An Analysis of Unreliability of Competence Information in Learning Networks and the First Exploration of a Possible Technical Solution

An Analysis of Unreliability of Competence Information in Learning Networks and the First Exploration of a Possible Technical Solution

Yongwu Miao, Peter Sloep, Hans Hummel, and Rob Koper
Centre for learning Sciences and Technologies,
Open University of the Netherlands
E-mail: (yongwu.miao, peter.sloep, hans.hummel, rob.koper)@ou.nl

Abstract: Automated competence tracking and management is crucial for an effective and efficient life-long competence development in learning networks. However, currently there is no systematic method to represent, measure, and interpret competence. In this paper, we analyze the problem of unreliability of competence information in learning networks. In tracking the development of competences in learning networks, a large amount of competence information can be gathered from diverse sources and diverse types of sources, which is subject to uncertainty and unreliable. This paper investigates information fusion technologies that may be applied to address the problem and that show promise as candidate solutions for achieving an improved estimate of competences by fusing (possibly inconsistent) information coming from multiple sources. This paper is intended to motivate educational technology researchers to learn more about information fusion, to perform studies with real and simulated data sets, and to apply in learning networks that may benefit from information fusion technologies.

1. Introduction

The advances of technologies enable self-directed learners to develop lifelong competences in learning networks (Koper et al., 2005). In order to support life-long competence development effectively and efficiently in learning networks, automatic competence tracking and management is crucial for determining learning goals, identifying competence gaps, seeking peers/partners, and offering appropriate learning opportunities.

However, tracking and management of competence is problematic. In theory, it is difficult to represent, measure, and interpret competence because competence is a very big subject complicated by very strong opinions and cultural traditions (Ostyn, 2005). In practice, no sufficient professionals serve for assessing competences of each lifelong learner in learning networks over time. As a non-expert in competence assessment, a lifelong learner may or may not evaluate a competence properly. In particular, somebody may intentionally not describe competences appropriately. As a consequence, the competence information captured in learning networks may be unreliable. The decisions and recommendations based on such unreliable competence information may be useless or make misleading.

In this paper, we will systematically analyze the problem of the unreliability of competence information in learning networks and explore technical solutions to solve the problem.

2. The Problem of Unreliability of Competence Information in Learning Networks

In this section, we analyze why competence information captured in learning networks may be unreliable. Figure 1 illustrates competence-relevant components (including actor, object, and software agent) in a learning network, actual competence (represented in oval which is the target to be detected and tracked by the system), competence information (represented in light blue rectangle which including competence source and competence record), their transformation (represented in arrow which are made by an actor or a software agent), and the main factors (illustrated beside the arrows) that influence the transformation. This section will explain in details.

