Abstract—In this paper, a review of prediction techniques suitable for ambient intelligence environments is presented. Prediction challenges in sensor networks are considered in two phases including pattern extraction and rule matching. The prediction techniques reviewed in this paper come from two main research areas, namely, data mining and soft computing techniques. Moreover, a statistical modelling technique based on Markov chain is also considered. In this paper, we identify the centralized and distributed techniques of both data mining and soft computing areas. In addition, we identify the distributed approaches that utilize computational power of sensors in an ambient intelligence environment. Moreover, we show that some techniques use compression, regression or fuzzy methods to reduce the size of the collected sensory data.

I. INTRODUCTION

With growing interest in the use of smart environments, a new generation of such environments has drawn attentions of many researchers [14]. The predictive environment as the third generation of smart environments will provide more intelligence capabilities in comparison with its former generations [15]. The predictive environment is an ambient intelligence environment that facilitates the interactions of occupants, devices and environment.

Energy saving, convenience of occupants, safety and security are the main objectives of smart environments. Despite the lack or minimum usage of sensors, the first and second generations of smart environments had some success to meet some of the above objectives. For instance, a bank equipped with an old security system in which employees can press a button to call the police in a bank robbery, is one of the first generation of smart environments. Second-generation smart environments are equipped with individual sensors or a sensor network. This generation of smart environments is also called an automatic environment. A smart building with automatic lights and heater control is an example of the second generation. New emerging third generation of smart environments also known as predictive environments have both manual and automatic control features; moreover, it can learn from environmental changes as well as behavioural patterns of occupants. Predictive environment collects data acquired from a sensor network. Collected data include variety of attributes, such as environmental changes and occupants interactions with environment. These data are used in a learning approach to make a predictive environment that can predict the occupancy of different areas as well as requirements and interests of occupants at different times. This predictive feature steps up the performance of energy saving approaches in a smart environment; in addition, it improves the convenience of occupants as well as security and safety.

Data collection and prediction are two challenges of predictive environments. The first challenge is due to the energy and bandwidth constraints in sensor network [12-13], but the second challenge which is the main focus of this article is a learning challenge in distributed sensor networks. Predictive environment can predict the next state of consecutive interactions with the use of the knowledge it has learnt from previously observed interactions. For instance, it can predict favourite light intensity of different occupants in a specific area of the environment at a specific time of a day. Prediction consists of first pattern extraction to identify sequences of actions, and then sequence matching to predict the next action in one of these sequences [1].

In this paper, a comprehensive review of available techniques suitable for the third generation of ambient intelligence environments is presented. Section II is an introduction to different prediction techniques. Section III is a review of data mining techniques and section IV is a review of soft computing techniques. A statistical model for prediction in predictive environments is considered in section V. Relevant conclusions and future works are drawn in the last section. Note that this is an ongoing research project, the ultimate goal of which is to build a predictive ambient environment.

II. PREDICTION TECHNIQUES IN PREDICTIVE ENVIRONMENTS

Prediction in predictive environments is a learning challenge in distributed sensor networks. There are variety of techniques for this learning challenge including data mining techniques and soft computing techniques. Some of these techniques are centralized approaches such as some data mining techniques and some of them are distributed approaches such as agent-based techniques. In addition, some of these approaches apply compression or fuzzy methods to reduce the amount of stored sensory data.
III. DATA MINING PREDICTION TECHNIQUES

In this section four data mining prediction techniques including case-based reasoning, lazy learning, Bayesian classification approach, and distributed voting approach are reviewed.

A. Case-Based Reasoning

Case-Based Reasoning (CBR) is a classification method that uses previous experiences to find a solution for current problem. It has two basic operations including case-generation and case-selection [2].

As a method of prediction, context-aware based case-based reasoning proposed in [3] is used as a method of pattern extraction of occupant’s behaviour in a predictive environment. In this method, the context in a smart home is classified into three dimensions, namely time, environment and person. In this method each case is represented as follows:

\[ \text{Case} = (\text{caseID, personID, habitID, environmentID, activeID, time}) \]

System framework proposed for implementing this method is shown in Fig. 1.

