Micro-Accreditation for Matching Employer e-Hire Needs

Zachary Zaccagni
Computer Science and Engineering
University of North Texas
Denton, United States of America
zacharyzaccagni@my.unt.edu

Aditya Paul
Computer Science and Engineering
University of North Texas
Denton, United States of America
aditya.c.paul@gmail.com

Ram Dantu
Computer Science and Engineering
University of North Texas
Denton, United States of America
ram.dantu@unt.edu

***Abstract*—This paper presents a novel way to help match employers' knowledge requirements with students' knowledge earned using blockchain's smart contracts to assure credentials and track student records. This decentralized approach proposes using the micro-accreditation of topics from the CAE framework to courses and associated tasks, while introducing a revolutionary idea of a blockchain-based peer-reviewed rigor score assignment. Our work and result metrics were completed in Ethereum and connected test networks. We concluded this new approach is mostly efficient and scalable depending on the network load, with faster transaction times when the miners are properly incentivized. Future work will include further fine-tuning of the transaction algorithms to improve time, as well as an investigation into a better consensus model for peer review and rigor determination.**

Keywords—blockchain, education, micro-accreditation, peer review, knowledge unit, topics, course, Ethereum, CAE, eHire, assured credentials, student

I. Introduction

We present an innovative approach to track student records through a school and help match employers' requirements to students' knowledge. The student's records are stored in an Ethereum-based blockchain network, along with a decentralized course storage system. Each course has a required set of Knowledge Units (KU) and micro-accredited topics mapped to assignments. The KU topics are defined by the Center of Academic Excellence (CAE) [13]. Each assignment element will be tagged with both a KU topic and a rigor score, determined by a peer-review process. Student assignments could also be evaluated on the blockchain via smart contracts. This would streamline the grading process and shorten the grading time, while reducing human error and bias. A customized algorithm is then used to calculate a proficiency score for each student for each KU topic, using student grades as well as the course rigor score within the parameters of the CAE framework. Rigor is defined as the level of difficulty some problem is in regards to the expected level of understanding of the problem-solver. If it is more difficult, it has greater rigor. We will focus only on the CAE framework, and in the future, this should be expanded to include other educational frameworks.

The process of a course's peer-review is handled by smart contracts on this blockchain. Each course assignment (e.g. homework, project, quiz, etc.) is segmented into elemental parts, and each elements' rigor score determined by peer-review. Once a consensus is reached, the course is made available to students. This technology will expedite a course's peer-review, helping set an overall rigor for the course and its assignments. This also matches the course rigor of one institute with another, helping to establish an immutable record of academic credentials and what was specifically learned by each student.

This student blockchain can then be accessed by prospective employers who can select employees based on the scores in these KU topics. Using the skills required for a possible opening, employers can specify the importance of a topic and match a student to an opening based on a compound score calculated from a student's proficiency in the topics required for the opening. Using this, it becomes much easier for employers to comb through student records and verify not only the authenticity of courses, but the rigor of each course and a student's success in a specific topic. Blockchain is an effective strategy to solve this problem in that each grade assigned can be modeled as a series of transactions and requires a consensus among both students and professors. The rules and regulations that are put into place in this system are outlined by the FERPA laws, and in turn, can be easily implemented in a smart contract. The auditability of the blockchain makes it easier for employers to review student transcripts, and in turn, make more educated decisions on hires, not only because of the immutability of the blockchain record, but also because of the rigor of the course. Results demonstrate that the system was successful in increasing the accuracy of hires through simulated data sets, and that it is efficient, as well as scalable.

II. Technology Background (Blockchain)

A blockchain consists of a distributed ledger similar to a linked list, in which all links are secured with hashes [1]. Each of the nodes in the linked list, known as blocks, are snapshots of the state of the network at the time at which the block was published. A copy of the blockchain is stored on each of the nodes on the network, and each time a block is published, it must be verified by the rest of the nodes on the network before it is added to the blockchain [3].

This project utilizes the Ethereum platform, a platform that uses a Proof of Work verification method for blocks. Whenever a transaction is to be completed, it is sent with a specific amount of ether as well, in order to compensate the miner, or the node executing the transaction, appropriately for the amount of computational power expended for executing the transaction.

