A Marking System Designed for Efficiency and Fragmented or Collaborative Assessments

Zhuhan Jiang and Jiansheng Huang

***Abstract*—While standard e-platforms do support a variety of teaching and learning activities in tertiary institutions to different degrees, they are typically designed to be "doable" and sometimes even theoretically so anyway, rather than taking efficiency as an equally high priority. In contrast, we developed a marking system to streamline the handling of marking allocations, irregular student transitions from class to class or from marker to marker, dynamic membership changes of student group work, instant creation and re-use of marker-defined feedback comments, along with other associated activities. In this work, we will however focus more on the functionalities and their design details on creating a marking system that especially well supports dynamically adaptable marking criteria and freely adjustable feedback data banks. The designing goal is to make this system as efficient as possible to the markers, as fair as possible to all the students, and as convenient as possible to the system operators.**

***Index Terms*—E-learning tools, marking system, optimized efficiency, fragmented or collaborative assessments, improved fairness.**

I. Introduction

In tertiary studies nowadays, modern technologies are already indispensable almost everywhere. While the effectiveness of e-learning and blended learning is still being widely investigated [1], the use of e-platforms such as Blackboard or various similar products can be found in almost all universities. We for simplicity will refer to all such platforms and their related tools as web course tools (WCT). Currently most WCTs are largely designed to achieve typical teaching and learning functionalities such as archiving learning materials, conducting online quizzes and managing student submissions, to name a few. Some forms of e-learning portals are often also provided for such WCTs. While evaluations on WCTs or portals [2]-[4] typically concentrate on what can be done and how they can be done, the consideration on the efficiency and effectiveness is largely left out or placed on a lower priority. For a core unit or a university service unit, aka a subject, each cohort can easily reach 3-5 hundred of students, and some can even push towards a thousand. This means markers for the submitted student work will often have to mark a large number of student submissions within a given period of time, and make the marking consistent and fair across the board at the same time. There are already some generic strategies [5]-[7] to make the marking more consistent and fairer, and to deal with assessments in open online courses [8]. While there can be automated marking [9] or even automated feedback systems [10] for certain types of disciplines or topics, the traditional way of marking with relevant written feedback is still irreplaceable. However, none seem to have touched on how to save markers' time or give markers a truly efficiency-oriented support at the level of software tools. The main aim of this work is thus to explore this missing part in depth, design and develop such a marking system that can reduce as much as possible the repetitive work in the marking process, and fast-track the navigation and comparison among different assessment items of the same or different student work.

One question one may contemplate now is why there need to be additional course tools when there are already a few popular WCTs available. This question is actually not that different in nature from asking why there should be more than one WCT on the market although the latter is also a legitimate question after all. First, most WCTs are designed for generic teaching and learning purposes or activities, with the instructors being assumed not versed as an IT professional. Hence learning materials are typically archived in a tree structure and most files are only in the form of Word, PDF or multimedia rather than, say, those more interactive HTML documents. Although the design paradigm, being object-oriented (OO) in general, is ideal for the software development and for the newbies to lay their hands on, it may not be the most efficient way to handle massive repetitive tasks such as marking the same assessment items online for many students. Massive number of mouse-clicking and page navigation do consume a great deal of marking time and drains the markers significantly. Our design here is thus to adopt a more intuitive and linear approach in the interface to save time while keeping the OO design only within the tool development. Second, the most effective and efficient solutions are more likely to arise from the field practitioners, the ones who are managing and doing the repetitive marking chores in this case. As a result of our developing, utilizing, and refining our e-tools for the marking, we hope we have also provided input to those major WCT vendors to further advance their products.

This work is organized as follows. We first in Section II consider the factors that impact on our general approach in designing a more efficient marking system, and lay out simple conventions that can help locate positions of interest in a feedback text. Section III then describes our strategies to deal with multi-platforms, fragmented and collaborative assessments, in comparison with other type of marking approaches. The technical design, implementation, and illustrations are left in Section IV. Finally Section V gives a

Manuscript received July 1, 2019; revised February 29, 2020.