As an example to show how this method works, assume a person with a specific personID goes to the lounge in a predictive building at 8pm and sets his/her favourite light intensity and temperature. In this case, system generates a new case in its database for this situation. As soon as the same person goes back to the same place at the same time, the system sets his/her favourite light intensity and temperature automatically as an existent case in the database matches this situation. As the case-selection in CBR is problematic, similarity calculation is used to overcome this problem in this method.

Context-aware based CBR is a centralized prediction technique. It stores all cases in a central database, but the case adaptation phase in its system framework reduces the number of cases should be kept in the data base. In addition, as a centralized approach, context-aware based CBR can not utilize computational power of sensory devices in sensor networks.

B. Lazy learners

Lazy learners or instance-based learners are learning techniques in which an instance is classified based on minimum distance classification. Lazy learners store all the training data samples. This may present difficulties when the learning is from very large datasets [2].

Modular approach in [4] is an example of lazy learners in a sensor network. Modular organization of the sensor network proposed in [4] addresses two main issues in mining sensor network data:

1. Minimization of communication effort with compression of aggregated data of each cluster
2. Extraction of high-level information from a massive dataset

In modular approach shown in Fig. 2, sensory devices are clustered in sensing units. Then a data compression technique such as Principal Component Analysis (PCA) is applied to the data received by aggregator nodes. Finally, the collected dataset is used by a lazy learning algorithm to produce a model of the mapping and then the dataset is discarded and only the model is kept.

Despite the utilization of computational power in aggregator nodes, modular approach is categorized as a less distributed technique. It minimizes the data should be kept in the data miner by applying compression and modelling techniques. On the other hand, it could be problematic in terms of robustness as the system may lose some data by applying these techniques.

C. Bayesian classifiers

Bayesian classifier is a method of classification based on probability distribution. In this approach, the classifier calculates the probability of being a member of different classes for each sample and predicts the class of the sample [2].

Bayesian belief network proposed in [5] is a novel distributed detection system for occupancy control of a building. Occupancy control plays an important role to overcome the challenge of behavioural pattern extraction in a predictive environment. The belief network proposed in [5] is constructed using three Passive Infra Red (PIR)
occupancy detectors and a telephone off-hook sensor for data acquisition in each office based on following rules:

1- The total number of occupants in all rooms is the sum of the numbers in each room [Fig. 3-A].
2- The number of occupants persists over time [Fig. 3-B].
3- Sensor measurement depends on the number of occupants [Fig. 3-C].
4- Each sensor may respond to occupancy in different ways depending on its status [Fig. 3-D].

The combination of all dependencies is illustrated in Fig. 4.

Fig. 3. Separate relations for constructing a belief network.

The number of occupants and their location in a building is determined by analyzing the acquired occupancy data in a belief network analysis framework.

As the occupancy control plays an important role in behavioural patterns extraction, the belief network was considered as a technique in this area. Belief network is based on a probability distribution. It is categorized as a centralized approach and it does not utilize the computational power of sensors. This technique is able to diagnose the sensor network because the status of each sensor is concerned in the belief network.

D. Distributed Voting Approach

Due to the distributed nature of sensor networks in ambient intelligence environments, implementing distributed algorithms for learning approaches becomes possible. Most of these algorithms use small computational power of individual sensors to construct a powerful learning approach in the whole network. Distributed voting algorithm proposed in [6] is one of these algorithms. In this algorithm a tree structure of sensors as small computing devices and a powerful computing device in the root of this tree is constructed to solve a classification problem. This tree structure is shown in Fig. 5.

Each sensor as a leaf of the tree uses neural network or decision tree approaches for local prediction. Due to the shortage of memory in sensory devices, all training data for different classes are stored in the root. During the learning process, each sensor receives a training data from the root. After training phase, each node can measure and classify one or more attributes in a local policy. Eventually, in a global prediction, the root receives local classification decisions from sensors and performs a global classification by applying a voting strategy.

Distributed voting approach is categorized as a distributed approach. In spite of the distributed nature of this technique, a huge training data is stored in the root. Utilization of sensor’s computational power is the most significant advantage of distributed voting approach.

