Using Recommendation to Explore Educational Video

Matthew Cooper, Jian Zhao, Chidansh Bhatt, David A. Shamma
FX Palo Alto Laboratory
Palo Alto, CA, USA
{cooper,zhao,bhatt,ayman}@fxpal.com

ABSTRACT

Massive Open Online Course (MOOC) platforms have scaled online education to unprecedented enrollments, but remain limited by their rigid, predetermined curricula. Increasingly, professionals consume this content to augment or update specific skills rather than complete degree or certification programs. To better address the needs of this emergent user population, we describe a visual recommender system called MOOCex. The system recommends lecture videos *across* multiple courses and content platforms to provide a choice of perspectives on topics. The recommendation engine considers both video content and sequential inter-topic relationships mined from course syllabi. Furthermore, it allows for interactive visual exploration of the semantic space of recommendations within a learner's current context.

CCS CONCEPTS

• **Computer systems organization → Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

KEYWORDS

ACM proceedings, LaTeX, text tagging

ACM Reference Format:
Matthew Cooper, Jian Zhao, Chidansh Bhatt, David A. Shamma. 2018. Using Recommendation to Explore Educational Video. In *Proceedings of ACM Intl. Conf. on Multimedia Retrieval (ICMR 2018).* ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Modern online education platforms, such as Coursera, edX, and Udacity, have become popular in recent years. These platforms allow for teaching at a distance and at scale by presenting educational materials as Massive Open Online Courses (MOOC). A course usually consists of a number of short videos, each targeting a specific concept. To achieve certain learning objectives, instructors commonly order the videos within a syllabus which may also group videos hierarchically into sections. However, the syllabus remains a *one size fits all* approach with a predefined curriculum, which contributes to courses' frequent low retention rates [20, 28].

While these services initially aimed to disrupt the higher education market, more recent studies show that professionals, rather than students, comprise an increasing portion of MOOC learners. These "lifelong learners" typically aim to achieve career growth rather than complete degrees or obtain certifications. As a result, they are less likely to adhere to the syllabus [28], and often have varied individual information needs and experience levels. It is critical to offer learners more flexible access to a broader range of content and perspectives (e.g., from multiple courses) [6, 15, 25].

ICMR 2018, June 2018, Yokohama, JAPAN

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

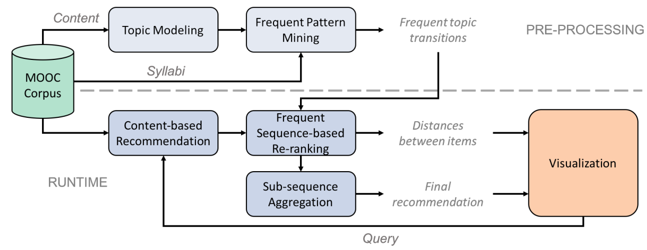

Figure 1: System architecture overview of MOOCex.

Although the syllabus can be too rigid, it encodes expert instructors' sequencing of topics to best aid comprehension. Indeed, the sequential character of educational video is an important cue for effectively exploring this content. Platforms such as KhanAcademy [14] provide an interactive knowledge (concept) map to enable more personalized navigation. However, concept maps are not well suited for sequential flow [19] and are neither scalable nor adaptive.

MOOCex aims to help learners effectively access MOOC content across courses and content platforms. We first aggregate this content across multiple courses, introducing the challenge of linking videos in which instructors cover related topics from different perspectives. Additionally, users must select among such related videos to address their information needs. While the course syllabus provides guidance *within* courses, we propose content-based recommendation and interactive visualization to facilitate navigation *across* courses.

MOOCex builds upon advanced data mining techniques, and recommends lectures by considering both videos' topics and sequential inter-topic relationships. MOOCex optionally recommends short sub-sequences of videos within courses, rather than individual videos, to provide additional depth around specific concepts and simplify learning. Unlike conventional user interfaces for recommendation, i.e., a ranked list or a set of ranked lists, MOOCex supports semantic visualization of recommended videos in users' current learning context, by projecting videos onto a 2D space annotated with topical regions and key phrases. This provides additional dimensions for learners to effectively explore related content and select what to watch next confidently.

