

Multilevel Visualisation of Topic Dependency Models for Assessment Design and Delivery: A Hypergraph Based Approach

Kendra M.L. Cooper¹ and Hassan Khosravi²

¹Independent Scholar, Canada

²University of Queensland, Australia

Abstract

The effective design and delivery of assessments in a wide variety of evolving educational environments remains a challenging problem. Proposals have included the use of learning dashboards, peer learning environments, and grading support systems; these embrace visualisations to summarise and communicate results. In an on-going project, the investigation of graph based visualisation models for assessment design and delivery has yielded promising results. Here, an alternative graph foundation, a two-weighted hypergraph, is considered to represent the assessment material (e.g., questions) and their explicit mapping to one or more learning objective topics. The visualisation approach considers the hypergraph as a collection of levels; the content of these levels can be customized and presented according to user preferences. A case study on generating hypergraph models using commonly available assessment data and a flexible visualisation approach using historical data from an introductory programming course is presented.

1 Introduction

Assessment remains a core educational activity, even as environments continue to evolve beyond the traditional classroom. Blended, flipped, and massive open on-line courses are supported by a wide range of assessment tools and techniques [1]. Instructors have many options for assessing the required topics of a course in a particular offering [2]. More traditional material includes homework assignments and examinations;

emerging material includes question repositories and games. This rich variety also introduces new challenges to educational stakeholders (e.g., students, instructors, administrators, education researchers) with respect to evaluating the coverage of assessment material and communicating achievements. Students may find it challenging to infer their strengths and weaknesses with respect to the topics and their relationships, which can impede their studies. Outside a classroom program administrators, course designers, course co-ordinators, and researchers also face challenges. Administrators find it challenging to compare the content and difficulty of formal assessments as well as students' outcomes across different offerings of a course. Course designers and co-ordinators find it challenging to ensure the required topics and their relationships (e.g. questions with a combination of topics *a*, *b* and *c*) have been assessed. Educational researchers need to compare the achievement results between control and experimental groups.

To address these assessment challenges, a new field, known as “Learning Dashboards” has emerged, which embraces learning analytics and educational data mining [3, 4, 5]. These dashboards help users to interactively explore and understand data sets through analysis and visualisation techniques. A variety of traditional plots and charts have been adopted (e.g., pie, box, histogram, radial) to visualise the achievements of students. The research typically considers independent (stand-alone) topics. Additional discussion on the related work is presented in Section 3.

As part of an on-going project, the authors of this paper have explored a collection of topic dependency models (TDMs) for assessment in which the relationships among topics are considered [6, 7]. The TDMs

DOI reference number: 10.18293/DMSVIVA2019-018

use a two-weighted undirected graphs foundation to formally represent and visualise a wide variety of assessment data for one or more topics (i.e., topics and their dependency relationship) to meet the needs of diverse stakeholders. The collection consists of a course reference model (to establish the topics and dependencies covered in a course), in addition to classroom models, both static and dynamic. As the work presented in this article leverages these results, a background section on the original TDM collection is presented in Section 1.

In this work, an alternative graph foundation is explored for the TDM collection: a two weighted hypergraph. This graph can explicitly visualise the n-ary topic coverage and achievement inherent in questions (i.e., a question can assess 1:N topics). However, visualisations of their more general n-ary hyperedge relationships may be more difficult to understand in the broader community. Here, a method to generate a two-weighted hypergraph model using commonly available assessment data and a flexible, approach to address the complexity of the graphs are presented. The visualisation approach is a multilevel filtering approach that considers the hypergraph as a collection of levels. The content of these levels can be customised and presented according to user preferences. For example, instead of being presented with the complete hypergraph, the users can select topics, achievement or number of responses (e.g., maximum, minimum, range of values) to view part of it. In addition, the visualisations of the levels can be presented in either a cumulative or accumulative mode. The new models are presented using an illustrative example. A case study that illustrates how the model can be applied to provide insight for instructors is presented. This study uses historical data from a first-year undergraduate level offering of a course on programming and engineering design at The University of British Columbia.

2 Background: The Original TDM Collection

A collection of TDMs has been presented in the author’s previous works [6, 7] to address some of the challenges stakeholders face that relate to the design, delivery and analysis of assessments. An overview of the collection is presented in Figure 1.

Stakeholders and Scenarios of Use. The wide variety of stakeholders interacting with these models (e.g., students, instructors, and so on) are shown at the top of the figure; they have roles inside and outside of the classroom. The earlier results present a preliminary scenario analysis that identifies 40 questions for the

stakeholders. For example, instructors interact with the models to explore questions such as:

- What topics and their relationships do I need to assess?
- What topics are covered each of the assessments?
- How well are the students performing on the topics?
- How well are the students performing compared to other (current or previous) classes?
- Who in the class may be at risk of failing the course?
- How have all the topics and their relationships been assessed?
- How much has the class improved over time on the topics?

Input Data. The input assessment data, shown on the left of Figure 1, includes results from current or previous classes. The data can be from a variety of formative (e.g., assignments) or summative (e.g., quizzes) assessments; the course data includes a specification of the course topics that need to be assessed.

