Individualization for Education at Scale: MIIC Design and Preliminary Evaluation

Christopher G. Brinton, *Student Member, IEEE*, Ruediger Rill, Sangtae Ha, *Senior Member, IEEE*, Mung Chiang, *Fellow, IEEE*, Robert Smith, and William Ju

Abstract—We present the design, implementation, and preliminary evaluation of our Adaptive Educational System (AES): the Mobile Integrated and Individualized Course (MIIC). MIIC is a platform for personalized course delivery which integrates lecture videos, text, assessments, and social learning into a mobile native app, and collects clickstream-level behavioral measurements about each student as they interact with the material. These measurements can subsequently be used to update the student's user model, which can in turn be used to determine the content adaptation. Recruiting students from one of our Massive Open Online Courses (MOOCs), we have conducted two preliminary trials with MIIC, in which we found (i) that the majority of students (70 percent) preferred MIIC overall to a one-size-fits-all (OSFA) presentation of the same material, (ii) that the mean level of engagement, when quantified as the number of pages viewed, was statistically higher (by 72 percent) among students using MIIC than among OSFA, and (iii) that the integrated multimedia learning features were generally favorable among the students (e.g., 87 percent found the videos helpful).

Index Terms—Personalized learning, adaptive educational systems, individualization, online learning, MOOC

✦

1 INTRODUCTION AND MOTIVATION

SUCCESSIVE innovations in distance learning have stretched the feasible length of separation between students in a given course. The most recent of these has been MOOCs [44], which have created global connectivity among users for learning. Platforms like Coursera, edX, and Udacity have become the subject of many debates as people explore the future of higher education [35].

There are now over a dozen MOOC platforms. Among them are a number of operational differences, but common across all are support for open content consumption, lecture videos with quizzes, homework assignments, and scalable student discussion forums, as well as the following two salient features: very large enrollments, but very low completion rates.

To illustrate the last two points, take Fig. 1, which shows empirical enrollment-completion data pairs for a variety of MOOCs offered on Coursera, edX, and Udacity [31]. Completion rate here is defined as the fraction of students who received a certificate at the end of the course. As one can see, it is rare to see more than 13 percent of students complete a MOOC.

These high attrition rates have been the focus of a number of recent studies (e.g., [12], [40], [42]). For MOOCs geared towards student enjoyment, some argue that completion is not the right measure of efficacy, since users may choose to survey or only focus on a subset of the material [33]. For MOOCs that are created by instructors to target serious students who are looking to obtain a certificate, some argue that large drop-off rates represent a fundamental debates about the long-term prospects of MOOC [22], [56].

We identify a number of reasons why these MOOC drop-off rates may occur [11]:

- *Asynchronous learning*. There is no common timetable or location, which makes it difficult for students to interact except through forums.
- *Small teacher-to-student ratios*. The number of students is orders of magnitude larger than the teaching staff. As a proxy, our statistical analysis in [12] showed an average ratio of 0.0035 over 73 courses, considering those who posted on the forums at least once.
- *Diverse demographics*. Coming from all over the world and from all different age groups, students have a diverse set of learning backgrounds and goals.

The presence of these challenges, among others, makes difficult for a standard OSFA course to be effective in a massively scaled learning scenario. In a traditional classroom, each student will have slightly different needs, necessitating the instructor to differentiate learning for each student individually [3]. This process is not scalable to an orders of magnitude larger student body further complicated by heterogeneity and asynchrony.

Our MOOC experience. Between 2012 and 2014, we instructed two MOOCs over six offerings. Our undergraduate course *Networks: Friends, Money, and Bytes* (*N:FMB*) was one of the six piloted by Princeton on Coursera in 2012. Through a pre-course survey in *N:FMB*, we found that: about half of the students were 30 years or older; one quarter did not have a college degree; only 30 percent were from the US, with 35 percent from Europe and Canada, and 35

- *C. Brinton and M. Chiang are with the Department of Electrical Engineering, Princeton University, Princeton, NJ, 08540 USA.*
  E-mail: {cbrinton, chiangm}@princeton.edu.
- *R. Rill and W. Ju are with Zoomi Inc., Princeton, NJ, 08540 USA.*
  E-mail: {ruediger.rill, william.ju}@myedwiser.com.
- *S. Ha is with the Department of Computer Science, University of Colorado Boulder, Boulder, CO, 80309 USA. E-mail: sangtae.ha@colorado.edu.*
- *R. Smith is with RAS Statistics LLC, Wilmington, NC, 28409, USA.*
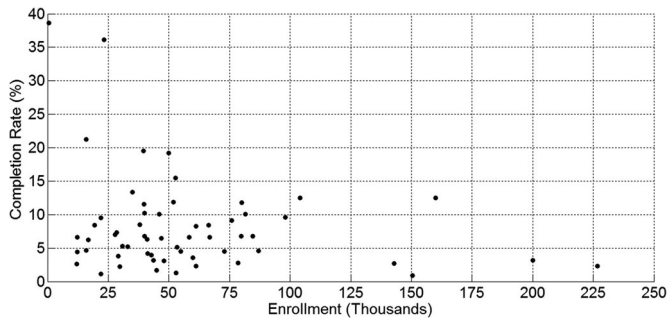 *E-mail: ras.statistics@gmail.com.*

Fig. 1. Empirical enrollment-completion data for 58 MOOCs offering certificates [31] shows that completion rates (i.e., the fraction of students receiving certificates) rarely exceed 13 percent.

percent elsewhere; and one third of the students were from backgrounds other than science, math, or engineering. Furthermore, through interacting with the students in the forums, we found that some were deterred because of their lack of knowledge or interest in mathematics, while others were tackling advanced material, emphasizing the diverse learning backgrounds of the student body.

To cater to those deterred, we created a second course, *Networks Illustrated: Principles Without Calculus* (*NI*), which explains the underlying concepts in *N:FMB* but with much simpler mathematics. In *NI*, many students complained that the material was too elementary, which caused some to lose interest, further emphasizing the diverse learning interests.

Basis of MIIC. This work presents an AES built to help overcome some of these challenges with MOOC. The basis and rationale behind this system is as follows:

First, AES have been shown to improve learning outcomes over that attainable from OSFA course delivery in traditional classrooms [2]. These systems generally define and continually update a user model (UM) based on a student's interaction with the AES [16], which is used to assist student navigation through the material (i.e., navigation adaptation) and/or to modify the presentation of the material itself (i.e., presentation adaptation). UM-based adaptation has the potential to improve the quality of distance education because of its ability to individualize learning to each student's needs and interests, especially in a setting as diverse as MOOC [7], [29].

Moreover, the inclusion of various modes of learning— such as video, audio, text, and graphics—into a course has been shown to be an effective instructional style (e.g., [45], [48]) because it gives students the ability to choose which of the modes they prefer, and it provides increased opportunity for cognitive reinforcement from different perspectives. We therefore believe it is beneficial for an AES to contain multiple learning modes, especially in distance learning where student needs are diverse. In particular, video lectures are important in e-learning because they most closely replicate the instructional style of a traditional lecture, with narration from the teacher, sequences of objects, and visuals presented to the student to help create a more personal learning environment [32], [57]. Another benefit of integrated learning modes is that an AES can use information collected as students interact with each mode to update the UM.

Finally, there has been a large increase in the popularity of tablet computers recently, with global sales increasing by 50 percent in 2013 [28]. Studies have indicated that students may prefer learning on mobile devices than on PCs [20], [37]. Moreover, there have been a number of recent studies which have shown mobile device users to prefer apps to browsers for computing tasks (e.g., [8], [54]); in particular, [8] found this to be true in the context of accessing course resources. AES development in native app format has a number of advantages in terms of device-side storage, document pre-loading, and a wider range of sensors (i.e., camera and accelerometer) to detect user interaction.

Overview of MIIC. Our AES is called the Mobile, Integrated and Individualized Course (MIIC), and possesses the properties outlined above. In particular:

- It integrates video, text, assessment, and social learning into a single platform, and is thereby built for full course delivery.
- It captures behavioral measurements, including clickstream data, about each student as they interact with the course material, including video-watching and pageview events, which can subsequently be used for adaptation.
- It is delivered as a native mobile app, as opposed to through a (mobile) web browser.

Organization. Section 2 will present a discussion and comparison to related AES. Then, Section 3 will describe the individualization framework for MIIC. Here, we will first outline our design process (Section 3.1), and then present the MIIC individualization used for the initial user trials as a special case (Section 3.2). Section 4 will overview the MIIC system architecture. Then, Section 5 will discuss the two preliminary user trials that have been conducted with MIIC using participants from MOOC, followed by next steps we have planned. We conclude the paper in Section 6.

2 RELATED WORK

Development of AES dates back to the early 1990s. Brusilovsky presented a taxonomy and summary in 1996 [16]. We will discuss some of the well-cited AES that have been developed since then, and direct the reader to [2], [15], [16] for more details.

ELM-ART [52] is a web-based AES which supports adaptive navigation through link annotation. The UM in ELM-ART is a multi-layered overlay, and is updated based on both knowledge inference from assessments and explicit user input. MIIC is different in this regard because it also supports presentation adaptation, and because it does not allow users to directly modify the UM.

AHA! [24] is another web-based adaptive system, where each page consists of a sequence of HTML fragments. Similar to MIIC, AHA! supports both navigation (through link annotation and hiding) and presentation (through conditional inclusion of fragments) adaptation. The UM in AHA! is based entirely on a user's browsing behavior, with fragments and pages being marked as desired or not based on pages visited previously. MIIC instead uses assessments to infer user knowledge of and/or tendency towards learning

concepts, with correlations with behavioral measurements to potentially enhance these inferences.