2 RELATED WORK

In this section, we review related work with a focus on video recommendation techniques as well as visual analytics systems for MOOC content and data.

2.1 Recommendation

Many approaches to video recommendation are content-based, for example, using features extracted via analysis of text transcripts or video frames. VideoReach implements recommendation by combining unimodal models for textual, visual, and aural information of videos [17]. TalkMiner is an educational video search system using OCR and lexical analysis of text displayed in video frames [1]. Topic modeling, such as Latent Dirichlet Allocation (LDA) [4], has also been applied [29]. However, none of these methods considered the sequential relationships between videos, or the explicit inter-video orderings of course syllabi.

Likewise, applications of sequential pattern mining techniques are generally distinct from video recommendation. Morales et al. [18] facilitates information discovery via educational hypermedia linking based on sequence mining of user logs. Agrawal et al. [2] proposed a system for linking web videos to supplement electronic textbooks, and argued that topic mining alone is insufficient. In other applications, sequential information has been used in recommending music [11], online products [5, 12], and travel itineraries [13].

Sequential organization of topics by experts in course syllabi can provide valuable information for educational content recommendation. However, we face the additional challenge that videos do not overlap across different courses. This contrasts with datasets used in previous user-centered work including, music playlists [11], movies (i.e.Netflix) or e-commerce products (i.e.amazon and [12]). To address this issue, we apply established sequential pattern mining techniques to analyze both global and local patterns in topic transitions exhibited in instructors' syllabi.

2.2 Information visualization

Many techniques have been proposed for visual analysis of data generated by MOOCs, such as user clickstreams and forum discussions [21]. One main topic in this area is to study learner behaviors [7, 24]. In addition, iForum provides another perspective for understanding learners via the analysis of the content and structure of MOOC forum threads [10].

In contrast with systems designed for instructors or analysts, we focus on recommendation and visual exploration of MOOC videos to benefit ordinary learners. One work that shares similar goals with ours is booc.io [23], which allows for visually exploring concept maps of instructional materials and following personalized learning plans. However, this approach is less scalable or flexible because of the required manual creation of the concept maps beforehand. Also, they focus on video exploration within a single course whereas we recommend videos sourced across multiple MOOC platforms.

3 SYSTEM DESCRIPTION

MOOCex consists of two components: a content-based recomemendation engine and a visual interface for video playback and semantic exploration. We describe each in the sections below.

3.1 Recommendation engine

While syllabi provide learners guidance within courses, we use recommendation to facilitate exploration of a multi-course corpus. We use sequential pattern mining to incorporate sequential information into the recommendation engine.

We assemble content from a number of courses, including videos, available text transcripts, and other meta-data. Generating recommendations across heterogeneous content requires a common data representation. Useful visual attributes such as whether the format is a classroom lecture, khan-academy style electronic ink, or slide-based video, etc. are largely captured by associated meta-data. Most platforms provide semantically rich closed caption transcripts enabling the identification of videos' prominent topics. We discover latent topics present across the collection using LDA [4], an established unsupervised topic modeling method.

Sequential pattern mining (SPM) algorithms identify prominent subsequences within a sequence database [9]. Denote the *topic signature* of the ith video by $V_i = \{k : P^{(i)}(z_k) \geq 0.1\}$ which is the discrete set of topic indices weighted at least 0.1 in the video's topic distribution. We construct an ordered sequence of video topic signatures according to each course's syllabus. These sequences are aggregated into a sequence database.

Our aim is to use the currently watched video to recommend videos covering concepts users are likely to *watch next*. However, frequent topic subsequence detection alone is not sufficient for prediction. Sequence mining addresses prediction by discovering rules of the form $X \Rightarrow Y$, where X and Y are two sets of topics. $X \Rightarrow Y$ indicates "topic(s) Y appear in the sequence after topic(s) X". The Top-K non-redundant sequential patterns (TNS) algorithm [9] detects sequential rules that reflect *global* analysis of inter-topic transitions. TNS eliminates rules that are deemed "redundant" (rules that are implied by other rules having the same support and confidence) to capture more varied sequences and automatically fine-tunes the minimum support parameter.