978-1-7281-4693-5/19/$31.00 ©2019 IEEE
DOI 10.1109/Blockchain.2019.00054

The amount of computational power used is determined by calculating the gas consumed by the transaction. The gas calculation algorithm is outlined in the Ethereum yellow paper [2]. Because the amount of gas consumed is determined by the number of opcodes as well as the amount of data being handled by a function, it is safe to say that the amount of gas consumed by a transaction is a measure if its efficiency.

A blockchain has three extremely important properties: immutability, auditability, verifiability. It is computationally hard to modify a block after a set number of blocks, auditable in that it represents a transparently traceable sequence of changes to the network, and verifiable in that all blocks must be verified before being appended to the blockchain. This ensures the validity of the block once appended.

Ethereum transaction rules are codified in a "smart contract", and can be accessed through a web3.js API, through which users can send transactions. A wallet to contain ether can be accessed through applications like the Mist Browser and Metamask and is also available for the user to access. Truffle, a development framework for Ethereum distributed applications, also supplies a wallet provider through its HD wallet. Ethereum provides three globally-accessible networks: Ropsten, and Rinkeby. Truffle also provides a network emulator, known as Ganache, for testing purposes.

III. Problem Description and Requirements

There are several problems with existing hiring methods and with students transferring to new schools.

A. *Some Common Problems:*

- The current hiring process is a lot of work, and requires many components, such as recruitment processes, rounds interviews, and paper cuts. Much of this is outsourced to expensive trusted third parties such as recruiters and human resources staff. Much of the complexity and confusion in finding the right candidate-match arises from these hiring parties being too removed from the actual work involved to fully understand the needs of the hiring manager. Currently, the hiring process lacks regular, organized, and agreed-upon definitions of knowledge and skills, encompassing both academic and the workforce.
- It is difficult to locate information on new grads in order to deduce their success in the workplace. Most info about them is contained in transcripts, which are difficult to interpret correctly from only course grades and school. This information is not normalized from one institution to another, and rarely reflects what specific knowledge was learned. Comparing courses between schools is very challenging. Accuracy is also an issue, since the skills required for a career is not matched to the courses accurately, as different skills are taught in different courses at different schools.
- It is also difficult to discern if any bias existed during grading. A student's grade for an assignment or overall course could have been influenced by bias (i.e. how good was the relationship between teacher and student), moods of the grader, general oversight or carelessness.
- Making mistakes in hiring is both risky and expensive, sometimes costing up to $60K [4][5].

Due to the above-listed reasons, it is of tantamount importance to create a system that efficiently and accurately reduces the risk involved in the hiring process through an effective matching algorithm, normalizes grades as well as the rigor of an institution, and eliminates middlemen involved in the process. Since the rigor is peer-reviewed, it reflects the most accurate standard in the score determination. In addition, combining rigor and KU topics creates a better solution for comparing courses from different academic institutions for school transfers. This helps eliminate the task of measuring what the student explicitly learned to match them successfully with their next course at their new school or next job. By connecting these topics both generally (to a course) and specifically (to each part of an assignment), we can compare courses' contents easily by examining both of their clearly defined Knowledge Units and topics. Then, to further weigh the value of similar courses, you can compare their rigor scores and determine if one course has a greater value and covers more material than the other.

B. *Requirements Considered:*

Blockchain is able to address many of the aforementioned problems, as its auditability, immutability, and verifiability serves to eliminate the need for recruiters, auditors, and other hiring staff, while the consensus-based nature of a blockchain based system enables us to normalize courses, and the system itself would make the hiring process cheaper due to the lack of third parties, as well as the reduced risk due to the streamlined matching algorithm.

1) *A solution for the e-Hiring match component must:*

- Eliminate the need for a third party. The distributed decentralized network can account for this, as the workload required by a third party is picked up by all of the nodes on the network. With verifiability, all blocks can be assumed to be valid, and all the transcript materials included are valid as well.
- Normalizes courses. The system does this through the properties of peer review and consensus, as it enables all nodes to come to an agreement as to the rigor of the