TANGOW [18] also features navigation (through link disabling and adaptive link sorting) and presentation adaptation, but differently than AHA!, the HTML pages are generated dynamically at runtime from content fragments. As a result, the author must specify the sequencing of subtasks as well as the features of each fragment. One drawback to this approach (i.e., having no path generation, see Section 3) may be that the author must label each separate fragment [36], rather than starting with static content blocks and tagging the modifications. TANGOW allows storage of quiz scores and visited pages, leaving it to the author to decide if/how these will be used for adaptation.

CoMoLE [37] is a Java-based AES that was built to support mobile delivery through a web browser, as opposed to MIIC which supports delivery through native app. It supports adaptive navigation by generating a list of recommended next activities, using (1) a rule-based filter which checks the context, features, and requirements of the activity against the UM, and (2) a Markovian filter which analyzes learning paths followed by similar users/groups. MIIC is not currently focused on UM updates based on similar users.

Learning styles. Many AES have been designed to support adaptation based on a user's inferred learning styles (LS) [15]:

WHURLE [13] is an XML-based adaptive learning environment on which different user models can be instantiated. It supports adaptive presentation, by removing chunks of lessons that are not valid for the current user, but not adaptive navigation. Omission of a particular UM makes WHURLE a flexible system, but may add burden on the designer who must specify it [36]. The authors have evaluated WHURLE using two different dimensions of the Felder-Soloman Inventory of Learning Styles [27]: WHURLE-HM [13], with a UM based on the visual–verbal dimension, and DEUS [14], based instead on the sequential–global dimension, and surprisingly found no significant effect in favor of LS adaptation.

LS-Plan [36] is a web-based AES with a UM based on four of the Felder and Silverman Learning and Teaching Style Dimensions [26]. This system supports adaptive navigation, and adapts by sequencing/re-sequencing the current learning path as opposed to MIIC which plans it one step at a time (see Section 3). The UM in LS-Plan is based heavily on assessment performance, but also uses lower and upper-bounds on the total time spent in a module to infer whether a user was on-task or not. Through experimentation, the authors found a statistically significant increase in the knowledge acquired from the adaptive modality.

Novelty of MIIC. AES that support mobile delivery via web browser have been developed [37], but none to our knowledge do so via native app. Also, MIIC presents multiple learning modes to users simultaneously. We are not aware of an AES with lecture videos, likely because most have been focused on acting as supplements to traditional classrooms [49]. Finally, MIIC collects more detailed behavior about user interaction than we have seen for other AES, including clickstream events of their pageviews (verified with device sensors) and video-watching behavior, because these can be used for individualization too.
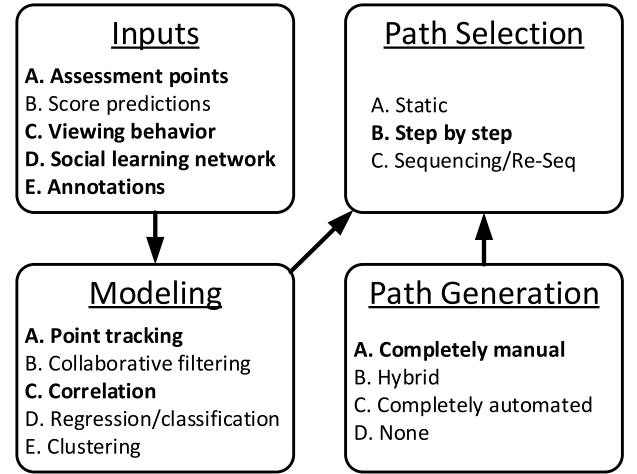

Fig. 2. Our AES design process. MIIC implements the subset of the options shown in bold.

3 MIIC INDIVIDUALIZATION FRAMEWORK

In this section, we will first present the general process we have been following in designing our AES. In doing so, we will discuss the options we have considered for each of its four modules. Subsequently, Section 3.2 will detail the individualization framework implemented for user trials as a special case.

3.1 AES Design Process

Our AES design process consists of specifying four modules: inputs, user modeling, path generation, and path selection, as illustrated in Fig. 2.

3.1.1 Inputs

This refers to the types of inputs that the AES collects. We identify four explicit types: assessment points, viewing behavior, social learning network (SLN) [11], and annotations. Additionally, pre-processing can be performed to give a richer and/or more useful set of inputs for the modeling stage. In particular, performance prediction [34] can be used to estimate a user's score on assessments she did not take.

3.1.2 User Modeling

This module consists of machine learning techniques that map the inputs to update a low-dimensional UM, which contains information about a student's current state of learning [16]. We refer to the dimensions of the UM as the *learning features* of the course, which guide the content adaptation based on user knowledge and/or similarity to them. The feature set $\mathcal{F}$ is typically author-specified; they can represent any of user "goals, knowledge, background, hyperspace experience, and preferences" [16]. We briefly discuss three possibilities:

Learning styles. The author could designate the features to be different LS preferences. These could be, for example, a subset of Felder and Silverman's Learning and Teaching Style Dimensions: sensing-intuitive, visual-verbal, sequential-global, and active-reflective [26]. There are a number of other theories as well, such as those proposed by Dunn and Dunn [25] and Honey and Mumford [30].

Acquired knowledge. The author could also interpret features as dimensions of existing knowledge, covering key
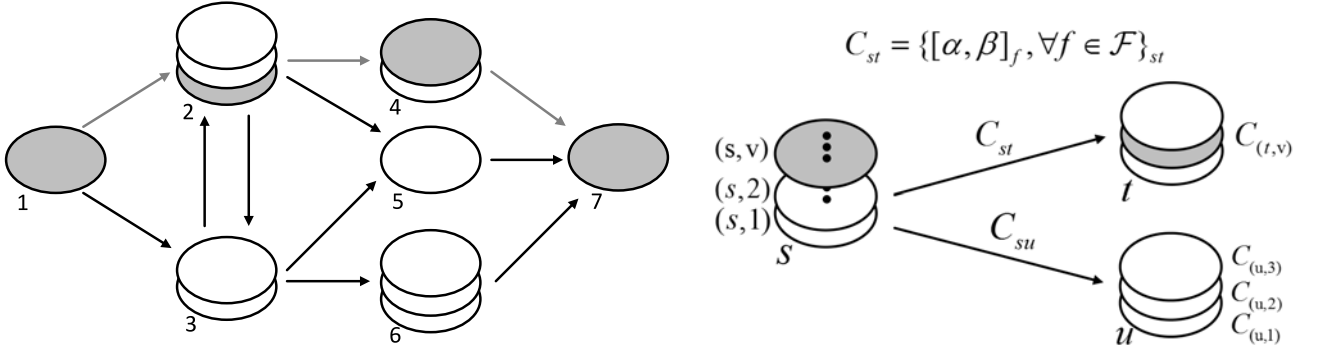
(a) Example lecture consisting of 7 segments, each with a different number of versions. Directional links denote potential transitions between segments.

(b) Illustration of transition terminology. Link constraints are specified by feature to limit the feasible region. Similar specifications are made for each version.

Fig. 3. Diagrams to illustrate the definitions of learning path, seg-vers, and navigational and content adaptation.

areas of the course. These would serve to track the knowledge acquired by the user while interacting with the course material, and could be very general in nature (e.g., "mathematical", "conceptual"), or more specific, even to the point of simply having one feature for each segment.

Domain background. Additionally, features could measure user background in the content domain, to indicate whether or not she satisfies prerequisites for certain sections of material.

One way to update the UM is through a score tracking system, where each answer choice in an assessment is associated with a number of points (possibly binary) for one or more features. This approach is taken in numerous developed systems because tests are the "most reliable source of evidence that a user has learned a concept" [52].

Beyond this, there are a number of algorithms one could use to map the inputs to the UM. For example, matrix factorization (MF), a type of model-based collaborative filtering [51], is a technique that has been applied to educational data to extract latent feature sets [5], [34]. In its simplest form, MF models each user i and quiz j in terms of a feature vector of dimension K, say $\mathbf{u}_i, \mathbf{q}_j \in \Re^K$, and seeks to minimize the prediction error $\mathbf{u}_i^T \mathbf{q}_j - s_{ij}$ of the actual score s_{ij} by optimizing the feature vectors across user-quiz pairs in a training set. Letting the matrix $\mathbf{Q} = [\mathbf{q}_j]$, Ref. [34] also gives a method for decomposing $\mathbf{Q}$ into a product of human-generated tags and concept-tag relations to enhance interpretability of the latent space.

One way to incorporate inputs besides assessments is through a large regression/classification problem that will compute correlations among them. An example is factorization machines (FM) [43], where each user-quiz pair is represented as a vector, say $\mathbf{x}^k \in \Re^D$ for pair k. The set of dimensions $\mathcal{D}$ contains all the possible attributes of the pair, which can take binary values, or real values, such as the percentage of the video the user completed. FM has been applied to educational data previously [50].

