Modeling Human Activity Semantics for Improved Recognition Performance

Eunju Kim[1] and Sumi Helal[1]

[1]Mobile and Pervasive Computing Laboratory
The Department of Computer and Information Science and Engineering
University of Florida, Gainesville, FL 32611, USA
{ejkim, helal}@cise.ufl.edu

Abstract. Activity recognition performance is significantly dependent on the accuracy of the underlying activity model. Therefore, it is essential to examine and develop an activity model that can capture and represent the complex nature of human activities precisely. To address this issue, we introduce a new activity modeling technique, which utilizes simple yet often ignored activity semantics. Activity semantics are highly evidential knowledge that can identify an activity more accurately in ambiguous situations. We classify semantics into three types and apply them to generic activity framework, which is a refined hierarchical composition structure of the traditional activity theory. We compare the introduced activity model with the traditional model and the hierarchical models in terms of attainable recognition certainty. The comparison study shows superior performance of our semantic model using activities of daily living scenario.

Keywords: Activity Recognition, Activity Modeling, Activity Semantic Knowledge, Generic Activity Framework, Accuracy and Certainty

1 Introduction

Human activity recognition (AR) is an essential technology of pervasive computing science because it can be applied to many practical applications including health care, eldercare or smart spaces [1][2]. In spite of the obvious importance of this technology, current AR technology has limited accuracy and further development is required for real world applications. It is because many human activities are so complex that their accurate recognition is a big challenge. To illustrate, when people perform an activity, it is performed in a variety of ways. Also it is often concurrent and interleaved with other activities.

To solve this problem, many activity models and activity recognition algorithms including probability-based or machine-learning based approaches have been developed to improve the recognition performance [11][12][13][14][15]. However, these approaches are not sufficient for practical applications because they do not adequately address complex, ambiguous or diverse human activities. Therefore, a new approach, which can capture and represent such unique characteristics of human activities more precisely, is necessary. Comprehensive understanding of human activities and careful activity modeling is especially important because other techniques including AR algorithm and AR system are based on the activity model and their accuracy is influenced from the activity model.

In this paper, we propose activity semantic knowledge and a knowledge-assisted activity modeling technique that utilizes the semantic knowledge for accurate modeling of human activities.

1.1 Motivation

Our motivation is developing a new activity modeling approach that can model real world human activities accurately. To achieve this goal, several challenges need to be addressed. In particular, the new approach should address the following characteristics of human activities.

Concurrent activities. People may be involved in actions corresponding to several activities at the same time. For example, people watch TV while talking to friends. These behaviors are not sequential, and therefore, an activity model needs to represent these characteristics of activities [3].

Interleaved activities. In a real world scenario, some activities may be interrupted by other activities before completion whereas some are not. For instance, if a friend calls while cooking, the resident may pause cooking and talk to the friend for a while. After talking, he/she continues to cook [3].

Ambiguity. Even though sensor detects user activities well, the exact interpretation may be difficult. For example, we cannot guarantee that a person really takes a medicine even though a sensor detects opening the medicine bottle because the sensor could be reporting on other related activities such as *checking whether bottle is empty* or *cleaning bottle* but not taking the medicine.

Variety. Humans can perform activities in a variety of ways. For example, there are multiple ways to eating such as *having a meal* or *having a snack*. Typical scenario based activity recognition is not enough to handle this variety.

Multiple subjects. More than one person could occupy the same space and perform activities together sometimes. An activity model needs to be capable of associating the detected activities with the resident who actually executed them [3].

1.2 Proposed Approach

As a solution to the aforementioned problems, we propose a new approach that utilizes activity semantic knowledge for modelling human activities. Activity semantics are the knowledge about the characteristics of activities and it can provide a variety of important information to express activities. For example, no other activity can be performed along with sleeping activity and the person will be in better condition after sleeping, the person will be lying down on the bed, and the next possible location of the person will be near the bedroom. This information constitutes the semantic knowledge of an activity. We named this kind of knowledge as activity semantic knowledge and utilized it for modelling activities. The major advantage of this modelling approach is that it can reduce uncertainty in an activity model and other AR technologies based on the activity model.

First, without prior semantic knowledge, activity models treat all activity components equally because it is difficult to know the differences. But a close look at human activities reveals that activity components have different roles. Some components are essential for an activity whereas some are trivial. Therefore, semantic activity knowledge is helpful for modelling activities more precisely.

Second, activity model can capture activity relationship more accurately. Without prior semantic knowledge, activity models may impose detection conditions that are too strict or too loose. For example, Hidden Markov Model (HMM) model is too strict, as it requires enumerating all possible orders of activities. Conditional Random Field (CRF) model is too loose in that it does not account for the order among activities [3][7][8]. However, if we have semantic knowledge of activities, we can account for the order in which actions are performed only if it is meaningful.

