

Semantic Link Network based Model for Organizing Multimedia Big Data

Chuanping Hu, *Member, IEEE*, Zheng Xu, *Member, IEEE*, Yunhuai Liu, *Member, IEEE*, Lin Mei, *Member, IEEE*, Lan Chen, and Xiangfeng Luo, *Member, IEEE*

Abstract—Recent research shows that multimedia resources “in the wild” are growing at a staggering rate. The rapid increase number of multimedia resources has brought an urgent need to develop intelligent methods to organize and process them. In this paper, the Semantic Link Network model is used for organizing multimedia resources. A whole model for generating the association relation between multimedia resources using Semantic Link Network model is proposed. The definitions, modules, and mechanisms of the Semantic Link Network are used in the proposed method. The integration between the Semantic Link Network and multimedia resources provides a new prospect for organizing them with their semantics. The tags and the surrounding texts of multimedia resources are used to measure their semantic association. The hierarchical semantic of multimedia resources are defined by their annotated tags and surrounding texts. The semantics of tags and surrounding texts are different in the proposed framework. The modules of Semantic Link Network model are implemented to measure association relations. A real data set including 100 thousand images with social tags from Flickr is used in our experiments. Two evaluation methods including clustering and retrieval are performed, which shows the proposed method can measure the semantic relatedness between Flickr images accurately and robustly.

Index Terms—Big Data, multimedia resources, semantic link network, multimedia resources organization

1 INTRODUCTION

RECENTLY, Big data is an emerging paradigm applied to datasets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time [10]. Various technologies are being discussed to support the handling of big data such as massively parallel processing databases [11], scalable storage systems [12, 33], cloud computing platforms [13, 34], and MapReduce [14, 35].

Understanding the semantics of multimedia has been an important component in many multimedia based applications. Manual annotation and tagging has been considered as a reliable source of multimedia semantics. Unfortunately, manual annotation is time-consuming and expensive when dealing with huge scale of multimedia data. Advances in Semantic Web [15] have made ontology another useful source for describing multimedia semantics. Ontology builds a formal and explicit representation of semantic hierarchies for the concepts and their relationships in video events, and allows reasoning to derive implicit knowledge. However, the semantic gap [16] between semantics and video visual appearance is still a challenge towards automated ontology-driven video annotation. With the rapid growth of video resources on the world-wide-web, for example, on YouTube¹ alone, 35 hours of video are unloaded every minute [2], and over 700 billion videos were watched in 2010. Vast amount of videos with no metadata have emerged. Thus automatically understanding raw multimedia solely based on their

visual appearance becomes an important yet challenging problem. Multimedia resources “in the wild” are growing at a staggering rate [1, 2]. The rapid increase number of multimedia resources has brought an urgent need to develop intelligent methods to represent and annotate them. Typical applications in which representing and annotating video events include criminal investigation systems [3], video surveillance [4], intrusion detection system [5], video resources browsing and indexing system [6], sport events detection [7], internet of things [8, 9], and many others. These urgent needs have posed challenges for multimedia resources management, and have attracted the research of the multimedia analysis and understanding. Overall, the goal is to enable users to search the related resources from the huge number of multimedia resources.

With the explosion of community contributed multimedia content available online, many social media repositories (e.g. Flickr², YouTube, and ZOOMR³) allow users to upload media data and annotate content with descriptive keywords which are called social tags. Flickr provides an open platform for users to publish their personal images freely. The principal purpose of tagging is to make images better accessible to the public. The success of Flickr proves that users are willing to participate in this semantic context through manual annotations [17]. Flickr uses a promising approach for manual metadata generation named “social tagging”, which requires all the users in the social network label the multimedia resources with their own keywords and share with others. The characteristics of social tags are as follows.

- Corresponding author: Zheng Xu is with the Third Research Institute of Ministry of Public Security, China. E-mail: xuzheng@shu.edu.cn
- C. Hu, Y. Liu, L. Chen, and L. Mei are with the Third Research Institute of Ministry of Public Security, Shanghai, China. E-mail: cphu@vip.sina.com, Yunhuai.liu@gmail.com, l_mei72@hotmail.com.
- X. Luo is with Shanghai University, China. E-mail: luoxf@shu.edu.cn

² www.flickr.com

³ www.zoomr.com

(1) **Ontology free.** The ontology based labeling defines ontology and then let users label the multimedia resources using the semantic markups in the ontology. Social tagging requires all the users in the social network label the multimedia resources with their own keywords and share with others. Different from ontology based annotation. There is no pre-defined ontology or taxonomy in social tagging. Thus the tagging task is more convenient for users.

(2) **User oriented.** The users can annotate images with their favorite tags. The tags of multimedia resources are determined by users' cognitive ability. To the multimedia resources, users may give different tags. Each multimedia resource may be with one tag at least, and each tag may appear in many different multimedia resources.

(3) **Semantic loss.** Irrelevant social tags frequently appear, and users typically will not tag all semantic objects in the image, which is called semantic loss. Polysemy, synonyms, and ambiguity are some drawbacks of social tagging.

In this paper, the Semantic Link Network (SLN) [18, 19, 20] model is used for organizing multimedia resources with social tags. Semantic Link Network is designed to establish associated relations among various resources (e.g., Web pages or documents in digital library) aiming at extending the loosely connected network of no semantics (e.g., the Web) to an association-rich network. Since the theory of cognitive science considers that the associated relations can make one resource more comprehensive to users [21], the motivation of SLN is to organize the associated resources loosely distributed in the Web for effectively supporting the Web intelligent activities such as browsing, knowledge discovery and publishing, etc. The tags and surrounding texts of multimedia resources are used to represent the semantic content. The relatedness between tags and surrounding texts are implemented in the Semantic Link Network model. The major contributions of this paper are summarized as follow.

(1) A whole model for generating the association relation between multimedia resources using Semantic Link Network model is proposed. The definitions, modules, and mechanisms of the Semantic Link Network are used in the proposed method. The integration between the Semantic Link Network and multimedia resources provides a new prospect for organizing them with their semantics.

(2) The tags and the surrounding texts of multimedia resources are used to measure their semantic association. The hierarchical semantic of multimedia resources are defined by their annotated tags and surrounding texts. The semantics of tags and surrounding texts are different in the proposed framework. The modules of Semantic Link Network model are implemented to measure association relations.

(3) A real data set including 100 thousand images with social tags from Flickr is used in our experiments. Two evaluation methods including clustering and retrieval are performed, which shows the proposed method can measure the semantic relatedness between Flickr images accurately and robustly.

(4) The relatedness measures between concepts are ex-

tended to the level of multimedia. Since the association relation is the basic mechanism of brain. The proposed Semantic Link Network based model can help the multimedia related applications such as searching and recommendation.

The rest of the paper is organized as follows. Section 2 gives the related work of social tags and semantic link network. The problem definition is introduced in Section 3. Section 4 proposes the method for measuring association of multimedia. Experiments are presented in Section 5. Conclusions are made in the last section.