3.1.3 Path Generation

The purpose of this module is to specify each of the *learning paths* a user may follow as a result of the adaptation logic. This logic will compare the UM to the properties of each path and select the one that best suits the user. We say that

each learning path consists of a sequence of *segments*; one can think of a segment (seg) as the smallest unit of knowledge presented before/after an assessment. A segment may also have a number of different *versions* (vers), corresponding to alternate presentations of the content. As such, we let (s, v) refer to ver v of seg s, but we will only use the ordered pair when it is necessary to distinguish between versions. Then, $S_u = ((s, v)_1, \ldots, (s, v)_n)_u$ denotes user u's learning path, which is the sequence of segment-versions (seg-vers) that she has visited.

For illustration, we can view a course as an author-defined network, where the nodes are content segments (with different versions) and the links are potential transitions between them. In Fig. 3a, we show an example with seven segments, where a link from s to t means that it is possible to transition to t once having finished with s. Shown is an example learning path $S_u = (1, (2, 1), (4, 2), 7)$. It is important to note the difference here between how navigation and presentation adaptation [16] are handled in our framework, which occur at the link level (e.g., direct guidance or annotation) and content level (e.g., collapsing/expanding or text emphasis), respectively. Navigation between segments encapsulates the former, while the choice of different versions refers to the latter.

Hence, it is necessary to (i) segment the content, (ii) generate the set $\mathcal{S}$ of learning paths, and (iii) specify the properties of the paths in terms of the learning features $\mathcal{F}$. For the trials in Section 5, we each perform of these manually, as will be explained. Other methods could automate a portion of this process for a given course. For example, if an author has completed (i), one could then recruit a set of users to interact with the content, monitoring their satisfaction and progress as they make their own adaptation decisions. Based on the paths chosen by the users who learned well, these actions could be hard-coded as paths for future users with similar UMs, thereby specifying (ii) and (iii). An alternative to these methods altogether is to have no set paths at all, by having the sections generate dynamically from content fragments based on the current UM, as in TANGOW [18].

3.1.4 Path Selection

The last module is the method to select the learning path for each user based on the UM. In a static regime, the path is fixed based on information acquired at the beginning

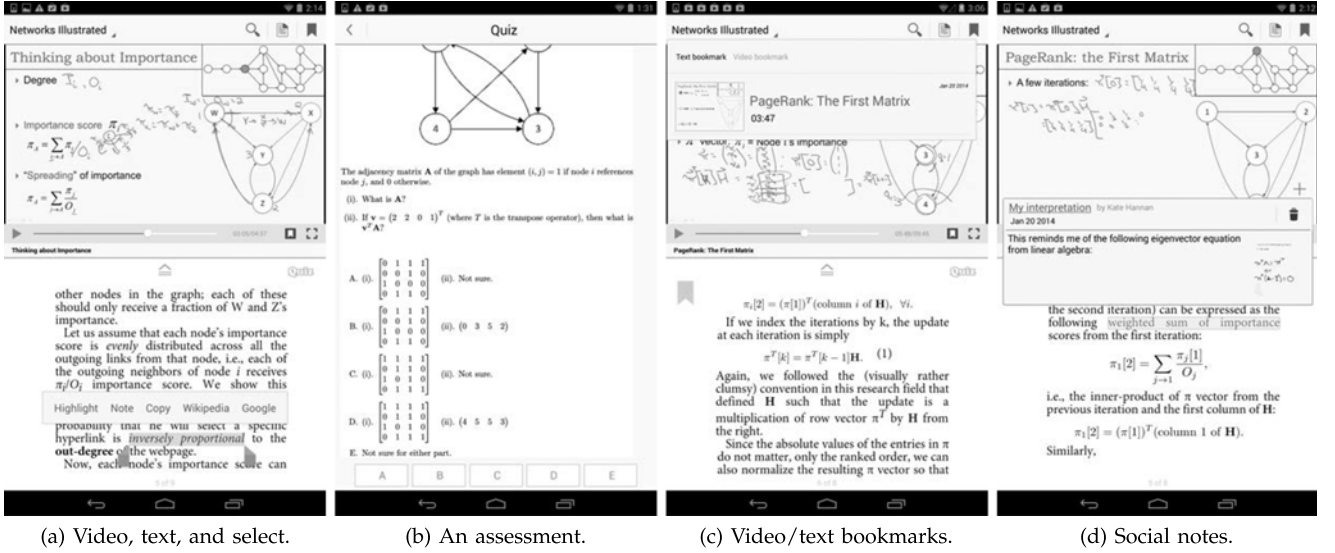| (a) Video, text, and select. | (b) An assessment. | (c) Video/text bookmarks. | (d) Social notes. |

Fig. 4. Illustration of how the different learning modes are currently integrated into MIIC. In (a), we show video and corresponding text, along with the ability to select text and add notes, perform external searches, and so on; in (b), an assessment is presented to the user; in (c), the video and text bookmark menu is selected, and a bookmark icon is on the page; and in (d), a note created by another user is selected on this user's tablet.

[13]. MIIC currently uses a step-by-step approach where the next seg-ver is determined at the end of the current one, so only the learning path up to the current point is known. Another alternative is sequencing/re-sequencing, as with LS-Plan [36], where at any given point a user is assigned to an end-to-end path, which will switch if another is found more suitable to the current UM.

3.2 Individualization for User Trials

The current MIIC individualization framework consists the subset of the AES design options bolded in Fig. 2. We have implemented three components that were tested through the user trials in Section 5: behavioral measurements, data analytics, and content/presentation adaptation. An extended algorithm for the data analytics component, which is embedded in our system but not yet tested through a trial at the time of writing of this paper, will be briefly discussed in Section 5.4.[1]

3.2.1 Behavioral Measurements

As users interact with MIIC, their behavior is monitored and subsequently uploaded to a server:

Viewing behavior. Viewing measurements are taken for video and for pages as a whole. The UI for the different learning modes is shown in Fig. 4a. The current position of the video, and the tags of the objects in the current page, are recorded with each touchscreen interaction. The interaction recorder that obtains these two types of viewing measurements will be explained in Section 4.

Quiz responses. The questions in MIIC currently take the form of radio-response multiple choice. Fig. 4b shows the standard assessment view. Each time a user answers a question, her response is recorded.

Notes and markings. MIIC allows the users to take and share notes, as well as place bookmarks on reading pages and in videos. Shown in Fig. 4c is the user menu for video and text bookmarks, and in Fig. 4d is the note sharing

aspect: the user can select a note made by another user and expand it to see it in full.

3.2.2 Data Analytics

For these trials, we took a simplistic approach to updating the UM and restricted ourselves to analyzing quiz responses, as is done by most other AES.

Learning features. Each segment s in a course is associated with a set of learning features $\mathcal{F}_s$, which is a subset of the features in $\mathcal{F}$, as discussed in Section 3.1. The purpose of the assessments within s is to test user proficiency with one or more of the features $f \in \mathcal{F}_s$. The content author will tag each segment with its corresponding features.

Feature weights. Let $q \in Q_s$ denote question q in the set of questions Q_s for segment s. We refer to w_{qf} as the weight of feature $f \in \mathcal{F}_s$ in question q. In general, w_{qf} can be any real number, and if feature f is not present in q then $w_{qf} = 0$.

Assessment grade. Let $c \in C_q$ be answer choice c within the set of choices for q, and let π_c be the (real-valued) points associated with choice c in question q. Upon completion of segment s, the points awarded to a student for feature f is mathematically given by:

$$\Pi_{sf} = \sum_{q \in Q_s} w_{qf} \left(\sum_{c \in C_q} \pi_c \times i_c \right), \qquad (1)$$

where i_c is 1 if choice c was selected and 0 otherwise.

UM update. Updating the user model here consists of aggregating (1) and storing the results for each feature. Letting p_f^a denote the current assessment performance of the user, then at the end of segment s, the UM is updated as follows:

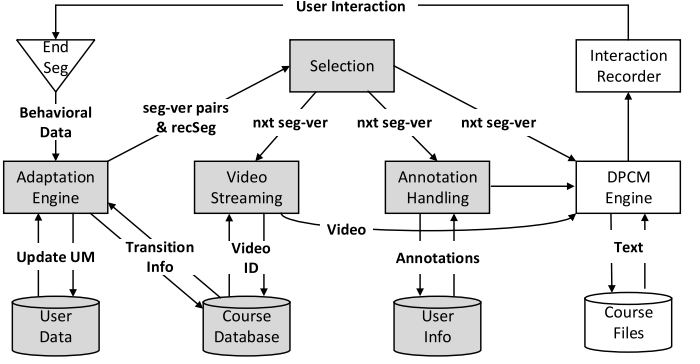$$p_f^a \leftarrow p_f^a + \Pi_{sf} \quad \forall f. \qquad (2)$$

3.2.3 Content/Presentation Adaptation

In the rest of this section we will explain how MIIC supports both navigation and presentation adaptation, referring to Fig. 3b for terminology.

---

1. Details of this extended algorithm are available in Section 6 of our online technical report [9].
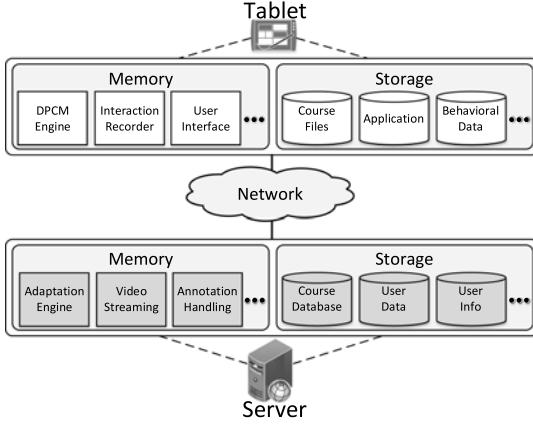
(a) Main components on the device and server.