The authors are with the School of Computing, Engineering and Mathematics, Western Sydney University, NSW, Australia (e-mail: z.jiang@westernsydney.edu.au, j.huang@westernsydney.edu.au).

doi: 10.18178/ijiet.2020.10.4.1372

conclusion.

II. Efficiency Factors

If one is to mark repetitively a large number of student submissions, a mere extra minute saved on each question can translate into many hours. Marking students' online-submitted work is a very pertinent example. A WCT will typically have support for assignment submissions, marking feedbacks and grading. The setting-up of the marking criteria and potential selectable marking feedback is tedious with such a WCT, and once set up, markers don't have much freedom to alter the structure. In fact, markers typically have to navigate or transit from one marking state to another as mandated by the workflow of the software tool. Although each step of such a transition, say, for the purpose of adding an extra comment, may take just a few clicks or page transitions, the time used will quickly add up. The result is that markers generally don't like using such marking tools despite of their posh interface look and the modular feel, and that is purely because the markers actually get much more drained in the process.

A. GUI Input vs Linear Text

A simple example to illustrate an input via GUI is not always efficient than an input in plain text is the typesetting of an article full of mathematical symbols and expressions. The typesetting of

$$\int_{\alpha}^{\beta}\sum_{n=1}^{9} f_n(x)dx = 1$$

in Microsoft Word, for instance, seems pretty simple and clear, but it takes much less time to achieve the same coding for TeX in terms of the following text

$$\int^\beta\alpha \sum_1^9 f_{n=1}(x)dx=0$$

as the objected-oriented input with Word requires a fair amount of time to locate different element objects and navigate among the index boxes. Although TeX does need to have an extra step of document compilation by a software, it is only a one-off step and hence the overall time required is much less, and is therefore more efficient for such cases when compared with the GUI-based, object-oriented, pleasantly organized and presented entry method provided by Word. This also partly explains why many scientists and mathematicians prefer using TeX or LaTeX to using Word.

B. Neat Posh Interface vs Over-Display

One of popular online marking approaches is to design a web-based form in which all questions or tasks in an assessment are displayed along with selectable standard marking comments as well as an optional space to enter feedback comment in free plain text format, see in Fig. 1 a snapshot from a marking sheet in a real delivered web subject. The screenshot on the top collects mutually exclusive choices into a dropdown menu for Task 3 and displays possible multiple features in checkboxes for Task 4, achieving the best presentational effect. Here a marker can select relevant features with marks automatically calculated, along with adding optional additional comments. This is a good design, on the first look, of having a superb professional interface, with a decent amount of automated calculations, and having most purposes catered for. However, behind the façade of the posh look, there are a few hidden aspects that are hindering a more efficient (faster) marking prospect.

