# OPTIMIZING RECOMMENDATIONS WITH AN RFID-BASED APPROACH FOR THE TAIWAN TRAINQUALI DOCUMENT MANAGEMENT SYSTEM

Chao-Tsong Fangtsou
National Taipei University
151, University Rd., San Shia District, New Taipei City 23741, Taiwan
ctfang@gmail.com

Rui-Yang Chen Aletheia University 32, Zhenli St., Danshui Dist., New Taipei City 25103, Taiwan a168.jacky@msa.hinet.net

Ming-Hung Yang
National Taipei University
151, University Rd., San Shia District, New Taipei City 23741, Taiwan s710036110@webmail.ntpu.edu.tw

#### **ABSTRACT**

Document management is a method to organize current knowledge and an appropriate document management method can decrease the time necessary to search for knowledge and also actively recommend a suitable document for a user thru an algorithm function. This paper proposes a document management system that combines RFID (Radio Frequency Identification) and an optimizing recommendation method. This system can immediately and intelligently obtain document information and provide suitable recommendation results based on a user's personal usage behavior and document characteristics and per document regulations of the TTQS (Taiwan TrainQuali System) standard. As such, the system furthers collaborative operation between organizations in document sharing. This paper describes the document management method in terms of an active response and communication platform that helps to consolidate and improve management efficiency and open up potential business opportunities.

**Keywords:** RFID-Based, Collaborative Filtering, TTQS Document Management, IoT, Optimizing Recommendation

## 1. INTRODUCTION

Along with the emergence of a knowledge economy, enterprises are increasingly focusing on knowledge asset management. Knowledge can be thought of as information that is accumulated, analyzed and classified, such that it becomes to a higher level of data. Because knowledge is an asset that an enterprise accumulates over time, it is important to both tap into this data and to not lose it. As such, it is an important for enterprises to conserve these knowledge assets. Knowledge can exist as a tangible document or digital record and is typically managed thru classification, marking and filing to maintain a link between users and document files. Physical document management tools normally include folders, filing cabinets, stickers or tags that classify or tag these documents with their content or name. These tools, however, do not offer a fast query or ready path to retrieve a document, especially when data increases to a large quantity. In fact, the subtle relationships between documents are hardly ever tapped. Although using electronic records to manage physical documents can achieve fast query, there is still much lacking in this approach. For example, the status of documents has to be reviewed by a human, where an actual document location by shelf number needs to be verified, and without much insight into the actual document contents. Additionally, physical documents may be placed incorrectly and often requires manual sorting and document checks. Importantly, there is little ability to predict and recommend selections based on user behavior or document-use<sup>1</sup>.

Other current document management tools and platforms still have some deficiencies in terms of collaborative operation in organizations cross-enterprise. For example, when enterprises transfer electronic files to each other, the files will have attached a physical corroboration document for authentication or certification. Currently, there are no document management systems that apply cross-enterprise to physical documents that addresses missing or lost physical documents when these items are transferred between enterprises<sup>2</sup>.

An intelligent document management platform that provides immediate status and actively retrieves relationships between documents to provide recommendations would be both convenient and value-added. As such, this research will construct an Internet of Things field thru the use of an RFID-based approach with an optimization recommendation algorithm to create an intelligent document management platform.

In this field, documents can reply to a query request and respond with its location and status, and also recommend related documents based on document characteristics or roles. This recommendation will improve the ability of the document management to add value. Additionally, all of the various roles of participants in an inter-organization using this system can receive document information with format consistency in information and electronic file delivery. The platform has an electronic authentication function that improves upon operational efficiencies over physical material-based collaboration.

A summary of these study purposes are provided below:

- Application of an RFID-based approach in a document-aware field to construct functions such as immediate documents status query or location search.
- Determination of TTQS document relative rules that reminds a user of other related documents that are available and authentication of users and files that are or not suitable for use.
- Introduction of an optimization recommendation algorithm for the document management system that can predict relevant documents based on user behavior.
- Construction of an inter-organization collaborative document management platform that reduces the time necessary for document transfer between enterprises, tracks document position, unifies file format and integrates the document transfer flow.

As the variety of possible documents can be complex and difficult to classify, the study scope is limited to documents produced in TTQS phases. This includes documents such as provisions, manuals, forms, lecture records, personnel information, etc.

## 2. LITERATURE REVIEW

# 2.1 Internet of Things

The Internet of Things (IoT) is a novel network architecture that links and communicates between physical objects. It is different from a traditional Internet that is linked with physical objects and an Internet that links via the TCP/IP protocol and server. In the IoT, objects communicate and exchange information with each other through radio-frequency identification, a wireless sensor network and an embedded chip. The Internet of Things must necessarily be the result of synergetic activities conducted in different fields, such as telecommunications, informatics, electronics and social science<sup>3,4</sup>. There are multiple possible application domains for an Internet of Things such as in energy management, transportation, logistics and healthcare.