(b) Flowchart depicting the interaction between the components.

Fig. 5. Architectural description of the current MIIC implementation. In (a), the key components coded and established for the user device and back-end server are shown. In (b), high-level device-server interaction logic is shown, which is used to store user data, update the UM, fetch the next segment, and render it on the device.

Adaptive navigation. Once the user has finished working in a segment, MIIC generates a recommendation as to which she should visit next. In the current MIIC implementation, these recommendations are not shown to the user; rather, they are used by the system to determine the next seg-ver to fetch. More generally, MIIC can implement a form *adaptive ordering*, where the potential next segments are ordered based on the current UM and shown to the user.

Each link specified by the author will have constraints on the current assessment grades (i.e., the p_f^a). Letting $\mathcal{F}_{st}$ denote the set of features used to constrain the transition between s and t, for each feature $f \in \mathcal{F}_{st}$ the author will specify a lower ($\alpha_{f,st}$) and upper ($\beta_{f,st}$) bound requiring $p_f^a \in [\alpha, \beta]_{f,st}$[2] for feasibility of the the transition to t. In Fig. 3b, these constraints are combined into a set C_{st}.

Considering all potential transitions from s, we obtain the set of recommended next segments as

$$\mathcal{R}_s = \left\{ t : p_f^a \in [\alpha, \beta]_{f,st} \ \ \forall f \in \mathcal{F}_{st} \right\}^3. \tag{3}$$

To determine the recommended next segment ρ_s, we consider three cases on $\mathcal{R}_s$:

- $\mathcal{R}_s = \emptyset$: This means that no transition is valid for the current UM. To avoid this problem, for each s the author will designate one segment d_s to be the default transition from s. In this case, $\rho_s = d_s$.
- $|\mathcal{R}_s| = 1$: Here, there is exactly one valid next segment t, and $\rho_s = t$ accordingly.
- $|\mathcal{R}_s| > 1$: The author should avoid this by choosing mutually exclusive constraints. If it arises, then the first valid seg u is chosen and $\rho_s = u$.

In general, ensuring that $|\mathcal{R}_s| = 1$ might be difficult for an author, depending on the complexity of the employed adaptation structure. Constraint based validation techniques for AES have been developed in the past [38], and we are currently investigating this for MIIC. For the initial trials in Section 5, we had no issues ensuring valid constraints manually.

Adaptive presentation. For each potential next segment, the most suitable version must be selected. The logic for this is similar to (3): if $\mathcal{F}_{(t,v)}$ is the set of features used to define constraints for version v of segment t, then the constraint for t to be feasible is that $p_f^a \in [\alpha, \beta]_{f,(t,v)} \ \ \forall f \in \mathcal{F}_{(t,v)}$. For a given segment, each version can have different properties in terms of content presentation, through the application of the following MIIC functions:

Replacing. Based on the UM, specific pieces of content can be replaced with others. For instance, one version may contain more images and less text than another.

Collapsing/expanding. Content can also be collapsed or expanded. For struggling students this can be useful to elaborate on explanation details/revision and hide advanced material. For advanced students, elaborate explanations can be hidden.

Emphasizing. Content pertaining to learning features that a user possesses strengths/weaknesses in can be emphasized. For text, this includes modifying the font/color or highlighting. This helps a student to focus on these areas for reinforcement or improvement.

4 MIIC SYSTEM IMPLEMENTATION

4.1 Device-Side Implementation

Users may access MIIC through a tablet computer. Our implementation of the first individualized learning environment on such mobile devices needed to be carefully designed for both efficient real-time adaptive rendering and scalable behavioral data collection. The main components residing on the device-side are illustrated at the top of Fig. 5a.

DPCM engine. Individualization must handle the process of dynamic content modification. Existing HTML rendering engines in mobile apps often rely on JavaScript, which, being an interpreted programming language, is too slow and inefficient for dynamic modification at this scale [23]. Issues with JavaScript become more severe in the case of platforms (such as iOS) that, for security reasons, disallow the use of Just-In-Time (JIT) compilation in third party applications. This increases the execution time of JavaScript.

The Dynamic Presentation and Content Modification (DPCM) engine in MIIC is instead a modified and optimized

---

2. This implies that p_f^a must be between $\alpha_{f,st}$ and $\beta_{f,st}$.

3. This implies that t is valid if $p_f^a \in [\alpha, \beta]_{f,st}$ for each feature f constraining the transition from s to t.

version of WebCore/WebKit. We have extended WebCore with a C++ API to allow for native access to the Document Object Model (DOM) as well as to the layout engine for the different types of content modifications that MIIC performs (see Section 3.2.3). Additionally, we improved the SVG rendering library, which is used to display math equations. Due to the size of WebCore (roughly 10M lines) as well as dependencies both within the library itself and with other frameworks, we spent a few months with this.

Interaction recorder (IR). The IR monitors user interaction with the video player and with the content on each page as a whole. For the video player, the time interval between every two successive VCR actions—play, pause, jump, end of video, or close app—is measured. The UNIX Epoch time, starting position, and interval duration are recorded in each case.

As for the page content, the time the user has spent viewing a page is recorded each time she switches the page or closes the app. We implemented a method to help check whether the user is viewing a page a given point in time. These take the form of four Boolean variables based on device sensors:

- *Last touch* ($\mathtt{TS}_{P_T}$). If a touchscreen interaction has occurred within the past P_T minutes, this is true.
- *Face detection* ($\mathtt{FD}$). If the person's face is detected in front of the device through the camera, this is true. The device pulls key frames from a continuous video stream to determine this.
- *Device angled* ($\mathtt{DA}$). If the accelerometer detects that the device is held on an angle, this is true. This determines whether the user has the tablet flat on a surface by checking if the acceleration in any of the three dimensions of the standard Cartesian coordinate system differs from Earth's gravitational acceleration.
- *Device movement* ($\mathtt{DM}_{P_D}$). If the accelerometer has detected device movement in the past P_D minutes, this is true. This checks if the user is holding the tablet.

Based on these, we define another Boolean variable *viewing page* ($\mathtt{VP}$) that is updated every 5 seconds. The following are the cases in which $\mathtt{VP}$ is true.

- $\mathtt{TS}_5 \wedge (\mathtt{FD} \vee \mathtt{DM}_1 \vee \mathtt{DA})$. If the user has touched the screen in the last 5 minutes, this is a good indication that she is focusing on the page. In addition, we require one other variable to be true for more continuous evidence; for instance, if the user walked away from the tablet, this condition would become false in quicker than 5 minutes. We choose to not lower P_T in case the user is reading without touching the screen.
- $\mathtt{FD} \wedge (\mathtt{DM}_1 \vee \mathtt{DA})$. Even if the the time since the last touch has exceeded 5 minutes, the user may still be viewing the page. What we require then is that they are in front of the tablet and that either of the accelerometer variables are true; otherwise, it is likely they are sitting with the tablet but engaging in off-task behavior.

Once the user switches the page, the UNIX Epoch time and counter duration are recorded as a pair. The counter duration measures time spent with any learning mode on the page, since $\mathtt{VP}$ will be true in all cases. The set of text objects (i.e., each paragraph, image, equation, and heading)

in the portion of the viewport that is currently visible is also recorded, to determine whether the page size was changed.

This IR logic was verified empirically prior to user trials. It is important to include it for data analytics, in order to reduce uncertainty associated with whether a user is currently on-task or not, as will be seen in Section 5 when quantifying engagement in terms of page views. Distinguishing between student intents (i.e., their actual behavior) and their actions (i.e., their apparent behavior) is currently an active area of research for intelligent tutoring systems [19].

Course files. The text and image content of the course, as well as questions and answer choices, are stored on the tablet in an EPUB container that conforms to the most recent specification (3.0 at the time of this writing). Each segment has its own universally unique identifier (UUID) and is written as a separate XHTML file, and every containing object is assigned a unique identifier as well. Different versions are created dynamically through tag logic to collapse/expand, replace, or highlight certain objects.

4.2   Server-Side Implementation

A server running Apache is currently used for the backend. The main components are shown at the bottom of Fig. 5a. To communicate with the server, devices require an Internet connection, and submit data using a REST API that sends HTTP POSTs with JSON objects as the body. Server side code was written in Python with the Django framework and JavaScript.

Adaptation engine. This engine has three functions: update the UM, determine the recommended next segment, and determine the potential next seg-vers. This corresponds to data analytics and content/presentation adaptation described in Section 3.

Video streaming. This implements HTTP streaming to the the native video player on the device.

The three main elements shown in server storage are implemented as tables in an SQLite database (DB).

The logic that is executed once the user has completed the current segment is outlined in Fig. 5b. First, the behavioral data collected with the IR is uploaded to the user data DB on the server, and any annotations made are uploaded to the user information DB. Then, the adaptation engine is fed with this, the UM, and the segment transition logic from the course DB. It returns an updated UM, $\mathcal{R}_s$, and the possible seg-ver pairs.

Once the selection is made (currently done automatically, but more generally could be driven by user input), the next seg-ver is fed to the video streamer, which will fetch the necessary video ID information from the course database and begin streaming to the device. Additionally, the annotation handling will check the user information DB for any markings the user has made in the segment previously, and will look at the social network identifiers of her "friends" (via Facebook) to check for shared notes. Finally, the DPCM engine will render the content on the screen.