IV. SOFT COMPUTING PREDICTION TECHNIQUES

In this section four soft computing prediction techniques namely, Reinforcement Learning, Fuzzy Rule-Based Learning, Adaptive Online Fuzzy Inference System and Mixture-Model are reviewed.

A. Reinforcement Learning

Reinforcement learning is a method of learning that learns the relation between input and output with trial and error. In this method, a function called reinforcement signal must be maximized [2]. Any significant difference between input signal and target signal is considered as a punishment; therefore, the value of reinforcement signal decreases. On the other hand, a slight difference between input signal and target signal is considered as a reward; hence, the value of reinforcement signal increases.

As an example of reinforcement learning technique, [7] proposes an intelligent lighting control in which a multi-agent system controls lights. This technique concerns varying lighting preferences of different users for different
tasks. Fig. 6 shows a physical space equipped with identification sensors, photo sensors and actuators.

Fig. 6. Physical space equipped with sensors and actuators.

In [7], reinforcement learning technique is used to train the agents. An agent uses user’s location and light readings as the state space for the reinforcement learner and attempts to take actions that lead to appropriate light settings. For example, the absolute difference between the light intensity sensed by an agent before and after the user action is used as a negative reinforcement or punishment. Also, if an agent turns a light on and the user turns it off then the agent receives a negative reinforcement. In contrast, if a person does not change anything the agent receives a positive reinforcement as a reward.

Due to the multi-agent nature of this technique, it is categorized as a distributed approach, but it does not utilize the computational power of sensory devices.

B. Fuzzy rule-based learning

Multi-agent framework proposed in [8] can be deployed in an intelligent building equipped with sensors and effectors. In this approach, each agent controls and learns about a small sub-region of the entire environment. In this technique, knowledge is represented by fuzzy rules and learning process is an unsupervised algorithm. In the learning process, inputs from sensors are sampled and transformed to fuzzy sets in a fuzzification phase. Then, the learning process compares the fuzzy inputs with stored fuzzy rules. Any significant difference between fuzzy inputs and stored fuzzy rules is considered as a punishment. On the other hand, slight difference between fuzzy inputs and fuzzy rule is a reward to the fuzzy rule. This technique is illustrated in Fig. 7.

For example, assume that an agent in the study room contains the following rule:

\[
\text{If time is 8pm and Bob is in study room then set the light intensity to 10}
\]

If as a sample, system recognizes that Bob is in study room and time is 8pm but the light intensity of Bob preference is 5, then the stored rule receives a significant punishment, or it may be replaced with a new rule.

This technique is a multi-agent technique and it is categorized as a distributed approach. Despite the distributed nature of this technique, it does not utilize the computational power of sensory devices. The most significant advantage of fuzzy rule-based learning is reduction of raw data by applying fuzzification mechanism.

C. Adaptive online fuzzy inference system

In [9], Adaptive Online Fuzzy Inference System (AOFIS) as a learning and control system is proposed. The authors do their experiments in the Essex intelligent dormitory as a test bed. AOFIS prediction approach contains three phases for learning and two phases for control and adaptation:

1. Monitoring the user’s interactions and capturing input/output data associated with their actions
2. Extraction of the fuzzy membership functions from the data
3. Extraction of the fuzzy rules from the recorded data
4. The agent controller
5. Life-long learning and adaptation mechanism

In the first phase, sensors take a snapshot from user’s action, as well as sensors readings before the user’s action. For instance, assume that the temperature of a space is 30 and user sets the air conditioner to 25. The system takes a snapshot from the both current temperature and user’s temperature preference. In the second phase, different techniques of clustering such as Fuzzy C-Means, Double Clustering, Agglomerative Hierarchical Clustering Approach and Quantification of Fuzzy Membership Functions are used to extract fuzzy membership function. With these techniques the accumulated user input/output data is categorized into a set of fuzzy membership functions which quantify the raw crisp values of the sensors and actuators into linguistic labels, such as normal, cold, or hot. In the third phase, the defined set of membership functions are combined with the existing user input/output data to extract the rules defining the user’s behaviours. The fuzzy rule extraction approach used by AOFIS is based on an enhanced version of Mendel Wang method that is a one-pass technique for extracting fuzzy rules from the sampled data. With extraction of membership functions and set of rules, the agent’s Fuzzy Logic Controller (FLC) becomes capable to capture human behaviours. Therefore, in the fourth phase, the agent monitors the state of the environment and affects actuators based on its learned FLC that approximates the preferences of the user. Finally, in the fifth phase, the agent adapts its existing rules or adds new rules based on the new preferences of the user. For example, if the user changes the settings of the environment, then the agent would adapt itself.
with new preferences. Five phases of AOFIS are illustrated in Fig. 8.