In an Internet of Things, however, there are some issues in terms of security and privacy of applications and include concerns such as data authentication, access control and client privacy. To address these concerns, we implement public-key cryptography, a key management system to satisfy security concerns. As part of the infrastructure, it will also be necessary to establish suitable provisions and regulations through international conferences, standards organizations, an intellectual property rights framework, and such to protect users in an Internet of Things<sup>5,6</sup>.

# 2.2 Radio-Frequency Identification

Radio Frequency Identification (RFID) is the one possible tool to construct an Internet of Things infrastructure. RFID is a popular wireless communication technology and has been developed for over sixty years. RFID has many positive characteristics, such as an ability to be identified, is quickly read, small in size, low cost, durable and offers large storage. Further, in comparison to barcodes, it quite suitable for attachments on objects to provide object tracking and status detection. Using RFID tags as a product sensor unit allows for immediate transfer of information to a back office system and for example can be used at a warehouse center to control inventory of less-popular items and yet allow for a rapid response to customers<sup>7,8</sup>.

The operational units of an RFID system includes tags, readers, middleware and a back office system. An RFID tag can be classified into active tags and passive tags. Passive tags cannot transfer signals independently and only broadcasts a signal when a reader provides energy to the tag. Active tags have batteries so that allow signals to be actively broadcast to a reader.

Tags are chips that record information and the composition can differ depending on how it will be used. A reader is a device that reads tag content and typically consists of a control unit and antenna, and can be fixed or handheld. Middleware and back office systems receive data from a reader.

During operation of a system, the middleware will filter and analyze received data and transfer this to the back-office system, according to the purpose and function of a back-office system tailored to various operational requirements. There are four parts to middleware functions: ordination RFID devices, data filtering, data collection and data path transfer, and integration of procedure management.

RFID technology has been applied to enterprises around the world. For example, Wal-Mart is a well-know U.S. enterprise that has asked its logistics providers to introduce RFID technology to facilitate coordination

with Wal-Mart. There are increasingly more cases of successful introduction of RFID technology<sup>2</sup>. In fact, many document management systems already use RFID technology. With the introduction of RFID technology, the status of a document becomes obvious and clear and it is possible to quickly track document location via the RFID technology and wireless sensors. RFID tags can store a large amount of information. It can also serve as identification to protect documents from unauthorized access. A document manager using the system also has the ability to determine document status and location via the position function of the RFID and can locate documents that may have otherwise been lost or gone missing during transfer procedures. Enterprises such as 3M and Book Tec have introduced RFID technology applications for document management<sup>2</sup>.

# 2.3 Document Management System

Document is a method to store current knowledge by recording data or words through paper documents or electronic files for future query. A document management system helps users to manage these documents thru various procedures like access, search and edit. In general, a document management system provides three basic functions for users: document access control, document sharing and document storage management.

A document management system can provide many advantages for both an individual and an enterprise. For example, an individual user can search for documents and collect information quickly without concerns about time or location. For an enterprise, operation efficiencies can come from effectively storing, retrieving and leveraging the knowledge and experiences accumulated from business cases. The ability to access work documents and integration with Web 3.0 concepts can provide added convenience and support more flexible work styles. Some research is looking at combining document management systems with collaborative filtering<sup>9,10</sup>. Lastly, security in document management systems is also a popular related field of study.

# 2.4 Optimizing Recommendation and Collaborative Filtering

Optimizing recommendations is the concept of creating a precise and logical resource allocation method. This consists of setting up a consistent index by keyword or quantifying the request of the entity requesting a resource or based on other characteristics of a resource. This type of system determines an ideal resource match thru this index. Collaborative filtering is one example of an optimization of recommendation method.

"Collaborative filtering" refers to a method that uses a group of users with the same characteristics, such as preferences or behavior patterns from their usage behavior or a rating on resources for recommendations to a decision maker. The advantage of a collaborative filtering system is that it allows users who know their preferences, but without experience, the ability to refer to similar users with experience and suggestions <sup>11,12,13,14</sup>.

There have two stages in the use of collaborative filtering to predict user preferences:

- Search for users with similar past value ratings.
- Use ratings to calculate a prediction index.

In addition, there are two approaches to collaborative filtering: a memory-based and a model-based approach. A memory-based method uses the data or records of other users who have experience and ratings to calculate the similarity between users <sup>15</sup>. This method determines which cluster that a user belongs to and as such provides recommendations. This is also sometimes called correlation-based collaborative filtering or nearest-neighbor collaborative filtering. A model-based approach will train rating data to develop a user behavior model that makes predictions about future user behavior. Compared with a memory-based approach, the model-based approach has a faster prediction time, but requires some data training time<sup>15</sup>.

# 2.5 Taiwan TrainQuali System

The Taiwan TrainQuali System (TTQS) is a standard proposed by the Bureau of Employment and Vocational Training (BEVT) Council of Labor Affairs in Taiwan<sup>16</sup>. This system ensures quality training across enterprises by maintaining quality and performance during staff training, and also by improving the ability and tendency of enterprises to hold training activities, with subsequent skill advancement of the national human resource pool and accumulation of training knowledge<sup>17,18</sup>.