5   PRELIMINARY USER TRIALS

Using a prototype of MIIC as an iOS mobile app, we conducted two initial trials in 2013. Our objective was to evaluate MIIC among students in our MOOC.
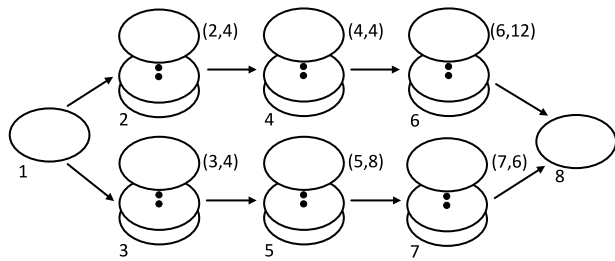
Fig. 6. Course structure for the MIIC employed in our student trials.

5.1 Authoring Process

We used material from our courses, *N:FMB* and *NI*, to convert two lectures to MIIC, one per trial. The videos and assessments were taken from the respective courses, and the text from our books [10], [21].

In architecting the features and transition logic, we set each MIIC lecture to present the most challenging content possible for each student, constrained by both her background knowledge in the prerequisite mathematics and her acquired knowledge at a given instant. An alternative would have been to set up these MIICs as intelligent tutors to bring everyone to the same level of understanding, by adapting the navigation through/around prerequisites. But without offering an incentive for participation, we decided to adapt to what students would want to learn. We did, however, structure each MIIC such that the key concepts were explained along any of the learning paths.

Course structure. Fig. 6 shows the structure we employed for MIIC in both of the trials. Beginning with two separate books and courses, we decided to split navigation into two paths: segments at the top (2, 4, and 6) tended to contain content from *N:FMB* and those on the bottom (3, 5, and 7) from *NI*, while the first segment was a combination and the last a summary. The number of versions shown for each segment here are specific to the second trial, though similar to the ones in the first.

User modeling. For both trials, multiple choice questions were presented at the end of the segments. The 3-5 questions occurring at the end of each of segs 1-5 determined how the UM was updated. Each was tagged with up to three learning features: concepts (C), mathematics (M), and examples (E); hence $\mathcal{F} = \{C, M, E\}$. From (1) in Section 3, we specified the w_{qf} as binary numbers, and the π_c as integers between 0 and 3. For brevity, we omit the exact values for these for each trial, but note that seg 1 covered concepts C and M, segs 2 and 3 covered C, M, and E, and segs 4 and 5 covered M and E in both cases.

Hence, the navigation decision in each trial upon completing seg 1 was dependent on p_C^a and p_M^a. In subsequent segments, each version corresponded to content being emphasized (with color) or collapsed/expanded. Text tagged as corresponding to mathematics or concepts were colored depending on the current feature performance: green for high, and red for low. Additionally, in segs 4 and 5, intermediate steps in numerical examples were hidden depending on p_M^a and p_E^a. And in segs 6 and 7, subsections corresponding to advanced material were expanded depending on the performance on all features.

The results of this modeling process, in terms of which learning paths were traversed, will be given for the second
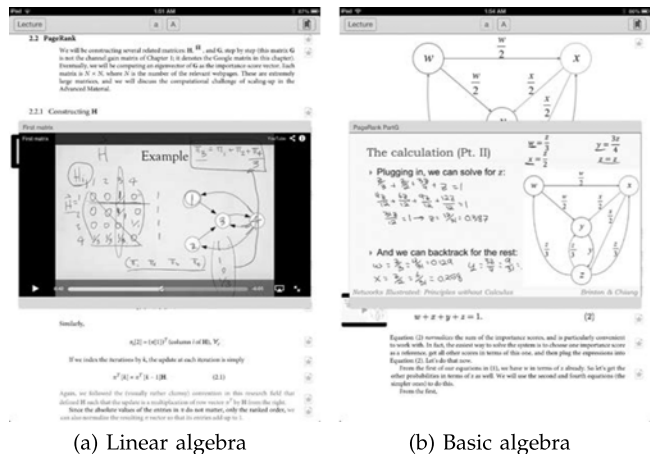

(a) Linear algebra   (b) Basic algebra

Fig. 7. Illustration of a page in two different segs for MIIC in the first trial. (a) is a from seg 2, which shows a more in-depth treatment of the subject. (b) is from seg 3, which only uses basic algebra to explain the fundamentals.

trial. The one-size-fits-all (OSFA) content in each trial consisted of content on the top navigation path in Fig. 6, with no version modifications.

5.2 Trial 1: Student Response

The first trial of MIIC was conducted in February 2013.

Research questions. The purpose of this study was to investigate two questions: (RQ1) Which features of MIIC are favorable among students, and which may need improvement? (RQ2) How does student experience compare between MIIC and OSFA?

Content. The content used here was a lecture on Google PageRank. For MIIC, Fig. 7 shows an example of the difference in content shown on two different learning paths; referring to Figs. 6 and 7a is from seg 2 and contains more advanced linear algebra, while Fig. 7b is from seg 3 and explains the same features but only using basic algebra. OSFA in this trial was chosen to be a standard PDF version of the material, the implications of which will be discussed further below.

Procedure. We announced the trial for iPad users concurrent with the release of the lecture on Coursera. Since this was the first time the software and its backend were used by students, we wanted to ensure the initial infrastructure could readily support the scale of the trial, so we restricted participation to the first 100 students who responded to our first come first serve email. These users received a download link to both the MIIC (.ipa) and OSFA (.pdf) files. In order to reduce bias in the sequence of presentation, we divided them into two groups: one was instructed to use MIIC and then OSFA, and the other was to do the opposite.

Questionnaire. Upon completion of these tasks, each participant was asked to fill out a 14-question multiple choice questionnaire. Five questions asked about the perceived usefulness of the learning modes and overall experience with MIIC, for RQ1, and another four asked about the MIIC versus OSFA comparison, for RQ2.

Forty-seven students filled out the questionnaire, and the 43 who indicated that they used both MIIC and OSFA are the focus of our analysis. This is much smaller than the MOOC enrollments cited in Section 1, because we limited
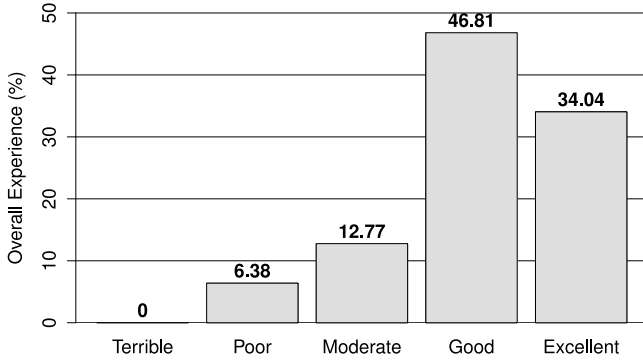
Fig. 8. Student rating of overall experience with MIIC, on a five point scale. Over 80 percent responded "good" or "excellent."

participation by design. These sample sizes are on the same order as the size of traditional classrooms on which many AES have been tested [2]. Also, since OSFA was a PDF document in this trial, strictly speaking, the comparisons made here for RQ2 are between delivery with mobile, integration, and individualization versus delivery lacking these features. For this reason, we attempted to target most of the questions towards a single aspect of our design.[4]

5.2.1   Results: MIIC Features (RQ1)

Lecture videos. One question asked about the usefulness of the integrated lecture videos. sixty-eight percent of students found this very useful, 19 percent found it somewhat useful, and 13 percent found it not useful.

External search. Another asked about the usefulness of selecting text and searching it on external platforms. thirty-six and 38 percent found this very and somewhat useful, while the other 25 percent found it not useful.

Social notes. Another asked about the usefulness of being able to take and share notes. Only 23 and 28 percent found this somewhat and very useful, respectively, while the remaining 49 percent found it not useful. One possible reason for this is the limited time the participants had to interact in the trial.

Text emphasis. Another question asked how well the text emphasis helped to direct users to important concepts. fifty-three percent found this very helpful, 32 percent found it somewhat helpful, and only 15 percent found it not helpful.

Overall experience. Finally, one question asked how the student would rate the overall experience with MIIC, on a five-level Likert scale [36]. The distribution is shown in Fig. 8: Thrity-eight (81 percent) responded excellent or good, and 9 (19 percent) responded moderate or poor.

5.2.2   Results: Comparing MIIC with OSFA (RQ2)

For each of these questions, participants were able to select (a) preference of MIIC, (b) preference of OSFA, or (c) indifference. A trinomial test described in [6] was used to determine whether there was a statistically significant difference for each question, using the number of positive (in favor of MIIC), neutral (no preference), and negative (in favor of

OSFA) responses. The four questions and their results, with significance evaluated at confidence levels of $\alpha = 0.05$ and 0.01, are as follows:

Difficult material. One of the questions asked which of the two contained excessive difficult material. Twenty-three (53 percent) felt that each was fine, another 15 (35 percent) felt OSFA had too much, and 5 (12 percent) felt MIIC had too much. The p-value on this test was 0.025, significant in favor of MIIC at $\alpha = 0.05$.

Simple material. Another question asked which contained too much simple material. Twenty-nine (67 percent) felt each was fine, another 9 (21 percent) felt OSFA had too much, and 5 (11 percent) felt MIIC had too much. The p-value of 0.285 was not significant.

Better understanding. Another asked which of the two led to better understanding of the material. Twenty-six (61 percent) were for MIIC, compared to only 10 (24 percent) for OSFA. The p-value of 0.008 was significant in favor of MIIC.