Topic Dependency Models. Based on the requests from stakeholders, the input data are selected and transformed into visual models. As illustrated in main block of Figure 1, the TDM collection consists of two kinds of models: Course Assessment Reference and Classroom Assessment Visualisation (static and dynamic). The Reference model establishes and communicates the required topics and their relationships for a course (e.g., a CS1 course is an introductory programming course covering variables, branching, loops, and so on). It provides a common foundation for all sections of a course offered over time (e.g., CS1 Section 001 Term 1 2019 Instructor A. Smith). The reference model helps to ensure the consistent coverage of topics by different instructors, clearly communicate the scope of the topics to students, and support administrative activities related to monitoring learning objective outcomes.

The Classroom Assessment Visualisation models present assessment data within a class and support comparisons of assessment data between classes at one point in time (static) and over time (dynamic). The assessment data are selectively visualised, e.g., for the whole class, individual students, specific topics, and so on. For example, the user can choose to visualise static models, in a side-by-side comparison, for a specific assessment (e.g., Assignment 3) for all of the students in two classes. Alternatively, the user can choose to visualise dynamic models to explore the progression of a class over time.

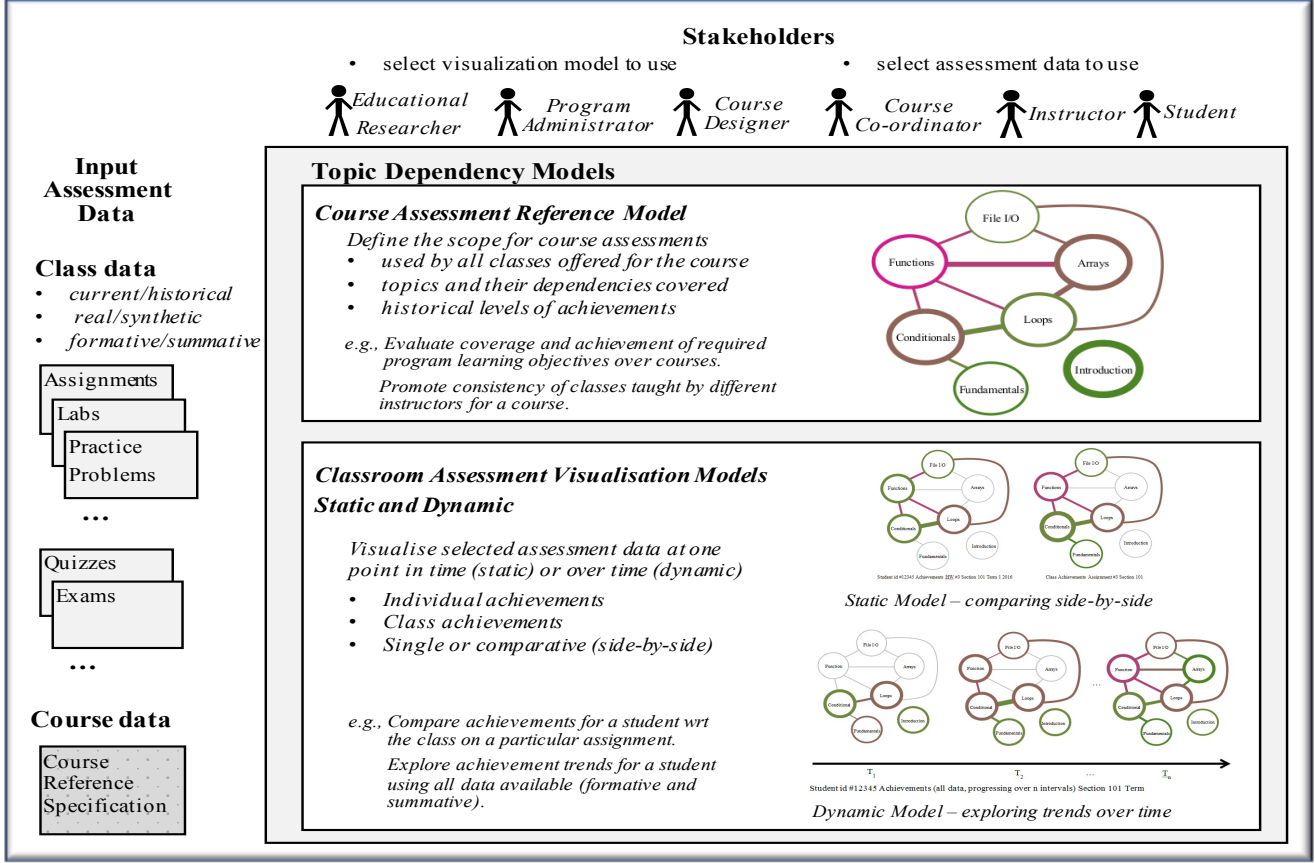


Figure 1: Original TDM model collection: two-weighted undirected graph foundation [6].

Graph Foundation of the TDMs The Reference and Classroom Assessment Visualisation models share a common graph foundation: a two-weighted undirected graph. The vertices represent topics in a course; the edges represent assessment material (e.g., questions) that address the topics the edges are related to. The weights are reflected in the visualisation using a colour palette (achievement) and width (coverage). As the graphs can only represent edges with up to two vertices, questions involving three or more topics must be redistributed in the models as combinations, in order to visualise them.