TTQS evaluates training quality based on a "PDDRO" standard that consists of five procedures with eighteen scored points. "PDDRO" stands for planning, design, delivery, review and outcome. These procedures thus capture all phases of enterprise training activities. Specific targets within each procedure are described below:

 Planning: In this procedure, enterprises examine whether or not the training target and the enterprise operation goals are consistent. Enterprises have to link both to make training objectives obvious and convey this to trainees.

- Design: According to training objectives and targets, design training courses and activities. Both course content and training objectives should be consistent. Definite and suitable course specifications and passing standards should be established.
- Delivery: Attention to whether the enterprises utilize resources efficiently during training course execution, such as with lecturers or equipment resources. Whether enterprises save and archive training documents, records or manuals can be reviewed or searched in the future.
- Review: Enterprises should evaluate every phase in training plan that are actually implemented and determine whether the targets and performance were consistent and completed.
- Outcome: In this phase, the results of the review phase are transferred into quality management manuals and all relative documents, such as textbooks, class records and trainee grades should be prepared by enterprise such that assessment committees can evaluate the training result from these documents. The committees generally use a TTQS scorecard to score. Every committee should have the same scales for evaluation to ensure consistency and outcomes objectivity<sup>19</sup>.

# 2.6 Fuzzy Theory

Fuzzy Theory was proposed by L.A. Zadeh in 1965. It is based on fuzzy set theory and integrated with concepts such as probability and membership function to explain the uncertainty of things. The object status in a set can be describe with belong and not belong. For example, there is one set,  $A = \{x=1,2,3 | x \in R \}$ , and we can divide all integers into two classes. Belong A set and not belong and transfer this to a binary status to mark. This partition method is called a crisp set that has two statuses, A and not A. For all integers, the scale of the A set is accurate and will not be affected by any element. Compared with a crisp set, a fuzzy set doesn't have an accurate scale and the partition result will be affected within the closed level of the elements in a set. The closed level of elements and fuzzy set is normally called a membership degree. This can be presented as a membership function form. The description of fuzzy set is belong: if there is a fuzzy set A, then A'=  $\{(x, \mu A'(x)) \mid x \in R \}$ . Where  $\mu A'$  is a membership function that calculates whether the membership degree to set A' of element x is greater than a threshold value to decide x is belong A' or not.  $\mu$ A'(x) means membership degree of element x to set A'. The membership function can be in different forms per user definition, such as with triangular membership functions, trapezoidal membership functions, a Z-type membership function or a Gaussian function. The number of partitions is determined by

experiential knowledge from the designer and adopted with corresponding membership functions<sup>20,21</sup>. Figure 1 shows the architecture of fuzzy theory.

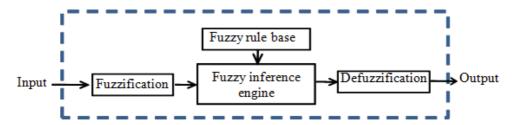


Figure 1. Fuzzy theory architecture

#### 3. RESEARCH DESIGN

This section presents the system architecture and the recommendation method. We introduce an overview of the document management system and then describe the recommended rules and document relationships, along with user behavior and the real-time status of documents. Finally, some examples of the system are provided.

# 3.1 System Architecture

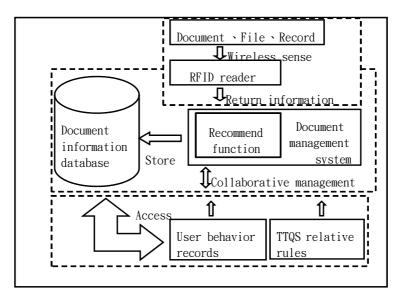


Figure 2. Platform architecture

This study proposes a platform for a document management function to improve document management work efficiency and to recommend relevant documents. This platform consists of two main processes. Firstly, we use RFID devices such as reader or antenna combine with RFID tags attached to

a document to construct an Internet of Things field. In the field, all TTQS documents can transfer data actively through a wireless sensor network and documents are managed thru this RFID-based approach. Secondly, a document recommendation function is constructed based on three components: recommendations per TTQS relation rules; recommendations per user behavior information integrated with collaborative filtering; and recommendations based on space or distance information collected by the Internet of Things. These three components comprise an operation to give a recommended value that then recommends document use order for a user. The platform architecture is show in Figure 2. The recommendation procedure is shown in Figure 3.

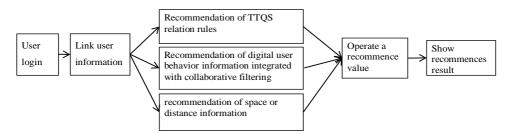


Figure 3. Recommendation procedure

#### 3.2 Recommendation of TTQS Document Relative Rules

#### 3.2.1 TTQS Document Relative Rules

TTQS operating procedures will follow the five PDDRO phases and every phase will produce related documents. Because of corresponding or an ordered relationship between the documents, this study creates a document relationship table from related literature that then uses these relationship rules to make recommendations.

Table 1 is an example of a document relationship table and the content consists of related documents with an example document per the five PDDRO phases.