The rest of this paper is organized as follows. In section 2, we discuss the traditional activity models such as activity theory based model and probabilistic graphical models and their limitations. The proposed approach is explained in section 3. A comparison and analysis are represented in section 4. Finally, section 5 concludes the paper.

2 Background

In this section, we describe traditional activity models. There are two popular approaches for modeling activities. One is activity theory based modeling; the other is probabilistic modeling such as Hidden Markov Model (HMM) or Conditional Random Field (CRF) model.

2.1 Activity Theory - Origin of Activity Modeling

Historically speaking, L. S. Vygotsk who was a psychologist during 1920s and 1930s founded the activity theory. Later, the activity theory was further developed by A. N. Leontjev and A. R. Lurija and coined the term “activity” [5][6]. Activity theory was first applied to human-computer interaction (HCI) in the early 1980s [5]. These days, it is applied implicitly or explicitly in a lot of activity recognition research.

The activity theory contains four components (subject, tool, objective, and outcome) [5][6]. A subject is a participant of an activity. An objective is a plan or common idea that can be shared for manipulation and transformation by the participants of the activity. Tool is an artifact a subject uses to fulfill an objective. Outcome is another artifact or activity that are result of the activity. Transforming the objective into an outcome motivates the performing of an activity. For example, having one’s own house is an objective and the purchased house is the outcome. Transforming an object into an outcome requires various tools.

As shown in Table 1, activity theory has a three-layered hierarchical structure and activity is composed of actions and an action is composed of operations [5].

Table 1. Hierarchical layers of an activity and an example of activity, action, and operation.

Levels	Related Purpose	Example of purpose
Activity	Motive	Completing a software project
Action	Goal	Programming a module
Operation	Conditions	Using an operating system

Activities are composed of cooperative actions or chains of actions. These actions are all related to the motive of an activity. Each action has a goal and consists of operations to reach the goal. Operation is a unit component and it depends on the faced condition where the operation performs. The detailed description of each level can be found in [4][5].

Even though activity theory is well known and is often used in activity recognition research, it has some limitations. First, the border between hierarchical layers is blurred. As described in [5], an activity can lose its motive and become an action, and an action can become an operation when the goal changes [5]. This unclear border makes automated activity recognition difficult because the change of motive of activity and goal of action are not easy to detect. Hence, it is necessary to find clearer ways to determine each layer.

Second, activity theory does not distinguish between tool and object. But, they are needed to be distinguished because the same item may be used as tool or object in several activities. In this case, the item has different meaning for each activity. For example, when a pan is used as a tool for cooking, it implies it contains food. On the other hand, if it is an object for washing dish activity, it means that it is an empty dish.

Last, some activities are too complicated to be represented by a single activity name. For instance, eating has several similar activities such as having a meal or having breakfast, lunch or dinner. Because the top layer is activity in activity theory, the layer includes everything. This makes AR system design cumbersome and difficult to conceptualize. This difference in granularity is not conducive to sharing or modularizing AR systems.

2.2 Probabilistic Activity Models

In probabilistic approach, human activities are continuously performed and each activity is a sequential composition of activity components such as motions, operations or actions according to a temporal sequence. According to this idea, several probabilistic models including Hidden Markov Model and the Conditional Random Field Model have been used to build an activity model because they are suitable for handling temporal data.

Hidden Markov Model (HMM). HMM is a probabilistic function of Markov chains based on the first order Markov assumption of transition [7]. The basic idea of Markov chain of order m is that the future state depends on the past m numbers of states. Therefore, for HMM based on the first order Markov assumption, the future state depends only on the current state, not on past states [7]. Also HMM is a model that is used for generating hidden states from observable data. HMM determines the hidden state sequence (y_1, y_2, ..., y_t) that corresponds to the observed sequence (x_1, x_2,

..., x_t) [3]. In activity recognition, hidden state is human activities and HMM recognizes activities from both sensor observation and previous activity according to the first order Markov chain. However, HMM is also a generative, directed graph model [3]. Generative model means that observation data is randomly generated. In other words, it should enumerate all possible random cases in the model. Directed graph is used capture order between states. Therefore, a generative and directed graph model in activity recognition implies it should find all possible sequences of observations.

However, many activities may have non-deterministic natures in practice, where some steps of the activities may be performed in any order. In practice, although many activities are concurrent or interleaved with other activities, HMM has difficulty in representing multiple interacting activities (concurrent or interleaved) [3]. Also HMM is incapable of capturing long-range or transitive dependencies of the observations due to its very strict independence assumptions on the observations. Therefore, enumerating all possible observation cases and orders is difficult for a practical system. Furthermore, missing an observation or an order will cause the HMM to produce errors in the model.

