Issues and Requirements for Bayesian Approaches in Context Aware Systems

Michael Angermann, Patrick Robertson

Institute of Communications and Navigation,
German Aerospace Center,
D-82234 Wessling/Oberpfaffenhofen,
Germany,
`firstname.lastname`@dlr.de

Abstract. Research in advanced context-aware systems has clearly shown a need to capture the inherent uncertainty in the physical world, especially in human behavior. Modelling approaches that employ the concept of probability, especially in combination with Bayesian methods, are promising candidates to solve the pending problems. This paper analyzes the requirements for such models in order to enable user-friendly, adaptive and especially scalable operation of context-aware systems. It is conjectured that a successful system may not only use Bayesian techniques to infer probabilities from known probability tables but learn, i.e. estimate the probabilities in these tables by observing user behavior.

1 Introduction

Context and context-awareness has become a major field of research in recent years. One of the reasons is that context-awareness is believed to be a promising solution for a couple of problems arising in pervasive computing scenarios [9, 13–15]. A well designed context model in conjunction with a framework that facilitates context management tasks such as collection, refinement, persistency, distribution, synchronization, provisioning and reasoning based on the model is a prerequisite for the fruitful use of context in any context-aware system [12].

Deriving contextual information about the system's entities, such as the user, the devices, the network or the environment is an essential, yet challenging task. This holds in particular if the contextual information is the outcome of a context refinement process (e.g. sensor fusion) and does not only reflect changes on primary aspects such as location, time, activity or identity.

A major issue in most context-aware systems is how to deal with uncertain, incomplete or ambiguous data. Incomplete contextual information can for instance be the result of temporarily or permanently absent sensors (e.g. due to energy deficiencies), whereas ambiguous context information can for instance be the result of two or more sensors observing the same entity but getting different values, such as two thermometers in the same room showing different temperatures (e.g. 19.1°C and 22.7°C). Both sensors add information to the knowledge about the context. However, since obviously their indicated values are subject to some

form of error, some uncertainty about the "true" temperature remains. To care about uncertainty and ambiguity must not be neglected in any serious context-aware system architecture. Many context modelling and retrieval architectures tend to over-simplify this uncertainty by assuming a perfect knowledge in combination with perfect inference. Only few approaches cover uncertainty, e.g. by using heuristic approaches such as attaching quality/validity meta-information to contextual information. Among them are the Extended ORM model [5], the ASC model [13] or the Cues [10].

Over the last 15 years, Bayesian networks have evolved as a major tool in a wide area of scientific disciplines requiring sound statistical analysis, automated reasoning or exploitation of knowledge hidden in noisy data [8]. These range from fields in medical research, genetics, insurance analysis, and fault handling to automation and intelligent user interaction systems. Bayesian networks (BN) combine techniques from graphical models with those from Bayesian analysis to provide a formal framework within which complex systems can be represented and analyzed. A BN encompasses a set of random variables that represent the domain of interest and the BN encodes many of the important relationships between these variables, such as causality and statistical dependence and independence. Specifically, their structure says something about the qualitative nature of these relationships whereas their network parameters encode the probabilistic relationships among the variables of interest. Past areas of work have focused on inference in BN [6] and estimation of network structure and parameters given real observed data (see e.g. [11, 2]). Current work is addressing numerous new areas of applications of BN, as well as special network structures such as dynamic Bayesian networks [7].

In this paper, we want to elaborate on the requirements for a probabilistic approach to cover uncertainty in context. Probabilistic approaches for context-modelling naturally lend themselves to a fruitful combination with Bayesian methods, in particular with Bayesian networks. In conjunction with other statistical techniques they have several advantages for data analysis. Following [3] this includes

- A Bayesian network readily handles situations where some data entries are missing.

- Given observed data, a Bayesian network can be used to learn causal relationships, and hence can be used to gain understanding about a problem domain and to predict the consequences of intervention.