2 RELATED WORK

Advances in Semantic Web have made ontology another useful source for describing multimedia semantics. The ontology builds a formal and explicit representation of semantic hierarchies for the concepts and their relationships in video events, and allows reasoning to derive implicit knowledge. In this section, the related work of the proposed model is given. The Semantic Web [15] is an evolving development of the World Wide Web, in which the meanings of information on the web is defined; therefore, it is possible for machines to process it. The basic idea of Semantic Web is to use ontological concepts and vocabularies to accurately describe contents in a machine readable way. These concepts and vocabularies can then be shared and retrieved on the web. In the Semantic Web, each fragment of the description is a triple, based on Description Logic. Thus, the implicit connections and semantics within the description fragments can be reasoned using Description Logic theory and ontological definitions. Earlier research work on the Semantic Web focused on defining domain specific ontologies and reasoning technologies. Therefore, data are only meaningful in certain domains and are not connected to each other from the World Wide Web point of view, which certainly limits the contributions of Semantic Web for sharing and retrieving contents within a distributed environment.

The Semantic Link Network (SLN) was proposed as a semantic data model for organizing various Web resources by extending the Web's hyperlink to a semantic link. SLN is a directed network consisting of semantic nodes and semantic links. A semantic node can be a concept, an instance of concept, a schema of data set, a URL, any form of resources, or even an SLN. A semantic link reflects a kind of relational knowledge represented as a pointer with a tag describing such semantic relations as cause Effect, implication, subtype, similar, instance, sequence, reference, and equal. The semantics of tags are usually common sense and can be regulated by its category, relevant reasoning rules, and use cases. A set of general semantic relation reasoning rules was suggested in [22] and [23]. If a semantic link exists between nodes, a link of reverse relation may exist. A relation could have a reverse relation. Relations and their corresponding reverse relations are knowledge for supporting semantic relation reasoning. SLN is a self-organized network since any node can link to any other node via a semantic link.

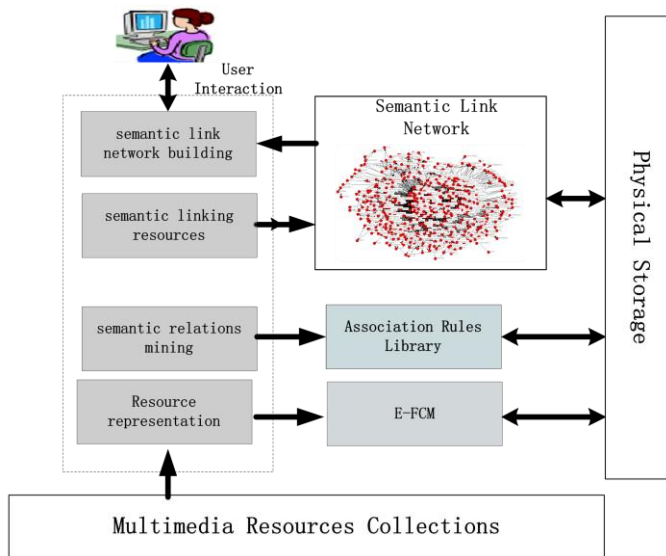


Figure 1. The basic mechanisms of the proposed model.

SLN has been used to improve the efficiency of query routing in P2P network [24], and it has been adopted as one of the major mechanisms of organizing resources for the Knowledge Grid. Pons has successfully applied the SLN to object prefetching and achieved a better result than other approaches [25].

3 THE SEMANTIC LINK NETWORK BASED MODEL

The tags and surrounding texts of multimedia resources are used to represent the semantic content. The relatedness between tags and surrounding texts are implemented in the Semantic Link Network model. In this section, the details of the proposed model are given. The basic definitions, representations, heuristics are introduced.

3.1 The Basic Mechanisms of the Proposed Model

SLN can be formalized into a loosely coupled semantic model for managing various resources. As a data model, the proposed model consists of the following parts, as shown in Fig. 1.

1) **Resources representation mechanism.** Element Fuzzy Cognitive Map (E-FCM) [19] is used to represent multimedia resources with social tags since it does not only reserve resources' keywords but also the relations among them.

2) **Resources storage mechanism.** Database/XML is used to store E-FCM since it is easy to define the mark-up elements.

3) **SLN generation mechanism.** Based on E-FCM and the association rules, ALN can be generated by machine automatically.

4) **Application mechanism.** SLN can be used for Web intelligence activities, Web knowledge discovery and publishing, etc. For example, when a user browses multimedia, other resources with semantic links to it can be recommended to the user.

Table 1. The variables and parameters used in the proposed computation model

Name	Description
f	A multimedia resource
t	A tag
$s(f)$	Tags set of a multimedia resource
$sr(t_1, t_2)$	Semantic relatedness of two tags
$sr(f_1, f_2)$	Semantic relatedness of two multimedia resources
$N(t)$	Page counts of a tag
$N(s(f))$	Set of page counts of a multimedia resource
$pos(t)$	Position information of a tag

3.2 The Basic Definitions

The three important definitions are defined firstly in this paper including the social tags set of a multimedia resource and the semantic relatedness between two multimedia resources.

Definition 1: Social tags set of a multimedia resource. The social tags (denoted by t) set of a multimedia resource f (denoted by $s(f)$) is a set of tags provided by users.

$$s(f) = \{t_1, t_2, \dots, t_{|s(f)|}\} \quad (1)$$

Definition 2: Semantic relatedness between tags. The semantic relatedness between tags (denoted by $sr(t_1, t_2)$) is the expected correlation of a pair of tags t_1 and t_2 .

Definition 3: Semantic relatedness between two multimedia resources. The semantic relatedness between multimedia resources (denoted by $sr(f_1, f_2)$) is the expected correlation of a pair of multimedia resources f_1 and f_2 .

The range of $sr(t_1, t_2)$ is from 0 to 1. A high value indicates that semantic relatedness between tags is more likely to be confidential.

3.3 The Basic Heuristics

Based on common sense and our observations on real data, five heuristics that serve as the base of the proposed computation model are given as follow.

Heuristic 1. Usually each tag of a multimedia resource appears only one time.

Different from writing sentences, users usually annotate a multimedia resource with different tags. For example, the possibility of using tags "apple apple apple" for an image is very low. Therefore, in this paper, we do not employ any weighting scheme for tags such as tf-idf [26].

Heuristic 2. The order of the tags may reflect the correlation against the annotated multimedia resource.