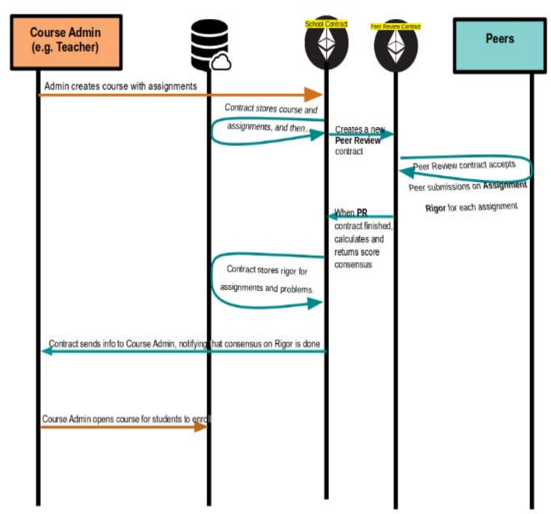

Fig. 1 Work Sequence for a Peer Review – This sequence was used as the base for peer review work

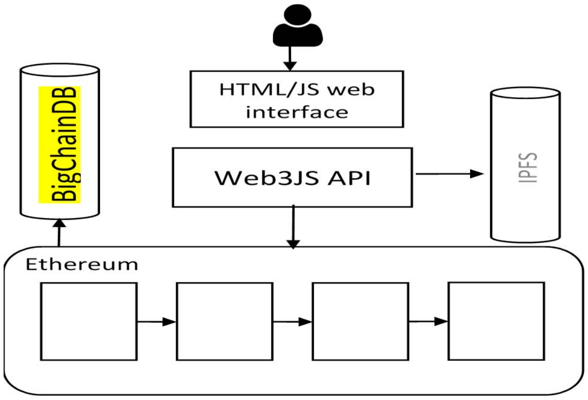

Fig. 2 Architecture of the System – A web-based dApp using web3.js to connect to the contract on the Ethereum network. The contract connects to the BigChainDB for adding and retrieving data that is stored off-chain. This architecture was not fully implemented and will need something similar in future work

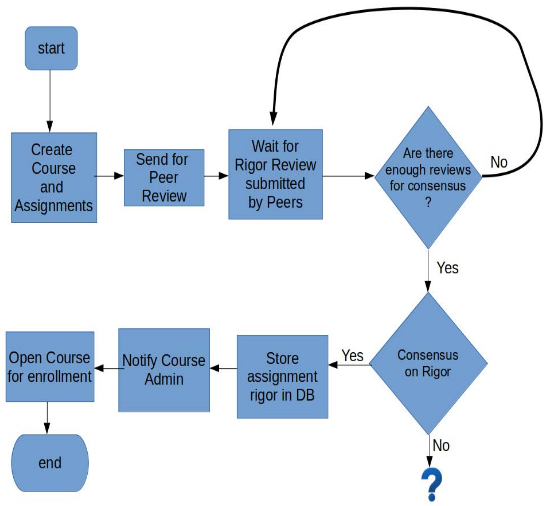

Fig.3 Work Flow for Peer Review – The work flow used in our simulations for peer reviewing. In future work, we may need a third party contract to handle times it doesn't come to consensus, as well as a new consensus model for rigor.

course itself. This will be factored into our grade calculations and skill calculations. The candidate's ability to succeed in a career is determined by the individual skills taught in the course, also determined by the peer review process.

- The system will reduce costs as there is no need to pay third parties and will also reduce the number of interviews and amount of effort required to hire a candidate, as by matching specific skills, the employer already recognizes that the candidate has demonstrated proficiency in the areas required for the career.

2) A procedure of the course creation and a peer-review component are as follows:

- Each course has a teacher and has multiple assignments. Each assignment can have multiple parts, which include connected KU topics, and a rigor score. The rigor of each part is peer-reviewed and determined through consensus. This is currently being done manually.
- Once complete, the course is made available for student enrollment as shown (Fig. 1).
- Each participant will have their own node on the blockchain. The course admin and peers will have different web interfaces for interacting with the smart contract, and different access to the underlying data.

C. *Architecture:*

The architecture for the system involves a traditional web-based dApp architecture, using web3.js to interface with an HTML webpage connected to Ethereum blockchain in the back end (Fig. 2). web3.js is also utilized to interface with IPFS, a decentralized filesystem, in order to store assignment files, as storing files on the blockchain directly is too expensive. One difference between the proposed design and the final design of the dApp is that BigChainDB, a blockchain-based decentralized database, was not utilized, in that there was not sufficient time to implement it. However, we address this later as a possible improvement.