- Because a Bayesian network has both a causal and probabilistic semantic, it is an ideal representation for combining prior knowledge and data.

- Bayesian statistical methods in conjunction with Bayesian networks avoid the overfitting of data and do not need to separate data into training and testing sets. They are also able to incorporate smoothly the addition of new data as it becomes available.

It is obvious, that these characteristics of Bayesian networks should qualify them perfectly to model uncertainty in a context-aware system. This is the rationale for working out the requirements and a generic architecture for including Bayesian networks into context-aware systems. We have identified four main issues for this Bayesian approach:

1. Infer contextual knowledge given a network and associated probabilities

2. Estimate network probabilities given observations

3. Deduce network structure given observations

4. Combine (few) data that is valid for a single individual with (much more) data valid for a larger group of users

This paper is organized as follows. In section 2 we discuss several issues to be considered when applying Bayesian information processing for context-aware systems. In particular we show how they result in system requirements. In section 3 we investigate the combination of knowledge derived from observing individual behavior with group behavior. This combination is considered to be essential for providing instant high perceived system performance for users coming freshly into the system.

2 Issues in Applying Bayesian Information Processing to Context-Aware Systems

The following list highlights in more detail the areas in context processing that should involve Bayesian techniques and the issues that arise when doing so.

1. **Human 'domain-expert' modelling of Bayesian Networks' structure and parameters.** In the Bayesian research community, it is well established that domain experts' knowledge can be valuable, if not essential contributions to the Bayesian process, mainly in specifying the structure of the underlying network. Therefore, a Bayesian context-aware system must provide an interface to such domain experts. The experts' contributions can be in the form of

   – identifying and configuring context variables suitable to the domain, this includes choice of suitable quantization levels and mapping to context ontologies, and choice of the total number of variables to be incorporated into the later network,

   – encoding causal relationships in a prior network [3],

   – defining priors on network structure,

   – using an initial network and defining equivalent sample size,

   – confirmation of learned structures, conditional probability distributions and a-priori probabilities,

– obtaining, and making explicit, new insights into the domain's features by analyzing learned structures and probability distributions.

In a limited sense, it is conceivable that the actual end user could be treated as an additional domain expert; at least he or she is the person who often has the best understanding of the personal domain. In traditional context aware systems the user is incorporated by allowing preferences or rules to be established that govern the system behavior. In a context system that uses Bayesian techniques, 'such preferences' could be enriched by letting the user state certain facts about his or her domain and translating these to network and parameter priors. One may even go as far as suggesting a hierarchy of domain experts in which the end user may play a certain role, in addition to, for instance, company representatives that govern the configuration of systems used in a corporate environment such as services for sales staff. Several, secondary requirements pertaining to integrity, quality control, privacy of user data and security arise from this potential involvement of others into the representation of a user's context.

2. **Automated learning of BN structure using complete or incomplete data** [2, 1]. By incomplete we mean the use of observed data sets where some variables' values are unknown, which will presumably be the case for many context systems where sensor data or other observations can be missing. This learning may apply to subsets of variables (i.e. with the goal of constructing a sub network) or the full set, depending on the domain complexity. It includes the identification of new hidden nodes to simplify the network structure and more accurately represent the domain at hand [16]. The result of automatic learning may be iteratively used to reinforce or expand the modelling done by hand in the first step.

3. **Automated learning of BN parameters** (a priori probabilities, conditional probabilities), using complete and incomplete data sets. Again, there may be an iteration to the first and second stages above. In a context systems application, such learning should ideally be a continuous process as new data is gathered.

4. **Inference given observations and updates to observations.** Depending on the network size and structure, suitable algorithms must be chosen, for instance belief propagation for tree structures [8], loopy belief propagation (see [17]) or join tree algorithms. This will may also apply optional consistency checks (non Bayesian) that the domain experts may define on individual variables or groups of variables (e.g. temporal constraints on location changes or checking whether certain combinations of variables' values simply cannot be the case).