Different tag reflects the different aspect of a multimedia resource. According to Heuristic 1, the weight of a tag against the image cannot be obtained. Fortunately, the order of the tags can be get since user may provide tags

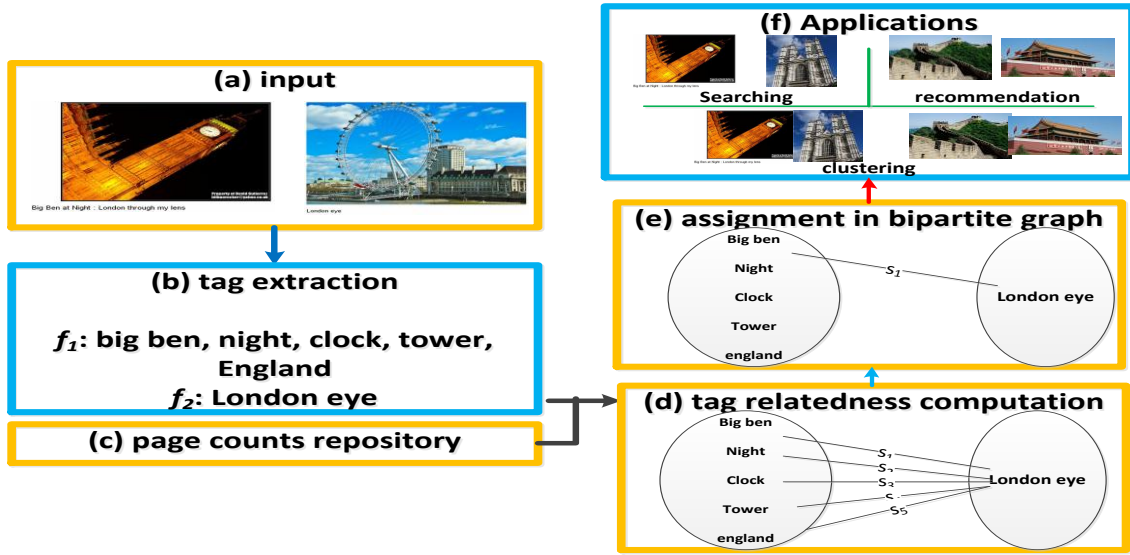


Figure 2. The illustration of the proposed method

one by one.

Heuristic 3. The number of tags of a multimedia resource may not relevant to the annotation correctness.

Different users may give different tags about the same multimedia resource. For example, users may give tags such as “apple iPhone” or “iPhone4 mobile phone” for a same image about iPhone. It is hardly to say which tag is better for annotation though the latter annotation has three tags.

Heuristic 4. Usually some tags may be redundant for annotating a multimedia resource.

Of course, users may give similar tags for a multimedia resource. For example, the tags “apple iPhone” may be redundant since iPhone is very semantic similar to apple.

Heuristic 5. Usually some tags may be noisy for annotating a multimedia resource.

Users may give inappropriate or even false tags for a multimedia resource. For example, the tags “iPhone” are false for an image about the iPod.

4 GENERATING THE SEMANTIC LINK

In this section, the computation model for generating the semantic link between multimedia resources is proposed. Based on the above five heuristics, the social tags provided by users are used in our computation model. Overall, the proposed computation model is divided into three steps.

(1) **Tag relatedness computation.** In this step, based on heuristic 1, all of the tag pairs between two multimedia resources are computed.

(2) **Semantic relatedness integration.** In this step, based on heuristics 3-5, the semantic relatedness between multimedia resources is computed.

(3) **Tag order revision.** In this step, based on heuristic 2, the multimedia resources relatedness on step 2 is revised.

Table 1 shows the variables and parameters used in the following discussion. Fig. 2 illustrates an overview of the proposed computation model.

4.1 Tag Relatedness Computation

According to definition 1, a multimedia resource can be represented as a set of tags provided by users. As for the semantic relatedness of a pair of multimedia resources, we can measure the semantic relatedness between tags of these multimedia resources. For example, two multimedia resources with tags “apple iPhone” and “iPod Nano”, we can measure the semantic relatedness between these tags. Since the number of each tag is usually one according to heuristic 1, the semantic relatedness between tags can be computed without considering their weight.

Many different methods of semantic relatedness measures between concepts have been proposed, which can be divided into two aspects [27]: taxonomy-based methods and web-based methods. Taxonomy-based methods use information theory and hierarchical taxonomy, such as WordNet, to measure semantic relatedness. On the contrary, web-based methods use the web as a live and active corpus instead of hierarchical taxonomy.

In the proposed computation model, each tag can be seen as a concept with explicit meaning. Thus, we use some equations based on co-occurrence of two concepts to measure their semantic relatedness. The core idea is that ‘you shall know a word by the company it keeps’ [28]. In this section, four popular co-occurrence measures (i.e., Jaccard, Overlap, Dice, and PMI [29]) are proposed to measure semantic relatedness between tags.

Besides co-occurrence measures, the page counts of each tag from search engine are used. Page counts mean the number of web pages containing the query q . For example, the page counts of the query ‘Obama’ in Google⁴ are 1,210,000,000⁵. Moreover, page counts for the query ‘ q AND p ’ can be considered as a measure of co-occurrence of queries q and p . For the remainder of this paper, we use the notation $N(p)$ to denote the page counts of the tag p in Google. However, the respective page counts for

⁴ <http://www.google.com>

⁵ The data was get in the data 9/28/2012

Algorithm 1: maxRel

Input: The tags set of two images f_1 and f_2 , which is $s(f_1)$ and $s(f_2)$
Output: The semantic relatedness of two images f_1 and f_2
for each $t_i \in s(f_1)$ /*page counts and position initial */
 $N(s(f_1)) \leftarrow N(t_i)$;
 $Pos(s(f_1)) \leftarrow Pos(t_i)$;
for each $t_j \in s(f_2)$
 $N(s(f_2)) \leftarrow N(t_j)$;
 $Pos(s(f_2)) \leftarrow Pos(t_j)$;
for each $t_i \in s(f_1)$
for each $t_j \in s(f_2)$
if ($t_i = t_j$) $sr(t_i, t_j) = 0$ /* pruning */
else $sr(t_i, t_j) = f(N(t_i), N(t_j))$ /*relatedness*/
return $maxRel(f_1, f_2) = f(Pos(t_i), Pos(t_j), sr(t_i, t_j))$;

the tag pair p and q are not enough for measuring semantic relatedness. The page counts for the query ‘ p AND q ’ should be considered. For example, when we query ‘Obama’ and ‘United States’ in Google, we can find 485,000,000 Web pages, that is, $N(Obama \cap United States) = 485,000,000$. The four co-occurrence measures (i.e., Jaccard, Overlap, Dice, and PMI) between two tags p and q are as follows:

$$Jaccard(p, q) = \frac{N(p \cap q)}{N(p) + N(q) - N(p \cap q)} \quad (2)$$

$p \cap q$ denotes the conjunction query ‘ p AND q ’.

$$Overlap(p, q) = \frac{N(p \cap q)}{\min(N(p), N(q))} \quad (3)$$

$\min(N(p), N(q))$ means the lower number of $N(p)$ or $N(q)$.