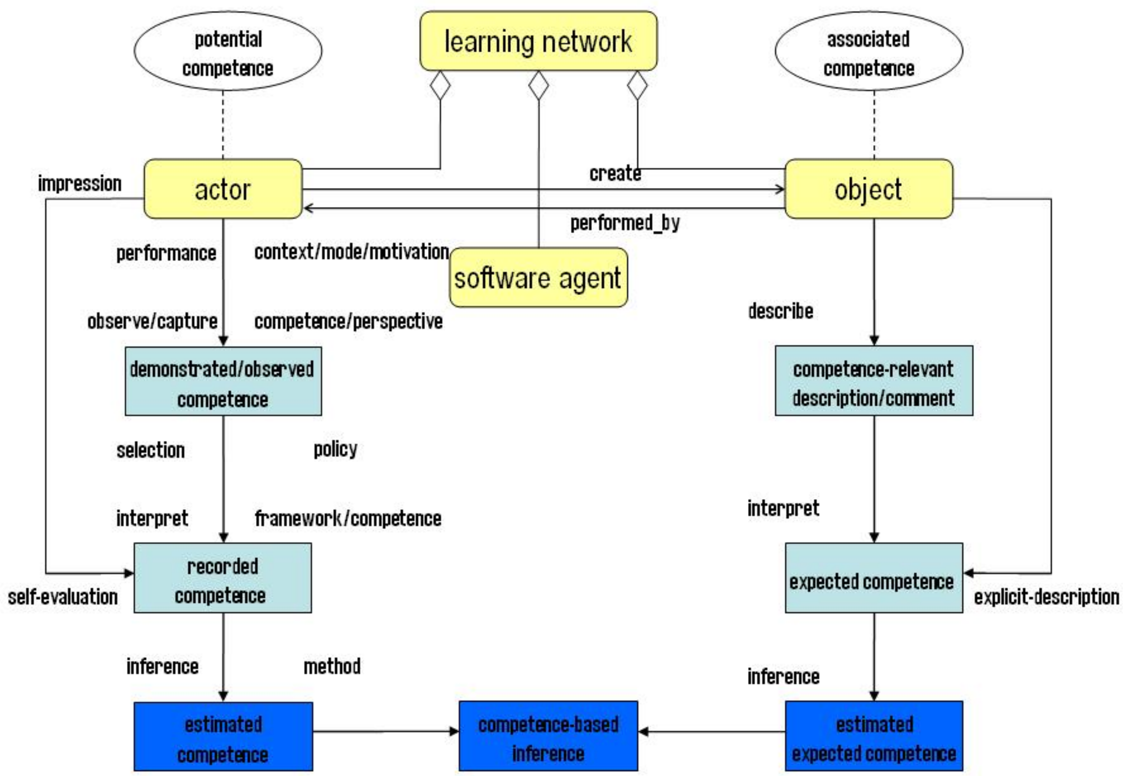

Figure 1: competence information and transformation

Competence is a latent attribute referring to an actor's (e.g., an individual, a team/group, or, an organization) underlying qualities and characteristics that lead to an effective performance. There is no systematic (objective) method to represent and measure potential competence like we represent and measure color and temperature. However, competence can be demonstrated and observed in a performance. The demonstrated competence can be captured as tangible source (as digital or non-digital evidence, which can be referenced persistently) or intangible source (as memory/impression, which can be recalled). In learning networks, various types of evidences can be captured such as a description of a performance (associated with a course, a task/activity, or a job), a product (e.g., an article, a design, and a response to a questionnaire), and an evaluation (e.g., a certificate, an evaluation of a response to a questionnaire, an analysis report of an article from a Latent Semantic Analysis (LSA) tool (van Bruggen et al., 2004)). It is important to note that evidence may or may not precisely reflect the potential competence. The competence owner may demonstrate a particular competence by performing tasks/activities with different characteristics under different situations (context) with different mode/motivation/attitude. On the one hand, the potential competence may be higher and lower than the demonstrated competence. On the other hand, a performance may or may not be precisely observed, recorded, and interpreted, because observers (or a software agent) may have different perspectives and measure methods, and may have higher/lower proficiency levels of necessary competences.

There may be a lot of evidences relevant to the same competence of an owner, which are originated from the same or/and different performances and captured by the same or/and different observers (or software agents). One or a set of evidences can be interpreted by actors (or software agents) as a competence record, which states that an actor has a known proficiency in a particular competence. For example, Sam's proficiency level of software development is "expert". However, the reliability of a competence record depends on which evidences are selected to create the competence record and how these evidences are interpreted. Various policies can be used to select evidences such as recently created evidences, certain types of evidences, and the evidences provided by particular actors or software agents. In addition, various competence frameworks and criteria may be used to interpret evidences. That is, the proficiency levels of a competence and corresponding indicators may be defined differently. Different communities of practice may map the components and/or facets of a competence in different ways (e.g., different roll-up patterns and weighting patterns). In addition, even though in the same community, different people may have different interpretations to the same evidence. The same person may have different interpretations to the same evidence at different time, or as his relevant competences are improved. Note that a competence record may be created by oneself in a self-evaluation or by someone else based on memory, an intangible source. In such a case, the reliability of competence records depends on whether the memory is good and how the impression is interpreted. In summary, there will be a huge amount of competence records about each competence of the owner in a competence tracking and management system if it captures and stores all relevant information in a long period of time.

As shown in Figure 1, a certain object such as course, task/activity, or job is associated with certain required/target competences. Like the potential competences of an actor, the associated-competences of the object can not be directly measured. However, it could be described as competence profiles as well. The problem is that different people may describe and interpret the same competence-relevant object differently. The competence profiles of a competence-relevant object may or may not be credible and trustworthy as well.

A competence tracking and management system can store all competence information such as competence evidences, competence records, and the relations to the owners, observers, interpreters, and the courses, tasks/activities, and jobs. They will be used to produce competence estimates. However, it is a challenge to produce an appropriate estimate of competence of an actor based on a huge amount of competence information, which may be inconsistent.

