multi Entity Bayesian Networks) [2] and its applications in several fields. The study carried out presents in-depth review of methods by which MEBN is being actively used in the scenarios which demand first-order logic probabilistic reasoning ability. Further, the study explores the applications of MEBN in inferring an outcome for the application domain. The concluding aspect of the study focuses on identifying the further improvements and extensions needed for probabilistic reasoning systems based on MEBN.

2. MULTI ENTITY BAYESIAN NETWORKS

In the last few decades, numerous approaches have been proposed which intend to integrate first-order representations with the probabilistic semantic graphical models which have been largely a sub discipline of artificial intelligence being referred as statistical relational learning [3]. MEBN is an extension to Bayesian network and first order logic, providing first order logic's expressive power. Similar to Bayesian networks MEBN rely on graphical representation in the form of directed acyclic graph of random variables.

The novelty in MEBN lies in its ability to represent the internal attributes of random variables and relation between such random variables, also MEBN specifies a first-order language for modeling probabilistic knowledge bases as parameterized fragments of Bayesian networks. MEBN are most expressive when it comes to first-order logic and probabilistic representation of knowledge bases [4]. The relation between groups of related random variables is expressed in terms of probability distribution information represented by MEBN fragments (MFrags).

An MFrag in MEBN consists of three types of random variables which are resident random variables, context random variables, and input random variables. Resident and input random variables represent the relationship between entities and their properties, Context random variables define the context under which probability distribution across an MFrag is valid. In short MFrags are key elements involved in capturing and representing Knowledge Bases in MEBN theory. Graphically all types of random variables are represented as nodes and arcs between nodes represent a dependency, an input random variable is represented as trapezoid node while the resident random variable is represented as rounded square node and context random variable is represented as a distorted Pentagon, as an example refer to figures 1,2 & 3. Formally an MFrag F is defined as

$$F = (C, I, R, G, D) \text{ where}$$

- *C is a finite set of values a context can take form as a value.*
- *I is a set of input random variables*
- *R is a finite set of resident random variables*
- *G is a directed acyclic graph representing the dependency between input random variables and resident random variables conditional on context random variables in one to one correspondence.*
- *D is a set of local conditional probability distributions where each member of R has its own conditional probability distribution in set D.*
- *Sets C, I, and R are pairwise disjoint.*

Given an MEBN theory for a specific problem to be solved under a given situation context, a bottom-up construction algorithm works on MFrags of MEBN theory and produces a situation specific Bayesian network (SSBN). The produced SSBN for a query on MEBN as a result of bottom-up construction algorithm is a Bayesian network, generated as a response to the query. Further inferencing of generated Bayesian network estimates the posterior probability for the query given to the MEBN theory.

The MEBN theory (MTheory) lets a user building ontology for a specific domain to judge a context within that domain in more realistic fashion by introducing uncertainty in the ontology. MTheory models classes, object properties and relationships in ontology to a set of MFrags. PR-OWL [5] and PR-OWL 2 [6] are the languages using MTheory to define probabilistic ontologies and generate inferences based on ontologies defined. PR- OWL and PR-OWL 2 add new elements to the OWL standard which lets the creation of MEBNs as part of an ontology modeling process and thus supporting probabilistic reasoning in ontologies.
Let us analyze a problem and construct MTheory. Consider a problem of a patient visiting a clinic with a fever and if it happens that the patient recently visited a flu epidemic affected region it is very likely that the patient's diagnosis may indicate he has a fever due to the flu. Otherwise, the patient may be having a fever due to some other reason.
In order to model the patient disease diagnosis problem into MTheory the first step is to construct ontology for the problem. The ontology for the patient diagnosis problem described above will have following classes, object properties and relationships in OWL syntax and semantics. It is important to note that in OWL object properties relate an object (individual) to object (individual) and data type properties relate objects (individual) to data values.