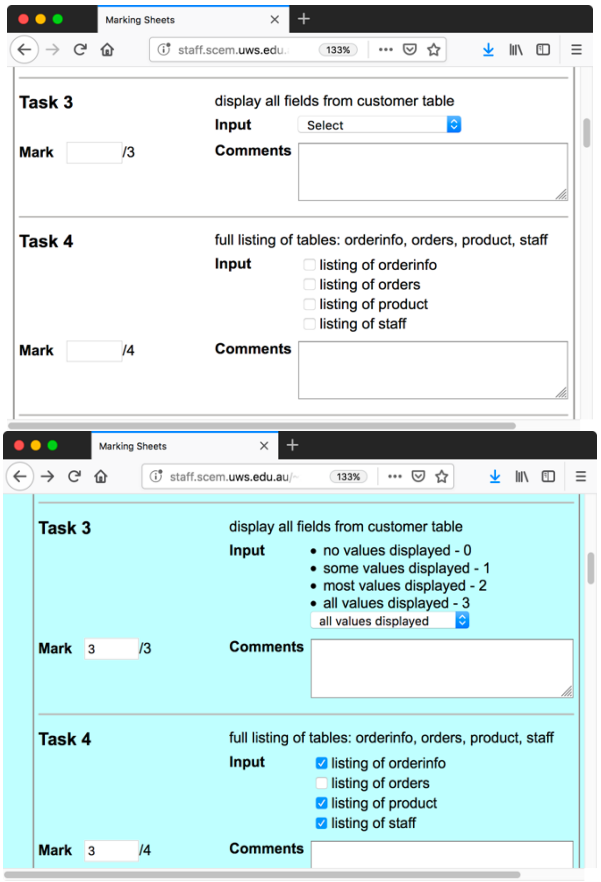

Fig. 1. Over-display with defaults for efficiency.

For simplicity, we will ignore the overhead of designing such a form as this is just one-off exercise for each new assessment, and one can also develop a software tool to help create such a form. Now we first observe that while the input dropdown box is neat and does fulfill properly its purpose, it irritates the markers as they cannot easily see *all* the available items in the selection box without having to click it open first. And if one browses over the whole marking sheet, clicking many such boxes just to view the other options is both time-consuming and irritating. This problem can be overcome if the combo items are also displayed in flat format or the dropdown box is replaced by exclusive radio button group. This "over-display" actually makes the marking more comfortable and efficient. Second, certain questions or sub-questions may receive a typical mark, such as 0 or full mark, for most students, asking a marker to keep entering those same typical marks costs time as well. An obvious solution here is to provide a default (predicted) mark for some or all questions, see the bottom of Fig. 1 for a modified marking system which solves this problem. Third, the space for entering optional comment is only allocated to each *whole* question rather than right below the most pertinent marking criterion or feature. This requires the marker to add necessary

positioning information in the comments rather than merely an academic comment. It will not make sense to attach such a textbox under every standard feature or description because it will be cluttering. Moreover, potential lengthier comments to include extra examples cannot be made fully visible directly on the form unless browsing further within the sub text window for the comments. This point will be further addressed later, and one of our main goals here is to flatten out this type of OO based design into a linearized text format with which all marks and comments can be rapidly and freely added, edited, and viewed in full anywhere. In a way our approach is to delegate the navigation among the labyrinth of OO objects to the software tool through our live text parser and the implemented context sensitivity.

C. Convention vs Configuration or OO

It is well-known that OO design has advantages over other forms of designs in software development, and one can use conventions to reduce the unnecessary flexibilities or apply configurations to alter a system setting to a particular convention. While potential overhead on the use of an OO design is more than compensated by its conceptual simplification and such an overhead has a negligible impact to a modern computer system. For repetitive chores by the human, such an overhead, such as a great many extra mouse clicks and page transitions, can easily accumulate into an intolerable amount. However, if one flattens out an OO design into a linear presentation in, say, the following form

```
MARKS  | ASSESSMENT ITEMS
========+================================
       | 3. Task 3  (3 points)
3      | Display all fields from customer table.
       | i.  No values displayed
       | ii. Some values displayed
       | iii. All values displayed
Comments in free textual format, taking up as many
lines as one needs, and placeable anywhere
       |    with screenshots
       | 4. Task 4 (4 x 1 points)
       | Full listing of tables: orderinfo, orders, etc.
0      | i.  Listing of orderinfo (1 point)
An optional comment to explain the loss of the mark

1      | ii. Listing of orders (1 point)
       | the rest is snipped
```

then it has a number of efficiency advantages compared with the object-oriented display: i) A holistic view with one-click to write marks or comments anywhere. ii) Textual comments can be inserted anywhere to indicate the relevance, and are visible in entirety and editable with most editing tools. iii) Marks are placed on the left of "|", and the absence of an "|" indicates it is just a comment rather than a feature or aspect that is expected of the student answer. iv) Marks and comments can be directly entered at any levels without having to keep clicking to enter or unfold multiple tiers of objects. v) Marks can be selectively entered at any level: a perfect answer requires to enter a single mark at the outmost layer for that question; a near perfect mark can be entered with a *deductive scheme* in that a full mark will first be given at the outmost layer, followed by inserting mark deductions, i.e. negative marks, where pertinent. A poorly completed question will have the marks entered with an additive scheme in that all individual marks are entered at the inner layers for sub-questions or individual goals or activities.