3. State of the Art

The problem of unreliability of competence information has not been sufficiently addressed currently. Ostyn (2005) explored to solve this problem by proposing a concept of distillation of competence information. According to his approach (see Figure 2), a confidence rating is introduced to qualify the competence evidence and competence record. The confidence rating is pre-determined according to a policy. For example, the results of a properly conducted 360 degree assessment are more credible than an assessment result from a supervisor, and in turn this result is more credible than that from a self-assessment or an online test on an unsecured computer somewhere on the Internet. The competence source or the competence record (called evidence record in the diagram) with the highest rating according to the policy will be selected as the competence estimate (called as competency record in the diagram) and other competence sources or the competence records will not be taken into account.

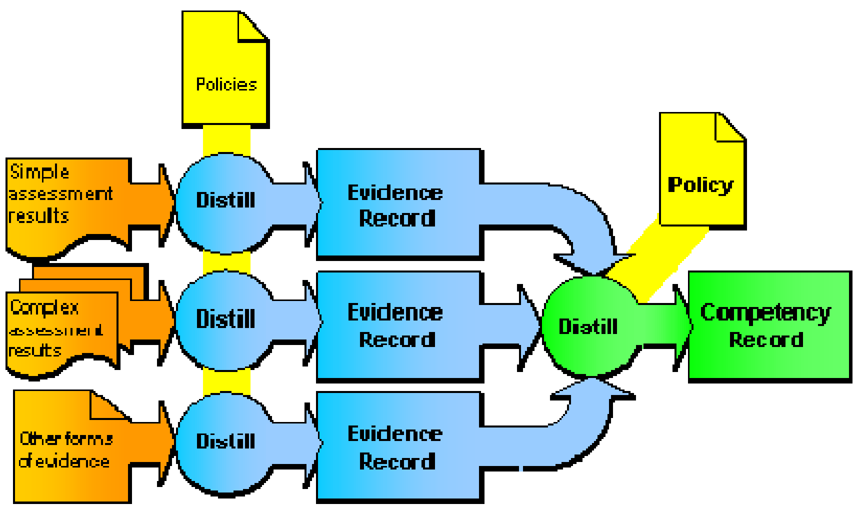

Figure 2 - Summary of the competency evidence distillation process (taken from Ostyn 2005)

However, it is not true that a pre-defined policy is suitable for all cases. For example, sometimes a self-assessment is more credible than an assessment result of his supervisor. Therefore, this approach can not effectively solve the problems. In this paper, we will investigate whether an information fusion approach is suitable for solving this problem.

4. Introduction of Information Fusion

The concept of information fusion (or data fusion) is easy to understand and the operation of information fusion by itself is not new. As stated in (Wald, 2001), the human being has the capability to use multiple senses to percept the environment. Rich information is acquired from various sensory organs such as eyes, nose, month, ears, tongue, and hands. In addition, a man has redundant sensors. Two eyes have slightly different viewing angles, making possible stereo vision and depth perception. If one eye is disabled, vision is still possible, though in a degraded mode. The brain processes the acquired information using additional sources of information: its memory, its experience, and its priori knowledge. Calling upon its reasoning capabilities, the brain "fuses" all available information to produce estimates about objects of interests, to assess situations, to make decisions, to update knowledge, and to direct actions.

However, information fusion, as techniques, is relatively new. It is multi-disciplinary by essence and is at the crossing of several sciences. According to (Wald, 1998; Wald, 1999), information fusion is “a formal framework in which are expressed the means and tools for the alliance of data originating from multiple and diverse sources”. Steinberg (2001) viewed information fusion as a process of combining data or information to estimate or predict entity states. The data range from numerical measurements to verbal reports. Some data cannot be quantified; their accuracy and reliability may be difficult to assess. Information fusion aims at achieving improved accuracies and more specific inferences that could not be achieved by the use of any single source alone (Hall & Llinas, 1997).

The information fusion offers some advantages (Waltz and Llinas, 1990):

- Robustness and reliability: The system is operational even if one or several sources of information are missing or malfunctioning,
- Extended coverage in space and time: The system can detect and trace the dynamic changes of the entities because a variety of distributed sensors can acquire information about the same entity at different time in different places,
- Improved confidence: The use of redundant and complementary information increases the certainty,
- Reduced ambiguity: More complete information provides better discrimination between available hypotheses,
- Providing a solution to process the vast amount available information for many complicated application systems.