$$Dice(p, q) = \frac{2 * N(p \cap q)}{N(p) + N(q)} \quad (4)$$

According to probability and information theory, the mutual information (MI) of two random variables is a quantity that measures the mutual dependence of the two variables. Pointwise mutual information (PMI) is a variant of MI (Equation (5)):

$$PMI(p, q) = \log \left(\frac{N * N(p \cap q)}{N(p) * N(q)} \right) / \log N \quad (5)$$

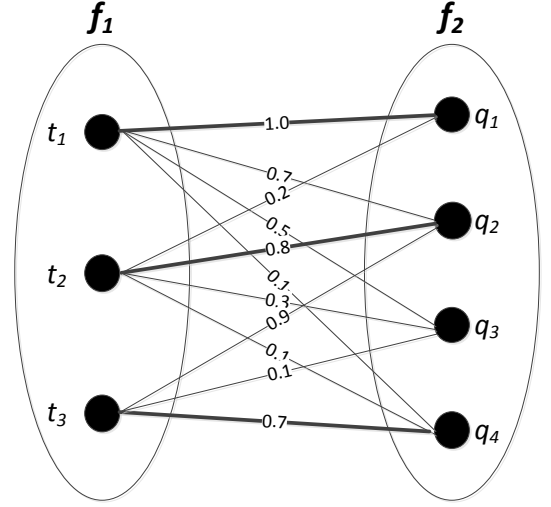
where N is the number of Web pages in the search engine, which is set to $N = 10^{11}$ according to the number of indexed pages reported by Google.

Through the equations 2-5, we can compute the tag relatedness as follows:

(1) Extracting the tags from two multimedia resources f_1 and f_2 , which are denoted as:

$$s(f_1) = \{t_1, t_2, \dots, t_{|s(f_1)|}\} \quad (6)$$

$$s(f_2) = \{t_1, t_2, \dots, t_{|s(f_2)|}\} \quad (7)$$



$$maxRel(f_1, f_2) = (sr(t_1, q_1) + sr(t_2, q_2) + sr(t_3, q_4)) / 3 = (1.0 + 0.8 + 0.7) / 3 = 0.83$$

Figure 3. Graphical representation of the assignment in bipartite graphs problem

(2) Issue the tags from f_1 and f_2 as the query to the web search engine (in this paper, we choose Google for its convenient API⁶), the page counts can be denoted as:

$$N(s(f_1)) = \{N(t_1), N(t_2), \dots, N(t_{|s(f_1)|})\} \quad (8)$$

$$N(s(f_2)) = \{N(t_1), N(t_2), \dots, N(t_{|s(f_2)|})\} \quad (9)$$

(3) Computing the semantic relatedness between each tag pair from f_1 and f_2 by equation 2-5. For example, if we use PMI to compute tag semantic relatedness, the equation can be:

$$sr(t_i, t_j) = \frac{\log \left(\frac{N * N(t_i \cap t_j)}{N(t_i) * N(t_j)} \right)}{\log N}, \quad t_i \in s(f_1) \wedge t_j \in s(f_2) \quad (10)$$

From the above steps, the tags relatedness can be computed, which is denoted as a triple $\langle t_i, t_j, sr(t_i, t_j) \rangle$. In the next section, we will give the detailed analysis for choosing the best measures from equations 2-5.

Overall, the page counts of each tag should be issued. Then the co-occurrence based measure is used to compute the semantic relatedness between tags. The reasons for using page counts based measures are as follow.

(1) **Appropriate computation complexity.** Since the relatedness between each tag pair of two multimedia resources should be computed, the proposed method must be with low complexity. Recently, web search engines such as Google provide API for users to index the page counts of each query. The web search engine gives an appropriate interface for the proposed computation model.

(2) **Explicit semantics.** The tag given by users may not be a correct concept in taxonomy. For example, users may give a tag ‘Bling Bling’ for a multimedia resource about a lovely girl. The word ‘Bling’ cannot be indexed

Table 3. The selected tags of group2 from Flickr

Group2	tags	tags	tags	tags	tags
Louis Vuitton	'Louis Vuitton' 'Keepall'	'Louis Vuitton' 'Alma'	'Louis Vuitton' 'Tivoli'	'Louis Vuitton' 'bolsas'	'LV' 'Multicolore'
Dior	'DIOR' 'lipstick' 'makeup'	'Dior' 'Diorskin Nude' 'Tan Sun Powder'	'Dior' 'Makeup' 'Palette'	'Dior' 'Addict 2'	'dior' 'jadore' 'perfume'
Gucci	'Gucci' 'Leather Belts'	'Gucci' 'Trainers'	'Gucci' 'Jolie Leopard' 'Orange'	'Replica' 'Gucci' 'Handbags'	'Gucci' 'Cruise'
Cartier	'Cartier' 'Pasha' 'Chronograph'	'CARTIER' 'Love Bracelet'	'Cartier' 'Santos Galbee'	'Calibre' 'Cartier'	'Cartier Watch' 'Tank Francaise'
Chanel	'Chanel' 'Coco Noir'	'chanel' 'chanel riva' 'Chanel nail polish'	'Coco Mademoi- selle'	'Chanel' 'No 5'	'Chance' 'Chanel'

Table 2. The detailed information of some selected groups

Group1	Average Tags per image	Group2	Average Tags per image
car	4.4	Louis Vuitton	3.1
phone	3.5	Dior	3.2
flower	2.2	Gucci	2.9
dog	5.6	Cartier	2.8
boat	3.1	Chanel	2.6

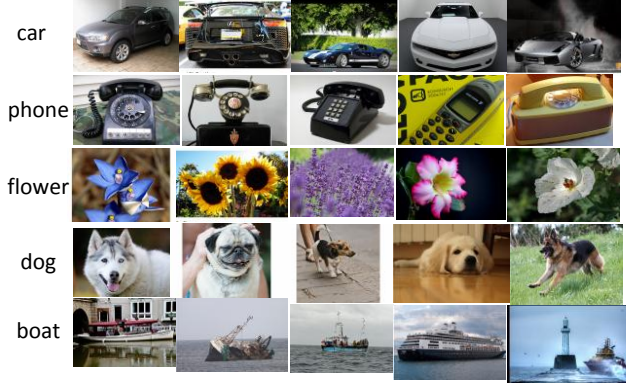


Figure 4. The selected images of group1 from Flickr

in many taxonomy such as WorldNet. The proposed method uses web search engine as an open intermediate. The explicit semantics of the newly emerge concepts can be get by web easily.

4.2 Semantic Relatedness Integration

In section 4.1, we compute the tag pair relatedness of two multimedia resources. Obviously, the tag pair relatedness of two multimedia resources f_1 and f_2 can be treated as a bipartite graph, which is denoted as

$$\begin{aligned}
 G &= (V, E) \\
 V &= \{f_1, f_2\} \\
 E &= \langle t_i, t_j, sr(t_i, t_j) \rangle, t_i \in s(f_1) \wedge t_j \in s(f_2)
 \end{aligned} \quad (11)$$

Based on the equation 11, we change the semantic relatedness integration of all tag pairs to the assignment in bipartite graph problem. We want to assign a best matching of the bipartite graph G.