Conditional Random Field (CRF). CRF is a more flexible alternative to the HMM, which relaxes the strict assumption of HMM [8]. In other words, CRF solves the problems of HMM by neglecting the order constraint. Like HMM, CRF is also used to determine a hidden state transition from randomly generated observation sequences. However, CRF is a discriminative model, which does not generate possible cases from the joint distribution of x and y. Therefore, CRF does not include arbitrarily complicated features of the observed variables into the model. Also, CRF is an undirected acyclic graph, flexibly capturing any relation between an observation variable and a hidden state [8]. Because CRF does not consider order, it considers only relationships such as state feature function (relationship between observations over a period of time and activities) and transition feature function (relationship between past activities and future activities). Even though CRF removes order constraint from an activity model, CRF could outperform HMM [15].

3 Semantic Activity Model

In order to recognize activities accurately, modeling activities precisely is essential because recognition performance will be limited unless activities are analyzed and represented accurately. To illustrate, the knowledge of components of activities, relationships between an activity and components, characteristics of activities, etc. are important for identifying activities. Therefore, this activity knowledge should not be overlooked when modeling activities.

However, as we mentioned in the previous section, both activity theory and probabilistic activity model have limitations and do not represent human activity precisely because they do not consider important activity semantic knowledge. For example, activity theory does not distinguish activity components clearly enough. HMM and CRF apply an order assumption that is too strict or too loose respectively. However, these modeling assumptions should be more adaptive in reality because there are some activities like an instruction, which consider order critical whereas some activities do not.

To solve this issue, the proposed activity semantics based model (Semantic activity model) incorporates substantial semantic activity information in the model. To exploit semantics, we first adopt a generic activity framework that extends the activity theory in section 2.1 because it provides a refined framework of activity model. We use daily living activities as demonstrative examples.

3.1 Generic Activity Framework

Generic activity framework has a hierarchical structure in which each layer of the structure consists of activity components. In total, there are eight primitive components in the framework as shown in Fig. 1. It is not necessary for every activity to contain all eight components as long as the activity is recognized clearly. For example, the walking activity does not require any object. Descriptions of the eight primitive components are summarized below and described in details in [4]:

Subject. A subject is an actor of the activity. Subject has an important role as an activity classifier especially when there are multiple people.

Time. This is the time when an activity is performed. It consists of start time and end time. We can also calculate the duration of an activity using time.

Location. Location is the place where an activity is performed. If an activity is performed in several places, location will have multiple values.

Motive. Motive is the reason or objective why a subject performs a specific activity.

Tool. Tool is an artifact that a subject uses to perform an activity. Tool provides essential information to classify activities. For example, a spoon or a fork is a tool for eating or cooking.

Motion. Motion is defined as the movement performed by a subject for handling tools. Motion explains what a subject does with a tool. For example, cutting and chopping are both performed using the same tool i.e. knife. The different motions associated with cutting and chopping can be used to differentiate between them.

Object. An object can also be any artifact like tool. But, object is the target of an activity whereas a subject uses a tool. Distinction between tool and object is important for accurate activity recognition because some artifacts are tool for an activity and object of another activity.

Context. Context provides information about the “vicinity” in which an activity is performed. Installed sensors directly find some contexts such as temperature or humidity. Other primitive components such as time or location contribute to finding more other contexts such as time to sleep, place to cooking, etc. On the other hand, some contexts like motive of an activity need some artificial intelligence techniques such as reasoning or inference to elicit them.

Fig. 1 shows a composition diagram of the generic activity framework. Rectangles are layers and ellipses are primitive components. According to the composition of components, the activity framework has a hierarchical structure. And the components of each layer are clearly defined. Brief description for each layer is given below (more details in [4]):

Sensors. Sensors are installed in the pervasive space (e.g. a smart home) to collect event information of the space. Based on the installed places of sensors, sensors are classified into four types: motion, tool, object, and context sensor.

Operation. Operation is a composition of tool and motion. The user operates tools with specific motion. For example, if computer is a tool, some hand or arm motion will be performed for typing a keyboard or using a mouse.

Action. Action is determined by combination of operation and object. For instance, if a user clicks a mouse to open a file, using a mouse is an operation and the file is an object and this combination is open file action.

Activity. Activity is a collection of actions. Activity may involve multiple actions and an action belongs to a subject. If a subject is different, we classify the activity separately. If multiple people collaborate for a same activity, the activity belongs to the multiple people.

Meta activity. A meta activity is a collection of activities. It is useful to use when an activity is complicated, in which case it can be composed of several simple activities. For instance, a meta activity *hygiene* is composed of *washing hands*, *brushing teeth* or *taking a bath*.