Relatively infrequent topic sets will be overlooked in the global analysis. The Top-K sequential pattern mining (TKS) algorithm [8] finds sequential patterns within a given minimum and maximum length such that specified items must appear within a defined allowed gap. For *local* analysis, we collect sequences that include each video's signature. We apply TKS with the constraint that each video's topic set appears within a distance of 3 to 6 in the sequence. We then apply TNS algorithm on these derived sequences to find significant sequential rules within this local data subset. By this design, each video's topic signature is described by sequential rules in the local analysis. Each sequential rule in the local analysis has a corresponding confidence score. This application of sequential pattern mining produces sets of prominent topic transitions describing the sequence database both globally and locally.

The recommendation engine first issues the currently watched video as the query against our baseline content-based recommendation system. This currently uses standard tf/idf vectorspace retrieval [16] based on the video transcripts with ranking according to cosine similarity. To emphasize results that users are more likely to want to *watch next* we have introduced a re-ranking method that incorporates scores emphasizing results consistent with topic transitions mined in the global analysis of the corpus. For global analysis, we retrieve N support and confidence score values, $\{(s_n, c_n)\}$ from the mined *global* sequential patterns with antecedent values matching V_q and consequent values matching subsets of V_r. The *GS score* is

$$\mathrm{Sim}_{\mathrm{GS}}(V_q, V_r) = \frac{1}{N}\sum_{n=0}^{N} c_n \frac{|V_r \cap V_q|}{|V_r|} + \frac{s_n}{D_G} \,. \tag{1}$$

D_G is the total number of global sequences. Additionally, we retrieve M additional support and confidence score values from mined *local* sequential patterns with antecedent matching a subset of V_q and consequent matching a subset of V_r. The *LS score* is

$$\mathrm{Sim}_{\mathrm{LS}}(V_q, V_r) = \frac{1}{M}\sum_{m=0}^{M} c_m \frac{|V_r \cap V_q|}{|V_r|} + \frac{s_m}{D_q} \tag{2}$$

D_q is total number of mined local sequences with antecedent matching any subset of V_q. These scores are uniformly fused with a topic similarity score derived from LDA to re-rank the initial pool of recommendations. For additional details, refer to [3].

3.2 Visualization interface

The MOOCex interface consists of three parts. The *Video Panel* is a media player for watching a selected video. The *Recommendation Panel* is where a learner can explore recommended videos and assess relationships between them, to inform their choice of a video to watch next. The *Configuration Panel* allows manipulation of basic parameters controlling the display of recommendations and the specific courses and videos in view.

The Recommendation Panel displays the current video, adjacent videos in its syllabus, and recommendations in a two-dimensional *Exploration Canvas* in the middle. Other videos in the current course appear in order on both sides (Figure 2-b), and are connected with gray arrows. Each video is represented as a circle with a number indicating its position in the course syllabus. Color hues indicate different courses, and color opacity indicates the rank of that video in the recommendation list (the lighter the color, the lower its rank).

In contrast to ranked result lists, we employ multidimensional scaling (MDS) [26] to position videos on the Exploration Canvas based on their pairwise similarities. In MDS, only the relative distance between items has meaning while the axes do not. Thus, we rotate the layout so that the direction of videos in the current course flows from left to right, aligning with the natural videos on either side. This rotation eases comprehension of the visualization. Zig-zags in longer video sequences occur, which cannot be completely averted. Because learners often focus on semantics in a local space, they typically do not include many neighboring videos in the Exploration Canvas. To minimize overlap of circles, we later apply a repulsive force between videos to obtain the final layout.

To help learners utilize the MDS layout, we perform an agglomerative clustering [27] of the videos, and split the Exploration Canvas into corresponding regions. Each region exhibits a relatively coherent set of topics. We use the agglomerative approach because it does not require prior knowledge of the number of clusters. The boundaries of the regions are shown as subtle white polylines, determined by aggregating Voronoi cells of videos in the same cluster.