The efficiency is essentially built on top of the adoption of the following layout conventions:

1) Each line of text belongs to the category of either a title if it contains a "|" character, or a comment if otherwise. Other relatively "rare" character such as '`' or '~' can also be used in place of '|', however '|' visually serves as a better vertical separator. All '|' on the same line apart from the leftmost one will be converted to its "blank" replacement '_' in the standardization so as to avoid the confusion.
2) A feedback title is the description of a question or a sub-question or any separate assessment aspect, and these are formally numbers in the format of #. or #.#.···#. where # are all positive integers, indicating respectively the depth of the sub-questions. For instance, '1.' refers to question 1, and '1.2.1' refers to subsub-question 1 of sub-question 2 of question 1. For simplicity, we also adopted shorthand with roman numbers and letter numbering. For example, 'iii.' under title of level '2.' is equivalent to '2.3.' and 'f)' under the level '2.3.' is equivalent to level '2.3.5', and is equivalent to '2.0.5' if directly under, say, level '2.'. In other words, the shorthand in roman numbers is designated to the 2nd level and the shorthand in English letters is designated to 3rd level.
3) Each full title has to start with a valid level index number or its shorthand representation, otherwise it is considered as the continuation of the previous title. A full (question or sub-question) title may consist of multiple textual lines in sequence, potentially interweaved with comment lines, see the red-colored part of the sample marking sheet in the above.
4) Some selected patterns containing '|' may be forced to be treated as comments. For example, a line containing "MARKS | ASSESSMENT ITEMS" on the left is always treated as a comment. This adds more flexibilities to the layout of the marking sheet. The very top of the marking sheet is considered to be already under the empty title of level '0.'.
5) On the left hand side of '|' of each title line, all alphabet letters will be ignored and essentially removed when calculating the total mark which is the sum of all the numbers on the left of '|' in each line.

Every marking sheet observing these design and authoring conventions can be parsed into arrays of titles and their corresponding relevant marks and comments, which can be utilized to regenerate the original marking sheet. In fact, this parsing-regenerating procedure will standardize any sheet to the standard format, and this would largely be equivalent to other WCT tools making everything object-oriented and thus forcing clients to navigate through many unnecessary layers of conforming activities. Our doctrine is that we will trust a client's conforming to all the conventions without having to go through those unnecessary layers, and will be able to convert any feedback file into the structured standard format

anyway if there is a need to do so. In the case of accidentally having multiple titles under the same title ID, the parsing can optionally choose to merge them or treat them as if they were separate IDs. The latter however better retains the original sheet format.

By now we can conclude for this section that marking systems designed for their professional look and uniformity across the disciplines and users are in reality often not the most efficient approaches for the markers.

III. Design Features

In this section, we will look at a few relevant aspects and features associated with our marking system, especially those that are not well supported by the current commercial WCTs on the market.

A. Fragmented Assessment Items and Marks

A good example of fragmented assessment items and their corresponding fragmented marks is the assessment of tutorial and practical activities, and their task completion. The typical characteristics are i) Small and irregular amount of marks are allocated to individual questions or tasks. ii) The marking takes place during the class activities, and the deadline is within the same class or is at the latest to be marked within the next class time. Student work must be done prior to the start of the next class time so as not to conflict with the current class activities other than the marking part. iii) The tut/practical classes are repeated for different groups of students and scheduled often at different time slots. iv) Students may have to move from one practical class to another in the earlier weeks and may have to attend a different practical class due to making up for the missing ones. Hence the marking system should support the cross-group marking and the marks are preferably instantly visible to the students, see the illustration in Section 4 for our real implementation.

