Utilizing Indirect Associations in Multimedia Semantic Retrieval

Hsin-Yu Ha, Shu-Ching Chen
School of Computing and Information Sciences
Florida International University
Miami, FL 33199, USA
{hha001,chens}@cs.fiu.edu

Mei-Ling Shyu
Department of Electrical and Computer Engineering
University of Miami
Coral Gables, FL 33124, USA
shyu@miami.edu

***Abstract*—Technological developments have lead to the propagation of massive amounts of data in the form of text, image, audio, and video. The unstoppable trend draws researchers' attention to develop approaches to efficiently retrieve and manage multimedia data. The inadequacy of keyword-based search in multimedia data retrieval due to non-existent or incomplete text annotations has called for the development of a content-based multimedia data management framework. Specifically, detecting high-level semantic concepts is one of the rapidly growing topics in this regard. In order to thoroughly identify semantic concepts in data which have different representations and are derived from different modalities, both positive and negative inter-concept correlations have been recently studied and explored to enhance the re-ranking performance. In this paper, an indirect association rule mining (IARM) approach is introduced to reveal the hidden correlation among semantic concepts. The effectiveness of IARM is evaluated by Multiple Correspondence Analysis (MCA). Furthermore, normalization and score integration are performed to achieve the optimal classification results. The TRECVID 2011 benchmark dataset is used to show the effectiveness of the proposed IARM factor in the re-ranking process.**

***Keywords*-Multimedia Data, Semantic Concept Detection, Indirect Association Rule Mining (IARM), Re-ranking, Concept Mining**

I. Introduction

With the increasing rate of digitization in industry, academia and among general public, efficient management of high-diversity multimedia data such as text, image, audio, and video poses a great challenge. In [1], Dragland claims that 90 % of the world's data were generated in the past two years, which makes it a great challenge to effectively retrieve the meaningful information from the large volume of data in different representations. Many researchers were thrilled to investigate a sufficient way to handle the huge amount of multimedia big data in terms of searching, browsing, indexing, etc. [2]–[10], but many challenges were still standing in the way. For example, it did not take long for the researchers to realize that due to non-existent or incomplete text annotations, the conventional keyword-based search was inadequate in retrieving multimedia data. Hence, content-based approaches were proposed [11]–[17] to better capture the semantic information through different types of low-level features. Specifically, many of these content-based approaches have been applied to improve multimedia semantic concept retrieval, whose goal is to identify high-level semantic concepts such as "dancing" and "forest" from data instances likes images, videos, or any complex multimedia data.

When facing multiple semantic concept retrievals, instead of bridging the semantic gap between low-level features and high-level semantic concept one at a time, it can be treated as a multi-label classification problem, which is solved at once by exploring the concept relations. Intuitively, most of the research work leveraged the positive inter-concept relationships [18]–[22], which means that if a concept is detected in one data instance, then there is a higher chance to identify another concept in the same data instance, such as the correlation between concept "sky" and concept "outdoor". On the other hand, negative correlations are also studied in [23]–[26] to explore the opposite correlations between concepts in enhancing the overall classification results. For example, the fact that a data instance contains the concept "outdoor" usually implies zero possibility of detecting concept "indoor" from the same data instance. Encouraged by the improvement of leveraging the direct concept correlation, indirect association rules among the concepts are explored in this paper. The goal is to reveal the implicit correlation when two concepts are rarely identified in the same data instance, but they are indirectly correlated through a mediator concept. For instance, the concept "basketball" and the concept "volleyball" might seldom co-occur in the same data instance, but they have a much higher chance of appearing together with the concept "gym". That is, we believe that there exists an indirect association between concept "basketball" and concept "volleyball", which is worth discovering and analyzing.

In this paper, a multimedia semantic retrieval framework which utilizes both negative correlations and indirect associations is proposed to refine the performance. An algorithm is developed to retrieve the indirect association rules (IAR) from the statistics information of the concept occurrences. The Association Affinity Network (AAN) mechanism [23] is extended in this paper to encompass both negative correlations and IARM correlations. In addition, two types of labels are defined and generated to estimate the posterior

978-1-4799-8688-0/15 $31.00 © 2015 IEEE
DOI 10.1109/BigMM.2015.89

probability of a positive IAR and a negative IAR toward the detected concepts.

The paper is organized as follows: In Section 2, the proposed framework is depicted for both training and testing processes, followed by the presentation and the in-depth discussion of major components. The experiments setup, evaluation criteria, the experimental results, and the corresponding discussion are all reported in Section 3. Finally, the conclusion is given in Section 4, which summarizes the performance of the proposed work.