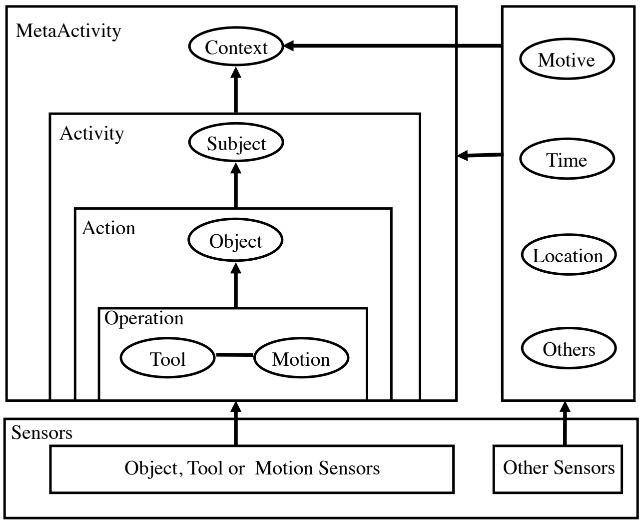

Fig. 1. Composition diagram of a generic activity framework. It is composed of several hierarchies and each hierarchical layer contains classifier components.

The hierarchical structure has several advantages. Firstly, it provides clear distinguish between layers so that user will not confuse operation, action and activity any more. Secondly, it makes the activity recognition system more tolerant to sensor environment change [4]. For instance, even if more sensors are inserted in the AR system, the upper layers in the hierarchy will not be seriously influenced from the change of sensor environment. Lastly, activity recognition using hierarchical structure is analogous to the way people recognize, so it is easier to design more natural and intuitive AR algorithm [4].

3.2 Semantic Activity Model

Even though a generic activity framework in section 3.1 describes the composition hierarchy of activity components, it is a general framework, which does not contain detailed activity semantic knowledge such as role of an activity component, constraint or relationship with other components. The activity semantics should be represented in an activity model because they are important for classifying an activity. For example, *eating* is composed of three actions such as *picking food*, *chewing food* and *swallowing food*. In this case, if only *picking food* and *chewing food* are detected, then it is not clear whether we consider *eating* is really performed or not because we are not sure the person completes the activity through chews and swallows the food or not. Activity semantics reduce these kinds of ambiguity. There are three activity semantics: dominance semantics, mutuality semantics and order semantics.

Dominance semantics is semantic information of vertical relationship between components in upper layer and lower layer in Fig. 1 (e.g. meta activity and activity, activity and action, or action and operation). In other words, components in upper layer (e.g. activity) are composed of the components in lower layer (e.g. action). In this hierarchical composition structure, the contribution of each action is different. Even though some actions are components of the same activity, some actions are dominantly essential component of the activity whereas some are not. According to the dominance, we classify them as key, unique, optional and conditioned components.

Key component. Key component is a mandatory component for identifying an activity. If an activity has multiple key components, all of them are required to agree with the activity. Otherwise, the activity is not considered performed. To illustrate, *swallowing* is a key action for *eating* because if people don't swallow food, it is not regarded *eating* is completely performed even though there are many other actions such as *picking food*, *scooping food* or *chewing food*.

Unique component. Unique component is a highly evidential component although it is not a key component. For instance, *chewing* is a unique action for *eating* because most *eating* requires chewing. However, *chewing* is not a key component because *chewing* can be omitted if food is soup.

Optional component. If a component is neither a key nor unique component, it is an optional component. It is possible to omit an optional component because it does not always affect activity classification. For example, *cutting food* is an action of *eating* but it may be omitted depending on the food. However, if optional component is detected, it increases the certainty of the recognition of an activity.

Conditioned component. If components should satisfy a specific condition for an activity, it is a conditioned component. For instance, *duration* is an example of condition for *sleeping* because it is highly unlikely that sleeping is performed if the duration is too short (e.g., a few minutes).

Mutuality semantics is semantic information of horizontal relationship between components at the same layer (e.g. meta activity and meta activity, activity and activity, action and action or etc). This semantic knowledge is used to determine whether multiple activities can be concurrently performed or not.

Concurrent component. If two or more components are performed together, they are in concurrent relationship. For example, *laundry* or *watching TV* is concurrent because while the washer is running, the user can watch TV at the same time.

Exclusive component. If an activity cannot be performed simultaneously with another activity, it is an exclusive activity. For example, *sleeping* is an exclusive activity because people cannot perform anything when they sleep.

Ordinary component. Ordinary components are partially exclusive and concurrent. If an activity is performed with a part of the body (e.g. human limb), the activity is both concurrent and exclusive. For example, when people eat food, they cannot sing a song at the same time. In this case, they are exclusive. But if the people take a walk, they can sing a song concurrently. Therefore, *sing a song* is both partially exclusive and concurrent.

Order semantics. Some activities like an instruction should follow a procedural sequence. However, many activities have flexible order or do not have any order. Therefore, the role of order among activity components should be considered depending on the activity.

No order. There is no specific order required between activity components. For example, actions for *eating* such as *cutting*, *picking* and *scooping food* does not have any order restriction.