B. Partial Statistics for Marks Consistency

When there are many duplicated tut classes run by quite a few different tutors who are also the corresponding markers, the marking consistency and fairness becomes a challenge. The grading system provided by our existing WCT, apart from the excessive time-consuming navigation, treats the whole teaching team uniformly without any discrimination. This is manifested in the following characteristics: i) Assessors or markers are typically granted the same access rights to marking all students, even though they may each be allocated a subset of the students for the marking. Markers may accidentally alter other students' marks without anyone's awareness. ii) Since a typical commercial WCT system doesn't track and log individual markers' activities, at least not at the users' level, an accidental editing on marks or feedback text is neither traceable nor recoverable for the overwritten content. The unit coordinator, or the instructor in charge, has no option but to place the same faith on everyone in his teaching team. iii) The breakdown of the marks statistics marker-wise is not directly available on a WCT and would have to be calculated via a third-party tool.

To overcome these shortcomings, our marking system tracks all grade changes and all major changes on marking feedback, easily traceable to the individual marker and time etc by the unit coordinator. Because of the tracking of the markers, the marking statistics can be calculated real-time so that each marker can view for instance what marks averages other markers are having up to the current time. This will allow markers to compare with each other and alter their own marking *instantly* so as to be more consistent with other markers, see the statistical entries in Section 4 for our real implementation. This will thus lead to a fairer grading to the students. Moreover, markers can also compare their own marking feedback with others' to see if further improvement can be done, and can also take on board other markers' better feedback comments for their own use.

C. Marking Rubrics vs Free Comments

Marking lubrics represent a set of generic descriptions which are often holistic and somewhat subjective. Although in principle some feedback comments can be predicted and incorporated into marking lubrics or the like, they are often less effective compared with giving a more relevant free-formatted comment. If all possible comments are to be itemized and incorporated into a marking system, it could turn into another nightmare as it would cost markers a lot of time to view and select right answer if there is one. While user-created feedback comments can explain everything back to the students, similar comments may have to be typed again and again, and it may not be easy to locate an earlier comment among all the previously completed marking feedback.

D. Feedback Creation and Insertion

Our solution is to allow individual markers to set up their own list of feedback items for each question or sub-question dynamically, and save them for a future use as well. This can be regarded as a semi-automated feedback system, sitting between the lubrics and the free-formatted comments. There are a number of features that can be further implemented to improve this system: i) The available feedback blocks can be dynamically ordered with the most frequent appearing on the top. ii) A mark may be suggested based on the marking of the previous students' work on this question who received a similar marking feedback. iii) An overall prediction may also be possible if some form of data-mining is utilized, although admittedly such a prediction is largely indicative and cannot be accurate.

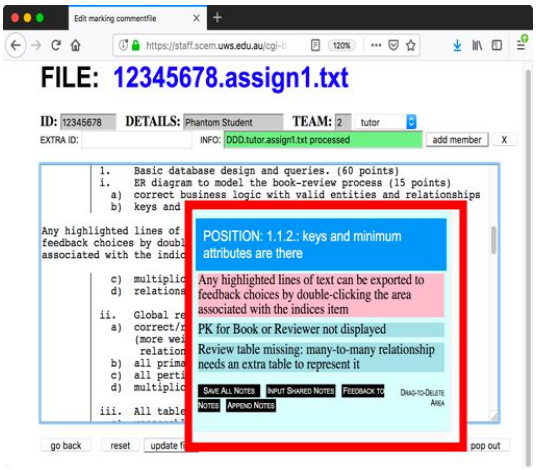

Fig. 2. Context-sensitive feedback.

Our main goal on this part is to provide a web-based plain text editor where all items or marking criteria are indexed by their IDs or shorthand IDs, and a double-click will bring up a feedback selection to pick according to where the cursor position is. If a new block of feedback text needs to be added to the feedback bank, just highlight the block before the double-click, as shown in Fig. 2. The strategy and the data flow for the design of such a system is described in the following diagram, where the entry starts with a double-click on the text being edited for the feedback. We note here the text in pink was originally selected before the double-click, after which the text in the editor gets deselected. The panes containing different feedback comments can be dragged to different ordering so that the most popular comments will appear on the top. Technically speaking, the most challenging part in this subsection is to create a parser for the feedback and to make use of Ajax and Json with Javascript (JS) for the interactivity, see Fig. 3 for the execution workflow.