II. Proposed Framework

Figure 1 and Figure 2 depict the training process and testing process of the proposed framework, respectively. As shown in Figure 1, the training process consists of three major components, namely "Multimedia Semantic Concept Detection", "Concept Correlation Mining", and "Dual Correlation Modeling". The "Multimedia Semantic Concept Detection" component mainly concerns the high-level process of building the classification models to detect the semantic concepts on multimedia data. From the beginning, the objective is to detect N high-level semantic concepts such as "Beach" and "Dancing" from the training process of a training dataset with M data instances. Low-level features are extracted to represent each training data instance and N binary content-based classification models are built as the concept detectors D_i, where $1 < i < N$. Finally, each detector outputs M ranking scores to indicate the probabilities of detecting the concept in the M data instances. The higher the ranking score, the better chance to identify the concept in the data instance.

As shown on the right side of Figure 1, both Integrated Correlation Factor (ICF) and conditional probability-based coarse filtering method are applied when performing negative correlation selection. A detailed process is described in [23]. IARM is proposed to reveal the hidden concept correlations from the formatted label matrix. After selecting only the conjunctive correlations between negative correlations and IAR corrections, the features extracted from the original training dataset are fed as the input to independently train two MCA-based weight estimation models for negative correlations and IAR corrections. Lastly, the "Dual Correlation Modeling" component combines two sets of weights and the ranking scores produced from the "Multimedia Semantic Concept Detection" component and normalizes them to better train the regression-based score integration model. Please note that the selected negative correlations, IAR correlations, two MCA-based weight estimation models, and the final regression models are all stored so that they can be applied to the testing data instances.

In Figure 2, the testing process starts with sending the testing dataset to each of the concept detectors to produce the testing ranking scores. After that, the same feature extraction method performed in the training process will be used to extract the same feature set from the testing instance. Two trained MCA-based weight estimation models take the extracted testing features to generate the weights for negative correlations and IAR correlations. At the end, the testing scores from the concept detectors and two different types of weights are normalized and sent to the trained regression models to generate the final re-ranked testing scores.

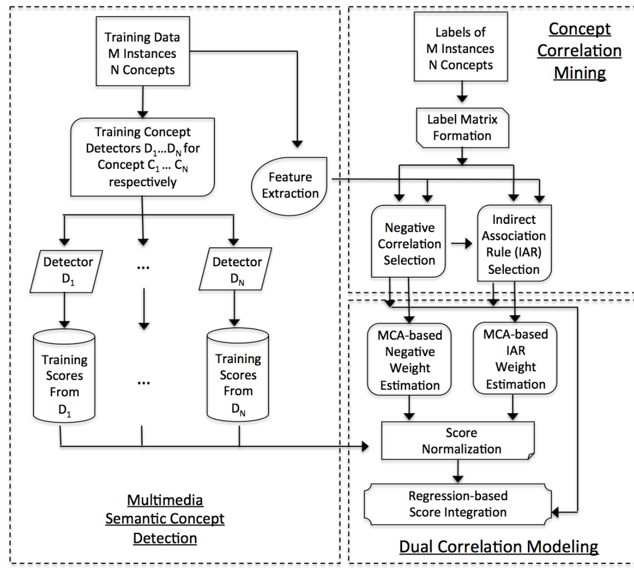

Figure 1: The proposed framework for adopting indirect association rules (IAR) in AAN (Training Process)

A. *Indirect Association Rules*

Indirect association rules (IAR) were first proposed by Tan et al. [27] for identifying a pair of items, x and y, which are rarely appeared together in the same transaction, but they both highly depend on a set of mediator item Med. The formal definition can be found at Definition 1.

Definition 1. ***Indirect Association Rules (IAR)***
An itemset pair $\{X, Y\}$ is indirectly associated through a mediator Med, if the following conditions hold:

1) $sup(\{X, Y\}) < itp_s$
2) *There exists a non-empty set Med such that:*
   - $sup(\{X\} \cup Med) \geq Med_s$, *and* $sup(\{Y\} \cup Med) \geq Med_s$
   - $dep(\{X\}, Med) \geq Med_d$, *and* $dep(\{Y\}, Med) \geq Med_d$

The threshold above are named itempair support threshold (itp_s), Mediator Support Threshold (Med_s), and Mediator Dependency Threshold (Med_d), respectively. In practice, it is subject to have $Med_s > itp_s$. When the rule is applied to discover the correlations among semantic concepts, a brief illustration is depicted in Figure 3. As shown in this figure, two concepts, C_X and C_Y, can rarely be identified in the