Classes:
- Patient
- Region
- Diagnosis

Properties:
- hasDiagnosis (Patient p)
    - Domain: Patient
    - Range: Diagnosis
- hasFluEpidemicPresent (Region r) | Range data type: Boolean
    - Domain: Region
    - Range: Boolean
- hasVisitedRegion (Patient p, Region r) | Range data type: Boolean
    - Domain: Patient X Region
    - Range: Boolean

Given, an ontology, constructing an MTheory requires defining a set of MFrags which capture the problem description in a Bayesian network of resident and input type of random variables conditional on context random variables. The MFrags constructed for the ontology defined for diagnosis of a patient will be as shown in figure 1, 2 & 3

MEBN are queried using statements similar to first order logic statements. Every query on MEBN involves the construction of situation specific Bayesian network (SSBN) from the set of MFrags belonging to the concerned MTheory. SSBN construction is an iterative bottom-up process on the MFrags having priori knowledge base of priori probability distribution across the resident random variables. The bottom up SSBN construction algorithm involves d separation and inferring an intermediate Bayesian network obtained from the set of Resident random variables for every iteration until the iterations are terminated. The final Bayesian network inference on SSBN produces the posterior probability outcome for the query. An interesting fact regarding MFrags is they can be instantiated several times for a specific use case of entity.

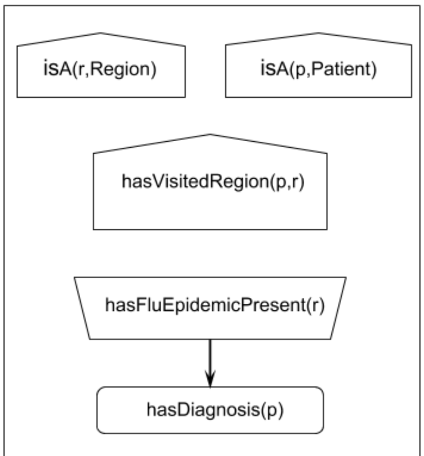

Fig 1: MFrag for class Patient

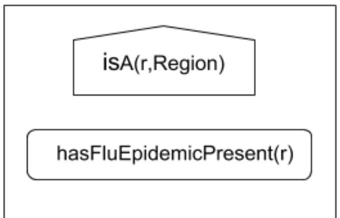

Fig 2: MFrag for class Region

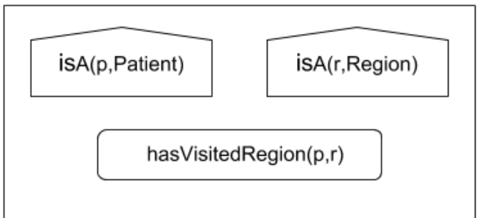

Fig 3: MFrag for a patient 'p' visiting a region 'r'

3. APPLICATIONS OF MEBN

The prime use of MEBN is to introduce uncertainty in ontologies in the form of probabilistic ontologies. A probabilistic ontology has several applications ranging from realistic modeling of a domain to building sophisticated intelligent systems based on semantic knowledge of the domains.
Soon MEBN has been quickly embraced by the researchers to model ontologies which would assist the reasoner built into the intelligent systems. The study puts forth some of the MEBN applications attempted by the researcher.
For user behaviour modeling the researchers modeled an entire user behavior and activity pattern into MEBN. The process of modeling MEBN for user's behavior was a spiral model like process which involved several iterations, experiments, and improvements over a period of time. High reliability and low rate of false alarms were achieved by the researchers in using MEBN to model user behavior analysis and detection in the context of internet security through simulated user actions. Researchers do point in their research that additional inclusion of several of the affecting factors and object properties might improve the reliability and robustness of their system [7].

To cater the need of knowledge representation for autonomous robots which are steadily taking places of humans in trivial tasks. The researchers highlighted the importance of probabilistic reasoning in the field of robotics and stressed for first-order probabilistic reasoning. A methodology BLN (Bayesian Logic Networks) similar to MEBN was proposed. MEBN was primarily used as a knowledge representation system for encapsulating uncertainty in the knowledge and as a first-order probabilistic reasoner for guiding robot's future action plans [8].

[12] The researcher attempted a hybrid approach of merging ontologies. The hybrid approach consisted use of MEBN and PROWL as a representing frameworks to represent uncertainty in ontologies. In devising their methodology for merging the ontologies consideration of temporal occurrence of events and concepts in ontologies was given the main focus.

[13] The preliminary work proposed a strategy using multivariate analysis to identify the classes for most active entities in a problem domain. The author suggests a sort of preprocessing and pre MEBN modeling tasks.