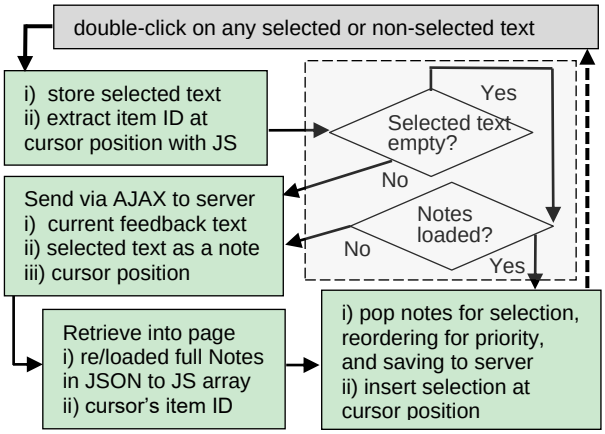

Fig. 3. Feedback notes selection and creation cycle via Ajax and Json.

More specifically, we first intercept a double-click and a right-click with JS within a TEXTAREA element which is the web editing window for the marking feedback, where the right-click is still reserved for the future addition of AI support. Then JS is utilized to extract current content, cursor position, assessment item ID etc and have these data posted to a parser on the web server. It is important that the content be first converted to MIME64 before the transmission so as to retain perfectly the special or other linguistic characters. The parser is written in PERL and will extract the complete data structure; it needs to be robust in that the it should be able to interpret the feedback content properly no matter how an author makes a "mess" of it. In particular, we added hidden indexing on potential duplicated item indices on the marking criteria or features so that the feedback regenerated from the extracted data structure remains very much the same in content and in order. The parser typically returns via Ajax the result, or the job to do in the user's browser, in JS along with the relevant data in Json, and must have the JS activated directly as the Ajax-returned JS won't be automatically executed as in loading a normal webpage. This system is not for experimental purposes; it is already in full and effective use.

IV. Implementations

In order to better illustrate our developed marking system, we opt to make use of essentially one of the actual unit deliveries. We will however scramble the student numbers and names phrase-wise. Even though these names will not look like real names, different original names will always lead to different masked names and vice versa. In other words, we lose nothing in terms of the realness under these masked names. Fig. 4 is a screenshot of an external portal to the unit website where everything is accessible in regard to a database unit. It supports all teaching, learning and management activities of which some are displayed by opening the dropdown box for the "tutor's selections". The main purposes for developing such an external portal is to make the delivery of large units more flexible and effective for the students, and more efficient and time-saving for the teaching staff.

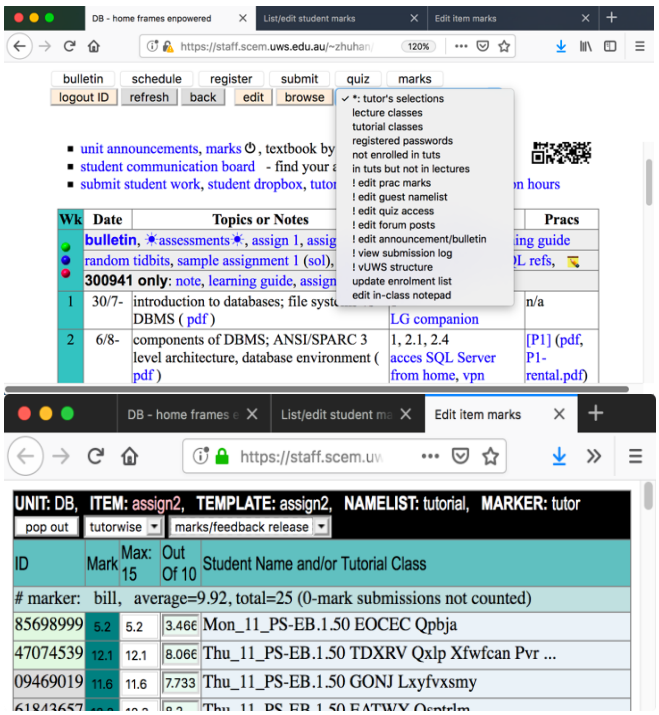

Fig. 4. Site portal and marks averages per marker.