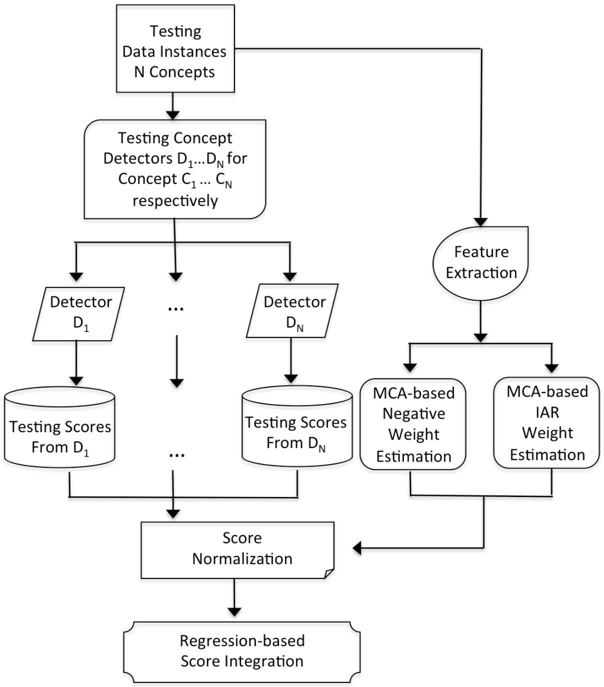

Figure 2: The proposed framework for adopting indirect association rules (IAR) in AAN (Testing Process)

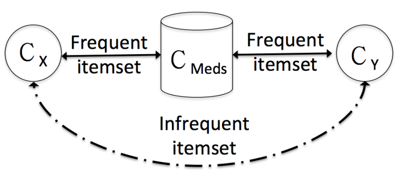

Figure 3: Applying IARM in Mining Concept Ontology

same data instance, but they both highly depend on the presence of a set of mediator concepts C_{Meds}.

Before describing how to incorporate the idea of IARM, it is necessary to introduce several definitions used throughout the paper.

Definition 2. ***Data Instance, Features, and Label***

A **data instance** is referred to as an image, a keyframe, or a video shot, depending on the content of the introduced dataset. In the experiment section, the TRECVID 2010 dataset is adopted to validate the proposed framework, where each data instance represents a keyframe of one video shot. Features are five well-known low-level features extracted from both training and testing datasets, including HAAR, CEDD, HOG, HSV, and YCBCR. Lastly, a label is the value of either 0 or 1 per instance to indicate whether the corresponding semantic concept exists in that instance.

Definition 3. ***Support and Confidence***

To calculate the support and confidence values, a combined label matrix must be formed (as shown in Table I), where each row represents a data instance and each column represents a concept label. In other words, each element in this matrix will indicate whether one data instance contains one semantic concept or not. Therefore, with the idea of association rule mining [28], each data instance can be considered as one transaction; while each concept is considered as one itemset. Let $C = \{C_1, C_2, ..., C_N\}$, TI be a set of all transactions where each transaction I is a set of items such that $I \subseteq C$, and $Occ(C_X)$ is the number of occurrences of C_X. Thus, for an association rule like $C_X \Rightarrow C_Y$, the support and confidence values can be calculated as shown in Equation 1 and Equation 2, respectively.

$$sup(C_X \Rightarrow C_Y) = \frac{Occ(C_X \cup C_Y)}{Number\_of\_TI} \quad (1)$$

$$conf(C_X \Rightarrow C_Y) = \frac{Occ(C_X \cup C_Y)}{Occ(C_X)} \quad (2)$$

Definition 4. ***Itemset Pair and Mediator***

IARM is introduced to discover the hidden correlation when concept X and concept Y seldom appear together in the same data instance, but they will usually be identified along with the mediator concept Med. Therefore, **Itemset Pair** is defined to include two concepts, e.g., X and Y, which rarely appear together and concept Med is the **mediator**.

Definition 5. ***Dependence: Interesting Ratio (IR)***

In addition to the confidence value, an interesting ratio is another perspective to further verify the significance of the retrieved rules. For example, if there is an indirect association rule, where the itemset pair is concept X and concept Y and the mediator concept is Med, an interesting ratio is introduced to ensure the following two conditions. First, concept X highly depends on the appearance of the mediator concept Med. Second, this IAR rule is not retrieved because of the high frequency of concept Med. The same thoughts should be also applied for concept Y. The interesting ratio between concept X and concept Med is calculated as shown in Equation 3.

$$IR(C_X \Rightarrow C_{Med}) = \frac{sup(C_X \cup C_{Med})}{sup(C_X) \times sup(C_{Med})} \quad (3)$$

The entire process of retrieving IAR correlations is described in Algorithm 1. At the beginning, the combined label matrix is the input and the set of the indirect association rules IAR, frequent 1-itemset FI, and frequent itemset pair FIP are all initialized as empty sets. The support of each concept is calculated and compared with the minimum support $minsup$ to find all the frequent 1-itemsets FI. The frequent itemset pair FIP is successively generated using all possible combinations of FI (as described in Algorithm 1, lines 2 to 7). For each frequent itemset pair, assuming it is represented as C_X and C_Y, only the support ratio less than the itempair support threshold itp_s will be selected since we are looking for the hidden correlation for the infrequent itemsets. Later, the possible mediator concept C_{Med} will be collected based on its support ratio and interesting ratio toward the selected infrequent itemset pair (as described in Algorithm 1, lines 8 to 16). The important thresholds including $minsup$, itp_s, Med_s, and Med_d are decided from the best performance run in the training process.