[14] Proposed a method named UMP-ST (Uncertainty Modeling Process for Semantic Technologies) which would assist general practitioners of MEBN in rightly modeling MEBN for their research problems concerned with (URP-ST) Uncertainty Reasoning Process for Semantic Technologies. UMP-ST is an iterative stepwise process which ensures the model keeps on refining and optimizing on each iteration of building MEBN for knowledge bases of any domain [15].

[16] The work attempted to detect network intrusions by modeling network intrusion domain knowledge into MEBN. The modeled MEBN had a better performance in terms of classification accuracy. The researcher incorporated a preprocessing step which involved discretization of continuous variables in the problem domain, thereby assisting in correct modeling of MEBN with entities responsible for network intrusion attempts. A similar attempt of modeling bandwidth depletion attacks by [17] achieved success in classifying the threats based on MEBN inferencing.

[18][19] Used MEBN in assisting military decisions of attacking, based on military strategies and associated knowledge bases. The researchers' work points out that MEBN assisted decision of attacking strategy would reap better outcomes than decision-based on intuitions of humans. The work highlights possible use of MEBN in crucial defense strategy decisions.

[20] Proposed a high level fuzzy Bayesian network based information fusion approach for merging knowledge representations based on MEBN. The work primarily focuses on presence and handling of inherent ambiguity and casualty aspect of knowledge representations. The proposed high-level approach, when implemented in collision warning systems, performed promisingly well in handling ambiguity [21][22]

[23] The contributed to state of art developments in MEBN by proposing MEBN learning methods in the context of predictive situation awareness environments. The proposed MEBN learning method is a hybrid method involving learning of discrete and continuous variables. Further by putting the method to test the researcher's earlier system PROGNOS, they concluded that MEBN learning methods are still in nascent stage and can not match the human way of modeling of MEBN.

MEBN has been used in Medical diagnosis and there are specialized tools developed to assist in modeling the MEBN correctly with inputs from experts in the concerned field [24] Further studies identify the benefits of using MEBN in medical diagnosis cases. Building an MEBN could prove to be greatly helpful in guiding future physicians and surgeons and patients [25]. [26] Successfully employed MEBN for identifying contract parameters breaches in Cloud Service providers SLA.

[27] Used MEBN to model intangible cultural heritage content, namely traditional dance and further demonstrated that in the cases which demand situation specific treatment with respect to the domain knowledge MEBN outperforms Bayesian networks.

SATCOM systems require frequent and sometimes rapid reconfigurations of systems to support dynamic user demands. The decision process usually involves a potentially wide range of competing factors and high degrees of uncertainty associated with those highly-dynamic and complex environments. [28][29] Built a functional causal modeling of SATCOM systems modeled on the basis of MEBN to assist in balanced decision making in highly dynamic environment.

Recent work [30] proposes a slightly different bottom up SSBN construction algorithm in MEBN to address the concerns of scalability when MEBN has large MFrags and resident nodes. The difference lies in using Bayes-Ball algorithm eliminating need to prune d-separated nodes thereby simplify the process of generating SSBN. Preliminary tests and experiments point out clear performance improvements in terms of time taken in generating SSBN by MEBN tool UnBBayes [31].

A very recent work which made a very promising use of MEBN [32] attempted to solve the challenge of heterogeneous underwater robots with a common understanding of information exchanged between robots. A special SWARM ontology modeled on basis of European SWARMs project effectively managed to inference the inherent uncertainty in harsh under water and maritime operations environment using PR-OWL.

4. CONCLUSION

It is quite evident after going through the work of researchers who used MEBN in the context of probabilistic reasoning that the combined first-order expressivity and Bayesian inferencing within MEBN prove to be very useful and promising for varied nature of solutions to the problems. Following are the highlights of MEBN based solution approaches adopted by the researchers.

1. MEBN has been Immensely useful in building probabilistic knowledge bases (ontologies)
2. MEBN enables knowledge fusion across knowledge bases through ontologies
3. MEBN's compatibility with OWL makes it firsthand tool to work with probabilistic ontologies
4. MEBN has been to date the most expressive first-order logic probabilistic reasoning method.

5. MEBN has been used meticulously to overcome several of limitations of Bayesian networks.

Given the highlights of MEBN Theory, following are the challenges reported and observed by the researchers who employed MEBN theory in their probabilistic reasoning methods.