We note that the colors on the ID indicate whether a student is in a group, or on his own, or not submitted. Many marking details are provided as tooltips. For example, leaving the cursor on a mark will indicate how many times this mark has been changed in the past and who is the marker who last updated the mark. This tutor-wise listing lists all the student entries per tutor and gives at top a statistical summary for the corresponding marker. Hence each marker can observe other peer markers' statistics and may thus adjust his or her own overall leniency accordingly on reading other markers' feedback as well. Each marker can also conveniently download all his feedback to the students via a single download hyperlink, see the red-colored "assign2" in Fig. 4. We deliberately disabled an automatic download for all markers' feedback although every marker can view and copy/paste everyone's feedback manually and one by one. This part proves surprisingly effective in bringing markers marking averages closer to one another's.

In order to protect the privacy of the relevant student

details, we masked the student IDs and names by mapping a word of digits into an encrypted word of digits of the same length, and likewise for a word of alphabet letters. Fig. 5 contains the screenshots of the marking system on a smart phone, and shows that it's very easy to enter or edit student marks for those fragmented assessment items. The screenshot on the right also illustrates the automatic and incremental student number matching when student ID is being entered. This makes the in-class marking easy as a breeze, a critical factor that makes in-class marking feasible.

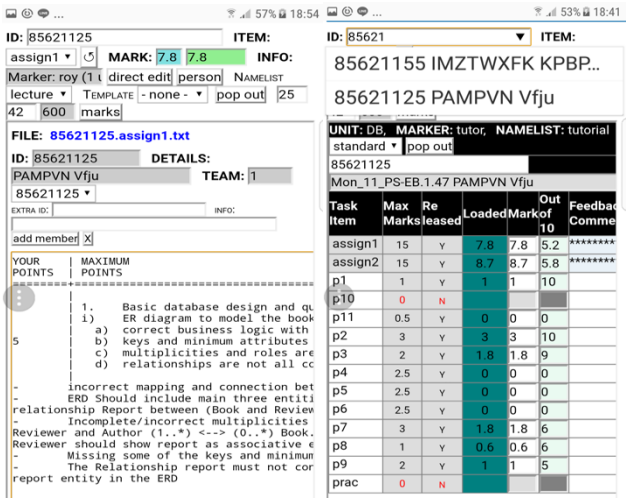

Fig. 5. Easiest to enter/edit marks or view marking comments on a smart phone.

To conclude this work, we note that, other than the feedback editing part that we covered in greater details in this work, our developed marking system also implemented many efficiency measures for speeding up the marking: i) While locating a student record can go through the student listing, campus wise or tutorial class wise, which is logically clean, it often involves an extra searching via a full student ID. Our system implements a live searching as the ID is being gradually entered, as the dropdown box shown on the top right of Fig. 5. This is implemented on every place where an ID may be entered. ii) The team membership for an assessment is implemented as linked files so that only one copy of feedback file will be created for all students within the same assignment team, and this is achieved through the use of a Linux operating system at the backend. iii) Marks can always be entered directly or in percentages. A modification in one format will instantly update the display in both formats. This feature is very handy for assessments that involve many small marks, and a marker's decision is mostly about an evaluation out of 10. iv) Marking feedbacks can be searched via a single phrase among the students marked by the current marker, returning a list of matched students along with the first few matched lines and their near-by text. This may help a marker pull back similar student work previously marked to check for potential similarity or consistency. v) Our system not only tracks all marks changes, which makes potential auditing possible, but also tracks major versions of feedbacks. Since a marker may do several file-savings before completing an assessment, it doesn't really make much sense to track all the updating on a written feedback file. Our approach here is to always log the previous version of a feedback document if it is currently being modified by a different marker, or if the previous version is at least 1 day ago.