In addition, we overlay frequent topical keywords extracted from the text transcripts of each video cluster to reveal contextual information of different regions in the MDS projection. To obtain discriminative keywords, we first employ the standard TF-IDF method and then re-weight the keywords based on terms in video titles, because the titles are created by human and provide high quality summarization. Next, we post-process the keywords for the video clusters to remove duplications. These keywords are placed using a force-directed layout, and can be hidden if users feel overwhelmed.

To facilitate exploration of videos, the Recommendation Panel displays auxiliary information on both sides (Figure 2-b). Videos that were recently visited and adjacent videos from the current course are shown in two vertical lists on the left. Similarly, recommended videos are shown on the right in a ranked list similar to the traditional approach. Interactive linking of the same video is provided as it is hovered over in the lists or in the Exploration Canvas. Meanwhile, a tooltip pops-up showing a set of important keywords extracted from the video transcript and title based on the RAKE algorithm [22] (Figure 2-d). Also, clicking any of the videos selects it as the current video and updates the visualization. Our goal here is to provide some semantic structure (not necessarily precisely) in the information space, allowing everyday users to better understand the MDS layout and confidently select a video to watch next.

4 CONCLUSION

Our corpus currently includes over four thousand videos from 41 MOOCs. MOOCex enables exploration of this content via recommendations that reflect sequential patterns mined from the collection of course syllabi. The visual interface facilitates effective exploration of the semantic relationships between videos within a course and the recommendation results. The prominence of inter-topic sequential information and inter-video semantics in our system empowers users to confidently select a video to watch next.

REFERENCES

[1] John Adcock, Matthew Cooper, Laurent Denoue, Hamed Pirsiavash, and Lawrence A. Rowe. 2010. TalkMiner: A Lecture Webcast Search Engine. In *Proceedings of the ACM International Conference on Multimedia (MM '10)*. 241–250. https://doi.org/10.1145/1873951.1873986

[2] Rakesh Agrawal, Maria Christoforaki, Sreenivas Gollapudi, Anitha Kannan, Krishnaram Kenthapadi, and Adith Swaminathan. 2014. *Mining Videos from the Web for Electronic Textbooks.* Springer International Publishing, 219–234. https://doi.org/10.1007/978-3-319-07248-7_16

[3] Chidansh Bhatt, Matthew Cooper, and Jian Zhao. 2018. SeqSense: Video Recommendation Using Topic Sequence Mining. In *MultiMedia Modeling*, Klaus Schoeffmann, Thanarat H. Chalidabhongse, Chong Wah Ngo, Supavadee Aramvith, Noel E. O'Connor, Yo-Sung Ho, Moncef Gabbouj, and Ahmed Elgammal (Eds.). Springer International Publishing, Cham, 252–263.

[4] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3 (2003), 993–1022.

[5] Keunho Choi, Donghee Yoo, Gunwoo Kim, and Yongmoo Suh. 2012. A Hybrid Online-product Recommendation System: Combining Implicit Rating-based Collaborative Filtering and Sequential Pattern Analysis. *Electronic Commerce Research and Applications* 11, 4 (July 2012), 309–317. https://doi.org/10.1016/j.elerap.2012.02.004

[6] Carleton Coffrin, Linda Corrin, Paula de Barba, and Gregor Kennedy. 2014. Visualizing Patterns of Student Engagement and Performance in MOOCs. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK '14)*. 83–92. https://doi.org/10.1145/2567574.2567586

[7] Franck Dernoncourt, Colin Taylor, Una-May O'Reilly, Kayan Veeramachaneni, Sherwin Wu, Chuong Do, and Sherif Halawa. 2013. MoocViz: A Large Scale, Open Access, Collaborative, Data Analytics Platform for MOOCs. In *Proceedings of NIPS Workshop on Data-Driven Education.*

[8] Philippe Fournier-Viger, Antonio Gomariz, Ted Gueniche, Espérance Mwamikazi, and Rincy Thomas. 2013. TKS: Efficient Mining of Top-K Sequential Patterns. In *Advanced Data Mining and Applications*, Hiroshi Motoda, Zhaohui Wu, Longbing