A matching is defined as $M \subseteq E$ so that no two edges in M share a common end vertex. An assignment in a bipartite graph is a matching M so that each node of the graph has an incident edge in M. Suppose that the set of vertices are partitioned in two sets f_1 and f_2 , and that the edges of the graph have an associated weight given by a function $f: (f_1, f_2) \rightarrow [0..1]$. The function maxRel: $(f, f_1, f_2) \rightarrow [0..1]$ returns the maximum weighted assignment, i.e., an assignment so that the average of the weights of the edges

is highest. Fig. 3 shows a graphical representation of the semantic relatedness integration, where the bold lines constitute the matching M.

Based on the expressing of the assignment in bipartite graphs, we have

$$\maxRel(f, f_1, f_2) = \begin{cases} \frac{\max \sum_{i \in I} s(t_i, t_j)}{|s(f_1)|}, & |s(f_1)| \leq |s(f_2)| \\ \frac{\max \sum_{i \in I} s(t_i, t_j)}{|s(f_2)|}, & |s(f_1)| > |s(f_2)| \end{cases} \quad (12)$$

$$I = [1..|s(f_1)|], J = [1..|s(f_2)|]$$

Using the assignment in bipartite graphs problem to our context, the variables f_1 and f_2 represent the two multimedia resources to compute the semantic relatedness. For example, that f_1 and f_2 are composed of the tags $s(f_1)$ and $s(f_2)$. $|s(f_1)| > |s(f_2)|$ means that the number of tags in $s(f_2)$ is lower than that of $s(f_1)$. According to heuristic 3, we divide the result of the maximization by the lower cardinality of $s(f_1)$ or $s(f_2)$. In this way, the influence of the number of tags is reduced, and the semantic relatedness of two multimedia resources is symmetric.

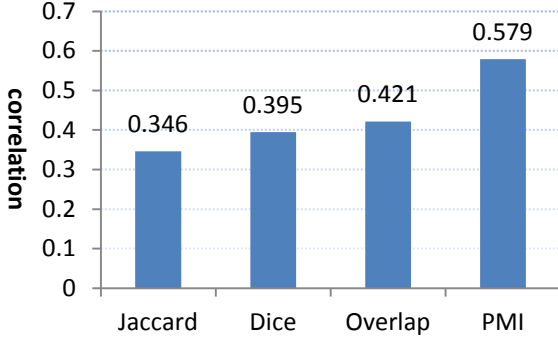


Figure 5. The correlation of four selected functions

Besides the cardinality of two tags set $s(f_1)$ and $s(f_2)$, the maxRel function is affected by the relatedness between each pair of tags. According to heuristics 4 and 5, the redundancy and noise should be avoided. In maxRel function, the one-to-one map is applied to the tags $s(f_1)$ and $s(f_2)$. Thus, the proposed maxRel function varies with respect to the nature of two multimedia resources.

Adopting the proposed maxRel function, we are sure to find the global maximum relatedness that can be obtained pairing the elements in the two tags sets. Alternative methods are able to find only the local maximum since they scroll the elements in the first set and, after calculating the relatedness with all the elements in the second set, they select the one with the maximum relatedness. Since every element in one set must be connected, at most, at one element in the other set, such a procedure is able to find only the local maximum since it depends on the order in which the comparisons occur. For example, considering the example in Fig. 3, t1 will be paired to q1 (weight=1.0). But, when analyzing t3 the maximum weight is with q2 (weight=0.9). This means that t2 can no more be paired to q2 even if the weight is maximum, since this is already matched to t3. As a consequence, t2 will be paired to q3 and the average of the selected weights will be $(1.0+0.3+0.9)/3=0.73$ which is considerably lower than using maxRel where the sum of the weights was $(1.0+0.8+0.7)/3=0.83$.

Overall, the cardinality of two tag sets is used to follow heuristic 3. The one-to-one map of tags pair is used to follow heuristics 4 and 5. The maxRel function is used to match a best semantic relatedness integration of two multimedia resources.

4.3 Tag Order Revision

According to heuristic 2, the order of tags should be considered to compute the semantic relatedness between two multimedia resources. Intuitively, the tags appearing in the first position may be more important than the latter tags. Some researches [30] suggest that people used to select popular items as their tags. Meanwhile, the top popular tags are indeed the “meaningful” ones.

In this section, the maxRel function proposed in section 4.2 is revised considering the order of tags. For example, the relatedness of tag pair with high position should be

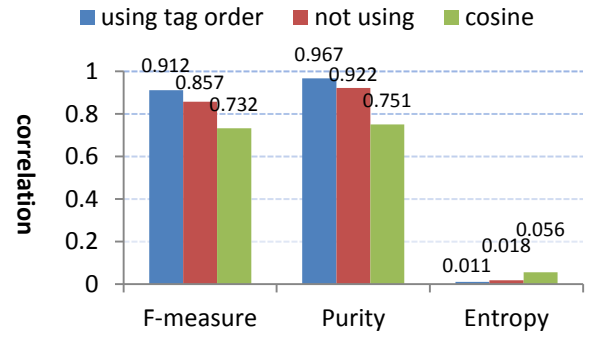


Figure 6. The clustering results of group 1

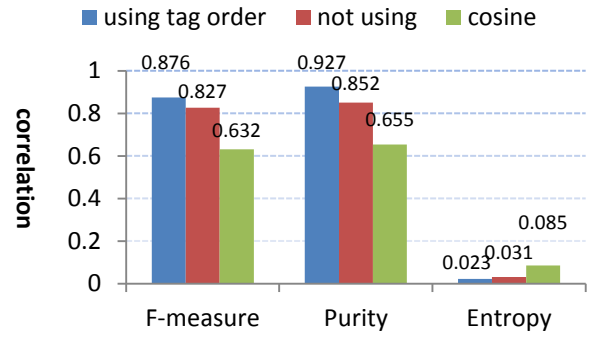


Figure 7. The clustering results of group 2

enhanced, which is summarized as a constrain schema:
Schema 1. Tag relatedness declining. This schema means that the identical tag pairs of two multimedia resources f_1 and f_2 should be pruned in maxRel function. In other words, the semantic relatedness of the same tag of two multimedia resources is set as 0.