<b>Example document: Future business direction and operational planning</b>					
documents of training department					
<b>Related documents</b> •Annual training goal, development planning and					
in planning phase	relative activities solution				
<b>Related documents</b>	<ul> <li>Documents of customers or market training</li> </ul>				
in design phase	required or professional abilities gap analysis •				
Related documents in delivery phase	•Learning experience, race records, practical training records or other comments of trainees in documents to show learning result •				
Related documents in review phase					
Related documents	•Final report				
in outcome phase	•Learning assessment regulations and records				

**Table 1.** An example of a document relationship table

This is an example of a document for "Future business direction and operational planning documents of the training department". When a user opens this file and the user is in a planning phase, the user will receive recommendations per TTQS document relation rules for documents such as "annual training goal", "development planning and relative activities solution", "documents of customers or market training required", "professional abilities gap analysis", "learning experience, race records, practical training records or other comments of trainees in documents to show learning result", "final report and learning assessment regulations and records". If the user leaves the planning phase, the system will check for whether or not these related documents were archived and give a warning.

## 3.2.2 Document Management System Roles

This study defines the roles of the TTQS operating process and based on roles makes document recommendations. The definitions are as below:

- 1. Trainee: a trainee indicates staff in training who were chosen by an enterprise, and can be either workers or managers. The possible documents for a trainee include attendance records, school textbooks, test papers, training performance and training regulations papers.
- 2. Trainee manager: a trainee manager is a supervisor from the original enterprise. The supervisor has to provide basic information about trainees and track student learning outcomes and status, and as such possible trainee manager documents include basic information about trainees, student learning outcomes and attendance records.

- 3. Enterprise for training: the enterprise for training is the department sending staffs for training. When the training department designs and plans courses or activities, they have to consider that the training results and operational goal of various enterprises may not be the same. As such, the enterprise for training has to provide the documents about operational planning, training goals and expected results.
- 4. Training department: the training department means the department providing training courses and activities. The training department has to consider the targets and expected results of an enterprise in regards to training and also follow TTQS regulations. As such, training department documents will include information about operational planning, training goals and expected results and documents that correspond to review rules.
- Lecturer: a lecturer will teach and explain in a subject area or on training techniques. Possible documents for lecturer include definitions, production and management rules of internal lecturers and attendance records.
- 6. The training location owner: a training location owner provides classroom or facilities for training. Possible documents for a training location owner includes information on safety certificates and rental details.
- 7. Committee: committees are comprised by experts who have been invited by the Bureau of Employment to review and score an enterprise for training and the training department during the TTQS review phase. As such, possible documents include scorecards, satisfaction feedback forms of trainees and lecturers and the training final report.

# 3.3 Digital Information Matching Method

The digital information matching method is characterized by the relationships between users and documents. Thus, we transform these relationships to quantitative indicators to evaluate the similarities of users, and to find groups that are similar with the primary user, thus giving recommendations to predict the behavior and preferences of the primary user per those of group users. For example, on a decision tree path, if two paths are similar in the procedures for making a decision, they will possibly be place on the same leaf node. As such, it is possible to determine that two similar results will also have two similar decision paths. If we use a user's behavior as a pattern and then proceed to find a similar group user, the user in this group will have some similar behaviors with the primary user. As

such, we can use the document-use record of a similar group user to predict and recommend documents to the current primary user.

Based on above statement, this study begins with recommendation procedures that are combined with collaborative filtering and a K-NN algorithm. This study uses the number of user access documents as a character, and the matching procedures are shown in Figure 4. Firstly, we converted the attribute into an input value of the fuzzy function, and split this fuzzy set to some necessary subsets. Then we design the membership function for every subset. Lastly, we classify the relationship between user and document to a fuzzy subset and in a corresponding calculation, we used the k-NN algorithm to find a similar user group for the primary user and provide a recommendation based on user behavior or preference.



Figure 4. User and document matching procedure

#### 1. Define variable

This study uses the distance of the user access records between users to calculate similarity. Table 2 is an example for a user, and the number in the table means document cumulative number of users used.

Table 2. Example	of document	t cumulative	number of	t user	used

User Item	User A	User B	User C
Item 1	4	2	4
Item 2	1	7	2
Item 3	0	1	3

#### 2. Design membership function and split fuzzy set

We used document cumulative number of users used. For example, we converted the number of used into the form of an input value by equation (1). This reflects the proportion of the number of users used and the documents used. We set the input value as I, users are  $u1 \cdot u2 \cdot u3...un$ , documents are  $d1 \cdot d2 \cdot d3...dm$ . Use (u, d) means the number of users that

use Document d. So the equation to calculate the relation level between User un and Document dm is:

$$I_{(u_n,d_m)} = \frac{Use(u_n,d_m)}{\sum_{x=1}^n Use(u_x,d_m)}$$
(1)

After calculation, we can convert Table 4 into the input value form that is shown in Table 3.