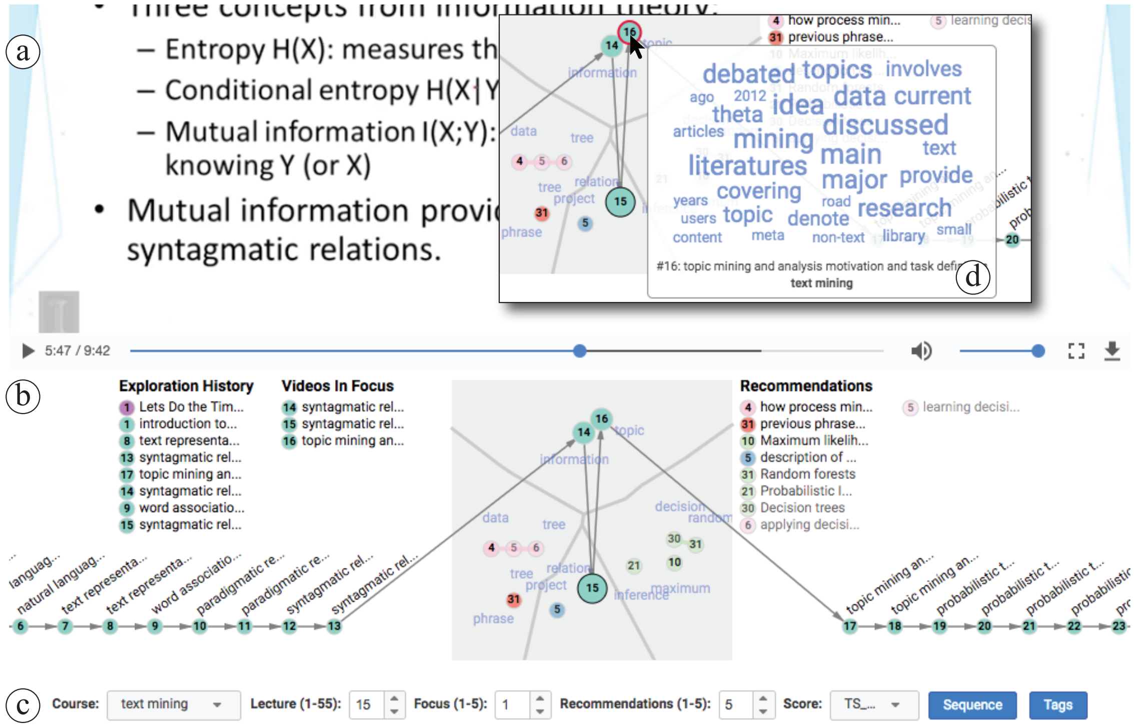

Figure 2: A learner is using MOOCex to explore recommendations for lecture "15" in *Text Mining*. The interface consists of three main components: a) a Video Panel (cropped), b) a Recommendation Panel, and c) a Configuration Panel, d) Hovering over a video initiates a tooltip with a tag cloud to describe the content of the video.

Cao, Osmar Zaiane, Min Yao, and Wei Wang (Eds.). Springer Berlin Heidelberg, 109–120. https://doi.org/10.1007/978-3-642-53914-5_10

[9] Philippe Fournier-Viger and Vincent S. Tseng. 2013. TNS: Mining Top-k Non-redundant Sequential Rules. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13)*. 164–166. https://doi.org/10.1145/2480362.2480395

[10] S. Fu, J. Zhao, W. Cui, and H. Qu. 2017. Visual Analysis of MOOC Forums with iForum. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 201–210. https://doi.org/10.1109/TVCG.2016.2598444

[11] Negar Hariri, Bamshad Mobasher, and Robin Burke. 2012. Context-aware Music Recommendation Based on Latenttopic Sequential Patterns. In *Proceedings of ACM Conference on Recommender Systems (RecSys '12)*. 131–138. https://doi.org/10.1145/2365952.2365979

[12] Ruining He, Wang-Cheng Kang, and Julian McAuley. 2017. Translation-based Recommendation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys '17)*. ACM, New York, NY, USA, 161–169. https://doi.org/10.1145/3109859.3109882

[13] S. Jiang, X. Qian, T. Mei, and Y. Fu. 2016. Personalized Travel Sequence Recommendation on Multi-Source Big Social Media. *IEEE Transactions on Big Data* 2, 1 (March 2016), 43–56. https://doi.org/10.1109/TBDATA.2016.2541160

[14] KhanAcademy Knowledge Map. Accessed in 2017. https://www.khanacademy.org/exercisedashboard. (Accessed in 2017).