3 Related Work

The learning analytics community continues to actively investigate approaches that support the exploration of learning activities by different stakeholders. With the increase in the use of educational technologies and the advancements in the areas of learning analytics and educational data mining, a new field, commonly known as “Learning Dashboards” has emerged to help make sense of data sets in learning and ed-

ucation [3, 4, 5]. A variety of visualisations such as bar charts [8], pie charts [9], histograms [10], box plots [11], radar graphs [12], and skill meters [13] have been adopted to show the achievements of students for independent (stand-alone) topics. [3] presents a systematic literature review on the use of learning dashboards. Based on the findings of this literature review, the use of graph-based visualisations in learning dashboards has not received much attention. In addition, a number of studies provide strong evidence that opening the model to learners, leading to the notion of Open Learner Models (OLMs) [14], can be effective in helping students learn [15]. The OLMs commonly use a set of individual topics as their underlying structure for modeling learners’ knowledge state, which ignores relationships among topics. An emerging new field applied methods from process mining and sequential data mining to educational data [16] to facilitate better understanding of the educational process. Educational process mining methods predominantly use graphs, but their main focus is on using activity logs to visualise students’ learning process in terms of the time, place, path, pace of learning activities.

Hypergraphs have been adopted as a foundation for data analytics and visualisations in a wide variety of domains including data warehousing, communication network analysis, geospatial metadata, and cellular biology networks. To support interactive queries in data warehousing systems, a framework for developing dashboards called Dashboard-by-example has been proposed [17]. This framework adopts hypergraph-based techniques to transform dissimilar, heterogeneous data into a homogeneous knowledge space of clusters and partitions. The framework is flexible, as the hypergraph-guided data linkages support the exploration and aggregation of data from multiple perspectives. For the analysis of network traffic traces, available in massive communication logs, a hypergraph based visualisation is proposed in [18]. The network traffic visualization approach utilises a frequent item set mining method to identify interesting traffic patterns in the large amount of data. The patterns are visualised as hypergraphs with explicit, multi-attribute relationships. A framework to comprehensively address all of the available geospatial metadata standards (i.e., documents) is proposed in [19]. Geospatial metadata describe geographic digital data resources such as earth imagery, geospatial databases and catalogues, and Geographic Information System files. The framework integrates hypergraphs and topic maps, representing the elements and their dependency relationships. The potential for adopting hypergraphs in the domain of cellular biology is introduced in [20]; recently, for example, [21] proposes a framework that adopts hypergraphs and associated hypergraphs to describe, analyse, and identify metabolic network alignments at the full genome level. These alignments are used to discover important similarities and differences between distinct molecular networks: they reveal mappings between components (topological, biological functional) across different networks.

To the best of our knowledge, a modeling and visualisation approach based on hypergraphs for educational assessment is not available at this time. The TDM_{MH} helps to address this gap in the literature.

4. Methodology

This section define the graph foundation (hypergraph, level), scenarios of use, and an illustrative example. The graph generation and visualisation algorithm are presented in Section 5, using the example. A case study demonstrating TDM_{MH} s based on data collected from the final exam of a large introductory course on programming is presented in Section 6.

4.1 Hypergraph Foundation of the TDM_{MH}

In this section, the formal definition for a two-weighted hypergraph and the concept of a level are presented.

Definition 1 A two-weighted, undirected hypergraph $G = (V, H)$, where V is the set of vertices representing the topics, and H is the set of hyperedges. A hyperedge $h \in H$ is represented as $h = (C, c_1, c_2)$, where C is the subset of the vertices being connected, and c_1 and c_2 represent the two weighted values for an edge. An edge involving only one node represents a self-loop.

In this work, the nodes are used represent topics; the hyperedges represent the assessment material that covers the topics. The weight c_1 represents the number of learning objects that are tagged with the topics for the hyperedge; and c_2 represents the represents the performance (e.g., achievement) on these learning objects. A simple example of a TDM_{MH} visualisation is illustrated in Figure 2. There are four topics in this example (A, B, C, and D); each topic is represented by a node. Topic B and Topic C both have a 1-ary hyperedge (i.e., a self loop). For Topic B, the edge has a dark pink colour and a wide width, reflecting the poor average performance (33%) on the questions answered and the large number of responses. The hyperedge for Topic C is a medium green colour and has a medium width; this represents the good average performance (75%) and a moderate number of responses to questions. A 2-ary hyperedge exists between Topics A and D. This edge is dark green and thin, in representing the excellent performance (100%) on the related questions and a small number of responses. A 3-ary hyperedge exists between Topics A, B, and D. This edge is a brownish pink colour and has a moderate weight, which represents the moderate average performance (50%) and a moderate number of responses to questions on these three topics.

The concept of a level is used in this work to support the visualisation. For example, in the simple example above, there are three levels. $Level_1$ is a subset of the hyperedges with one vertex (i.e., h_1, h_2), $Level_2$ is a subset of the hyperedges with two vertices (i.e., h_3), and $Level_3$ is a subset of the hyperedges with three vertices (i.e., h_4). More generally, $Level_i$ is the subset of hyperedges involving i vertices.