Table I: Combined Label Matrix

	C_1	C_2	...	C_K	...	C_N
Instance 1	1	0	...	0	...	0
Instance 2	0	0	...	0	...	1
...	...	...	...	0	...	...
Instance i	0	0	...	0	...	0
...	...	...	...	...	...	...
Instance M	0	1	...	0	...	0

Algorithm 1: IARM Concept Correlation Retrieval

input : Combined Label Matrix $M \times N$, where M represents the number of data instances and N represents the number of concepts
output: IAR - A set of indirect association rules

1 $IAR \longleftarrow \emptyset$; $FI \longleftarrow \emptyset$; $FIPair \longleftarrow \emptyset$;
2 **for** *Each Concept* C_i, $i \leftarrow 1$ **to** N **do**
3   **if** $sup(C_i) > minsup$ **then**
4     $FI \longleftarrow C_i$
5   **end**
6 **end**
7 $FIPair \longleftarrow Combine(FI)$
8 **for** *Each* $FIPair(C_X, C_Y) \in FIPair$ **do**
9   **if** $sup(C_X, C_Y) < itp_s$ **then**
10   **for** *Each Concept* C_{Med}, $M \leftarrow 1$ **to** $Num(FI)$ **do**
11   **if** $sup(C_X \cup C_{Med}) \geq Med_s$ **and** $sup(C_Y \cup C_{Med}) \geq Med_s$ **and** $IR(C_X \Rightarrow C_{Med}) \geq Med_d$ **and** $IR(C_Y \Rightarrow C_{Med}) \geq Med_d$ **then**
12         $IAR \longleftarrow (C_X, C_Y, C_{Med})$
13   **end**
14   **end**
15   **end**
16 **end**

B. Integrate with Association Affinity Network (AAN)

The prototype of AAN was initially proposed in [29], called Concept Association Network (CAN). It starts with applying association rule mining (ARM) to select significant association links and capture the strong associations among different concepts. Next, CAN gradually improved with more essential factors such as negative correlation selection, estimated weight represented the posterior probabilities of correlations, and made it to what an AAN is. Inspired by the idea of AAN and other research work related to association rule mining (ARM) [30]–[34], which motivates us to introduce IAR in exploring the hidden concept correlations.

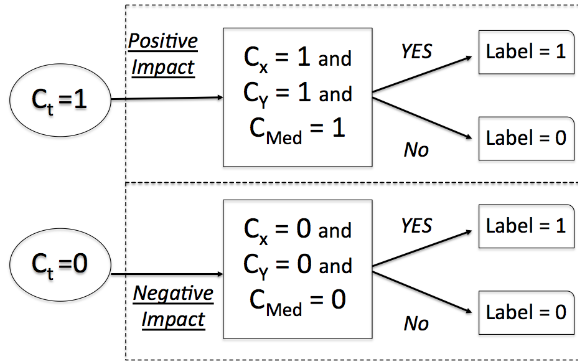

Figure 4: Two Types of IAR Label Generation

1) MCA-based IAR Weight Estimation: In conjunction with the negative correlations introduced in [23], for a target concept C_t, the same methodology of calculating the probability of detecting a positive target concept is applied for IAR. Let IAR consist of concept C_X, concept C_Y, and mediator concept C_{Med}, and F_i indicate the observed features for data instance i. If either C_X or C_Y is the target concept C_t selected from the negative correlations, then $P(C_t^1|F_i)$ can be used to represent the probability that i is negative, given F_i. With the assumption of IAR mentioned earlier, it can be expanded as shown in Equation 4.

$$\begin{aligned} P(C_t^1|F_i) &= P(C_t^1|C_{IAR}^0, F_i)P(C_{IAR}^0|F_i) \\ &\quad + P(C_t^1|C_{IAR}^1, F_i)P(C_{IAR}^1|F_i) \\ &= P(C_t^1, C_{IAR}^0|F_i) + P(C_t^1, C_{IAR}^1|F_i) \end{aligned} \tag{4}$$