5. **Utility theoretic decision making.** It is clear that in order to reduce the burden on developers or providers of the actual end-user context aware services, and also to improve efficiency, such services will not only exploit

context variables' estimated values or their likelihoods, but would like to delegate to the Bayesian processing layer certain decisions based on cost functions and the estimated likelihoods of context variables. An example is a service that wants to influence the choice of a communications network used to transport its application data, based on certain context conditions.

6. **Value-Of-Information evaluation** to gain information on
   - which sensors' readings would be of value
   - which additional (possibly non-Bayesian and more or less costly) context inference would be of value
   - which remote context elements should be accessed in a federated system, possibly over a costly network

   Such methods must minimize global costs and take into account the aggregate benefit (or cost) of acquiring the new context data (or of not obtaining it) and the costs that are incurred to acquire the context. Suitable interfaces to services are again needed that indicate the value to the service of knowing values of context variables - possibly to a certain accuracy, and also interfaces to a context broker that allow it to indicate costs of determining certain context values - including sensor values and items of only remotely available context. These costs will, of course, depend on the state (including observational status) of other variables in the BN.

7. **Combine (a smaller amount of) data that is valid for a single individual with (much more) data collected from observing a larger group of users.** This difficult task will be addressed in the next section.

8. **Enforce privacy restrictions on users' personal data.** The exploitation of context information is always a balance between users' privacy and benefits gleaned from exploiting personal data. In a system where such information is being gathered not only for the benefit of an isolated user in his isolated BN, but also for many other users, and furthermore may require the inspection of human domain experts, the privacy issue become more pressing. The architecture, design and implementation of all components will need to ensure that this problem is being addressed satisfactorily. These considerations will also have implications on point 6, where, for example, pieces of context pertaining to location may only be available in raw form on a local device and the user would not like them to be passed continuously to a central server.

9. **Scalability and performance issues.** This problem is related to the last three in the sense that it is expected that each active user will own a particular instantiation of a personal BN and that context information may not always be available centrally. With potentially many millions of active users, and the period of an individual's 'activity' in pervasive computing environments being extended to almost 24 hours a day, hosting many such networks

will pose memory and computational challenges. It may become necessary to host various parts of a BN on different computing nodes, and take into consideration the topographic availability of context. For instance, a person may carry location sensors and continuous context transfer to a central server might be too ex-pensive or undesired, e.g. for privacy concerns; in this case, local pre-processing on a local sub network could reduce network costs and/or preserve privacy.

An object oriented design approach with proper encapsulation, providing separation of above functions from their actual implementation, may possibly allow future developments in this rapidly evolving field to be incorporated more easily.

For context-aware systems to become key components of commercially viable services, several interfaces and functionalities are necessary to enable a complete *Bayesian context processing "chain"*, which would be the result of implementing the all major components related to steps 1-7 above.

3 Deriving Individual User Behavior from Group Behavior

Among the most attractive features of the Bayesian approach is its capability to use data, gained from observing a large user community (*collective*), as information about an individual, whose behavior may even not have been observed at all. This is facilitated by deriving suitable a priori distributions ("priors") for the parameters of the individual's Bayesian network. In the domain of context aware applications this seems particularly important since the potential scope of observations of each particular individual is much smaller than the observation of many users and in fact, a system begins with no observations for a new user. Nevertheless, as observations accrue, they should be incorporated fairly with the prior knowledge obtained from the collective data.

An important issue is the identification of suitable clusters to distinguish certain features of users. These clusters effectively define sets of users, e.g. "all male users", "all users of age below 18" years, "all users spending more then 100$ per month on mobile phone charges". Suitable a priori distributions for the individual are then computed by selecting the sets he or she fits into and adding up the probability distributions of all of each set's members and multiplying the sums obtained for all fitting sets.

The influence of these priors gradually declines with increasing availability of observations upon the individual. Fig. 1 illustrates the necessary steps from including domain experts' knowledge to the derivation of information about individual user behavior.