4.2 Scenarios of Use

As with the original TDM models, the users can select the assessment data of interest. For example, they may need to explore data for formative or summative

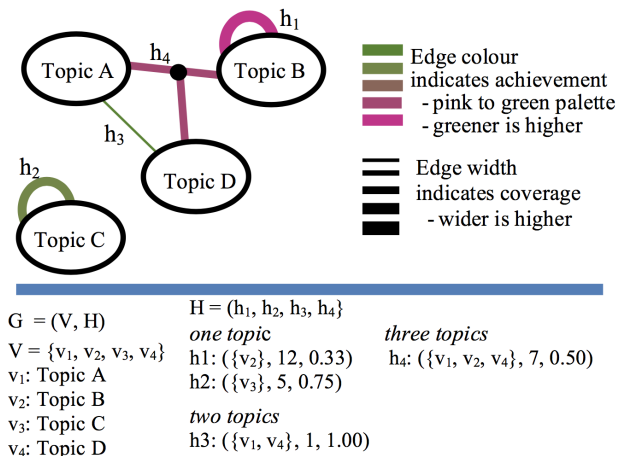


Figure 2: Example of a two-weighted hypergraph: definition and visualisation

assessments (e.g., one or more assignments or examinations); the data can be from one or more classes. Here, scenarios of use include filtering by topic(s), or constraints involving achievement or coverage. Additional filters can be added in a straightforward way. Currently, the scenarios focus on using the static models; the reference and dynamic models for the TDM_{MH} are planned in future work. Two examples, derived from the scenario analysis in [6], are described below.

Topics. How are the students doing on a specific topic? One or more topics of interest can be selected in order to identify topics that may need additional attention. For example, an instructor can select one of many topics covered in a course; the levels are filtered to emphasize the topic(s) of interest.

Achievements. What topics have poor achievement? A constraint for visualising a particular achievement level can be selected. For example, an instructor can select to view the minimum achievement scores over all topics in order to identify those that need additional attention; the levels are filtered to emphasize the achievement scores of interest.

The users can choose to combine filters. In other words, they can filter the visualisations based on one or more constraints such as selecting both a topic and an achievement constraint, e.g. highest performance involving topic T_1 . The results are presented on a level-by-level basis (a level is described in Section 4.1).

When visualising the results, the user can also choose to view the results in either the cumulative or accumulative mode. In the accumulative mode, only the results for one level are presented. In the cumulative mode, the results up to and including the current level of interest are presented.

4.3 Illustrative Example

An example based on six students, three formative assignments (five questions per assignment), and six topics has been defined by the authors. The following constraints are considered for creating the illustrated example:

- The questions must span a range of one to four topics.
- There must be a range in the individual student achievements in the responses (low to high scores).
- There must be a range in the number of responses to the questions.
- There must be a range in the average scores for the questions (0%-100%).
- Assessments on groups of topics demonstrate a range of achievements (very poor to very strong); one or more questions address the groups of topics.
- The multilevel visualisation algorithm with a list of filtering options (e.g., selected topic(s), achievement, cumulative/accumulative mode) can be clearly illustrated.

Table 1 presents a sample data set meeting these constraints. As input to the TDM_{MH} generation algorithm, the data are formatted into two CSV files: (1) A student-question-answer (SQA) file that contains the student identifier, question identifier, and score (correct/incorrect), and (2) a question-topic (QT) file that contains the question identifier and topics (i.e., tags) it addresses.

5 Generating and Visualising TDM_{MH}

In this section, the approach to generating and filtering a TDM_{MH} is presented. Section 5.1 presents methods for generating the graph and Section 5.2 explores filtering mechanisms to address the complexity of hypergraph visualisations.

5.1 Generating TDM_{MH}

This section demonstrates how commonly available input data (student achievements/grades for specific questions and the mapping from the questions to the course topics) are transformed into a TDM_{MH} . High-level code and notation are presented in Algorithm 1. The algorithm consists of three high level steps: create working dictionaries and matrices; define the TDM graph elements (vertices and edges); and visualise (i.e., plot) the TDM graph.

Q	Topic List	Student ID					
		S_1	S_2	S_3	S_4	S_5	S_6
Assignment 1							
Q_1	T_1	-	1	-	-	0	1
Q_2	T_3	1	1	1	1	1	1
Q_3	T_4	0	1	0	1	1	1
Q_4	T_1, T_2	1	1	0	1	1	0
Q_5	T_1, T_4	1	-	-	0	1	0
Assignment 2							
Q_6	T_4	0	0	1	0	0	1
Q_7	T_1, T_4	-	0	1	1	0	0
Q_8	T_4, T_5	0	1	0	0	0	1
Q_9	T_1, T_4, T_5	0	0	0	0	-	1
Q_{10}	T_1, T_2, T_4, T_5	0	0	0	1	0	1
Assignment 3							
Q_{11}	T_1, T_4	1	-	0	1	-	1
Q_{12}	T_2, T_6	0	0	0	0	0	0
Q_{13}	T_1, T_2, T_6	-	-	1	1	-	1
Q_{14}	T_1, T_2, T_4, T_5	-	0	1	0	0	1
Q_{15}	T_2, T_4, T_5, T_6	-	0	0	1	1	1

Table 1: Illustrative example: questions, topic lists, and student responses. The value 1 indicates the question is answered correctly; 0 indicates it is answered incorrectly; and - indicates it is not attempted.

The first six steps of Algorithm 1 transform the data in the SQA.csv and QT.csv files into working dictionaries and matrices. In Steps 1, 2, and 3 of the algorithm, three dictionaries are created, $QDict$, $SDict$, $TDict$, to map array and matrix indices to question, student, and topic identifiers. In Steps 4, 5, and 6, three working matrices are created: T , A , and R . The information on topics assigned to each question is represented in matrix T , in which $t_{ij} = 0$ indicates that question i is not tagged with topic j and $t_{ij} = 1$ indicates that question i is tagged with topic j . The correctness of the answers provided by the users are represented in matrix A , where $a_{ij} = 0$ indicates that user i has answered question i incorrectly, $a_{ij} = 1$ indicates that user u has answered question i correctly, and $a_{ij} = n$ indicates that user u has not attempted question i . Matrix R is used to keep track of attempted questions, where $r_{ij} = 0$ if user i has not attempted question i and one otherwise.