Prefer overall. The last asked which of the two the user preferred overall. Thirty (70 percent) were in favor of MIIC, compared to only 9 (21 percent) for OSFA. The p-value was less than 0.001, significant in favor of MIIC.

5.3   Trial 2: Student Engagement

The second trial was conducted in September 2013.

Research questions. The purpose of this study was to investigate two more research questions: (RQ3) Which learning paths do students of MIIC traverse as a result of the user modeling process? (RQ4) Do students using MIIC have a higher level of engagement compared with those using OSFA?

Procedure. Three points distinguish the procedure of this trial from the first: (1) the content was Cellular Power Control; (2) OSFA was given as an integrated mobile app, making the only difference between MIIC and OSFA the lack of adaptation; and (3) each participant was only given either MIIC or OSFA, and was unaware of which she received.

Endpoints for engagement. In general, engagement is difficult to quantify, being defined as "the amount of physical and psychological energy that the student devotes to the academic experience" [4], [17]. In the end, we chose total page count as the main endpoint for engagement to investigate RQ4 (a similar endpoint was chosen in [53]). The reason for focusing on pages is two-fold, referring to the discussion in Section 4.1: (1) the pagecount timer captures the total time spent with any learning mode on the given page, and (2) the IR logic helps reduce uncertainty in the recorded times. The fact that the same measurement is used for both MIIC and OSFA also makes the comparison more fair. The reason for using total count rather than time spent is that viewing for a longer time is ambiguous; it could mean higher engagement or more confusion. To account for differences arising from users changing page size, we used total object count a second endpoint.

5.3.1   Results: Learning Paths (RQ3)

We will first give an overview of the learning paths traversed by the students, to give the direct results of the user modeling process outlined in Section 5.1. Here, we focus

---

4. Another approach may have been to make OSFA a multimedia e-Book (i.e., MIIC without individualization), as is done in the second trial, though it is not clear whether students would prefer an eBook to a textbook (see e.g., [55]).

Section(s)	Version Encoding (Base 2)
2 & 3	$(C_{emph}\ M_{emph})_2$
4 & 5	$(M_{hid}\ C_{emph}\ M_{emph})_2$
6	$(OL_{exp}\ MR_{exp}\ C_{emph}\ M_{emph})_2$
7	$(SH_{exp}\ OL_{exp}\ C_{emph}\ M_{emph})_2$

Fig. 9. Binary encoding of the version numbers for different sections in Fig. 6.

Learning Path	Frequency
$(2, (11)_2), (4, (000)_2), (6, (1000)_2)$	1
$(2, (11)_2), (4, (111)_2), (6, (1010)_2)$	1
$(2, (11)_2), (4, (111)_2), (6, (1111)_2)$	1
$(3, (00)_2), (5, (000)_2), (7, (0000)_2)$	1
$(3, (00)_2), (5, (010)_2), (7, (0010)_2)$	1
$(3, (10)_2), (5, (010)_2), (7, (0010)_2)$	2
$(3, (11)_2), (5, (000)_2), (7, (0010)_2)$	2
$(3, (11)_2), (5, (010)_2), (7, (0010)_2)$	4
$(3, (11)_2), (5, (011)_2), (7, (1111)_2)$	1

Fig. 10. Learning paths (omitting segs 1 and 8) among the 14 MIIC users who answered all the questions.

only on the users who were given MIIC, since OSFA had only a single path.

Version encoding. Referring to Fig. 6, different segment numbers were assigned the binary encoding given in Fig. 9 to describe the adaptive presentation. Each bit in the second column is a variable specific to the different sections:

C_{emph} and M_{emph}: These denote emphasis of conceptual and mathematical content, respectively. When set to 1, the color is green, and 0 means it is red.

M_{hid}: This denotes hiding extra example steps. When set to 1, they are hidden, and when 0 they are not.

OL_{exp}, MR_{exp}, and SH_{exp}: These denote expanding the advanced material subsections "Open Loop Power Control," "Matrix Representation," and "Soft Handoffs." When set to 1, they are expanded, and when 0 they are collapsed.

Note that the reason the version counts in Fig. 6 are not 2^x with x the number of version variables is that some combinations are not possible.

Analysis. The 24 MIIC users who proceeded far enough to answer the questions at the end of seg 1 are the subject of our analysis here. Of them, 7 (29 percent) were navigated to seg 2 while 17 (71 percent) went to 3, meaning that the majority of students received the less difficult (*NI*) path. Fourteen (58 percent) completed all questions on their respective paths, with a total of 9 distinct learning paths out of the 74 possible considering all combinations. These paths are shown in Fig. 10, along with the number of users for each.

The encoded variables range from a student who was navigated to the top path and had all variables 1 (third row) to a student who went to the bottom and had all variables 0 (fourth row). This corresponds to a range from the most to least advanced presentations possible, underscoring the heterogeneous demographic of the participants. The most common learning path was taken by four users (eighth row). While they were initially proficient in both concepts and math for their level ($C_{emph} = M_{emph} = 1$ in seg 3), as they moved through, they remained so in concepts ($C_{emph} = 1$ in segs 5 and 7) but began to struggle with math ($M_{hid} = 0$ in seg 5 and $M_{emph} = 0$ in segs 5 and 7). Additionally, no advanced material was shown to them ($SH_{exp} = OL_{exp} = 0$ in seg 7).

By including additional navigation paths in Fig. 6, it may have been possible to split users who were on the same initial paths further. This may have been beneficial for those 70 percent who were initially navigated to seg 3. To investigate this, we considered the subset of this 70 percent that were on the borderline of navigation to seg 2 (roughly speaking, having achieved $\geq 70\%$ of the required points). Only 5 of these 24 satisfied this criteria, which is to small to make statistical claims with, but it is surprising that of these 5, 3 completed the lecture. The 60 percent finishing rate of this group is roughly the same as the 58 percent rate of the 24 participants as a whole. This means that there was no evidence that having an

additional navigation path for this group would have helped their completion rate.

5.3.2 Results: Page and Object Count (RQ4)

Data handling. Since many users did not traverse far into the lecture and a number of the entries constituted a short duration more in line with browsing than studying, two filters were created: (1) an entry in the database was only considered valid if the elapsed time was at least 10 seconds; and (2) only users who reached seg 2/3 were considered, since these were the users who experienced the effect or lack of adaptation. Combined with the IR logic, the 10 second cutoff was a second precaution taken to discount entries most likely associated with off-task browsing or skipping through the material. Including it was seen to help discount a number of users with this apparent behavior.

There were 44 users who satisfied these criteria: 25 in the MIIC group and 19 in the OSFA group. Fig. 11 gives boxplots of the two endpoints by group. In Fig. 11a, the mean (standard deviation) for MIIC is 10.76 (5.95) pages, compared to 6.26 (4.92) for OSFA; in Fig. 11b, these values are 74.12 (39.45) for MIIC compared to 46.68 (41.93) for OSFA. The distribution for MIIC is visibly shifted to the right, suggesting a higher level of engagement when quantified in terms of page and object counts.

Analysis. Since Shapiro-Wilk tests [46] detected significant departures from normality, non-parametric tests were preferred over the standard t-test. The Wilcoxon rank sum test [47] is a nonparametric procedure which is more sensitive to differences between central tendencies than others;
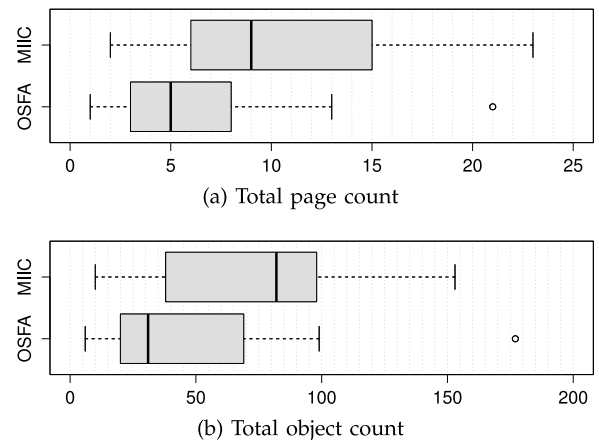

Fig. 11. Distribution of the (a) pages and (b) objects accessed by students in the OSFA and MIIC groups. The distribution for MIIC is shifted to the right, indicating higher engagement (with the exception of one outlier).

	Total page count	Total object count
p-value	0.009**	0.015*
95% CI	(1.00, 8.00)	(7.00, 57.00)
HLE	5.00	30.00

Fig. 12. p-values, 95 percent confidence interval (CI), and Hodges-Lehmann estimate (HLE) of the difference between the MIIC and OSFA groups for the distributions in Fig. 11a and 11b, respectively, using the Wilcoxon test.

we therefore we employed this as our primary method, with a continuity correction to the discrete distribution of the test statistic. Using this test, we computed (1) a two-sided p-value for testing the null hypothesis of no difference, with significance evaluated at $\alpha = 0.01$ or $\alpha = 0.05$, (2) a 95 percent confidence interval (CI) estimate for the shift in location, and (3) a Hodges-Lehmann estimate (HLE) of the shift; the HLE is an estimate of the shift in location parameter based on the Wilcoxon test.

The results are given in Fig. 12:

Significance testing. For pages, a p-value of 0.009 was obtained, significant in favor of MIIC. For objects, the p-value was slightly higher (0.015) but still significant.

Confidence interval. For pages, the difference between the means was between 1 and 8 with 95 percent confidence. For objects, it was between 7 and 57.