To statistically quantify the impact of the IAR toward the target concept with the observed low-level feature values, two conditional probabilities, e.g., $P(C_t^1, C_{IAR}^0|F_i)$ and $P(C_t^1, C_{IAR}^1|F_i)$, are produced and summed up as $P(C_t^1|F_i)$. Two types of labels are redefined and generated based on the retrieved IAR correlations as shown in Figure 4. Afterward, the new labels along with the observed features are used to train the MCA-based weight estimation models for IAR. The upper side in Figure 4 describes the positive IAR impact toward the target concept C_t. Given a positive target concept, e.g., $C_t = 1$, the new label has value 1 if all the concepts included in the IAR are positive, and the new label has value 0, otherwise. The lower side in Figure 4 depicts the negative IAR impact toward the target concept. With a negative target concept, e.g., $C_t = 0$, the label's value is 1, if all concepts in the IAR are negative and the label's value is 0 for other cases.

Multiple Correspondence Analysis (MCA)-based model is selected to estimate these two probabilities. Originally, MCA was extended from the standard correspondence analysis to analyze the correlation among variables. Later, it has demonstrated its competence in enhancing multimedia retrieval research topics through capturing the correlations among high-level semantic concepts and low-level features [35]–[37], and modeling posterior probability [38]–[40].

2) Score Normalization and Regression-based Score Integration: Given the output generated from the target concept detectors, related concept detectors, and MCA-based weight estimation models, the effectiveness of using a negatively correlated concept to detect a target concept was modeled in [23].

In this paper, the idea of revealing the indirect association rules among the concept correlation network is introduced. Hence, a detection matrix DM can be formed where the first three vectors are target concept detector DM_t, related concept detectors DM_r, the negative correlation, which is between target concept and related concept, modeled by the MCA-based weight estimation DM_{nw}. Two more vectors are added at the end to represent the indirect association rule detector DM_{iar} and the corresponding weight estimated by the MCA-based methodology DM_{iw}. Therefore, each row DM^i can be represented by a row vector $[1, DM_t^i, DM_r^i, DM_{nw}^i, DM_{iar}^i, DM_{iw}^i]$. A likelihood function is formulated accordingly as shown in Equation 5. θ is the parameter vector composed of $[\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T$. In Equation 5, $C^i = 1$ indicates the label of a data instance is positive and $C^i = 0$ means the data instance is labeled as negative. m is the total number of data instances.

$$L(DM;\theta) = \prod_{i=1}^{m} (g(DM^i\theta))^{C^i} \cdot (1 - g(DM^i\theta))^{1-C^i} \\ where\ g(x) = \frac{1}{1+e^{-x}} \tag{5}$$

$$J(DM;\theta) = -logL(DM;\theta) + \lambda||\theta||_2 \\ subject\ to\ \theta_1 \geq 0, \theta_2 \leq 0, \theta_3 \leq 0, \theta_4 \geq 0, \theta_5 \geq 0. \tag{6}$$

Here, the variable θ_1 is the indicator of the positive target concept, and thus it is subject to be greater than or equal to zero. θ_2 and θ_3 are introduced to better estimate the negative correlation so that they should be both less than or equal to zero. Finally, θ_4 and θ_5 consider the impact on the positive target concept of having the indirect association rules or not, and therefore they are both set to be greater than zero. The variable $lambda$ is adopted in the cost function to avoid the possible overfitting problem.

III. Experiments

A. Dataset

The dataset "IACC.1.B" prepared for the TRECVID 2011 semantic indexing task [41] is adopted as a benchmark dataset to evaluate the classification results among different methods. The labels of the 346 high-level semantic concepts are provided through a collaborative annotation activity hosted by NIST [42] and the concept list can be found with detailed definition in [41]. It is a collection of videos with a total duration of 200 hours, and each video lasts between 10 seconds and 3.5 minutes. The detection scores were generously provided by the Shinoda Lab at the Department of Computer Science at Tokyo Institute of Technology [10], whose group achieved the top performance at the TRECVID 2011 Semantic Indexing Task.

Table II: Dataset statistics information

Dataset	IACC.1.B
TRECVID Year	2011
No. Concepts	346
No. Training Instances	144774
No. Testing Instances	137327
Average Positive No. Instances	408.42
Average P / N Ratio	0.003

Table V: Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

B. Evaluation Criteria

The well-known measurement method called Mean Average Precision (MAP) is used. To calculate and understand the MAP value, a derivation process is described as follows,

First, **Precision** is an accuracy evaluation method, and is derived from the confusion matrix as shown in Table V. Confusion matrix is widely used in machine learning and data mining areas to visualize the classification results in table-layout fashion and based on it, precision can be calculated as shown in Equation 7. It demonstrates the fraction of retrieved instances that are relevant, where a high precision value indicates a lower false positive rate.