Steps 7 and 8 in Algorithm 1 establish the graph vertices and hyperedges. $VList$ stores the list of the vertices of the TDM_{MH} graph and $HList$ stores the list of its hyperedges. The coverage and competency associated with an edge are both computed using T , A and R within the $ComputeH$ function. The coverage (Cov) associated with a hyperedge among a set of vertices $V = v_j, \dots, v_k$ is computed using the following formula:

Algorithm 1 Generating a TDM_{MH}

Require: $SQA.CSV$, $QT.CSV$, $filters$

Create dictionaries and matrices for efficient indexing

- 1: $QDict \leftarrow CreateQDict(SQA.CSV)$
- 2: $SDict \leftarrow CreateSDict(SQA.CSV)$
- 3: $TDict \leftarrow CreateTDict(QT.CSV)$
- 4: $T \leftarrow CreateT(QT.CSV, QDict, TDict)$
- 5: $A \leftarrow CreateA(SQA.CSV, SDict, QDict)$
- 6: $D \leftarrow CreateD(SQA.CSV, SDict, QDict)$

Compute the Graph Elements: Vertices and Edges

- 7: $VList \leftarrow ComputeV(TDict)$
- 8: $HList \leftarrow ComputeH(T, A, D, TDict)$

Create and visualise the Graph

- 9: $TDMStatic \leftarrow CreateTDM(VList, HList)$
 - 10: $Visualise(TDMStatic, filters)$
-

$$Cov(V) = \sum_{i \in QDict} t_{ij} \times \dots \times t_{ik} \sum_{u \in SDict} d_{ui} \quad (1)$$

The outer summation loops through all of the questions. A question i contributes to $cov(V)$ only if it is tagged with exactly the topics that are included in V (i.e., $t_{ij} \times \dots \times t_{ik} = 1$). For example, a question tagged with topics (T_1, T_2, T_3) attempted by nine learners contributes $\frac{9}{1+0} = 9$ to the coverage of $Cov(\{T_1, T_2, T_3\})$.

The achievement ($Achv$) associated with a hyper-edge between a set of vertices $V = v_j, \dots, v_k$ is computed using the following formula:

$$Achv(V) = \frac{\sum_{i \in QDict} t_{ij} \times \dots \times t_{ik} \sum_{u \in SDict} a_{ui}}{Cov(V)} \quad (2)$$

A question i contributes to $Achv(V)$ only if it is tagged with exactly the topics that are included in V (i.e., $t_{ij} \times \dots \times t_{ik} = 1$). The extent of the contribution is determined by the the number of learners who have correctly answered the question, which is computed by $\sum_{u \in SDict} a_{ui}$. The numerator of this formula computes the number of times questions tagged with exactly topics that are included in V . Dividing this number by $Cov(V)$ produces the probability of correctly answering questions on topics in V . Note that Cov and $Achv$ of self-loops can also be computed via the same two formulas by using $V = v_j$.

Figure 3 represents the data which are stored in dictionaries $QDict$, $SDict$ and $TDict$ and matrices T , A and R .

QDict:

{‘Q₁’:0, ‘Q₂’:1, ‘Q₃’:2, ‘Q₄’:3, ‘Q₅’:4, ‘Q₆’:5, ‘Q₇’:6, ‘Q₈’:7,
‘Q₉’:8, ‘Q₁₀’:9, ‘Q₁₁’:10, ‘Q₁₂’:11, ‘Q₁₃’:12, ‘Q₁₄’:13, ‘Q₁₅’:14}

SDict:

{‘S₁’:0, ‘S₂’:1, ‘S₃’:2, ‘S₄’:3, ‘S₅’:4, ‘S₆’:5}

TDict:

{‘T₁’:0, ‘T₂’:1, ‘T₃’:2, ‘T₄’:3, ‘T₅’:4, ‘T₆’:5}

Topics (6)						Questions (15)														
Questions (15)	Students (6)	Matrix A																		
		n	1	0	1	1	0	n	0	0	0	1	0	n	n	n				
		1	1	1	1	n	0	0	1	0	0	n	0	n	0	0				
		n	1	0	0	n	1	1	0	0	0	0	0	1	1	0				
		n	1	1	1	0	0	1	0	0	1	1	0	1	0	1				
		0	1	1	1	1	0	0	0	n	0	n	0	n	0	1				
		1	1	1	0	0	1	0	1	1	1	1	0	1	1	1				
		Matrix R																		
		Questions (15)																		
		Students (6)	Matrix R																	
			0	1	1	1	1	1	0	1	1	1	1	1	0	0	0			
			1	1	1	1	0	1	1	1	1	1	0	1	0	1	1			
			0	1	1	1	0	1	1	1	1	1	1	1	1	1	1			
			0	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
			1	1	1	1	1	1	1	1	0	1	0	1	0	1	1			
1	1		1	1	1	1	1	1	1	1	1	1	1	1	1					

Figure 3: The data represented in Table 1 are organized into two input files: SQA.csv and QT.csv. The data are loaded, transformed into three dictionaries: *QDict*, *SDict* and *TDict* (Steps 1, 2 and 3 in Algorithm 1). The dictionaries and the input files are used to create three working matrices: T, A, and R (Steps 4, 5 and 6 in Algorithm 1). These matrices are used to create the TDM_{MH} model.