Distribution shift. The HLE of the difference for pages was 5, and 30 for objects. These shifts are large when considering the maximum counts from students in each case (23 pages, 153 objects). The percent increase in mean from OSFA to MIIC was 71.8 percent for pages 58.8 percent for objects.

5.4 Discussion: Key Messages and Next Steps

In investigating these research questions, we found:

- MOOC students using MIIC tended to have higher engagement than those using OSFA, when quantified in terms of page counts.
- MOOC students responded favorably to most of the features of MIIC (e.g., lecture videos and text emphasis), but not to the social learning aspect.
- MOOC students favored course delivery via MIIC to OSFA on a few dimensions, including overall preference and better understanding.

As a result of this, and also of the various perceived limitations identified for each trial, the following are the next steps we have begun to explore:

Additional data analytics. First and foremost are additional analytics. One way is to analyze the data collected from each separate learning mode in MIIC (rather than collectively, as is the case with total page count), which can each serve as a different proxy of engagement. We are planning additional trials to obtain data for this.

Another way is to implement more advanced techniques for updating the user model based on data collected from each learning mode, using the methods outlined in Section 3.1. To this end, we briefly describe the user modeling algorithm that was not tested in the preliminary trials but that is currently implemented in MIIC.[5] This algorithm relates a user's video behavior to her performance with a given

learning feature, which is accomplished by finding and updating the Pearson correlation coefficient [39] between performance on assessments the user has completed and the time the user spent watching the videos corresponding to these assessments for each feature. A composite performance measure, combining quiz scores with the video-watching behavior scaled by this correlation coefficient, is updated each time a user completes a segment, and is in turn used to determine the next seg-ver.

There are two potential benefits of having these two measures of performance. First is that performance can be updated even if the user has chosen to skip an assessment (i.e., by using the watching behavior score), which will be particularly useful in a situation like MOOC where quiz responses may only be optional. Second is that with additional information, the effect of the noise associated with guessing correctly and slipping behavior (i.e., answering incorrectly when the user actually knows the information, see e.g., [41] for a discussion) can be reduced. One of our next steps is to evaluate these potential benefits through additional trials.

Additional courses. These trials only include MOOC users from our own courses. In order to evaluate it in a more general setting, we are working with other authors to transform their content to MIIC format and run further trials. An example of this is with instructors from our own non-profit online education platform "3 Nights and Done" [1].

Additional metrics. In working with additional authors, we will change our endpoints to reflect the measure of efficacy in the given setting. In particular, for a class with strict learning goals, we can treat incremental performance as a primary endpoint [36].

Platform additions. We are working on extending the social learning in MIIC to include discussion forums, due to the poor user experience reported with the social notes from the first trial. Also, we are extending MIIC to other platforms besides iOS.

6 CONCLUSION

Scaling up effective learning is challenging. Moving from one-size-fits-all to a truly individualized experience can now be realized with the recent advances in mobile app programming and in data analytics. In this paper, we presented the design, implementation and preliminary evaluation of MIIC, an AES that delivers video, text, assessment, and social learning to users through a mobile native app that can automatically individualize content at a fine-granularity based on behavioral data collected about each user. The first such mobile app that was deployed to users around the world, MIIC collects behavioral measurements about each user as they interact with the course material, which can subsequently be used to drive the adaptation engine. We presented the results from our first two user trials with MIIC, which were conducted by recruiting participants from MOOC, and showed, for example, that these students tended to have higher engagement (when quantified as page and object counts) when using MIIC than when using OSFA. We have also identified next steps that we have begun to explore with MIIC, such as additional data analytics, evaluating our currently implemented user modeling algorithm, and testing with content from authors in other fields.

5. For the formal description of this algorithm, the reader is referred to Section 6 of our online technical report [9].

ACKNOWLEDGMENTS

REFERENCES

[1] (2013, Jan). 3 nights and done. [Online]. Available: http://www.3nightsdone.org

[2] Y. Akbulut and C. S. Cardak, "Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011," *Comput. & Educ.*, vol. 58, no. 2, pp. 835–842, 2012.

[3] A. N. Alutu, "The guidance role of the instructor in the teaching and learning process," *J. Instr. Psychol.*, vol. 33, no. 1, pp. 44–49, 2006.

[4] A. W. Astin, "Student involvement: A developmental theory for higher education," *J. College Student Pers.*, vol. 25, no. 4, pp. 297–308, 1984.

[5] Y. Bergner, S. Droschler, G. Kortemeyer, S. Rayyan, D. Seaton, and D. E. Pritchard, "Model-based collaborative filtering analysis of student response data: Machine-learning item response theory," in *Proc. 5th Int. Conf. Educ. Data Mining Soc.*, 2012, pp. 95–102.

[6] G. Bian, M. McAleer, and W.-K. Wong, "A trinomial test for paired data when there are many ties," *Mathem. Comput. Simul.*, vol. 81, no. 6, pp. 1153–1160, 2011.

[7] Á. F. Blanco, F. J. García-Peñalvo, and M. Sein-Echaluce, "A methodology proposal for developing adaptive cmooc," in *Proc. First Int. Conf. Technol. Ecosystem for Enhancing Multiculturality*, ACM, 2013, pp. 553–558.

[8] K. Bowen and M. D. Pistilli. (2012). Student preferences for mobile app usage. *Research Bulletin)(Louisville, CO: EDUCAUSE Center for Applied Research, forthcoming). [Online]. Available: http://www.educause.edu/ecar*

[9] C. Brinton, R. Rill, S. Ha, M. Chiang, R. Smith, and W. Ju. (2014, Oct). Miic design and preliminary evaluation: Online technical report. [Online]. Available: http://www.princeton.edu/cbrinton/MIIC_techReport.pdf

[10] C. G. Brinton and M. Chiang, *Networks Illustrated: 8 Principles Without Calculus*, Mendham, NJ, 07945, EdWiser Scholastic Press, 2013.

[11] C. G. Brinton, M. Chiang, "Social learning networks: A brief survey," in *Proc. Conf. Inf. Sci. Syst.*, 2014.

[12] C. Brinton, M. Chiang, S. Jain, H. Lam, Z. Liu, and F. Wong, "Learning about social learning in MOOCs: From statistical analysis to generative model," *To appear, IEEE Trans. Learning Technol.*, 2014, DOI: 10.1109/TLT.2014.2337900.

[13] E. Brown, T. Brailsford, T. Fisher, A. Moore, and H. Ashman, "Reappraising cognitive styles in adaptive web applications," in *Proc. 15th Int. Conf. World Wide Web*, 2006, pp. 327–335.

[14] E. Brown, T. Fisher, and T. Brailsford, "Real users, real results: Examining the limitations of learning styles within aeh," in *Proc. Eighteenth Conf. Hypertext and Hypermedia*, 2007, pp. 57–66.

[15] E. J. Brown, T. J. Brailsford, T. Fisher, and A. Moore, "Evaluating learning style personalization in adaptive systems: Quantitative methods and approaches," *IEEE Trans. Learn. Technol.*, vol. 2, no. 1, pp. 10–22, Apr. 2009.

[16] P. Brusilovsky, "Methods and techniques of adaptive hypermedia," *User Modeling and User-Adapted Interaction*, vol. 6, nos. 2-3, pp. 87–129, 1996.

[17] R. M. Carini, G. D. Kuh, and S. P. Klein, "Student engagement and student learning: Testing the linkages," *Res. Higher Educ.*, vol. 47, no. 1, pp. 1–32.

[18] R. M. Carro, E. Pulido, and P. Rodríguez, "Tangow: Task-based adaptive learner guidance on the www," in *Proc. Second Workshop Adaptive Syst. User Model. World Wide Web*, 1999, pp. 49–57.

[19] S. Cetintas, L. Si, Y. P. Xin, and C. Hord, "Automatic detection of off-task behaviors in intelligent tutoring systems with machine learning techniques," *IEEE Trans. Learn. Technol.*, vol. 3, no. 3, pp. 228–236, Sep. 2010.

[20] M.-P. Chen and J.-C. Yen, "An evaluation of learners' satisfaction toward mobile learning," in *Proc. 6th WSEAS Int. Conf. Appl. Comput. Sci.*, vol. 6, Citeseer, 2007, pp. 382–388.

[21] M. Chiang, *Networked Life: 20 Questions and Answers*. Cambridge, United Kingdom: Cambridge University Press, 2012.

[22] D. Clow, "Moocs and the funnel of participation," in *Proc. Third Int. Conf. Learn. Analy. Knowledge*, 2013, pp. 185–189.

[23] D. Crawford. (2013, Jul.) Why mobile web apps are slow. [Online]. Available: http://sealedabstract.com/rants/why-mobile-web-apps-are-slow/

[24] P. De Bra, D. Smits, and N. Stash, "Creating and delivering adaptive courses with aha!" in *Innovative Approaches for Learning and Knowledge Sharing*. New York, NY, USA: Springer, 2006, pp. 21–33.

[25] R. Dunn, "Understanding the Dunn and Dunn learning styles model and the need for individual diagnosis and prescription," *Reading, Writing Learn. Disabil.*, vol. 6, no. 3, pp. 223–247, 1990.

[26] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Eng. Educ.*, vol. 78, no. 7, pp. 674–681, 1988.