V. Conclusion

We proposed and developed a full-fledged working system for efficiently managing and marking fragmented or collaborative assessments, and for significantly saving markers' time there. We used a simple intuitive presentation for the tree-structured assessment feedbacks to achieve the goal of mostly "just one click away to everything", including adding or selecting context-sensitive pool of feedback comments. This system also lays the foundation to having it further developed into an AI-based assessment system in future, where assessment marks and feedbacks can perhaps be predicted to a good extent too.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Z Jiang and J Huang conducted the research and paper writing together and interactively while Z Jiang did most of program implementation and data collection, and all had approved the final version.

References

[1] A. M. Nortvig, A. K. Petersen, and S. H. Balle, "A literature review of the factors influencing e- learning and blended learning in relation to learning outcome, student satisfaction and engagement," *Electronic Journal of e-Learning*, vol. 16, no. 1, pp. 46-55, 2018.

[2] S. S. Noesgaard and R. Ørngreen, "The effectiveness of e-learning: An explorative and integrative review of the definitions, methodologies and factors that promote e-learning effectiveness," *Electronic Journal of e-Learning*, vol. 13, no. 4, pp. 278-290, 2015.

[3] S. M. Imran and B. A. Malik, "Evaluation of e-learning web-portals," *Journal of Library & Information Technology*, vol. 37, no. 3, pp. 205-212, 2017.

[4] B. Fetaji and M. Fetaji, "E-learning indicators: A multi-dimensional model for planning and evaluating e-learning software solutions," *Electronic Journal of e-Learning*, vol. 7, no. 2, pp. 1- 28, 2009.

[5] S. Scott, C. F. Webber, J. L. Lupart, and D. E. Scott, "Fair and equitable assessment practices for all students," *Assessment in Education: Principles, Policy & Practice*, vol. 21, no. 1, pp. 52-70, 2014.

[6] Z. Jiang and J. Huang, "Bias reduced designation of inhomogeneous assessors on repetitive tasks in large numbers," *International Journal of e-Education, e-Business, e-Management and e-Learning*, vol. 2, no. 3, pp. 176-182, 2012.

[7] K. Willey and A. Gardner, "Improving the standard and consistency of multi-tutor grading in large classes," *ATN Assessment Conference*, Sydney, 18-19 November, 2010.

[8] W. Admiraal, B. Huisman, and P. Pilli, "Assessment in massive open online courses," *Journal of e-Learning*, vol. 13, no. 4, pp. 207-216, 2015.

[9] Z. Jiang and J. Huang, "A fast and effective design and implementation of online programming drills," in *Proc. International Conference on Frontiers in Education: Computer Science and Computer Engineering*, 2012, pp. 314-320.

[10] T. Barker, "An automated individual feedback and marking system: An empirical study," *Electronic Journal of e-Learning*, vol. 9, no. 1, pp. 1-14, 2011.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

Zhuhan Jiang received the B.Sc. from Zhejiang University in 1982, and the Ph.D. from the Victoria University of Manchester, Institute of Science and Technology, UK, in 1987. He is currently affiliated with University of Western Sydney, in the School of Computing, Engineering and Mathematics. His relevant research interests include mathematical modeling and algorithms, efficient e-learning tools and assessment technologies, web based security and applications, as well as image and video processing.

Jiansheng Huang received the B.E. and M.E. from Hefei University of Technology in 1982 and 1984 respectively, the MSc from the University of New South Wales in 1997 and the PhD from the National University of Singapore in 1999. Currently he is working with the School of Computing, Engineering and Mathematics. His relevant research interests include information systems and security, power system operation and protection.