IV. Existing Work

- Learning Machine [6]. Using Blockcerts, their technologies are currently used by the Massachusetts Institute of Technology in order to verify MIT degrees and transcripts. However, this is different from our application, as it does not have an effect on hiring specifically.
- Chronobank [7] operates similarly, in that it streamlines background checks, and is more structured towards hiring by loading transcripts on a blockchain. However, this application is not oriented towards matching students to specific jobs.
- In "Connecting decentralized learning records: a blockchain based learning analytics platform" [9], the authors discuss storing normalized student record data on the blockchain, regardless of origin. However, since the "Secure Box" transforms all the data from each LMS (Learning Management System) to uniform data records for each student, the cost may not be scalable.
- "Hierarchical interactions between Ethereum smart contracts across Testnets" [10] presents the idea of a "Custodian Contract" that spurs and manages new contracts. This is similar to the role of the Course contract in this project. Using the Course contract as a custodian that generates new smart contracts between students and assignments is a fitting and seemingly correct approach.
- "Data-driven Generation of Rubric Criteria from an Educational Programming Environment" [11] relates to auto-grading systems, specifically for programming. The paper discusses non-blockchain related solutions.

V. Detailed Design

The system participants are professors, students, and employers. The purpose of the inclusion of the students in the system is to for them to complete coursework so that professors can assign grades to them. Professors are tasked with assigning grades and peer-reviewing courses, while employers are tasked with posting job opportunities. Once students graduate, they can apply for career opportunities, at which point all job applications are ranked by a compounded skill score calculated from scores and rigor.

A. *The application follows the following methodology for course peer-review (Fig 3):*
- It is assumed that the course exists to enroll in and is currently set up with assignments and associated parts.
- The course administrator starts the process in the workflow, and the smart contract ends it once peer review consensus on rigor has been determined.
- An admin calls a transaction for Peer Review. This generates a new PeerReview contract between the admin and the peers. The admin furnishes the completed course, and the peers will provide their analysis of rigor for each assignment. Then, a consensus is made on the rigor.
- Once the peer review is complete, the course is open for either admin review or for student enrollment. The transaction with the course data is added to the blockchain, while any files for the course are submitted to IPFS, and references are stored on the blockchain.

B. *The application follows the following methodology for an enrolled student:*
- A student enrolls in the course, completing assignments, with their grades appended to a list of grades.
- Once a student has completed their coursework, they can apply for course credit. If the instructor deems the student's grades sufficient, the professor approves the credit application, with the student's associated skill scores are calculated and are stored. A new course block is appended to the student's transcript blockchain.
- Once a student has enough credits, they apply for graduation, approved by at least 3 professors to graduate.
- A student then applies for jobs, and the match score is calculated by the algorithm mentioned above.

VI. Experimental Methodology and Steps to Goal

System was tested for operational success and efficiency.

A. *Tests for simulating hiring data are conducted as follows:*

A1. DATA SETS AND SIMULATION: We utilized data sets and simulating hiring patterns. The sets were publicly accessible grade distributions from three different institutions (UNT, UT Austin, UC Berkeley), hiring statistics for one company (Amazon) found through LinkedIn statistics. Rigor for this test was taken from CSRankings.org's CS rankings where s is the given rigor rating, and r is the ranking given.

$$s=50+50100-r100 \quad (3)$$

The success of this test was determined by the variance from the real statistics for last year. This can be tested on Ganache, the Truffle framework's built-in Ethereum emulator.

A2. TESTING FOR EFFICIENCY: The efficiency is determined by the amount of gas consumed by the algorithm. The amount of gas consumed by a function is proportional to its efficiency and the amount of computational work required. We measure the gas consumed by each function on the Ropsten network, in order to provide an appropriate testbed for the application, simulating the network load at the time of deployment.

A3. TESTING FOR SCALABILITY: The test for scalability can be conducted by measuring the increase in gas price as the number of participants increases as well. This shows the decrease of efficiency of the program as the number of users increases and will also demonstrate exactly how the system responds to higher loads, and for which specific functions. This is useful, in that it may reveal potential optimizations for future reference. Ganache was used for this test, as it is infeasible to test using many accounts on Ropsten.