Strong order. Some activity requires that activity components (e.g. actions) should be performed in a specific order always. For instance, in case of *sleeping* and *waking-up*, *waking-up* comes immediately after *sleeping* because people perform another activity before waking up.

Weak order. For many activities, their action components are performed according to a flexible order, which is not mandatory or strict. For example, usually *eating* is performed after *cooking*, but there are exceptions to this order depending on several situations.

Skip chain order. When an activity is interleaved, other activity components may be performed between two ordered activity components. To illustrate, *eating* is usually performed immediately after *cooking*, but sometimes we can do other activities between them.

The Fig. 2 shows the modeling notations of each semantic. A component can have multiple semantics. For example, if a component is unique and exclusive, it is represented with both bold dotted line and filled circle. Also we can find that there are multiple elements like “x,y,z” in some circles whereas there is only “x” in other circles in Fig. 3. The circle that contains multiple elements is a compound component and the circle with an element is an elementary component.

Elementary component. An elementary component such as optional component or ordinary component can have only one element.

Compound component. A compound component contains either single or multiple elements. Even though each individual sub-component of a compound component is optional or ordinary component, there are some cases their combination have stronger semantics. For example, *picking food* or *scooping food* is an optional and an ordinary component because we can have food without picking if we can have the food by scooping and vice versa. However, we need to do one of them while eating food. Even though *picking* and *scooping* is less evidential compare to *chewing* and *swallowing*, missing both *picking* and *scooping* will reduce the probability of the activity. In this case, we create a compound component with multiple optional or ordinary components.

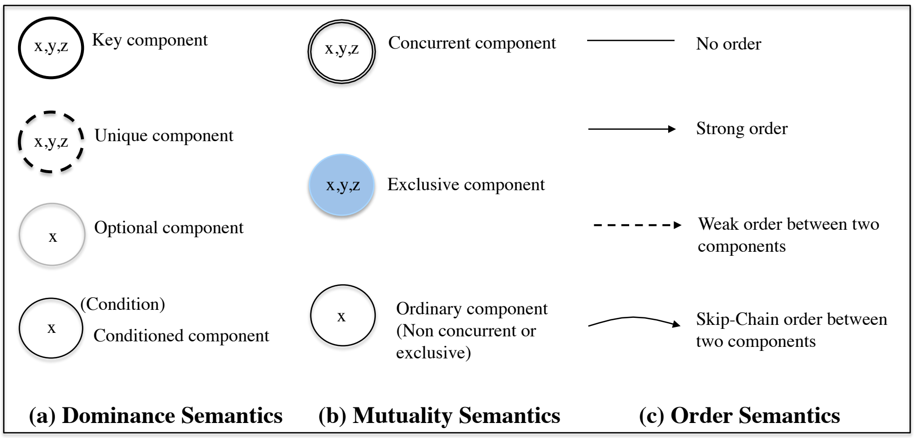

Fig. 2. Notations of semantic components. Dominance semantics and mutuality semantics are represented as nodes whereas order semantics are represented as edges.

Fig. 3 is an example of semantic activity modeling of daily living activities. In this example, *sleeping* is a unique activity and it also exclusive activity where as *watching TV* is a concurrent activity. *Scooping* and *picking* are compound key components of *eating*. There is a skip chain relationship between *preparing meal* and *having a meal* because it can be interleaved by other activities.

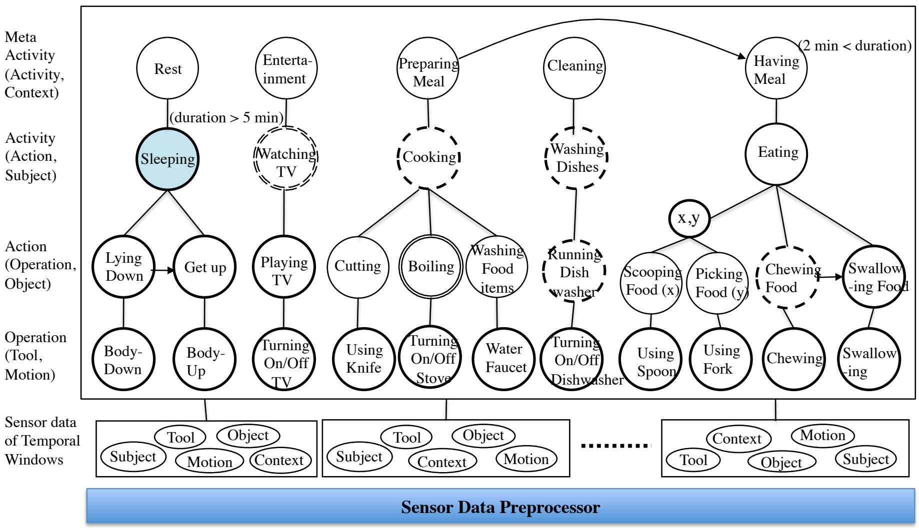

Fig. 3. An example of semantic activity modeling of daily living activities. Semantic components are represented on a hierarchical composition structure, which is based on the generic activity framework.