Table III: MAP values at different number of instances retrieved for IACC.1.A

Frameworks	Top10	Top20	Top40	Top60	Top80	Top100	Top500	Overall
RAW	0.4508	0.4084	0.3576	0.3137	0.2738	0.2441	0.1305	0.1910
DASD	0.4827	0.4020	0.3340	0.3113	0.2786	0.2431	0.1222	0.1778
AAN	0.8626	0.7355	0.6054	0.5588	0.5105	0.4729	0.3397	0.4478
AAN + IAR (Proposed)	0.8820	0.7710	0.6343	0.5876	0.5451	0.4945	0.3757	0.5123

Table IV: MAP values at different number of instances retrieved for IACC.1.A using three-fold cross validation

Fold Number	Top10	Top20	Top40	Top60	Top80	Top100	Top500	Overall
Fold1	0.8723	0.7810	0.6188	0.5846	0.5541	0.5007	0.3719	0.5075
Fold2	0.8935	0.7533	0.6443	0.5907	0.5367	0.4867	0.3688	0.4935
Fold3	0.8801	0.7786	0.6397	0.5876	0.5444	0.4962	0.3864	0.5103
Overall	0.8820	0.7710	0.6343	0.5876	0.5451	0.4945	0.3757	0.5123

- **Precision**

$$Precision = \frac{TruePos}{(TruePos + FalsePos)} \tag{7}$$

- **Average Precision and Mean Average Precision**
  Average precision (AP) and mean average precision (MAP) are two metrics extended from precision, as defined in Equation 8 and Equation 9, respectively. **Average Precision** at K is used to evaluate the top K ranked results, where $\#(TopR)$ represents the number of data instances which are correctly classified as positive instances among the top R retrieved data instances, $R = 1...K$. A higher AP value means more relevant results are ranked earlier than the irrelevant ones.

$$AP(K) = \frac{1}{K}\sum_{R=1}^{K}\frac{\#(TopR)}{R} \tag{8}$$

  Mean Average Precision is used to validate the ranked results for more than one concept, where TC is the total number of concepts and $AP_C(K)$ is the average precision at K for concept C. It can also be used to represent the overall performance for a three-fold cross alidation experiment.

$$MAP(K) = \frac{\sum_{C=1}^{TC} AP_C(K)}{TC} \tag{9}$$

C. Experimental Results

To evaluate the proposed framework, it was compared with three different frameworks. First, the original ranking scores without any modifications were indicated as “RAW”. Second, the domain adaptive semantic diffusion “DASD” proposed in [24] was applied. Third, the association affinity network with only the negative correlation proposed in [23] was indicated as “AAN”. The last one is the proposed framework, which is indicated as “AAN + IAR”.

The MAP values at different numbers of retrieved data instances are reported for each framework as shown in Table III. The last column represents the MAP values calculated while considering all the testing data instances. All the results are the average MAP values of a three-fold cross validations. The comparisons between “RAW” and “AAN” show the importance of mining negative concept correlations. Tao et al. has explained two possible reasons why “AAN” has higher MAP values against “DASD” in [23]. One is the selection of significant negative concept correlations and the other is the accuracy of posterior probability estimation. Most importantly, the proposed framework produced the highest MAP in various retrieved levels among all the frameworks, which can be explained in two-fold. First, using IAR correlations is able to dig out the valuable correlations from the infrequent concept itemsets, which are those concepts rarely being identified together in the same data instance. Second, applying IAR correlations is able to identify interesting negative correlations, because $P(C_t^1, C_{IAR}^0|F_i)$ and $P(C_t^1, C_{IAR}^1|F_i)$ comprehensively consider the IAR's positive and negative impacts toward selected negative correlations from AAN.

In Table IV, the steadiness of the proposed method can also be reflected from the MAP values generated for each fold. There are no major differences among the classification results for three folds, which shows the robustness of the proposed method. In addition, all the folds can perform close to 50% MAP values when considering the whole testing dataset.

IV. Conclusion

In this paper, the idea of indirect association rule mining (IARM) is introduced into a semantic concept detection framework for multimedia semantic retrieval. First, a novel algorithm is proposed to retrieve significant IAR correlations based on the statistic information of the semantic concept labels. Two types of newly defined labels are used to train the weight estimation models for generating the posterior probability between the IAR and the target positive concepts. Lastly, the IAR correlation model is incorporated with the negative correlations to refine the final ranking scores through the explicit normalization and regression-based model designed for dual correlations. From the experiments, the proposed framework achieved the highest classification results against other related work demonstrate the strength in two folds. First, thoroughly exploring the indirect semantic concept correlations can enhance the classification results for semantic concept retrieval from a large amount of multimedia data. Second, discovering IAR correlation is a good combination with the existing negative-based correlation framework because of its capability of detecting the interesting negative correlations.

V. Acknowledgment

This research was supported in part by NSF HRD-0833093 and CNS-1126619.