Steps 9 and 10 of Algorithm 1 create and visualise the graph. Figure 4 illustrates the resulting model, without any filtering. It has six vertices as the questions in this data set are tagged with six topics: T_1 , T_2 , T_3 , T_4 , T_5 , and T_6 . The coverage of the hyperedges are computed using Formula 1. For example, the hyperedge with id h_5 has coverage of 13, which consists of contributions of 4 from question 5, 5 from question 7, and 4 from question 11. The achievement of the hyperedges are computed using Formula 2. For example the achievement of the hyperedge with id h_5 is computed as $\frac{7}{13} = 0.54$.

5.2 Visualising TDM_{MH} (static model)

In Step 10 of Algorithm 1, the TDM_{MH} can be presented in either whole or part according to the filters the user selects. To address the complexity of the model, the topology of the graph is used; the illustrative example (with six nodes) provides six lev-

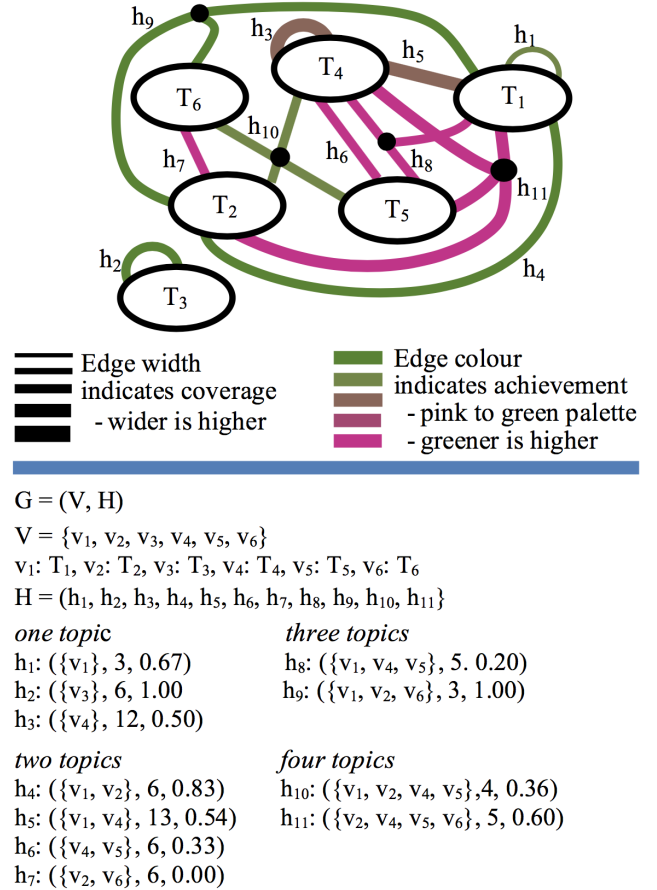


Figure 4: Complete TDM_{MH} for the data set provided in Table 1 using Algorithm 1 (no filtering). The example does not have questions involving five or six topics.

els. Table 2 illustrates the levels before any filtering is done. For example, $Level_1$ is the subset of hyperedges $\{h_1, h_2, h_3\}$.

If the user does not select any filters, then all levels are displayed by default, including all topics, achievements, and coverage of the learning objects as illustrated in Figure 4. Alternatively, the selected filters are applied to each level. The selected elements are presented using the original TDM colour palette; other items are greyed out. $Level'_i \leftarrow filter(Level_i, t, a, n)$, where the parameters t , a , and n represent optional constraints for the topics, achievements, and number of learning objects. For example, if a user requests to filter with respect to topic T_1 , the resulting filtered levels are illustrated in Table 2. In this case, for example, the filtered $Level'_i$ is the subset of hyperedges $\{h_1\}$, as only h_1 contains T_1 .

After the levels are filtered, they are ready to be vi-

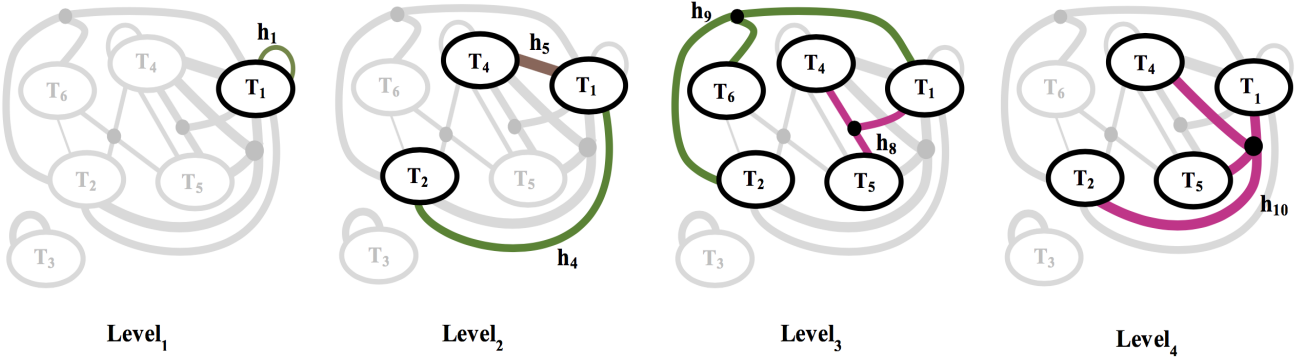


Figure 5: Multilevel visualisation: one topic is selected (T_1) as a filter (accumulative mode).