**Table 3.** Example of the proportion of users used and the document used

User	User A	User B	User C
Document 1	0.4	0.2	0.4
Document 2	0.1	0.7	0.2
Document 3	0	0.25	0.75

In this study, the degree of correlation between the user and the document is divided into three fuzzy sets such as "high", "medium" and "low" with a triangle, Z-type and S-type membership function. The "high", "medium" and "low" fuzzy sets means the relationship degree between users and documents. When we know that a specific user often used a specific document, then the relationship between the user and document was higher through equation (1). These three membership functions of the fuzzy subset are shown in equation (2), (3) and (4) with the similar levels of "high", "medium" and "low".  $\mu_{low}$ ,  $\mu_{medium}$ ,  $\mu_{high}$  stand for the input value  $I(u_n, d_m)$  of the relationship between the user and document. The relationship between the user and document can be classified into "high", "medium" or "low" level by membership degree number. The definitions of the three relation levels are shown in Table 4. Figure 5 is membership function graph.

$$\mu_{low}(x) = \begin{cases} 1 & \text{if } x \leq 0.25 \\ 0.5 - x & \text{if } 0.25 \leq x \leq 0.5 \\ 0 & \text{if } x \geq 0.5 \\ 0 & \text{if } x \leq 0.25 \\ 0.5 - 0.25 & \text{if } 0.25 \leq x \leq 0.5 \\ 0.75 - x & \text{if } 0.25 \leq x \leq 0.5 \\ 0.75 - x & \text{if } 0.5 \leq x \leq 0.75 \\ 0 & \text{if } x \geq 0.75 \end{cases}$$

$$\mu_{high}(x) = \begin{cases} 0 & \text{if } x \le 0.5 \\ \frac{x - 0.5}{0.75 - 0.5} & \text{if } 0.5 \le x \le 0.75 \\ 1 & \text{if } x \ge 0.75 \end{cases} \tag{4}$$

**Table 4.** Definitions of three relationship levels

Level	Definition	Explanation
Low	$\mu_{low}$ gets highest membership degree of the input value $I_{(u_n,d_m)}$ between user and document in three membership functions.	The proportion of user u access document d times in total document d access times is less relatively.
Medium	$\mu_{medium}$ gets highest membership degree of the input value $I_{(u_n,d_m)}$ between user and document in three membership functions.	The proportion of user u access document d times in total document d access times is normal relatively.
High	$\mu_{high}$ gets highest membership degree of the input value $I_{(u_n,d_m)}$ between user and document in three membership functions.	The proportion of user u access document d times in total document d access times is more relatively.

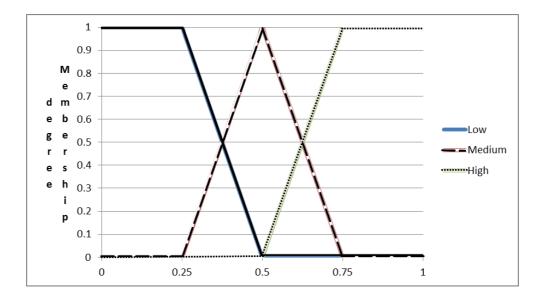


Figure 5. Membership function graph

#### 3. Calculate the membership degree with primary user

We can convert user used record into a fuzzy level when we obtain a membership degree to find the preference of the user and document. We can use the input values in Table 3. This is then used to make a calculation with the membership function and the result is shown in Table 5.

User	User A	User B	User C
Document 1	Low	Low	Medium
Document 2	High	High	Low
Document 3	Low	Low	Medium

**Table 5.** Correspondence fuzzy result of user and document

#### 4. Find similar user group by K-NN algorithm

The relationship between all users and documents are shown in Table 5 and the similarity is presented by vector. The membership degree of two users A, B with a document is B(UserA, UserB), if the membership degree of A, B is identical then we give it a value "0". If the membership degree of A, B is different then we give it a value "1" and we will obtain a vector between any two users Vector(A, B) = (B1, B2, ..., Bm). The vector length DT (A, B) means the distance between User A, B and it also means similarity. If the DT value is higher, then the similarity of User A, B is less. On the contrary, the similarity of User A, B is higher. The distance formula is shown in equation (5).

$$DT(A \cdot B) = \sqrt{\sum_{x=1}^{m} [B_{d_x}(User_A, User_B)]^2}$$
 (5)

For example for Table VII, the membership degree of User A and User B for document 1 is  $B_1(User_A, User_B) = 0$ , the membership degree of User A and User C for document 1 is  $B_1(User_A, User_B) = 1$ , and so on. Such that vector (A, B) = (0, 0, 0), vector(A, C) = (1, 1, 1). Therefore, DT(A, B) = 0,  $DT(A, C) = \sqrt{3}$ . When we get the distance of any two users, we can choose a similar user with the K-NN algorithm. If we use K=1 to choose with the example above for User A, the similarity of User B is higher than User C for User A and we can predict the preference of User A by User B's behavior.

#### 5. Predict and recommend by similar user's behavior

In this phase, we can use similar user group behavior that we get above to make a recommendation to the primary user. According to the frequencies that a user accesses a document, this recommendation process can make recommendations to a user based on previous access to various documents.