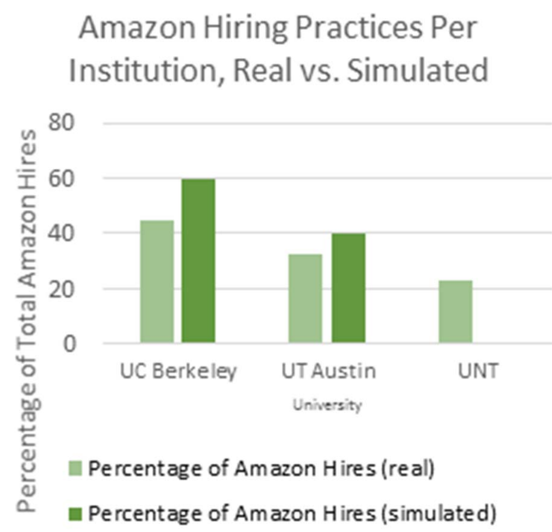

Fig.4 Hiring Practices comparison between real and simulated data. Though UC Berkeley and UT Austin simulated results are similar to data set, UNT Simulated results were zero due to the algorithm's weight on rigor. The algorithm will need to be adjusted

B *The tests for simulating a course peer review are conducted like so:*

B1 SETTING UP: Two courses were manually created with 3-4 assignments and 5 parts to each assignment. This data was used to load these courses and assignments into the contract when starting the dApp, establishing a base data set to start with. Since assignments were available, the Admin/Teacher could perform actions like closing the course, calling for peer review, viewing the course. At this point, all assignments are exact answers and had no rigor assigned. Metamask plugin was used to help simulate different peers. Metamask managed the "handshake" agreement for transactions of the contract.

B2 SIMULATION OF PEERS AND STUDENTS: Using Metamask to create peers, students, and a course admin, and after the test data was loaded, the dApp was run as the admin, calling for peer review, while recording the performance metrics for transaction time and gas cost. We compared multiple peer reviews both on a stand-alone simulated network, and the Ropsten Test Network for transaction times. It was also used to simulate students enrolling and completing assignments.

VII. Results and Discussion

A. *Metrics used*

The metrics utilized in our analysis of hiring include those reserved for efficiency, scalability, and operational success. The metrics utilized in our analysis of peer-review include gas cost, transaction times, and the effect of the gas price (incentive for the miners to do the work) on transaction time. These metrics and their testing methodologies are outlined above.

- In order to measure success, a representative dataset was gathered and simulated from Amazon's LinkedIn hiring data utilizing data sets from UC Berkeley, UT Austin, and UNT, with rigor rankings from CSRankings.org.
- To measure efficiency, the amount of gas consumed by each transaction run on the network was measured, as gas is a measure of the efficiency of a function.
- To measure scalability, the amount of gas consumed as the number of entities accessing the network was measured, in this case, the students. We track this using the Ganache emulator and use this to measure how different transactions respond to a high network load.

B. *Data, Graphs, and Discussion*

1) Simulated and real hiring practices of Amazon based on the data sets.

This graph (Fig 4) shows that the current algorithm is somewhat capable of distinguishing hiring patterns, and that the new graduates' hiring rates from UT Austin and UC Berkeley are like that of the given data set. However, it is observed that the data for UNT in the simulated run is completely zeroed out. The error for this measurement can be attributed to the algorithms extra weight on the rigor of an academic program. consequently, the current data set is not capable of properly representing the state of real-world hiring, and because of this, is insufficient to properly deduce the success of the hiring algorithm

2) The efficiency of running transactions, the second graph (Fig. 5), shows the growth of transaction gas cost as the number of entities on the network increases. The only transaction whose gas cost grows is the Approve-Credit transaction. This is due to the de-allocation of memory containing the student data structure. We can say the system is somewhat scalable.

3) The third graph (Fig 6) shows transaction time for peer-review. By default, when you call a transaction, the gas price is 1. The gas price (GP) is the incentive for a miner to work, the higher the price is, the faster the transaction will happen. Changing GP from 1 to 50 significantly affected transaction times, as the higher GP incentivized it further. Fine-tuning is needed to find the optimal price.