The application of information fusion in technical systems requires mathematical and heuristic techniques from fields such as probability and statistics, Bayesian decision theory, plausibility theory, pattern recognition, fuzzy logic, neural network, expert systems, cognitive psychology, information theory, and decision theory. The functional application of information fusion is grounded in mathematical theories which are beyond the scope of this paper. The interested reader is referred to (Hall, 1992; Waltz, 1990; and Varshney, 95) for a detailed mathematical discussion. Information fusion is useful for several objectives such as detection, recognition, identification, tracking, change detection, and decision making. These objectives are encountered in many application domains such as defense, robotics, medicine, space, transportation, and weather forecast.

In order to have a better understanding of data fusion technologies, we brief introduce one of its applications in military with Wireless Sensor Networks (WSN), a special type of ad hoc network composed of a large number of nodes equipped with different sensor devices (Akyildiz et al., 2002; Nakamura et al., 2007). In comparison with large and powerful sensors, which are usually deployed in positions far from the battlefield and are definitely the targets being attacked by the opposing forces, the sensors in a WSN is small and inexpensive with limited sensing, computation, and communication ability. They are prone to failures and the information received from a single sensor may or may not be credible and trustworthy. They are different types of sensors such as seismic, low sampling rate magnetic, thermal, visual, infrared, acoustic sensors and radar, which are able to monitor a wide variety of ambient conditions. They can constantly monitor the status of friendly troops, the condition and the availability of the equipment and the ammunition in a battlefield. They can closely watch for the activities of the opposing forces and some valuable, detailed, and timely information about the opposing forces and terrain can be gathered. They can detect and track targets of the opposing forces (such as tanks, planes, and missiles) and can be incorporated into guidance systems of the intelligent ammunition. As the operations evolve and new operational plans are prepared, new sensor networks can be deployed anytime if necessary.

5. Competence Information Fusion

Generically speaking, some objectives such as detection, recognition, tracking, change detection, and decision making will be encountered to automatically track competence development in learning networks. Because of the limitation in size, this paper briefly analyzes similar characteristics of wireless sensor networks and learning networks from perspectives of application of information fusion technologies. Then we discuss one of important technical issues to solve the problem of unreliability.

In a wireless sensor network applying in military, the targets to be detected and tracked are objects such as tanks, planes, and missiles. An object has properties such as size, shape, and color and attributes such as position, direction, and velocity. There exist actual data if the object is moving in the battlefield. However, it is difficult to precisely measure the properties and attributes in the battlefield, where many factors (e.g., distance, perspective, bad natural conditions, and military operations) influence the measurement. In particular, the object may be with a designed shape, special material, and equipments to pretend and hide it from being detected. In a learning network, the object to be detected and tracked is the life-long learner with a set of competences. Each competence has an actual proficiency level at a given time according to a certain criteria. As mentioned, it is difficult to precisely measure the competences because many factors influence the accuracy of the competence records. In a wireless sensor network, a detected object is represented as a set of measurements, or attributes, or rules describing the object, completely or not. The goal is to produce an estimate of the values of properties and attributes, which are as closed as possible to the actual states, and then to make a correct judgment about the object. In a learning network, a competence profile is used to represent all competences. Each competence profile item can be represented as an estimate of competence. In a wireless sensor network, a sensor is a measurement device, and an imprecision value is usually associated with its observation. In addition, the sensing capability of a node is restricted to a limited region. Moreover, a given type of sensors can only perceive certain properties of the target. In a learning network, life-long learners and software agents (e.g., LSA tools and assessment simulators) measure competence. The capability of an agent (a human being or a tool) is restricted and different agents may have diverse abilities and bias. In a wireless sensor network, the data gathered by sensors are more or less credible and trustworthy. In order to overcome sensor failures, technical limitations, spatial and temporal coverage problems, multiple sensor nodes (with various types) will be deployed fully covering a region of interest. Each sensor obtains a partial view of a target under observation in a certain location at a certain time. These pieces of view can be fused into a continuously changed trace of the target. The redundant observations and measurements of multiple sensors can be fused to obtain more accurate data. Different types of sensors can perceive different properties of the target and the complementary information can be fused to produce a complete perception. In a learning network, a given competence can be evaluated by oneself, peers, experienced people, and software tools based on a certain performance from certain aspects at a certain time. There may or may not be credible and trustworthy. As the actor works within a learning network for a period of time, massive competence information about the actor will be captured. Why don't we apply information fusion technologies to produce a more accurate estimate of the competence and to obtain a continuous trace of competence development by fusing all competence information in a learning network?