# 3.4 The Recommendation Process in an Internet of Things

Because of application via an Internet of Things and the RFID-based approach, the objects in this field are intelligent, such that they can transfer status and information through a wireless system and replace traditional manual processes. This application also helps in making recommendations regarding real-time information to calculate recommendations that were previously not considered, such as document real-time distance, document real-time status, etc. While an object has to be seen and keyed-in manually before being in the system, subsequently the Internet of Things can automatically capture object information anytime. This makes it possible to obtain real-time information to make suitable user recommends. In this section, we will design calculation and recommendation procedures with real-time information.

#### 3.4.1 Document Real-Time Distance

Sometimes documents will be held by different people or departments, and the physical distance between a user and documents can influence the decision of user as to whether or not they want the documents. Therefore, we considered distance as an important attribute in a recommendation.

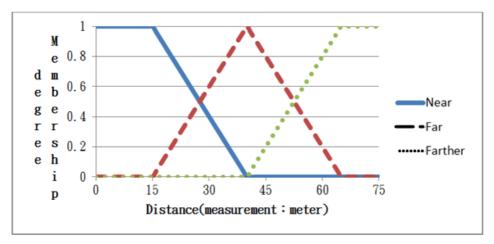
#### 1. Define variable

In this study, we use the distance between a user and document as an input and distance is measured by meter. Table 6 is an example of distance between a user and document.

User	User A	User B	User C
Item 1	40	60	150
Item 2	120	40	20
Item 3	60	80	40

#### 2. Design membership function and split fuzzy set

Because the definition of the document in this study is TTQS training, the training locations are usually a classroom, office and archive facilities. These may be outside of a user's current building. Therefore, the levels of distance between user and document are divided into three types: "near", "far", and "farther". The "near" location reflects a classroom, the "far" location is in the building and the "farther" location is outside of the building. These three membership functions of a fuzzy subset are shown in equations (5), (6) and (7) with the similar level of "near", "far", and "farther".  $\mu_{low}$ ,  $\mu_{medium}$ ,  $\mu_{high}$  stands for the distance between the user and document with a membership degree in these three fuzzy subsets. Figure 6 shows a membership function graph.



**Figure 6.** Membership function graph for the distance between user and document

$$\mu_{low}(x) = \begin{cases} 1 & \text{if } x \leq 15 \\ \frac{40 - x}{40 - 15} & \text{if } 15 \leq x \leq 40 \\ 0 & \text{if } x \geq 40 \\ \frac{x - 15}{40 - 15} & \text{if } 15 \leq x \leq 40 \end{cases}$$

$$\mu_{medium}(x) = \begin{cases} 0 & \text{if } x \leq 15 \\ \frac{x - 15}{40 - 15} & \text{if } 15 \leq x \leq 40 \\ \frac{65 - x}{65 - 40} & \text{if } 40 \leq x \leq 65 \\ 0 & \text{if } x \geq 65 \end{cases}$$

$$(6)$$

$$\mu_{high}(x) = \begin{cases} 0 & \text{if } x \leq 40 \\ \frac{x-40}{65-40} & \text{if } 40 \leq x \leq 65 \\ 1 & \text{if } x \geq 65 \end{cases} \tag{8}$$

After designing the membership function, we can convert the actual distance between the user and document thru a sensitive technique that converts the input value to the fuzzy membership function in a calculation to get the fuzzy level. Table 7 is the result after input of the variables in Table 6 into the membership function.

Document	r User A	User B	User C
Document1	Far	Farther	Farther
Document2	Farther	Far	Near
Document3	Farther	Farther	Far

**Table 7.** Document distance calculation results

#### 3. Calculate the membership degree with primary user

We can get the membership degree of a document and user through Table 7 and present the distance between them by fuzzy level. Because the distance will be considered in the recommended calculation, we can give it a weight based on three levels of the distance fuzzy level that can be calculated for the recommendation.

#### 3.4.2 Real-time Document Status

#### 1. Define variable

Through the Internet of Things, object status can be detected in real-time, according to different statuses that the recommendation system will change prior to document order. In this study, the statuses of documents are defined with "on board", "borrowed", "reviewed", "edited" and "unusual". The statuses are explained below:

- On board: The document is on board or in the archive and available for lending to the the user.
- Borrowed: The document has been borrowed by another user and is not in the enterprise operation process.
- Reviewed: The document has been borrowed by another user and is in the enterprise operation process.

- Edited: The document has been borrowed by another user and is being editing.
- Unusual: The document is lost or may be damaged.

## 2. Design recommendation level

According to the document status design above, this study classifies these statuses into three levels. The numbers from low to high represents a prior order. The recommended level will be a part of the total recommendation formula when constructing a complete recommendation model. Table 8 is a classification table of the document status and levels.

**Table 8.** The recommended order of document real-time status

Recommended level	1	2	3
Document real-time status	On board	Borrowed Reviewed Edited	Unusual

#### 4. CITATIONS

## 4.1 Design and Implementation of Internet of Things

In the past, the relationship between documents had to queried with electronic records thru manual instructions in an e-document management systems, such as per document attribute or document life cycle. In this system, document storage location is determined by deployment of antennas and via an RFID reader that allows users to quickly search for many documents. The documents will also indicate their location via active RFID tags. Figure 7 show the document management operation model in an Internet of Things concept.