4) Transaction times for the Assignment Grading. The fourth graph (Fig. 7) shows the transaction times for the Assignment Grading. Similar to the Peer Review times, increasing the GP to 50 significantly lowered the times. For example, at GP=1, addAssignment function times started at 1.1 million ms (~18 minutes) to a much more reasonable 102775 ms (~1.5 minutes) at a GP of 50. More research is needed to find a good balance for speed and gas price.

Once a course and its assignments are complete, peer-reviewed with rigor calculated, and opened to students, outside entities will be able to compare courses between different institutions (academic or other). This establishes a fuller picture to gain better insight into a course, the course parts, what is to be learned, and especially how well it will be learned (i.e. the established rigor of the material taught). As students take these courses, potential employers will be able to use the system to help match what they are needing in a position with students

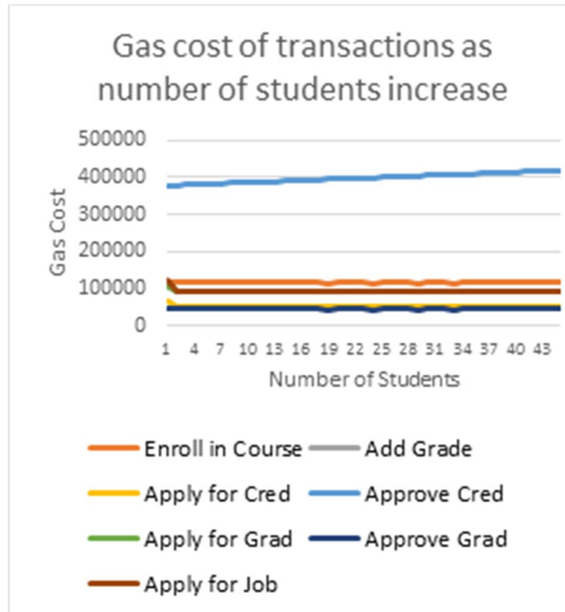

Fig. 5 Gas Cost versus the Increase of Students on the Network - As the network load increased, most transactions at this level stayed approximately the same. The Approve Cred transaction saw an increase in gas cost due to increasing needs to de-allocate

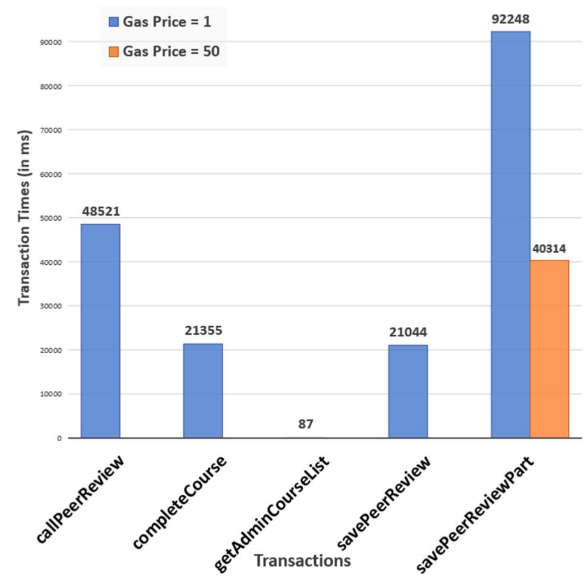

Fig. 6 Peer Review Transaction Cost – The transaction times for calling different transactions (i.e. functions) during a Peer Review. This peer review model was simplified and does not yet calculate rigor, though requires manual entry from each peer for each part of an assignment. The algorithm will change in the future and will better reflect actual peer reviews. In this chart, the Gas Price indicates the incentive to miners to do the work. What was learned in this is that increasing the Gas Price increases the incentive, and subsequently transacts faster on the network.

who have proven successful in gaining that specific knowledge. This may prove useful in finding and filling the many positions left unfilled due to a perceived skills-gap. Tod Beardsley stated recently, "If you're only looking at … the top ten universities in the U.S. then yes, there are hardly any candidates." [12] Our approach provides a possible way to open widely the pool of candidates to fill those positions by assuring their knowledge credentials, so employers can better understand what a student has accurately learned, regardless of school reputation

VIII. Limitations

Limitations for this project in testing include the testing data sets utilized. As stated earlier, they are not a representative sample of the current state of hiring. In order to address this, future tests must include data sets that are taken from the company itself in order to properly model hiring processes.