We add a decline factor to the maxRel function, and the detailed steps are:

(1) According to the maxRel function in section 4.2, the best matching tag pairs are selected, which is denoted as:

$$\maxRel(f_1, f_2) = \sum sr(t_i, t_j), t_i \in s(f_1) \wedge t_j \in s(f_2) \quad (13)$$

Of course, the selected tag pairs are the best matching of the bipartite graph between multimedia resources f_1 and f_2 ;

(2) Computing the position information of each tag, which is denoted as $Pos(t_i)$

$$Pos(t_i) = \frac{|s(f)| + 1 - i}{|s(f)|}, t_i \in s(f) \quad (14);$$

(3) Add the position information of each tag to the equation 13, which can be seen as a decline factor:

$$sr(f_1, f_2) = \sum Pos(t_i) * sr(t_i, t_j) * Pos(t_j), \quad t_i \in s(f_1) \wedge t_j \in s(f_2) \quad (15)$$

(4) Of course, similar to maxRel function, equation should divide the result of the maximization by

$$sr(f_1, f_2) = \frac{\sum Pos(t_i) * sr(t_i, t_j) * Pos(t_j)}{\sum Pos(t_i) * Pos(t_j)} \quad (16)$$

Table 4. The comparison of the cut-off point precision between the proposed method and Flickr

Cut-off point	Louis Vuitton	Gucci	Dior	Chanel	Cartier
P^1	100%	100%	100%	100%	100%
$P^1(\text{Flickr})$	100%	100%	0	100%	100%
P^5	100%	100%	100%	100%	100%
$P^5(\text{Flickr})$	80%	60%	60%	60%	80%
P^{10}	100%	100%	100%	100%	100%
$P^{10}(\text{Flickr})$	90%	70%	70%	80%	80%

We also consider the example in Fig. 3. According to equation 16, the semantic relatedness is revised as $(1 \cdot 1 \cdot 0 \cdot 1 + \frac{2}{3} \cdot 0.8 \cdot \frac{3}{4} + \frac{1}{3} \cdot 0.7 \cdot \frac{1}{4}) / (1 \cdot 1 + \frac{2}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{4}) = 0.92$.

Besides adding decline factor to the maxRel function, we also add a constrain schema: identical tag pruning.

Schema 2. Identical tag pruning. This schema means that the identical tag pairs of two multimedia resources f_1 and f_2 should be pruned in maxRel function. In other words, the semantic relatedness of the same tag of two multimedia resources is set as 0.

The above schema is used to ensure the relatedness measures of two multimedia resources. If we do not prune the identical tag pairs of two multimedia resources, the proposed method will be transformed to the similarity measures. For example, the cosine similarity between two tags is to find the number of identical elements of two vectors. The overall algorithm of the proposed computation mode is presented in algorithm 1.

5 EXPERIMENTAL RESULTS

In this section, we evaluate the results of using the proposed method for relatedness measurement. In section 5.1, we introduce the data set for the evaluation. In section 5.2, we determine to use the co-occurrence function for tag relatedness measures. In section 5.3 and 5.4, clustering and retrieval are used for evaluate the proposed method. In section 5.5, some discussions about the experimental results are given.

5.1 The Data Sets

We choose Flickr groups as the resources for building data sets. Users on online photo sharing sites like Flickr have organized many millions of photos into hundreds of thousands of semantically themed groups. These groups expose implicit choices that users make about which images are similar. Flickr group membership is usually less noisy than Flickr tags because images are screened by group members. We download 100 thousand images with tags from 100 groups. The tags of these images are extracted. Each group means a concept. Thus, if the proposed method can do well in these groups, we may say that it can measure the semantic relatedness between Flickr images accurately and robustly.

Table 2 gives the detailed information of some selected group of the data set. Some selected images from some groups are shown in Fig. 4. Table 3 gives some selected tags.

Query Top 5 searching result of the proposed method

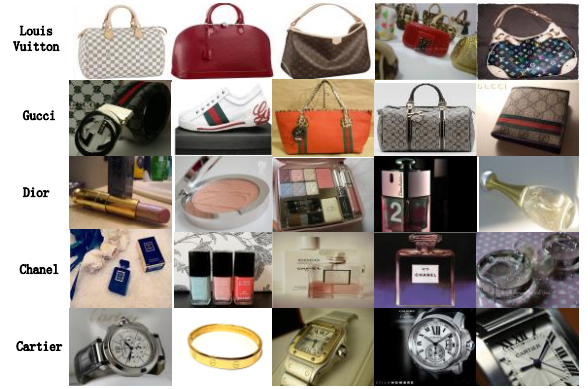


Figure 8. The top five searching results of the proposed method

Query The top 5 results from Flickr

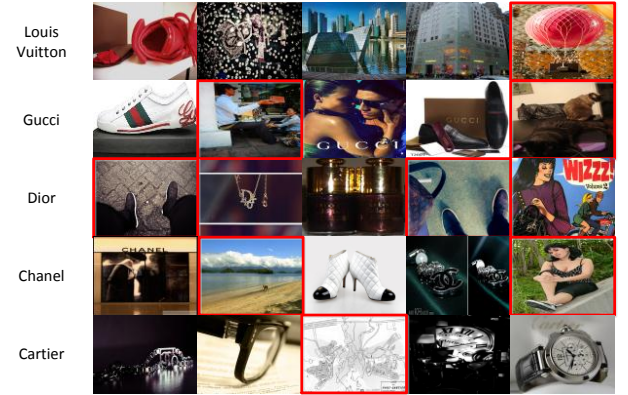


Figure 9. The top five searching results from Flickr

group of the data set. Some selected images from some groups are shown in Fig. 4. Table 3 gives some selected tags.

5.2 Relatedness Function Selection

In the section 4.1, four co-occurrence measures (i.e., Jaccard, Overlap, Dice, and PMI) are given for relatedness measures between tags. In [31], Rubenstein and Goode-nough proposed a dataset containing 28 word-pairs rating by a group of 51 human subjects, which is a reliable benchmark for evaluating semantic similarity measures. The higher the correlation coefficient against R-G ratings is, the more accurate the methods for measuring semantic similarity between words are. Fig. 5 gives the correlation coefficient of four functions against R-G test set. From Fig. 5, we can say that PMI performs best on relatedness measures for its highest correlation coefficient. Thus, in the latter experiments, we select PMI as the relatedness measures between tags.

5.3 Evaluation on Image Clustering

In this section, we evaluate the correctness of using tag order. In section 4.3, we add the position information of each tag to the semantic relatedness measures. The tags with high position are treated as the major element for semantic relatedness measures. We evaluate the using of tag order by the clustering task. We employ the proposed semantic relatedness of images into K-means [32] clustering model. Since the K-means model depends on the initial points, we random



Figure 10. The searching interface of the developed tool.

select core points 100 times. We evaluate the effectiveness of document clustering with three quality measures: *F-measure*, *Purity*, and *Entropy* [32]. We treat each cluster as if it were the result of the proposed method and each class as if it were the desired set of images. Generally, we would like to maximize the *F-measure* and *Purity*, and minimize the *Entropy* of the clusters to achieve a high-quality document clustering. Moreover, we compare the clustering results between the proposed method using tag order or not. Fig. 6 and Fig. 7 give the clustering results of group1 and group2 data sets. From Fig. 6 and Fig. 7, we can conclude that:

(1) **The proposed method performs better than cosine based clustering.** This result can be obtained from Fig. 6 and Fig. 7. The three metrics including *F-measure*, *purity*, and *entropy* of the proposed method are better than cosine based clustering. This may be caused by the inherent feature of the proposed method. The proposed method is based on the semantic relatedness other than the co-occurrence of the cosine based clustering. If the tags of two images are not overlapped, the cosine based clustering may be unavailable.