Level ID	Level Membership	Level' Membership (T_1)
$Level_1$	$\{h_1, h_2, h_3\}$	$\{h_1\}$
$Level_2$	$\{h_4, h_5, h_6, h_7\}$	$\{h_4, h_5\}$
$Level_3$	$\{h_8, h_9\}$	$\{h_8, h_9\}$
$Level_4$	$\{h_{10}, h_{11}\}$	$\{h_{10}\}$
$Level_5$	\emptyset	\emptyset
$Level_6$	\emptyset	\emptyset

Table 2: Levels in the illustrative example graph: before and after filtering (e.g., topic T_1).

sualised. The user can choose to present them in either an accumulative or cumulative mode. In the accumulative mode, only the hyperedges for one level are presented at a time. The first four levels are illustrated in 5 in the accumulative mode. In the cumulative mode, the presentation for a level includes the hyperedges for the current levels and the previous levels. A preliminary layout algorithm is used in the prototype tool at this time; however, adopting a more effective proposal from the literature is planned in the next step.

This straightforward visualisation example based on filtering the levels for one topic can be extended to combinations of one or more topics, achievements, and coverage. The options to highlight part of the graph, in either cumulative or accumulative mode, provide a flexible approach for users. However, to easily access more detailed information (e.g., exploring specific achievement values for hyperedges), additional research is needed to extend the TDM_{MH} visualisations.

6 Case Study

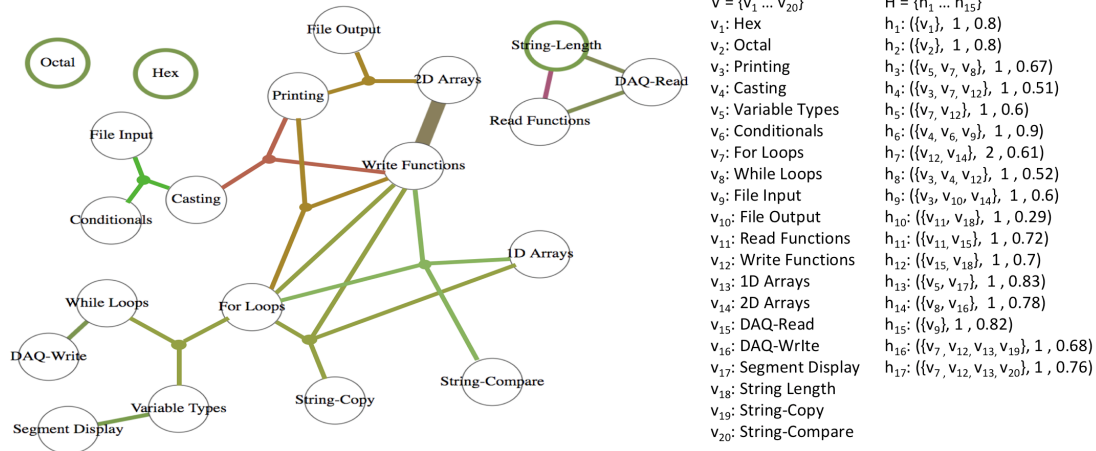
This section presents a case study that demonstrate the applications of multilevel hypergraph TDMs. This case study is based on data collected from a first-year undergraduate level offering of a course on programming and engineering design at The University of

British Columbia. This offering of the course had 377 students and was held during the Fall of 2016. The course covers many topics that are generally included in an introductory course on programming and engineering design in 8 modules: number conversions, programming fundamentals, conditionals, loops, file I/O, functions, arrays, strings, and DAQ systems. Functions, strings and DAQ systems received two weeks of lecture time; all of the other modules received roughly one week of lecture time.

We use data collected from the final exam of the course, which is captured via the Gradescope platform [22], a system for the on-line assessment of handwritten exams, are used. The final exam of this offering had 17 independent sub-questions that formed a total of eight main questions. The questions are tagged using 20 concept-level topics, defined by the instructors, based on the 8 modules that are covered in this course. For example, the Strings Module is further decomposed into three topics: String-Length, String-Copy, and String-Compare, providing a finer level of granularity.

Figure 6 shows the hypergraph TDM for this exam. It indicates that the exam adequately covered all of the modules of the course and that all modules except Conversions (including the Hex and Octal topics) were highly connected to one another. The exam included questions that had only a single topic (e.g., h_1 on Hex), two topics (e.g., h_{14} on While loops and DAQ-write), three topics (e.g., h_9 on Printing, File output and 2D arrays) and four topics (e.g., h_{17} on For loops, 1D arrays, Write functions and String-compare).

The further decomposition to concept-level tags provides insights which might have not been possible to gauge using module-level tags. For example it shows that students did well on questions on String-Copy and String-Compare but not so well on questions on String-Length. As another example, further decomposition



makes it scalable and efficient. The visualisation helps to address the complexity of a hypergraph through a flexible, multilevel approach. A case study that illustrates how TDM_{MH} can be applied to provide insight in the context of a large university course is presented.