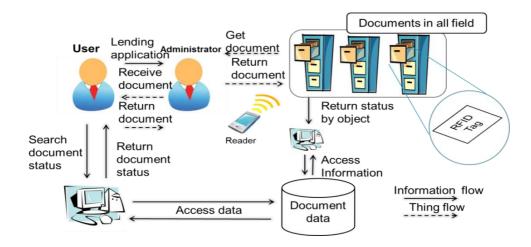
The Internet of Things framework includes three layers: application, network and sensing. In this study, we will attach the tags on documents and at various service units such as middleware to filter and analyze data. The model of an Internet of Things infrastructure and back-end system is shown in Figure 8.

# 4.2 System authentication

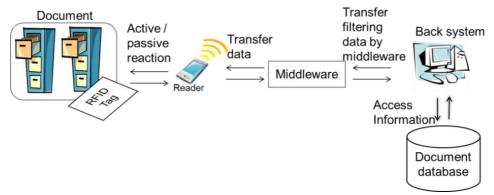
This study compared existing document systems with the novel document management system integrating an RFID-based approach and/or collaborative filtering functions. These comparisons were done to ensure

that our proposed management method would improve the efficiency of document management processes. We compared the following:

- Collaborative filtering without a wireless sensor network
- A wireless sensor network without collaborative filtering.



**Figure 7.** The Internet of things model



**Figure 8.** Document information transfer in an Internet of Things

## 4.2.1 Using Collaborative Filtering Management Only

Using collaborative filtering to manage documents can make predictions about a target user with a user group that is similar to the target user<sup>22</sup>[21]. Existing document management systems that are integrated with collaborative filtering help a user to find the documents that are suitable with user's requirement or behavior and decreases time costs for a user. Thus, the application of collaborative filtering can improve search efficiency, but can't help a user to obtain agent actual document. The collaborative

filtering method provides information on documents needed by the user, but without any way to readily obtain the documents. Our study proposed a document management model that provides recommendations on documents needed and also provides document location information via RFID. This model can improve physical document management and access.

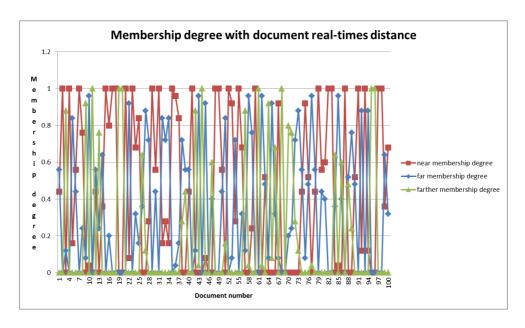
## 4.2.2 Using Wireless Sensor Network Only

The RFID technique has been applied to logistics and storage management for many years. The ability to indicate location, track or identify can help to improve management efficiency. There also have many examples of document management and related products utilizing RFID tags, such as with sticky notes or staples<sup>2</sup>. But existing RFID document management systems lack of some functions that would help a user determine what documents they need. Therefore, we have proposed a document management system that integrates an RFID-based approach and recommendation method to provide user with a solution that provides both physical and logical information.

#### 4.2.3 System Implementation

In this study, we use the PHP programming language and a MySQL database to implement a document management system integrated with a RFID-based approach, a recommendation optimization method and the collection of operating data for verification. The key performance indicators of a document management operation are the number of related documents that a user can reach and the time required for the file transfer process.

Table 9 is an example of 100 documents at random distances. Figure 9 shows the calculated result of 100 documents in terms of membership degree with real-time document distance by the membership degree function as proposed in section 3.4.



**Figure 9.** An example for membership degree with document real-time distance

### 5. PERSPECTIVE

This study proposed a document management model integrated with an RFID-based approach and recommendation optimization method to provide functions such as query in real-real time, active response and across-enterprise collaboration, that also recommends suitable documents for users based on TTQS relationship rules, user behavior and document real-time status. This model actually improves the availability of documents. In this study, we expect this novel document management model can provide a document management system that addresses issues between authentication, and a physical document and information system. As such, this novel document management system can improve document management efficiency and uncover potential business opportunities.

Number	Distance								
1	29	11	70	21	6	31	13	41	13
2	12	12	29	22	38	32	36	42	62
3	62	13	59	23	6	33	33	43	41
4	4	14	31	24	23	34	36	44	66
5	36	15	6	25	19	35	2	45	38
6	26	16	20	26	56	36	16	46	2
7	10	17	10	27	43	37	19	47	55
8	21	18	15	28	33	38	47	48	4
9	63	19	69	29	1	39	51	49	10
10	39	20	67	30	26	40	29	50	26
Number	Distance								
51	44	61	65	71	59	81	25	91	12
52	15	62	41	72	47	82	11	92	37
53	17	63	27	<b>73</b>	43	83	13	93	5
54	33	64	63	<b>74</b>	29	84	56	94	37
55	3	65	42	<b>75</b>	17	85	39	95	68
56	23	66	57	<b>76</b>	27	86	55	96	69
<b>57</b>	62	<b>67</b>	17	<b>77</b>	41	<b>87</b>	9	<b>97</b>	14
58	41	68	70	<b>78</b>	29	88	52	98	7
59	34	69	12	<b>79</b>	11	89	46	99	31
60	1	70	60	80	26	90	27	100	23