References

[1] A. Dragland, "Big Data for better or worse," http://www.sintef.no/home/corporate-news/Big-Data–for-better-or-worse/, 2013, [Online; accessed 22-May-2013].

[2] S.-C. Chen, R. L. Kashyap, and A. Ghafoor, *Semantic models for multimedia database searching and browsing*. Springer Science & Business Media, 2000, vol. 21.

[3] M.-L. Shyu, C. Haruechaiyasak, S.-C. Chen, and K. Premaratne, "Mining association rules with uncertain item relationships," *Computers and Industrial Engineering*, vol. 34, no. 1, pp. 3–20, 1998.

[4] S.-C. Chen, S. H. Rubin, M.-L. Shyu, and C. Zhang, "A dynamic user concept pattern learning framework for content-based image retrieval," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 36, no. 6, pp. 772–783, 2006.

[5] X. Li, S.-C. Chen, M.-L. Shyu, and B. Furht, "An effective content-based visual image retrieval system," in *IEEE International Conference on Computer Software and Applications Conference, (COMPSAC)*, 2002, pp. 914–919.

[6] M.-L. Shyu, S.-C. Chen, M. Chen, C. Zhang, and K. Sarinnapakorn, "Image database retrieval utilizing affinity relationships," in *ACM International Workshop on Multimedia Databases*, 2003, pp. 78–85.

[7] S.-C. Chen and R. Kashyap, "Temporal and spatial semantic models for multimedia presentations," in *International Symposium on Multimedia Information Processing*, 1997, pp. 441–446.

[8] H.-Y. Ha, S.-C. Chen, and M. Chen, "FC-MST: Feature correlation maximum spanning tree for multimedia concept classification," in *IEEE International Conference on Semantic Computing (ICSC)*, 2015, pp. 276–283.

[9] Y. Yang, J. Song, Z. Huang, Z. Ma, N. Sebe, and A. G. Hauptmann, "Multi-feature fusion via hierarchical regression for multimedia analysis," *IEEE Transactions on Multimedia*, vol. 15, no. 3, pp. 572–581, 2013.

[10] N. Inoue and K. Shinoda, "A fast and accurate video semantic-indexing system using fast map adaptation and gmm supervectors," *IEEE Transactions on Multimedia*, vol. 14, no. 4, pp. 1196–1205, 2012.

[11] M. Chen, S.-C. Chen, and M.-L. Shyu, "Hierarchical temporal association mining for video event detection in video databases," in *IEEE International Conference On Data Engineering Workshop*, 2007, pp. 137–145.

[12] S.-C. Chen, M.-L. Shyu, C. Zhang, and M. Chen, "A multimodal data mining framework for soccer goal detection based on decision tree logic," *International Journal of Computer Applications in Technology*, vol. 27, no. 4, pp. 312–323, 2006.

[13] S. Chen, S. Sista, M.-L. Shyu, and R. L. Kashyap, "Indexing and searching structure for multimedia database systems," in *Electronic Imaging*. International Society for Optics and Photonics, 1999, pp. 262–270.

[14] M.-L. Shyu, S.-C. Chen, M. Chen, and C. Zhang, "A unified framework for image database clustering and content-based retrieval," in *ACM International Workshop on Multimedia Databases*, 2004, pp. 19–27.

[15] L. Lin, G. Ravitz, M.-L. Shyu, and S.-C. Chen, "Effective feature space reduction with imbalanced data for semantic concept detection," in *IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing (SUTC)*, 2008, pp. 262–269.

[16] F. C. Fleites, H.-Y. Ha, Y. Yang, and S.-C. Chen, "Large-scale correlation-based semantic classification using mapreduce," *Cloud Computing and Digital Media: Fundamentals, Techniques, and Applications*, p. 169, 2014.

[17] Y. Yang, H.-Y. Ha, F. C. Fleites, and S.-C. Chen, "A multimedia semantic retrieval mobile system based on HCFGs," *IEEE MultiMedia*, vol. 21, no. 1, pp. 36–46, 2014.

[18] Q. Zhu, L. Lin, M.-L. Shyu, and S.-C. Chen, "Effective supervised discretization for classification based on correlation maximization," in *IEEE International Conference on Information Reuse and Integration (IRI)*, 2011, pp. 390–395.

[19] H.-Y. Ha, F. C. Fleites, S.-C. Chen, and M. Chen, "Correlation-based re-ranking for semantic concept detection," in *IEEE International Conference on Information Reuse and Integration (IRI)*, 2014, pp. 765–770.

[20] J. Costa Pereira, E. Coviello, G. Doyle, N. Rasiwasia, G. R. Lanckriet, R. Levy, and N. Vasconcelos, "On the role of correlation and abstraction in cross-modal multimedia retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 521–535, 2014.