1. Define priors on structure or prior collective BN or an initial collective BN
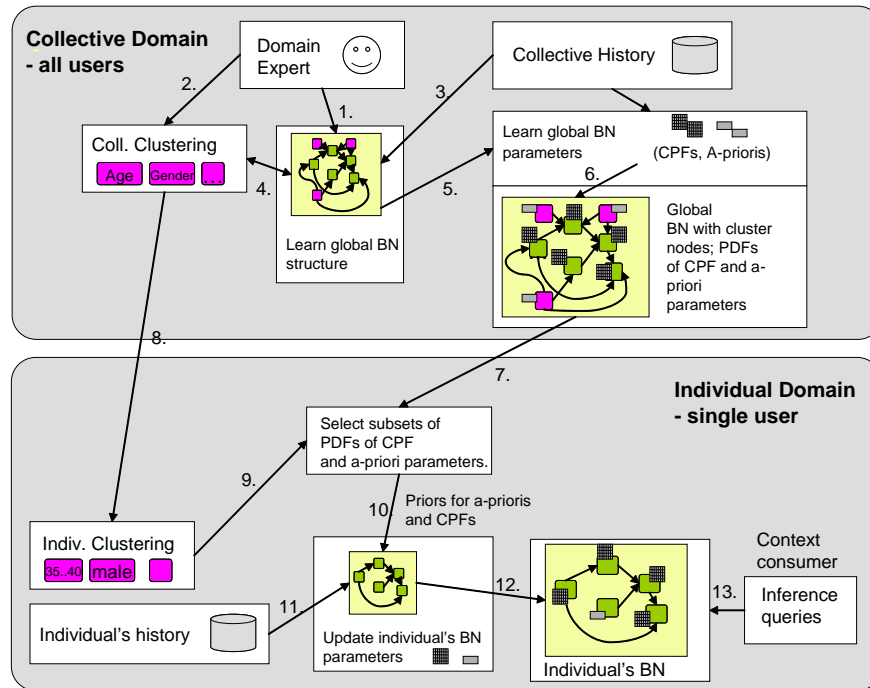
2. Define cluster variables

Fig. 1. Necessary steps from including domain experts' knowledge to the derivation of information about individual user behavior

3. Update collective BN using collective data

4. Update clusters if this is beneficial

5. Update collective BNs parameters using collective data

6. Provide a valid collective BN as a result of steps 1-5

7. Clone this collective BN to become the individual's BN

8. Determine cluster variables' values for the individual (e.g. from individual's preferences or subscription data)

9. Using the clusters, identify in the collective BN which variables these cluster choices have effects upon

10. Using 7 and 9, compute priors of the parameters of conditional probability functions (CPF) and a-priori probabilities, respectively; for instance assuming an unrestricted multinomial distribution model [3]

11. Using the individual's personal data history continually update the parameters of her BNs CPFs and a-priori probabilities

12. Construct the individual's BN using estimates of the individual's BN parameters based on step 11

13. Perform inference and/or decisions based on the individual's BN

We can see that the actual transfer of knowledge obtained from observations of collective user behavior to individual user behavior is performed in steps 7 and 8, when the collective BN is cloned to become the individual's BN and the values of the individual's cluster variables are determined.

4   Conclusions

In this short paper we have addressed several issues we consider relevant for the development of future context-aware systems that become capable of a) properly operating on the ambiguous/uncertain context of real world applications and b) scaling towards potentially very large numbers of users (e.g. several millions, such as in today's mobile phone networks).