**Table 9.** An example for 100 documents with random distances

## 6. REFERENCES

- [1] C.C. Tseng, A Study of RFID Document File Management System. I-Shou university Department of Information Management thesis, 2011.
- [2] K.Y. Tsai, L.R. Chen, and Y.M. Huang, The application of RFID in document management research. GS1 Taiwan report, Winter issue, p42-45, 2009.
- [3] L. Atzori, A. Iera, and G. Morabito, The internet of things: A survey. Computer networks, 54(15), p2787-2805, 2010. http://dx.doi.org/10.1016/j.comnet.2010.05.010.

- [4] M.C. Domingo, An overview of the Internet of Things for people with disabilities. *Journal of Network and Computer Applications*, 35(2), p584-596, 2012. http://dx.doi.org/10.1016/j.jnca.2011.10.015
- [5] R.H. Weber, Internet of Things New security and privacy challenges. *Computer Law & Security Review*, 26(1), p23-30, 2010. http://dx.doi.org/10.1016/j.clsr.2009.11.008.
- [6] R.H. Weber, Accountability in the Internet of Things. *Computer Law & Security Review*, 27(2), p133-138, 2011. http://dx.doi.org/10.1016/j.clsr.2011.01.005.
- [7] K.C. Yeh,R.S. Chen, and C.C Chen, Intelligent service-integrated platform based on the RFID technology and software agent system. *Expert Systems with Applications*, 38(4), p3058-3068, 2011. http://dx.doi.org/10.1016/j.eswa.2010.08.096.
- [8] R.S. Chen, *Introduction to RFID and application*. New Taipei City, Chuan Hwa Publishing Ltd, 2010.
- [9] C.P. Wei, C.S. Yang, and H.W. Hsiao, A collaborative filtering-based approach to personalized document clustering. *Decision Support Systems*, 45(3), p413-428, 2008. http://dx.doi.org/10.1016/j.dss.2007.05.008.
- [10] D.R. Liu, C.H. Lai, and H. Chiu, Sequence-based trust in collaborative filtering for document recommendation. *International Journal of Human-Computer Studies*, 69(9), p587-601,2011. http://dx.doi.org/10.1016/j.ijhcs.2011.06.001.
- [11] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, Recommender systems survey. *Knowledge-Based Systems*, 46(July), p109-132, 2013. http://dx.doi.org/10.1016/j.knosys.2013.03.012.
- [12] J. Bobadilla, A. Hernando, F. Ortega, and J. Bernal, A framework for collaborative filtering recommender systems. *Expert Systems with Applications*, 38(12), p14609-14623, 2011. http://dx.doi.org/10.1016/j.eswa.2011.05.021.
- [13] P.C. Chung, *Contextual Personalized Wearing Recommendation System.* National Cheng Kung University Department of Engineering Sciemce thesis, 2009.
- [14] S.F. Han, *E-newsletter for Transaction Matching by Collaborative Filtering A Case of e-Marketplace in the Food Industry.* Chung Yuan Christian University Department of Information Management thesis, 2003.
- [15] T.H. Dao, S.R. Jeong, and H. Ahn, A novel recommendation model of location-based advertising: Context-Aware Collaborative Filtering using GA approach. *Expert Systems with Applications*, 39(3), p3731-3739, 2012. http://dx.doi.org/10.1016/j.eswa.2011.09.070.
- [16] Bureau of Employment and Vocational Training (BEVT) Council of Labor Affairs. *Taiwan TrainQuali System*. Retrieved on April 4, 2012,

- from http://ttqs.evta.gov.tw/.
- [17] Investors in People. *Welcome to Investors in People*, Retrieved on April 28, 2012, from http://www.investorsinpeople.co.uk/Pages/Home.aspx.
- [18] W.T. Lin, W.C. Liao, C.F. Li, P.W. Cheng, and C.Y. Kung, The research in TTQS TrainQuali System develop strategy. *Employment Security*, 8(2), p68-75, 2009.
- [19] Y.H. Liu, A Study of the Evaluation Indicators for Training Quality System: The Case of Taiwan TrainQuali System(TTQS). National Kaohsiung University of Applied Sciences Department of human resource development thesis, 2009.
- [20] C.W. Chen, *The application of genetic algorithms in fuzzy controller design*. National Central University Department Of Mechanical Engineering thesis, 2005.
- [21] H.Y. Ho, A Document Retrieval and Recommendation System Based on Fuzzy Theory—Using Software Requirement Documents as A Study Case. National Cheng Kung University Department of Computer Science thesis, 2008.
- [22] H.C. Huang. A Content via Collaboration Approach to Text Filtering Recommender Systems. National Sun Yat-sen University Department of Information Management thesis, 2006.