4 Comparison and Analysis

In this section, we compare the proposed Semantics and Generic Activity Framework based Model (S-GAM) with a Traditional Activity Model based on activity theory (TAM) and the Generic Activity Framework based Activity Model without semantics (GAM). We used a daily living activity scenario. To establish an activity scenario for the comparison, we used the eldercare scenarios of daily living described in [2] instantiated with a real activity dataset provided by University of Amsterdam [15]. The eldercare scenarios in [2] describe 33 different daily living activities of the elderly. The activity data set in [15] is records of activities of daily living performed by a man living in a three-bedroom apartment for 28 days. We named this dataset the Amsterdam dataset. We chose seven activities, which are common in both [2] and [15]. In terms of sensors, we assume the same sensor environment with the Amsterdam dataset are used to see how S-GAM performs in real situations. We add a Bed sensor according to the scenario in [2]. Table 2 lists the seven activities and their components.

Table 2. Activity list collected from the Amsterdam dataset. It shows meta activities, activities, actions, operation tools, objects and related semantics for each activity.

Meta Activity	**Activities (Location)**	**Action**			**Semantics**
		Operation	**Tool**	**Object**	
Rest	Sleeping (Bedroom)	-Going to the bedroom -Lying down	 Bed	Bedroom door	Mutuality: Exclusive activity
Hygiene	Taking a bath (Bathroom)	-Taking a shower -Washing face		Bathroom door Restroom door	Mutuality: Exclusive activity
	Using the toilet (Restroom)	-Opening a restroom door -Pressing a toilet flush		-Restroom door -Toilet flush	Dominance: -Key: Toilet flush -Unique: Restroom door
Preparing a meal -Breakfast -Lunch -Dinner	Cooking (Kitchen)	-Preparing food items -Heating food	 -Microwave -Pan	-Groceries -Refrigerator -Freezer	Weak order: -Taking food items -> Heating food
Drinking	Drinking	Taking drink	Cup	Refrigerator	Dominance: Key: Cup
Cleaning	Washing dishes		Dishwasher	-Pans -Cup -Dishes	Mutuality: Concurrent activity
Going out	Leaving the house	Opening front door		Front door	Dominance: -Key: Front door Mutuality: -Exclusive activity

Our activity scenario consists of six meta activities, seven activities, eleven actions, five tools and eleven objects. It also shows the activity semantics of each activity. In Table 2, we can see that many artifacts are used as tools or object in activities. Especially when an artifact is used in several activities, TAM is difficult to recognize activities accurately because it regards artifacts as tools only. For example, in Table 2, an artifact pan is a tool for cooking and it is also an object for washing dishes. Since TAM does not distinguish tool and object, sensing pan is not sufficient to determine which activity is performed. In contrast, GAM and S-GAM consider the usage of artifacts as both tool and object. Especially, S-GAM can recognize activities more accurately because it classifies activities using activity semantics. For example, sensing tools such as pan or microwave and objects like groceries usually mean that *cooking* is performed in GAM model. However, food items should be prepared before turning the microwave on if it is a cooking. Otherwise, it is unlikely the microwave is for cooking. GAM does not check this order semantic that is necessary for accurate activity recognition.

To compare the accuracy of the three activity models, we measured the uncertainty incurred by each model under the same activity scenario. Certainty factor is very effective evaluative analysis used in several areas such as diagnostics and medicine [9]. We briefly define Certainty Factor below.

CF(H, E): CF is a certainty factor from hypothesis H influenced by evidence E [9]. The value of certainty factor ranges from -1(very uncertain) to +1(very certain) through zero (neutral).

$$\mathrm{CF(H, E) = MB(H, E) - MD(H, E)} \tag{1}$$

MB(H, E): MB is the measure of increased belief in hypothesis H influenced by evidence E [9]. p(H) and 1-p(H) are the probabilities of that hypothesis being true or false respectively. p(H|E) is a probability of hypothesis given E. If the evidence, E, is very strong, p(H|E) will equal to 1 and p(H|E) - p(H) will be also close to 1 - p(H) and MB will be close to 1 and certainty factor will increase. On the other hand, if the evidence is very weak, then p(H|E) - p(H) is almost zero, and the uncertainty remains about the same with MD (H, E). The function max is used to normalize the MB value positive (between 0 and 1).

$$\mathrm{MB(H, E)} = \begin{cases} 1 & \text{if } \mathrm{p(H) = 1} \\ \dfrac{\max\left(\mathrm{p(H \mid E), p(H)}\right) - \mathrm{p(H)}}{1 - \mathrm{p(H)}} & \text{otherwise} \end{cases} \tag{2}$$