(2) **The schema on using of tag order is effective.** This result can also be obtained from Fig. 6 and Fig. 7. The three metrics including *F-measure*, *purity*, and *entropy* of using tag order are highest. The position information reflects the importance of each tag. The proposed method emphasizes the tags with high order, which raises the performance on images clustering.

(3) **The proposed method is robust in different data sets.** The proposed method performs well in group1 and group2 data set. It is worth noting that the difference between the proposed method and cosine method of group2 is higher than that of group1. The reason of that is due to the semantic correlation of group2 being stronger than group1. In other words, the performance of the proposed method relies on the semantic correlation of classes in data sets. The stronger the semantic correlation between classes of data, the better the proposed method performs.

5.4 Evaluation on Image Searching

In this section, we evaluate the proposed method query-based image searching task. Five queries from group2 are

selected as the test set including "Louis Vuitton", "Gucci", "Chanel", "Cartier", and "Dior". These queries are searched in Flickr. The top 50 images are obtained as the data set. Moreover, we remove the queries on the tags of each image. For example, the tag "Cartier" of the top 50 images is removed of the query "Cartier". The reason for that operation is that the proposed method is based on the semantic relatedness other than co-occurrence. We choose cut-off point precision to evaluate the proposed method on image searching. The cut-off point precision (P^n) means that the percentage of the correct result of the top n returned results. We compute the P^1 , P^5 , and P^{10} of the group2 test set. Table 4 lists the comparison of the cut-off point precision between the proposed method and Flickr. Fig. 8 and Fig. 9 give the top 5 results of the five test queries from the proposed method and Flickr⁷. Especially, we put the red rectangle to the wrong search results in Fig. 9. From the experimental results, we can conclude that:

(1) **The proposed method performs better than Flickr.** In table 4, the P^1 , P^5 , and P^{10} of the proposed method are higher than Flickr. The experimental results prove the correctness of the proposed method on image searching task.

(2) **The proposed method is effective on image searching task.** In Fig. 8 and Fig. 9, we compare the top 5 returned results by the proposed method and Flickr. It is obviously that the returned results from Flickr are rough. Some returned images are irrelevant to the given query. For example, in Fig. 9, almost 40% searching results are incorrect.

(3) **The proposed method can handle the relatedness searching problem.** The proposed method can measure the semantic relatedness of two images robust and correctly. In Fig. 8, the tags of the search results do not contain the search query, which is different from the traditional co-occurrence based search mechanism.

(4) **The proposed method can support the faceted exploration of image search.** Faceted exploration of search results is widely used in search interfaces for structured databases. Recently the faceted exploration is also appearing in on-line search engine in the form of search assistants. The proposed method can measure the semantic relatedness of two images. Given the search queries, we can select the related images for faceted search.

6 APPLICATIONS

Content-based image retrieval (CBIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases. "Content-based" means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata and this produces a lot of garbage in the results. Also having humans manually

enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results.

The proposed SLN based model can be used for video searching. The ontology based video searching is similar to CBIR, which also focuses on the content of the videos. Fig. 10 gives the searching interface of the developed tool based on the proposed SLN based model. From Fig. 10, the searching procedures for a user are as follow.

(1) Ontology based queries. Different from web search engines, the proposed SLN based video search constricts the searching method. Users can only select the defined attributes or concepts as the searching queries.

(2) Associated videos suggestion. Since the video resources are organized by their association relation, the associated videos can be suggested to the users.

7 CONCLUSION

Recent research shows that multimedia resources “in the wild” are growing at a staggering rate. The rapid increase number of multimedia resources has brought an urgent need to develop intelligent methods to organize and process them. In this paper, the Semantic Link Network model is used for organizing multimedia resources. Semantic Link Network (SLN) is designed to establish associated relations among various resources (e.g., Web pages or documents in digital library) aiming at extending the loosely connected network of no semantics (e.g., the Web) to an association-rich network. Since the theory of cognitive science considers that the associated relations can make one resource more comprehensive to users, the motivation of SLN is to organize the associated resources loosely distributed in the Web for effectively supporting the Web intelligent activities such as browsing, knowledge discovery and publishing, etc. The tags and surrounding texts of multimedia resources are used to represent the semantic content. The relatedness between tags and surrounding texts are implemented in the semantic Link Network model. The data sets including about 100 thousand images with social tags from Flickr are used to evaluate the proposed method. Two data mining tasks including clustering and searching are performed by the proposed framework, which shows the effectiveness and robust of the proposed framework.

ACKNOWLEDGMENT

This work was supported in part by the National Science and Technology Major Project under Grant 2013ZX01033002-003, in part by the National High Technology Research and Development Program of China (863 Program) under Grant 2013AA014601, 2013AA014603, in part by National Key Technology Support Program under Grant 2012BAH07B01, in part by the National Science Foundation of China under Grant 61300202, and in part by the Science Foundation of Shanghai under Grant 13ZR1452900.