[27] R. M. Felder and B. A. Soloman. (2012, Sept). Index of learning styles. [Online]. Available: http://www.ncsu.edu/felder-public/ILSpage.html

[28] A. France-Presse. (2014, Jan). Tablet sales jumped 50 percent in 2013. [Online]. Available: http://gadgets.ndtv.com/tablets/news/tablet-sales-jumped-50-percent-in-2013-idc-477029

[29] P. A. Henning, F. Heberle, A. Streicher, A. Zielinski, C. Swertz, J. Bock, and S. Zander, "Personalized web learning: Merging open educational resources into adaptive courses for higher education," in *Proc. Personalization Approaches Learn. Environ.*, p. 55, 2014.

[30] P. Honey and A. Mumford, *Learning Styles Questionnaire*, Organization Design and Development, Incorporated, 1989.

[31] K. Jordan. (2013). Mooc completion rates: The data. [Online]. Available: http://katyjordan.com/MOOCproject.html

[32] J. Kim, P. J. Guo, D. T. Seaton, P. Mitros, K. Z. Gajos, and R. C. Miller, "Understanding in-video dropouts and interaction peaks in online lecture videos," in *Proc. First ACM Conf. Learning@ Scale Conf.*, 2014, pp. 31–40.

[33] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses," in *Proc. Third Int. Conf. Learn. Anal. Knowl.*, 2013, pp. 170–179.

[34] A. S. Lan, A. E. Waters, C. Studer, and R. G. Baraniuk, "Sparse factor analysis for learning and content analytics," *J. Machine Learning Res.*, vol. 15, pp. 1959–2008, 2014.

[35] T. Lewin, "Online classes fuel a campus debate," *The New York Times*, vol. 16, p. A16. Jun. 2013.

[36] C. Limongelli, F. Sciarrone, M. Temperini, and G. Vaste, "Adaptive learning with the ls-plan system: A field evaluation," *IEEE Trans. Learn. Technol.*, vol. 2, no. 3, pp. 203–215, Sep. 2009.

[37] E. Martín and R. M. Carro, "Supporting the development of mobile adaptive learning environments: A case study," *IEEE Trans. Learning Technol.*, vol. 2, no. 1, pp. 23–36, Apr. 2009.

[38] M. Melia and C. Pahl, "Constraint-based validation of adaptive e-learning courseware," *IEEE Trans. Learn. Technol.*, vol. 2, no. 1, pp. 37–49, 2009.

[39] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.

[40] I. Nawrot and A. Doucet, "Building engagement for mooc students: introducing support for time management on online learning platforms," in *Proc. Companion Publ. 23rd Int. Conf. World Wide Web Companion*, International World Wide Web Conferences Steering Committee, 2014, pp. 1077–1082.

[41] Z. Pardos and N. Heffernan, "Using HMMs and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset," *J. Mach. Learn. Res.*, 2011, http://people.csail.mit.edu/zp/papers/pardos_JMLR_in_press.pdf

[42] L. Perna, A. Ruby, R. Boruch, N. Wang, J. Scull, C. Evans, and S. Ahmad. (2013, Dec.). The life cycle of a million mooc users. [Online]. Available: http://gse.upenn.edu/pdf/ahead/perna_ruby_boruch_moocs_dec2013.pdf

[43] S. Rendle, "Factorization machines," in *Proc. IEEE 10th Int. Conf. Data Mining*, 2010, pp. 995–1000.

[44] L. Ryan. (2013, Jan.). Moocs are on the move: A snapshot of the rapid growth of moocs. [Online]. Available: https://efmd.org/index.php/blog/view/250-white-paper-moocs-massive-open-online-courses

[45] V. D. Schmittmann, I. Visser, and M. E. Raijmakers, "Multiple learning modes in the development of performance on a rule-based category-learning task," *Neuropsychologia*, vol. 44, no. 11, pp. 2079–2091, 2006.

[46] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.

[47] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL, USA: CRC Press, 2003.

[48] N. Sonwalkar, "The first adaptive mooc: A case study on pedagogy framework and scalable cloud architecture part i," in *MOOCs Forum*, vol. 1, no. P, 2013, pp. 22–29.

[49] D. Szafir and B. Mutlu, "Artful: Adaptive review technology for flipped learning," in *Proc. SIGCHI Confer. Human Factors in Comput. Syst.*, 2013, pp. 1001–1010.

[50] N. Thai-Nghe, L. Drumond, T. Horváth, and L. Schmidt-Thieme, "Using factorization machines for student modeling," in *Proc. UMAP Workshops*, 2012, http://ceur-ws.org/Vol-872/fact-mod2012_paper_4.pdf

[51] A. Toscher and M. Jahrer, "Collaborative filtering applied to educational data mining," in *Proc. KDD Cup*, 2010, http://pslcdatashop.org/KDDCup/workshop/papers/KDDCup2010_Toescher_Jahrer.pdf

[52] G. Weber, and P. Brusilovsky, "Elm-art: An adaptive versatile system for web-based instruction," *Int. J. Artif. Intell. Educ.*, vol. 12, pp. 351–384, 2001.

[53] G. Weber and M. Specht, "User modeling and adaptive navigation support in www-based tutoring systems," *Int. Centre Mech. Sci.: Courses Lect.*, pp. 289–300, 1997.

[54] S. H. R. Wong, "Which platform do our users prefer: website or mobile app?" *Reference Serv. Rev.*, vol. 40, no. 1, pp. 103–115, 2012.

[55] W. D. Woody, D. B. Daniel, and C. A. Baker, "E-books or textbooks: Students prefer textbooks," *Comput. Educ.*, vol. 55, no. 3, pp. 945–948, 2010.

[56] D. Yang, T. Sinha, D. Adamson, and C. P. Rosé, "Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses," in *Proc. NIPS Data-Driven Educ. Workshop*, 2013, http://lytics.stanford.edu/datadriveneducation/papers/yangetal.pdf

[57] D. Zhang, L. Zhou, R. O. Briggs, and J. F. Nunamaker Jr, "Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness," *Inf. & Manage.*, vol. 43, no. 1, pp. 15–27, 2006.

Christopher G. Brinton (S'08) received the BS degree in electrical engineering from The College of New Jersey (valedictorian and summa cum laude), Ewing Township, NJ, in May 2011, and the master's degree in electrical engineering from Princeton University, Princeton, NJ, in May 2013. He is currently working toward the PhD degree in the Department of Electrical Engineering at Princeton University. His research focus is in adaptive learning systems, machine learning for educational data mining, and social learning networks. He is a MOOC instructor and co-author of a textbook on social and technological networking.

Ruediger Rill is currently the Director of Software Engineering at Zoomi Inc, Winnipeg, Canada. He has managed, planned, and executed various software projects as a freelancer and has been selling his proprietary software for over 6 years. His free software products targeting the iOS jailbreak community have received over 1.1 million downloads, and those receiving royalties have in excess of 100,000 sales.

Sangtae Ha (S'07, M'09, SM'12) received the PhD degree in computer science from North Carolina State University, Raleigh, NC, in 2009. He is currently an Assistant Professor in the Department of Computer Science at the University of Colorado at Boulder, CO, with a joint appointment in the Interdisciplinary Telecommunications Program. He began his postdoctoral appointment at Princeton University, Princeton, NJ, where he co-founded the EDGE Lab as its first Associate Director in 2009 and led its research team as an Associate Research Scholar from 2010 to 2013. He serves as an Associate Editor for the *IEEE Internet of Things* (IoT) Journal and an Editorial Board for IEEE ComSoc Technology News. He is a senior member of the IEEE and was a general co-chair of the Smart Data Pricing workshop in INFOCOM 2013. He co-founded DataMi in 2012 and served as the founding CTO, and is a technical consultant to a few startups.

Mung Chiang (S'00, M'03, SM'08, F'12) is the Arthur LeGrand Doty Professor of Electrical Engineering at Princeton University, Princeton, NJ. His research on communication networks received the 2013 Alan T. Waterman Award from the US National Science Foundation, the 2012 Kiyo Tomiyasu Award from IEEE, and various young investigator awards and paper prizes. A TR35 Young Innovator Award recipient, he created the Princeton EDGE Lab in 2009 to bridge the theory-practice divide in networking by spanning from proofs to prototypes, resulting in several technology transfers to industry and two startup companies. He is the Chairman of the Princeton Entrepreneurship Advisory Committee and the Director of the Keller Center for Innovations in Engineering Education. His MOOC in social and technological networks reached about 200,000 students since 2012 and lead to two undergraduate textbooks and he received the 2013 Frederick E. Terman Award from the American Society of Engineering Education. He was named a Guggenheim Fellow in 2014.

Robert Smith received the MS degree in mathematics from Syracuse University, Syracuse, NY, and the PhD degree in statistics from the University of North Carolina at Chapel Hill, NC. He has worked for 29 years on the design, analysis and reporting of clinical trials, first at CIBA-Geigy, then at Bristol-Myers Squibb, and now as a private consultant.

William Ju received the MD degree with Alpha Omega Alpha honors from the School of Medicine, University of Pennsylvania, Philadelphia, PA, and the AB degree with Phi Beta Kappa honors from Princeton University, Princeton, NJ. He is a co-Founder of Zoomi Inc, Winnipeg, Canada. He is physician-scientist who has been President and CEO of Follica, Inc. and COO at PTC Therapeutics, Inc. Before entering industry, he was a Senior Staff Fellow in basic research in the Laboratory of Cellular Oncology at the National Cancer Institute, National Institutes of Health. He has served on the Board of Trustees at Caldwell College, the Pingry School, and the Peck School.