MD(H, E): measure of increased disbelief on hypothesis H influenced by evidence E [9]. If the evidence, E, is very strong, p(H) – min(p(H|E), p(H)) will equal 0 and MD will be 0. On the other hand, if the evidence is very weak, then p(H) – p(H|E) is almost p(H), and the uncertainty will be close to 1. The purpose of function min is to make the MD value positive.

$$\mathrm{MD(H,\,E)} = \begin{cases} 1 & \text{if } \mathrm{p(H)} = 0 \\ \dfrac{\mathrm{p(H)} - \min\left(\mathrm{p(H \mid E)},\, \mathrm{p(H)}\right)}{\mathrm{p(H)}} & \text{otherwise} \end{cases} \tag{3}$$

To find MB(H, E) and MD(H, E), the probabilities of hypothesis p(H) and the conditional probability p(H|E) need to be determined. For calculating the probabilities, we enumerate 77 possible cases based on Table 2. For example, to find the probabilities for *sleeping*, we found 6 possible cases with 2 components (bedroom door and bed) and one semantic (if the activity is exclusively performed or not). Then the number of all possible cases is three (detecting bedroom door, bed and both bedroom door and bed) for each semantic case. In TAM, *bedroom door* and *bed* are equally treated as artifacts. In GAM, *bedroom* is an object and *bed* is a tool for lying down. S-GAM adds semantic information in the GAM model. We counted activities for each of the evidences. Table 3 shows an example of *Sleeping*. Sum of probability is calculated using the addition law of probability that is the probability of A or B is the sum of the probabilities of A and B, minus the probability of both A and B. The probabilities of other activities are calculated similarly.

Table 3. The probabilities of *sleeping* activity according to the models.

	Evidence (E)	p(H and E)	p(E)	p(H|E)	Sum of p(H|E)
TAM	Artifacts	2	6	0.33	0.33
GAM	Tool	1	4	0.25	0.25
	Object	1	4	0.25	0.44
S-GAM	Tool	1	4	0.25	0.25
	Object	1	4	0.25	0.44
	Semantics	2	7	0.29	0.63

Table 4 represents the conditional probability of hypothesis H given E for every activity. We can observe some semantics are highly evidential whereas some are not according to how much the semantic contribute for identifying activities. However, S-GAF has higher probability overall because it is based on GAF.

Table 4. The probabilities of hypothesis H given evidence E for each model, p(H|E).

Activity	TAM	GAF (Tool)	GAF (Object)	GAF (Tool and Object)	Semantics	S-GAF (GAF and Semantics)
Sleeping	0.33	0.25	0.25	0.44	0.29	0.60
Taking a bath	0.38	0.00	0.38	0.38	0.43	0.64
Using the toilet	0.29	0.00	0.40	0.40	1.00	1.00
Cooking	0.35	0.40	0.40	0.64	0.72	0.90
Drinking	0.06	0.25	0.04	0.28	0.33	0.52
Washing dishes	0.16	0.88	0.19	0.90	0.21	0.92
Leaving the house	0.5	0.00	0.50	0.50	0.14	0.57

Using the estimated probabilities, we computed the certainty factor. Fig. 4 shows the certainty factor for each activity.

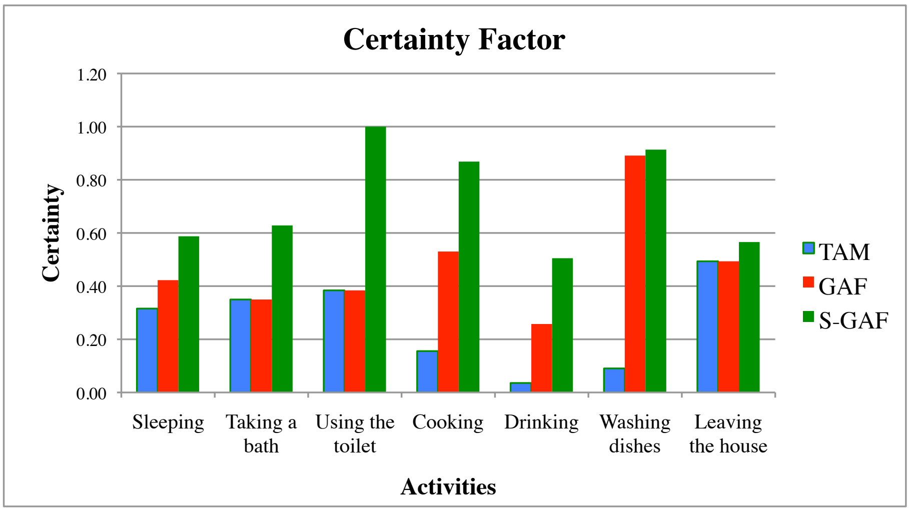

Fig. 4. Uncertainty according to activities. It compares uncertainties of TAM, GAF and S-GAF for each activity.