A number of directions are planned for the next steps in this research, including exploring the TDM_{MH} dynamic model, refining the prototype tool, and additional validation studies, in particular a user study to compare the strengths and limitations of the original TDM and TDM_{MH} . In addition, highly interactive 3D visualisations of the TDM_{MH} may also be of great value to the diverse stakeholders who are involved with the design and delivery of assessments.

References

- [1] C. Englund, A. D. Olofsson, L. Price, Teaching with technology in higher education: understanding conceptual change and development in practice, *Higher Education Research & Development* 36 (1) (2017) 73–87.
- [2] P. Griffin, E. Care, *Assessment and Teaching of 21st Century Skills*, 2015.
- [3] B. A. Schwendimann, M. J. Rodriguez-Triana, A. Vozniuk, L. P. Prieto, M. S. Boroujeni, A. Holzer, D. Gillet, P. Dillenbourg, Perceiving learning at a glance: A systematic literature review of learning dashboard research, *IEEE Transactions on Learning Technologies* 10 (1) (2016) 30–41.
- [4] V. Argumentation, Software tools for collaborative and educational sense-making, Paul A. Kirschner, Simon J. Buckingham Shum and Chad S. Carr (Eds.) Springer-Verlag: London (2003) 117–135.
- [5] Y. Park, I.-H. Jo, Development of the learning analytics dashboard to support students’ learning performance, *Journal of Universal Computer Science* 21 (1) (2015) 110.
- [6] H. Khosravi, K. Cooper, Topic dependency models: Graph-based visual analytics for communicating assessment data, *Journal of Learning Analytics* 5 (3).
- [7] K. Cooper, H. Khosravi, Graph-based visual topic dependency models: Supporting assessment design and delivery at scale, in: *Proc of the 8th International Conference on Learning Analytics and Knowledge*, 2018, pp. 11–15.
- [8] S. Dawson, L. Macfadyen, E. F. Risko, T. Foulsham, A. Kingstone, Using technology to encourage self-directed learning: The collaborative lecture annotation system, in: *Proc of the conference of the Australian Society for Computers in Tertiary Education*, 2012.
- [9] A.-M. Tervakari, K. Silius, J. Koro, J. Paukkeri, O. Pirttila, Usefulness of information visualizations based on educational data, in: *Global Engineering Education Conference*, 2014, pp. 142–151.
- [10] D. Leony, A. Pardo, L. de la Fuente Valentín, D. S. de Castro, C. D. Kloos, Glass: a learning analytics visualization tool, in: *Proc of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, pp. 162–163.
- [11] D. Albert, A. Nussbaumer, C. Steiner, Towards generic visualisation tools and techniques for adaptive e-learning, in: *Proc of the 18th International Conference in Computers in Education*, 2010, pp. 61–65.
- [12] M. May, S. George, P. Prevot, Travis to enhance online tutoring and learning activities: Real-time visualization of students tracking data, *Interactive Technology and Smart Education* 8 (1) (2011) 52–69.
- [13] S. Bull, B. Ginon, C. Boscolo, M. Johnson, Introduction of learning visualisations and metacognitive support in a persuadable open learner model, in: *Proc of the Sixth International Conference on Learning Analytics & Knowledge*, 2016, pp. 30–39.
- [14] S. Bull, J. Kay, Open learner models, in: *Advances in Intelligent Tutoring Systems*, 2010, pp. 301–322.
- [15] R. Bodily, J. Kay, V. Aleven, I. Jivet, D. Davis, F. Xhakaj, K. Verbert, Open learner models and learning analytics dashboards: a systematic review, in: *Proc of the 8th International Conference on Learning Analytics and Knowledge*, 2018, pp. 41–50.
- [16] A. Bogarín, R. Cerezo, C. Romero, A survey on educational process mining, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8 (1).
- [17] D. Hoang, B. Nguyen, M. Tjoa, Dashboard by-example: A hypergraph-based approach to on-demand data warehousing systems, in: *Proc of the 2012 IEEE International Conference on Systems, Man, and Cybernetics*, 2012, pp. 1853–1858.
- [18] E. Glatz, S. Mavromatidis, B. Ager, X. Dimitropoulos, Visualizing big network traffic data using frequent pattern mining and hypergraphs, *Computing* 96 (1) (2014) 27–38.
- [19] G. Rajaram, H. C. Karnatak, S. Venkatraman, K. R. Manjula, K. Krithivasan, A novel computational knowledge-base framework for visualization and quantification of geospatial metadata in spatial data infrastructures, *GeoInformatica* 22 (2) (2018) 269–305.
- [20] S. Klamt, U.-U. Haus, F. Theis, Hypergraphs and cellular networks, *PLoS computational biology* 5 (2009) e1000385.
- [21] T. Shen, Z. Zhang, Z. Chen, D. Gu, S. Liang, Y. Xu, R. Li, Y. Wei, Z. Liu, Y. Yi, X. Xie, A genome-scale metabolic network alignment method within a hypergraph-based framework using a rotational tensor-vector product, *Scientific Reports* 8.
- [22] A. Singh, S. Karayev, K. Gutowski, P. Abbeel, Grade-scope: a fast, flexible, and fair system for scalable assessment of handwritten work, in: *Proc of the 4th Conference on Learning@ Scale*, 2017, pp. 81–88.