If we want to develop an automatic competence tracking and management system, we will face a formidable set of hurdles, all of which need to be taken. This paper discusses only one of important technical issues concerning the unreliability of competence information. In general, fusion requires appropriate weighting of information based on the quality of the source of the information. A credibility model is needed to characterize the quality of information based on the source and the circumstances under which the information is collected. In information gathering, it is necessary to rate separately the quality (reliability, degree of trustworthiness) of both the source that produces the record and the content of the record itself. In practice, if the source is judged 'unreliable', the record is essentially discarded. If the source is judged 'reliable', then the content of the record is evaluated to decide how much trust should be given to it. Usually, a computational model of the quality of the information is used to compare and analyze data by using prior information, evidences, and opportunities for learning from data. If the conflict is small, it means the record fits with previous opinions, and seems thus to reinforce them. If the conflict is large, it means that the content of the record clashes with the previous opinions. It is needed to find out the origin of the clash, and try to resolve it. For example, if it is proved that the record is created by one who trends to over grade certain competences or the record is originated from a performance, on which most records were with lower ratings, the record will not be taken into account and the credibility of the actor and the performance will be re-assigned. However, if the record is produced by one, which is quite credible to assess this kind of competences, the credibility of the records and the sources which were used to develop the previous opinions will be re-checked. That is, the fusion process results in a revision or an update of the current belief function. Because there are very complicated inter-relationships among the competence information in a learning network, one change may trigger a sequence of changes.

A large variety of models and algorithms have been proposed in the literature to solve the problems. More models and algorithms will be developed in information fusion community in the future. We feel that the problem in learning networks may be more complicated than that in traditional application domains because the “sensor” node is usually human being.

6. Summary

We systematically analyzed the problem of unreliability of competence information gathered in learning networks. In order to address the problem, we briefly introduced information fusion as a technique that may help us solve the problem we are bound to encounter once we implement automatic competence tracking and management in learning networks. We promote to launch research before information fusion can begin to deliver on this promise. We feel that a great deal of research is needed to introduce, implement, and leverage the concept of competence information fusion in order to make an organizational impact.

Acknowledgement

The work on this paper has been sponsored by the TENCompetence Integrated Project that is funded by the European Commission's 6th Framework Programme, priority IST/Technology Enhanced Learning. Contract 027087

References

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey, Computer Networks, 38 (2002) 393–422.

Bass, T. 2000. Intrusion detection systems and multisensor data fusion. Communications of the ACM, v.43 n.4, p.99-105, April 2000. ACM Press.

Hall, D. (1992). Mathematical Techniques in Multisensor Data Fusion, Boston, MA: Artech House.

Hall, D.L. & Llinas, J. (1997). An introduction to multisensor data fusion, Proceedings of the IEEE, vol. 85, no. 1, pp. 6-23.

Koper, R., Giesbers, B., Van Rosmalen, P., Sloep, P., Van Bruggen, J., Tattersall, C., Vogten, H., & Brouns, F. (2005). A Design Model for Lifelong Learning Networks. Interactive Learning Environments, 13(1-2), 71-92.

Nakamura, E. F., Loureiro, A. A. F., and Frery, A. C. (2007). Information fusion for wireless sensor networks: methods, models, and classifications. ACM Computing Surveys, 39, 3, Article 9, 55 pages.

Ostyn, C. (2005). Distilling Competency Information. White Paper, http://www.ostyn.com/rescompetency.htm.
Steinberg, A.N., (2001). Data Fusion Systems Engineering, IEEE Aerospace and Electronic Systems Magazine, vol. 16, no. 6, pp. 7-14.

van Bruggen, J., Sloep, P., van Rosmalen, P., Brouns, F., Vogten, H., Koper, R., and Tattersall, C., (2004). Latent semantic analysis as a tool for learner positioning in learning networks for lifelong learning, British Journal of Educational Technology, Vol 35, No 6, pp 729–738.

Varshney, P. (1995) .Distributed Detection and Data Fusion. Springer-Verlag, New York, NY.

Wald, L. (2001). Introduction to Information Fusion, Available at: http://www.data-fusion.org/article.php?sid=64

Wald, L. (1998). A European proposal for terms of reference in data fusion. International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 7, 651-654.

Wald, L., (1999). Some terms of reference in data fusion. IEEE Transactions on Geosciences and Remote Sensing, 37, 3, 1190-1193.

Waltz, E. and Llinas, J. (1990). Multisensor Data Fusion. Artech House.