REFERENCES

- [1] “Cisco Visual Networking Index: Forecast and Methodology, 2009–2014,” Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/whitepaper_c11-481360_ns827_Networking_Solutions_White_Paper.html, accessed in 11.11.2013.
- [2] “Great Scott! Over 35 Hours of Video Uploaded Every Minute to Youtube,” The official YouTube blog, 2013. Available: <http://youtube-global.blogspot.com/2010/11/great-scott-over-35-hours-of-video.html>, accessed in 11.11.2013.
- [3] L. Wu and Y. Wang. The process of criminal investigation based on grey hazy set. *2010 IEEE International Conference on System Man and Cybernetics*, pp.26-28, 2010.
- [4] L. Liu, Z. Li, and E. Delp. Efficient and low-complexity surveillance video compression using backward-channel aware wyner-ziv video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(4):452-465, 2009.
- [5] J. Zhang, M. Zulkernine, and A. Haque. Random-forests-based network intrusion detection systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 38(5):649-659, 2008.
- [6] H. Yu, C. Pedrinaci, S. Dietze, and J. Domingue. Using linked data to annotate and search educational video resources for supporting distance learning. *IEEE Transactions on Learning Technologies*, 5(2):130-142, 2012.
- [7] C. Xu, Y. Zhang, G. Zhu, Y. Rui, H. Lu, and Q. Huang. Using webcast text for semantic event detection in broadcast sports video. *IEEE Transactions on Multimedia*, 10(7):1342-1355, 2008.
- [8] Y. Liu, Y. Zhu, Lionel M. Ni, and G. Xue. A Reliability-Oriented Transmission Service in Wireless Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, 22(12): 2100-2107, 2011.
- [9] Y. Liu, Q. Zhang, and Lionel M. Ni. Opportunity-Based Topology Control in Wireless Sensor Networks. *IEEE Transactions on Parallel and Distributed Systems*, 21(3): 405-416, 2010.
- [10] M. Wigan and R. Clarke. Big data’s big unintended consequences. *Computer*, 46(6):46-53, 2013.
- [11] D. Yuan, Y. Yang, X. Liu, W. Li, L. Cui, M. Xu, and J. Chen. A highly practical approach towards achieving minimum datasets storage cost in the cloud. *IEEE Transactions on Parallel and Distributed Systems*, 24(6):1234-1244, 2013.
- [12] X. Zhang, C. Liu, S. Nepal, S. Pandev, and J. Chen. A privacy leakage upper-bound constraint based approach for cost-effective privacy preserving of intermediate datasets in cloud. *IEEE Transactions on Parallel and Distributed Systems*, 24(6):1192-1202, 2013.
- [13] X. Liu, Y. Yang, D. Yuan, and J. Chen. Do we need to handle every temporal violation in scientific workflow systems. *ACM Transactions on Software Engineering and Methodology*, early access, 2013.
- [14] X. Zhang, T. Yang, C. Liu, and J. Chen. A scalable two-phase top-down specialization approach for data anonymization using mapreduce on cloud. *IEEE Transactions on Parallel and Distributed Systems*, early access, 2013.
- [15] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):34-43, 2001.
- [16] H. Ma, J. Zhu, M. Lyu, and I. King. Bridging the semantic gap between image contents and tags. *IEEE Transactions on Multimedia*, 12(5):462-473, 2010.
- [17] L. Wu, X. Hua, N. Yu, W. Ma, and S. Li. Flickr Distance: A Relationship Measure for Visual Concepts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):863-875, 2012.
- [18] H. Zhuge. Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning. *IEEE Transactions on Knowledge and Data Engineering*, 21(6):785-799, 2009.

- [19] X. Luo, Z. Xu, J. Yu, and X. Chen. Building Association Link Network for Semantic Link on Web Resources. *IEEE transactions on automation science and engineering*, 8(3):482-494, 2011.
- [20] H. Zhuge. Semantic linking through spaces for cyber-physical-socio intelligence: A methodology. *Artificial Intelligence*, 175:988-1019, 2011.
- [21] Z. Xu, X. Luo, and L. Wang. Incremental building association link network. *Computer systems science and engineering*, 26(3):153-162, 2011.
- [22] H. Zhuge. Interactive Semantics. *Artificial Intelligence*, 174:190-204, 2010.
- [23] H. Zhuge. The Knowledge Grid – Toward Cyber-Physical Society, World Scientific Publishing Co., Singapore, 2012. 2nd Edition.
- [24] H. Zhuge, X. Chen, X. Sun, and E. Yao. HRing: A structured P2P overlay based on harmonic series. *IEEE Transactions on Parallel and Distributed Systems*, 19(2):145-158, 2008.
- [25] A.P. Pons. Object Prefetching Using Semantic Links. *ACM SIGMIS Database*, 37(1):97-109, 2006.
- [26] R. Schenkel, T. Creelius, M. Kacimi, S. Michel, T. Neu-mann, J. Parreira, and G. Weikum. Efficient top-k querying over social tagging networks. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 523-530, 2008.
- [27] G. Salton, A. Wong, and C. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613-620, 1975.
- [28] Z. Xu, X. Luo, J. Yu, and W. Xu. Measuring semantic similarity between words by removing noise. *Concurrency and Computation: Practice and Experience*, 23(18):2496-2510, 2011.
- [29] R. Firth. A synopsis of linguistic theory 1930-1955. In *Studies in Linguistic Analysis*. Philological Society: Oxford, 1957.
- [30] M. Vojnovic, J. Cruise, D. Gunawardena, and P. Marbach. Ranking and suggesting popular items. *IEEE Transactions on Knowledge and Data Engineering*, 21(8):1133-1146, 2009.
- [31] H. Rubenstein and B. Goodenough. Contextual correlates of synonymy. *Communications of the ACM*, 8(10):627-633, 1965.
- [32] M. Steinbach, G. Karypis and V. Kumar. A Comparison of Document Clustering Techniques. *KDD Workshop on Text Mining*, 2000.
- [33] L. Wang and S. Khan. Review of performance metrics for green data centers: a taxonomy study. *The journal of supercomputing*, 63(3):639-656, 2013.
- [34] L. Wang, D. Chen, et al. Towards enabling cyber infrastructure as a service in clouds. *Computer & Electrical Engineering*, 39(1):3-14, 2013.
- [35] L. Wang, J. Tao, et al. G-Hadoop: MapReduce across distributed data centers for data-intensive computing. *Future Generation Computer Systems*, 29(3):739-750, 2013.

Parallel and Distributed Systems, IEEE Journal of Selected Areas in Communications, IEEE Trans. on Mobile Computing, IEEE Trans. on Vehicular Technology etc.



Lin Mei received his PhD. degree from Xian Jiaotong University, China. He is currently working in the third research institute of ministry of public security, China. He is the Dean Professor of the department of internet of things.



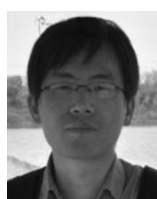
Lan Chen is a PhD. Candidate of Beihang University, China. She is currently working in the third research institute of ministry of public security, China.



Xiangfeng Luo is a professor in the School of Computers, Shanghai University, China. Currently, he is a visiting professor in Purdue University, USA. His main research interests include Web Wisdom, Cognitive Informatics, and Text Understanding. He has authored or co-authored more than 50 publications and his publications have appeared in IEEE Trans. on Systems, Man, and Cybernetics-Part C, IEEE Trans. on Automation Science and Engineering, IEEE Trans. on Learning Technology, etc. He has served as the Guest Editor of ACM Transactions on Intelligent Systems and Technology. Dr. Luo has also served on the committees of a number of conferences/workshops, including Program Co-chair of ICWL 2010 (Shanghai), WISM 2012 (Chengdu), CTUW2011 (Sydney) and more than 40 PC members of conferences and workshops.



Chuanping Hu received his PhD. degree from Tongji University, China. He is currently working in the third research institute of ministry of public security, China. He is the Dean Professor of the third research institute of ministry of public security.



Zheng Xu was born in Shanghai, China, in 1984. He received the Diploma and Ph.D. degrees from the School of Computing Engineering and Science, Shanghai University, Shanghai, in 2007 and 2012, respectively. He is currently working in the third research institute of ministry of public security, China. His current research interests include semantic Web and Web mining



Yunhuai Liu is a professor in the third research institute of ministry of public security, China. He received the PhD degrees from Hong Kong University of Science and Technology (HKUST) in 2008. His main research interests include wireless sensor networks, pervasive computing, and wireless network. He has authored or co-authored more than 50 publications and his publications have appeared in IEEE Trans. on