[21] S. Nikolova, J. Boyd-Graber, and C. Fellbaum, "Collecting semantic similarity ratings to connect concepts in assistive communication tools," in *Modeling, Learning, and Processing of Text Technological Data Structures*. Springer, 2012, pp. 81–93.

[22] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, and N. Ayache, "Learning semantic and visual similarity for endomicroscopy video retrieval," *IEEE Transactions on Medical Imaging*, vol. 31, no. 6, pp. 1276–1288, 2012.

[23] T. Meng, Y. Liu, M.-L. Shyu, Y. Yan, and C.-M. Shu, "Enhancing multimedia semantic concept mining and retrieval by incorporating negative correlations," in *IEEE International Conference on Semantic Computing (ICSC)*, 2014, pp. 28–35.

[24] Y.-G. Jiang, J. Wang, S.-F. Chang, and C.-W. Ngo, "Domain adaptive semantic diffusion for large scale context-based video annotation," in *IEEE International Conference on Computer Vision*, 2009, pp. 1420–1427.

[25] R. Hong, M. Wang, Y. Gao, D. Tao, X. Li, and X. Wu, "Image annotation by multiple-instance learning with discriminative feature mapping and selection," *IEEE Transactions on Cybernetics*, vol. 44, no. 5, pp. 669–680, 2014.

[26] Y.-G. Jiang, Q. Dai, J. Wang, C.-W. Ngo, X. Xue, and S.-F. Chang, "Fast semantic diffusion for large-scale context-based image and video annotation," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 3080–3091, 2012.

[27] P.-N. Tan, V. Kumar, and J. Srivastava, *Indirect association: Mining higher order dependencies in data*. Springer, 2000.

[28] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD Record*, vol. 22, no. 2, 1993, pp. 207–216.

[29] T. Meng and M.-L. Shyu, "Leveraging concept association network for multimedia rare concept mining and retrieval," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2012, pp. 860–865.

[30] J. R. Smith, M. Naphade, and A. Natsev, "Multimedia semantic indexing using model vectors," in *IEEE International Conference on Multimedia and Expo (ICME)*, vol. 2, 2003, pp. II–445.

[31] K.-H. Liu, M.-F. Weng, C.-Y. Tseng, Y.-Y. Chuang, and M.-S. Chen, "Association and temporal rule mining for post-filtering of semantic concept detection in video," *IEEE Transactions on Multimedia*, vol. 10, no. 2, pp. 240–251, 2008.

[32] A. Babashzadeh, M. Daoud, and J. Huang, "Using semantic-based association rule mining for improving clinical text retrieval," in *Health Information Science*. Springer, 2013, pp. 186–197.

[33] V. Nebot and R. Berlanga, "Finding association rules in semantic web data," *Knowledge-Based Systems*, vol. 25, no. 1, pp. 51–62, 2012.

[34] J. Ke, Y. Zhan, X. Chen, and M. Wang, "The retrieval of motion event by associations of temporal frequent pattern growth," *Future Generation Computer Systems*, vol. 29, no. 1, pp. 442–450, 2013.

[35] H.-Y. Ha, F. C. Fleites, and S.-C. Chen, "Content-based multimedia retrieval using feature correlation clustering and fusion," *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, vol. 4, no. 2, pp. 46–64, 2013.

[36] L. Lin, G. Ravitz, M.-L. Shyu, and S.-C. Chen, "Video semantic concept discovery using multimodal-based association classification," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2007, pp. 859–862.

[37] L. Lin and M.-L. Shyu, "Weighted association rule mining for video semantic detection," *Methods and Innovations for Multimedia Database Content Management*, p. 12, 2012.

[38] T. Meng and M.-L. Shyu, "Automatic annotation of drosophila developmental stages using association classification and information integration," in *IEEE International Conference on Information Reuse and Integration (IRI)*, 2011, pp. 142–147.

[39] L. Lin, C. Chen, M.-L. Shyu, and S.-C. Chen, "Weighted subspace filtering and ranking algorithms for video concept retrieval," *IEEE MultiMedia*, vol. 18, no. 3, pp. 32–43, 2011.

[40] H.-Y. Ha, F. C. Fleites, and S.-C. Chen, "Building multi-model collaboration in detecting multimedia semantic concepts," in *IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom)*, 2013, pp. 205–212.

[41] A. F. Smeaton, P. Over, and W. Kraaij, "Evaluation campaigns and trecvid," in *ACM International Workshop on Multimedia Information Retrieval*, 2006, pp. 321–330.

[42] S. Ayache and G. Quénot, "Video corpus annotation using active learning," in *Advances in Information Retrieval*. Springer, 2008, pp. 187–198.