We suggest that domain experts need user interfaces to set up and continuously maintain the structure and eventually certain a priori probabilities of the Bayesian networks. It is proposed that implementors of services need clearly specified and standardized interfaces not only to query for context values, but also for obtaining such (meta-)information as the "value-of-information", probabilities of context values or decision-theoretic computation of utilities. Interfaces to low-level sensors appear necessary which, again, do not only convey the actual values but also meta-information such as error distributions or a priori distributions. Proper algorithms for continuous automatic learning of the BNs' structure and parameters as well as inference of higher-level context have to be in place and well encapsulated inside of the context-subsystems. The transfer from observation made on the collective of users to individuals is crucial to create a favorable user experience from the beginning. In order to protect privacy it appears desirable to keep the observations that can be directly ascribed to an individual under the individual's direct control, e.g. on the personal device. However, despite this and and other efforts for privacy protection it may not be possible to develop an "inherently safe" context-aware system that can absolutely prevent the misuse of context information for activities that may be considered as privacy violating or discriminating by the individual user. A Bayesian perspective on the privacy protection problem and the possibilities to infer context that is considered private from context that is considered public, reveals that misuse can hardly be prevented by technological measures, alone. Instead, a combination of careful software and system design efforts and sensible legal frameworks may be necessary to break the ground for wide acceptance and successful operation of large-scale context-aware systems and services.

References

1. Chickering, D., Heckerman, D.: Efficient Approximations for the Marginal Likelihood of Bayesian Networks with Hidden Variables, Machine Learning, (1997), Vol. 29, pp. 181-212.
2. Cooper, G., Herskovits, E.: A Bayesian Method for the Induction of Probabilistic Networks from Data, Machine Learning, (1992), Vol. 9, pp. 309-347.
3. Heckerman, D.: A Tutorial on Learning with Bayesian Networks. Technical Report MSR-TR-95-06, Microsoft Research, Redmond, Washington, 1995. Revised June 96
4. Heckerman, D., Mamdani, A., Wellman, M.: Real World Applications of Bayesian Networks, Communications of the ACM, (1995), Vol. 38.
5. Henricksen, K., and Indulska, J. Modelling and Using Imperfect Context Information. In WorkshopProceedings of the 2nd IEEE Conference on Pervasive Computing and Communications (PerCom2004), (Orlando, FL, USA, March 2004), pp. 33-37.
6. Lauritzen, S., Spiegelhalter, D. J.: Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. Journal of the Royal Statistical Society Series B 50:157-224 (1988).
7. Murphy, K.: Dynamic Bayesian Networks: Representation, Inference and Learning. Ph.D. Dissertation, UC Berkeley, Computer Science Division, 2002.
8. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo, CA (1988).
9. Satyanarayanan, M.: Pervasive Computing: Vision and Challenges. IEEE Personal Communications Magazine. Vol. 8. No. 4. (2001) 10-17
10. Schmidt, A., Beigl, M., and Gellersen, H.-W. There is more to Context than Location. Computers and Graphics 23, 6 (1999), 893-901.
11. Spiegelhalter, D., and Lauritzen, S.: Sequential Updating of Conditional Probabilities on Directed Graphical Structures, Networks, Vol. 20, (1990), pp. 579-605.
12. Strang, T., Linnhoff-Popien C.: A Context Modeling Survey. Workshop on Advanced Context Modelling, Reasoning and Management as part of UbiComp 2004 - The Sixth International Conference on Ubiquitous Computing, Nottingham/England (2004)
13. Strang, T.: Service-Interoperability in Ubiquitous Computing Environments. PhD thesis, LMU Munich, October 2003
14. Want, R., Hopper, A., Falcao, V., Gibbons, J.: The Active Badge Location System. ACM Transactions on Information Systems. Vol. 10. No 2. (1992) 91-102
15. Xynogalas, S., Chantzara, M., Sygkouna, I., Vrontis, S., Roussaki, I., Anagnostou, M.: Context Management for the Provision of Adaptive Services to Roaming Users, IEEE Wireless Communications. Vol. 11. No 2. (2004) 40-47
16. Boyen, X. , Friedman, N.,Koller, D.: Learning the structure of complex dynamic systems,UAI, 1999.
17. Murphy, K.P., Weiss, Y., Jordan, M: Loopy belief propagation for approximate inference: an empirical study, in Proceedings of Uncertainty in AI, 1999.