Fig. 4 shows that S-GAF has higher certainty for all activities and GAF model has higher or comparable certainty to that of TAM. This is an obvious result because S-GAF and GAF models provide more evidence than TAM and GAF models respectively. If no tool is used for an activity like *using the toilet* or *leaving the house*, GAF and TAM show comparable certainty. We also can see that *cooking*, *drinking*, and *washing dishes* in TAM have low certainty compared to other activities because their tools or objects have low evidential certainty. The low evidential certainty may be attributed to artifacts such as a pan or a cup being used in multiple activities. Also, we can observe that the certainty of S-GAF for *using the toilet* and *cooking* have significant difference with other models. It is because the semantic information for *using the toilet* and *cooking* are more activity centric compared to other activities. For example, mutuality semantic is applied for several activities such as sleeping, *taking a bath* or *leaving the house*. On the other hand, the order semantic for *cooking* is only for the *cooking* activity and it is not applied for another activity of the scenario in Table 2. Therefore, the semantic is highly evidential.

5 Conclusion

Accurate activity modeling is important for increasing activity recognition (AR) performance because AR model affects other AR techniques, which are based on the AR model. However, the characteristics of human activities such as complexity, ambiguity or diversity make accurate activity modeling very challenging. In order to address the challenges, we propose a new activity modeling technique, which is based on both generic activity framework and activity semantic knowledge. The generic activity framework is a refinement of the classical activity theory. And the proposed approach adds meaningful semantic knowledge to the generic activity framework for representing activities more precisely. A major advantage of the proposed approach is that it can represent real world activities accurately by using the eight components of our generic activity framework along with the activity semantics introduced in this paper. This advantage implies reducing a great deal of uncertainty that may inherently exist in the activity model. Therefore, our modeling technique does increase the performance of activity recognition.

References

1. Helal, A., King, J., Zabadani, H., Kaddourah, Y.: The Gator Tech Smart House: An Assistive Environment for Successful Aging. In: Advanced Intelligent Environments, H. Hagrass, (eds), Springer Verlag (2008)
2. Mann, W., Helal, S.: Smart Technology: A Bright Future for Independent Living. In: The Society for Certified Senior Advisors Journal, vol. 21, 15-20 (2003)
3. Kim, E.J., Helal, S., Cook, D.: Human Activity Recognition and Pattern Discovery.: IEEE Pervasive Computing vol. 9, n. 1, 48-52 (2010)
4. Kim, E.J., Helal, S.: Revisiting Human Activity Frameworks In: 2nd International ICST Conference on Wireless Sensor Network Systems and Software - S-Cube, (2010)
5. Kuutti, K.: Activity theory as a potential framework for human-computer interaction research. In: B.A Nardi (eds.), Context and consciousness: Activity theory and human-computer interaction. Cambridge, MA: MIT Press (1996)
6. Davydov, V. V., Zinchenko, V. P., Talyzina, N. F.: The Problem of Activity. In: the Works of A. N.Leontjev, Soviet Psychology, vol. 21, n. 4, pp. 31-42 (1983)
7. L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2):257–286 (1989)
8. C. Sutton and A.McCallum, An introduction to conditional random fields for relational learning, In: L. Getoor and B. Taskar, (eds), Introduction to Statistical Relational Learning. MIT Press, 2006.
9. David Irwin J: The industrial electronics handbook, pp. 829. IEEE Press (1997)
10. Kusrin, K.: Question Quantification to Obtain User Certainty Factor in Expert System Application for Disease Diagnosis. In: Proceedings of the International Conference on Electrical Engineering and Informatics, pp. 765-768 (2007)
11. Pentney, W., et al.: Sensor-Based Understanding of Daily Life via Large-Scale Use of Common Sense. In: Proceedings of AAAI '06, Boston, MA, USA (2006)
12. Surie, D., Pederson, T., Lagriffoul, F., Janlert, L.-E., Sjolie, D.: Activity Recognition using an Egocentric Perspective of Everyday Objects. In: UIC (2007) Proceedings of IFIP 2007 International Conference on Ubiquitous Intelligence and Computing. LNCS, vol. 4611, pp. 246–257. Springer, Heidelberg (2007)
13. Lefre, P.: Activity-based scenarios for and approaches to ubiquitous e-Learning. In: Personal and Ubiquitous Computing, vol. 13, n. 3, pp. 219-227 (2009)
14. Pentney, W., et al.: Sensor-Based Understanding of Daily Life via Large-Scale Use of Common Sense. In: Proceedings of AAAI '06, Boston, MA, USA (2006)
15. Kasteren, T. Noulas, A. Englebienne, G., Krose, B.: Accurate Activity Recognition in a Home Setting. In: Proceedings of the Tenth International Conference on Ubiquitous Computing (Ubicomp), Korea, pp. 1-9 (2008)