![Fig. 8. Five phases of AOFIS.](image)

Due to the use of fuzzy membership functions, the amount of data should be kept in AOFIS technique is reduced. This technique can be used as either centralized or multi-agent approaches.

### D. Mixture-Model

In [10], data collected from motion detectors are used to determine four attributes including the location of occupant, the start time of being in the location, the length of the time spent in it, and the activity level of the occupant in the location. In the training phase, a mixture-model makes a cluster of each activity. In this method an activity is recognized based on the time spent in a specific location and activated sensors during this activity. For instance, the activity of putting a book in the library needs less time rather than checking email in the same room. In addition, a different set of sensors will be activated during the activity of putting a book in the library rather than checking email in the same room.

The mixture-model is a combination of different methods including event estimation, self organizing maps, and k-means clustering. The power of the mixture-model is due to its capability to distinguish different mixed activities. For instance, the activities of putting a book and checking email can occur simultaneously. Simultaneous activities make it more difficult to identify them. For example, it would be difficult to recognize that which activity has fired a sensor. The mixture-model concerns the time spent in a location and fired sensors in it to calculate the probability of each trained activity. Finally, the more probable activities are expected ones. Fig. 9 is an illustration of clustering with mixture-model.

![Fig. 9. Clustering in a Mixture-model.](image)

### V. STATISTICAL MODELLING PREDICTION TECHNIQUES

Statistical approaches namely Markov model are also considered in ambient intelligence environment. Markov model is a statistical method of modelling that uses Markov chain to define a process. In a Markov chain next state of the system only depends on the present state. Transition probability between two states in a Markov chain is represented by a transition matrix. Fig. 10 illustrates a simple example of a Markov chain, as well as its transition probability matrix.

![Fig. 10. Markov chain and transition probability matrix.](image)

Markov chain is used in [11] to model daily activity of elderly people living alone in a predictive home. In this approach, first of all, a profile transition probability matrix from observed sensory data for each elderly person is generated and stored in a database. Then, during a daily activity, a test transition probability matrix is generated. Minor differences between profile and test matrices with an acceptable tolerance shows that the health status of the elderly person is not changed. In contrast, any significant statistical difference between these two matrices can be considered as an abnormal health status of the elderly person.

The significance of a statistical model such as Markov chain in modeling ambient intelligence environments is due to its simplicity and capability of representing systems with multiple transitions. As an example, occupancy detection as an important application in an ambient intelligence environment is multiple transition and can be modeled as a Markov chain.

### VI. CONCLUSIONS

Prediction as the key feature of an ambient intelligence environment was the main focus of this paper. We showed that the prediction problem in an ambient intelligence environment is mostly the pattern extraction problem in a distributed sensor network. Then, we considered different prediction techniques from two areas including data mining and soft computing as solutions for this problem. In addition, a statistical approach for modelling behavioural pattern in sensor networks was reviewed.

Due to the distributed nature of sensor networks, it becomes possible to apply distributed prediction approaches to sensor networks. Distributed voting approach shown as a prediction technique in data mining area is one of them. In contrast, small computational power of individual sensors makes it difficult to execute complicated prediction techniques with their huge training datasets. Therefore, the push is to less distributed techniques or multi-agent
Adaptation and control will be the last phase of building intensity sensors, pressure sensors and humidity sensors. Received from sensors including temperature sensors, light will be concerned for storing the vast amount of data we will apply a simple method of prediction such as control will be considered in the next phase. In this phase, network in our predictive environment. The occupancy we will need a fusion of suitable appropriate methods of techniques for prediction purposes. In this case, applying a fusion of different techniques is useful. For example, occupancy control can be done by simpler approaches such as Bayesian classifiers. On the other hand, in temperature or occupancy detection through sensor belief networ...