1. Scalability is still an issue in MEBN
2. Though there has been a standard uncertainty modeling process in semantic technologies (UMP-ST) the process of building MEBN is still by the large a manual process and requires human inputs in modeling and building a reasonably accurate and reliable probabilistic ontologies using MEBN.
3. Absence of standard approach and methods to let MEBN's MFrags to dynamically evolve and optimize as the ontologies and knowledge bases keep updating
4. No comprehensive, reliable and widely recognized method available for learning MEBN from knowledge bases of the domains.

Researchers are steadily working in making MEBN a mainstream first-order probabilistic reasoning method. MEBN is gaining wide acceptance among various researchers and being adopted as first-order probabilistic reasoning tools in sub disciplines of artificial intelligence in an attempt to provide solutions to real world artificial intelligence applications.

5. FUTURE DIRECTIONS

The following are the immediate needs for MEBN to be adopted as a mainstream first-order probabilistic reasoning tools

1. Development of reasonably accurate and reliable MEBN learning methods.
2. Improvements and newer MEBN inference methods have to be developed to cater to the need of scalability in supporting real world's vast and huge domain knowledge bases.
3. Integration methods for MEBN reasoning in Deep Learning environments to assist deep learning methods in grasping the semantic information in domain knowledge.

REFERENCES

[1] Wang, Pei. "The limitation of Bayesianism." Artificial Intelligence 158.1 (2004): 97-106.

[2] K. B. Laskey, "MEBN: A language for first-order Bayesian knowledge bases," Artif. Intell., vol. 172, no. 2–3, pp. 140–178, 2008.

[3] L. Getoor and B. Taskar, Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning). The MIT Press, 2007.

[4] Howard, Catherine, and Markus Stumptner. "A Survey of Directed Entity-Relation--Based First-Order Probabilistic Languages." ACM Computing Surveys (CSUR) 47.1 (2014): 4

[5] Da Costa, Paulo Cesar G., Kathryn B. Laskey, and Kenneth J. Laskey. "Pr-owl: A bayesian ontology language for the semantic web." Proceedings of the 2005 International Conference on Uncertainty Reasoning for the Semantic Web-Volume 173. CEUR-WS. org, 2005.

[6] Carvalho, Rommel N., Kathryn B. Laskey, and Paulo CG Costa. "PR-OWL 2.0-bridging the gap to OWL semantics." Proceedings of the 6th International Conference on Uncertainty Reasoning for the Semantic Web-Volume 654. CEUR-WS. org, 2010.

[7] G. A. AlGhamdi, K. B. Laskey, E. J. Wright, D. Barbara, and K. Chang, "Modeling insider user behavior using multi-entity Bayesian network," in 10th International Command and Control Research and Technology Symposium, 2008, vol. 4444, no. 703.

[8] D. Jain, L. Mosenlechner, and M. Beetz, "Equipping robot control programs with first-order probabilistic reasoning capabilities," 2009 IEEE Int. Conf. Robot. Autom., pp. 3626–3631, 2009.

[9] R. N. Carvalho, P. C. G. Costa, K. B. Laskey, and K. C. Chang, "PROGNOS: Predictive situational awareness with probabilistic ontologies," 2010 13th Int. Conf. Inf. Fusion, pp. 1–8, 2010.

[10] R. N. Carvalho, R. Haberlin, P. C. G. Costa, K. B. Laskey, and K. C. Chang, "Modeling a probabilistic ontology for Maritime Domain Awareness," 14th Int. Conf. Inf. Fusion, pp. 1–8, 2011.

[11] K. B. Laskey, P. C. G. Costa, and T. Janssen, "Probabilistic ontologies for multi-INT fusion," Front. Artif. Intell. Appl., vol. 213, pp. 147–161, 2010.

[12] M. D. Mas, "Ontology Temporal Evolution for Multi-Entity Bayesian Networks under Exogenous and Endogenous Semantic Updating," CoRR, vol. abs/1009.2, pp. 1–20, 2010.

[13] H. Bouhamed, A. Rebai, T. Lecroq, and M. Jaoua, "Data-organization before Learning Multi-Entity Bayesian Networks Structure," pp. 305–308, 2011.

[14] R. N. Carvalho, M. Ladeira, and L. Weigang, "Probabilistic Ontologies Incremental Modeling Using UnBBayes," 2011.