[15] Juho Kim, Philip J. Guo, Daniel T. Seaton, Piotr Mitros, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Understanding In-video Dropouts and Interaction Peaks Inonline Lecture Videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14)*. 31–40. https://doi.org/10.1145/2556325.2566237

[16] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval.* Cambridge University Press.

[17] Tao Mei, Bo Yang, Xian-Sheng Hua, and Shipeng Li. 2011. Contextual Video Recommendation by Multimodal Relevance and User Feedback. *ACM Transactions on Information Systems* 29, 2, Article 10 (2011), 24 pages. https://doi.org/10.1145/1961209.1961213

[18] C Romero Morales, AR Porras Pérez, S Ventura Soto, C Hervás Martınez, and A Zafra. 2006. Using sequential pattern mining for links recommendation in adaptive hypermedia educational systems. In *Current Developments in Technology Assisted Education*, Vol. 2. 1016–1020.

[19] Joseph D Novak and D Bob Gowin. 1984. *Learning how to learn.* Cambridge University Press.

[20] D.F.O. Onah, J. Sinclair, and R. Boyatt. 2014. Dropout Rates Of Massive Open Online Courses: Behavioural Patterns. In *Proceedings of International Conference on Education and New Learning Technologies.* 5825–5834.

[21] H. Qu and Q. Chen. 2015. Visual Analytics for MOOC Data. *IEEE Computer Graphics and Applications* 35, 6 (Nov 2015), 69–75. https://doi.org/10.1109/MCG.2015.137

[22] Stuart Rose, Dave Engel, Nick Cramer, and Wendy Cowley. 2010. Automatic keyword extraction from individual documents. *Text Mining: Applications and Theory* (2010), 1–20.

[23] M. Schwab, H. Strobelt, J. Tompkin, C. Fredericks, C. Huff, D. Higgins, A. Strezhnev, M. Komisarchik, G. King, and H. Pfister. 2017. booc.io: An Education System with Hierarchical Concept Maps and Dynamic Non-linear Learning Plans. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 571–580. https://doi.org/10.1109/TVCG.2016.2598518

[24] Conglei Shi, Siwei Fu, Qing Chen, and Huamin Qu. 2015. VisMOOC: Visualizing video clickstream data from Massive Open Online Courses. In *Proceedings of IEEE Pacific Visualization Symposium (PacificVis)*. 159–166. https://doi.org/10.1109/PACIFICVIS.2015.7156373

[25] R. Spiro and J. Jehng. 1990. Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In *Cognition, education and multimedia: Exploring ideas in high technology*, D. Nix and R. Spiro (Eds.). Lawrence Erlbaum.

[26] Warren S. Torgerson. 1952. Multidimensional scaling: I. Theory and method. *Psychometrika* 17, 4 (Dec 1952), 401–419. https://doi.org/10.1007/BF02288916

[27] Ian H. Witten, Eibe Frank, and Mark A. Hall. 2011. *Data Mining: Practical Machine Learning Tools and Techniques* (3rd ed.). Morgan Kaufmann.

[28] Saijing Zheng, Mary Beth Rosson, Patrick C. Shih, and John M. Carroll. 2015. Understanding Student Motivation, Behaviors and Perceptions in MOOCs. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '15)*. 1882–1895. https://doi.org/10.1145/2675133.2675217

[29] Q. Zhu, M. L. Shyu, and H. Wang. 2013. VideoTopic: Content-Based Video Recommendation Using a Topic Model. In *Proceedings of the IEEE International Symposium on Multimedia.* 219–222. https://doi.org/10.1109/ISM.2013.41