[15] R. N. Carvalho, K. B. Laskey, and P. C. G. Da Costa, "Uncertainty modeling process for semantic technology," PeerJ Comput. Sci., vol. 2, p. e77, 2016.

[16] C. Ewell, "Detection of Deviations From Authorized Network Activity Using Dynamic Bayesian Networks," 2011.

[17] A. Boruah, "A Probabilistic Approach to Detect and Prevent Bandwidth Depletion Attacks," Int. J. Comput. Appl., vol. 150, no. 5, pp. 42–49, 2016.

[18] Z. Yun, Z. Cheng, L. Ting, Z. Weiming, and L. Zhong, "A COG Analysis Model of System-of-Systems (SoS) Based on Multi-Entity Bayesian Networks (MEBN)," 2012.

[19] Wang, Hao-Ran, et al. "Tactical Air Target Intention Recognition Based on Multi-Entities Bayesian Network." Huoli yu Zhihui Kongzhi 37.10 (2012): 133-138.

[20] K. Golestan, F. Karray, and M. Kamel, "High level information fusion through a fuzzy extension to multi-entity bayesian networks in vehicular ad-hoc networks," IEEE 16th Int. Conf. Inf. Fusion, pp. 1180–1187, 2013.

[21] K. Golestan, R. Soua, F. Karray, and M. S. Kamel, "A model for situation and threat/impact assessment in vehicular ad-hoc networks," Proc. fourth ACM Int. Symp. Dev. Anal. Intell. Veh. networks Appl. - DIVANet '14, pp. 87–94, 2014.

[22] K. Golestan, B. Khaleghi, F. Karray, and M. S. Kamel, "Attention Assist: A High-Level Information Fusion Framework for Situation and Threat Assessment in Vehicular Ad Hoc Networks," IEEE Trans. Intell. Transp. Syst., vol. 17, no. 5, pp. 1271–1285, 2016.

[23] C. Y. Park, K. B. Laskey, P. Costa, and S. Matsumoto, "Multi-Entity Bayesian Networks Learning In Predictive Situtation Awareness," Proc. 18th …, vol. 4444, no. 703, pp. 1–20, 2013.

[24] Cypko, M., et al. "User interaction with MEBNs for large patient-specific treatment decision models with an example for laryngeal cancer." Int J CARS 9.Suppl 1 (2014).

[25] Sarkar, I. N. "Web-tool to Support Medical Experts in Probabilistic Modelling Using Large Bayesian Networks With an Example of Hinosinusitis." (2015).

[26] O. Jules, A. Hafid, and M. A. Serhani, “Bayesian network, and probabilistic ontology driven trust model for SLA management of Cloud services,” 2014 IEEE 3rd Int. Conf. Cloud Networking, CloudNet 2014, pp. 77–83, 2014.

[27] G. Chantas, A. Kitsikidis, S. Nikolopoulos, K. Dimitropoulos, S. Douka, I. Kompatsiaris, and N. Grammalidis, “Multi-entity Bayesian networks for knowledge-driven analysis of ICH content,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8926, pp. 355–369, 2015.

[28] T. Martin, K. C. Chang, X. Tian, G. Chen, T. Nguyen, K. D. Pham, and E. Blasch, “A probabilistic situational awareness and reasoning methodology for satellite communications resource management,” IEEE Aerosp. Conf. Proc., vol. 2015-June, pp. 1–12, 2015.

[29] X. Tian, G. C. Ift, T. Martin, K. C. C. Gmu, T. N. Cua, K. Pham, and E. B. Afrl, “Multi-entity Bayesian network for the handling of uncertainties in SATCOM Reconfiguring SATCOM Resources,” vol. 9469, pp. 1–11, 2015.

[30] L. L. Santos, R. N. Carvalho, M. Ladeira, and L. Weigang, “A new algorithm for generating situation-specific Bayesian networks using Bayes-Ball method,” CEUR Workshop Proc., vol. 1665, pp. 36–48, 2016.

[31] https://sourceforge.net/projects/unbbayes/ (accessed 12th June 2017)

[32] Li, Xin et al. “SWARMs Ontology: A Common Information Model for the Cooperation of Underwater Robots.” Sensors (2017).