An additional concern is that the system needs to protect data from non-participant/non-permissioned eyes. As a blockchain system is fully transparent, all parties, malicious or otherwise, have access to the assignments and answers.

Finding a way to accommodate students with special needs is especially challenging using a blockchain. Smart contracts don't lend themselves to multiple tracks of required coursework.

Lastly, there are still logistics and scalability concerns. Can a blockchain handle the real-world number of peers in a feasible way (cost and time)? It is possible that research is needed on "chunking" a blockchain consensus for massive-node systems.

IX. Conclusions and Future Work

- Efficient. most transactions with the exception of a few, fell beneath a safe threshold. The exceptions can be fixed through implementation on a different framework and it is possible for further optimize of the system.
- Scalable. Again, all transactions had O(1) complexity when varied with network load, except for credit. The scalability of the entire system suggests needing a different blockchain network, using fewer resources with faster transactions. Incentivizing faster transaction times still leaves us with an undesirably slow process. The Etherium network, currently, is not suitable for our future work needs. Though Vitalik Buterin (Etherium) is working to release a new version [8], Ethereum 2.0 is not available, so we will move to another open source blockchain with a new model for consensus.
- Outside the above scalability issues, success has been demonstrated through our earlier investigation of hiring rates from UT Austin and UC Berkeley, as the relative hiring ratios for these two institutions was quite close. However, the algorithm places quite a bit of weight on rigor, which requires further investigation.
- Further investigation is needed to find the relationship between the rigor of an institution and hiring rates. For testing, future work includes testing with cleaner data sets in order to fully replicate real-world hiring scenarios.

X. References

[1] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system", [Online]. Available: https://bitcoin.org/bitcoin.pdf.

[2] G. Wood, "Ethereum: a secure decentralised generalised transaction ledger", Ethereum Project Yellow Paper, vol. 151, pp. 1-32, 2014.

[3] Dejan Vujičić, Dijana Jagodić, Siniša Ranđić, "Blockchain Technology, Bitcoin, and Ethereum: A Brief Overview" 17th International Symposium INFOTEH-JAHORINA, 21-23 March 2018

[4] SHRM, "Average Cost-per-Hire for Companies Is $4,129, SHRM Survey Finds," *SHRM*, 19-May-2017. [Online]. Available: https://www.shrm.org/about-shrm/press-room/press-releases/pages/human-capital-benchmarking-report.aspx. [Accessed: 11-Dec-2018].

[5] L. Frye, "The Cost of a Bad Hire Can Be Astronomical," *SHRM*, 11-Apr-2018. [Online]. Available: https://www.shrm.org/resourcesandtools/hr-topics/employee-relations/pages/cost-of-bad-hires.aspx. [Accessed: 11-Dec-2018].

[6] "Learning Machine," *Learning Machine*. [Online]. Available: https://www.learningmachine.com/. [Accessed: 11-Dec-2018].

[7] "Hiring, Rebuilt.," *ChronoBank.io*. [Online]. Available:https://www.chronobank.io/. [Accessed: 11-Dec-2018].

[8] "Buterin: Proof-of-Stake Ethereum Is Really No Longer So Far Away", *Finance Magnates* [Online] Available: https://www.financemagnates.com/cryptocurrency/news/buterin-proof-of-stake-ethereum-is-really-no-longer-so-far-away/

[9] "Connecting decentralized learning records: a blockchain based learning analytics platform" (Patrick Ocheja,Brendan Flanagan, Hiroaki Ogata) 2018-03-07 ACM

[10] "Hierarchical interactions between Ethereum smart contracts across Testnets" (Yao-Chieh Hu, Ting-Ting Lee,Dimitris Chatzopoulos,Pan Hui) CryBlock'18 Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems 2018-06-15 ACM

[11] "Data-driven Generation of Rubric Criteria from an Educational Programming Environment" (Nicholas Diana, Michael Eagle, John Stamper, Shuchi Grover, Marie Bienkowski,Satabdi Basu) 2018. International Conference on Learning Analytics & Knowledge, March 7–9, 2018

[12] "The 11 biggest issues IT faces today", *CIO Magazine* [Online] Available: https://www.cio.com/article/3245772/the-12-biggest-issues-it-faces-today.html

[13] "CAE-CD Knowledge Units", IAD.gov [Online] Available: https://www.iad.gov/NIETP/CAERequirement

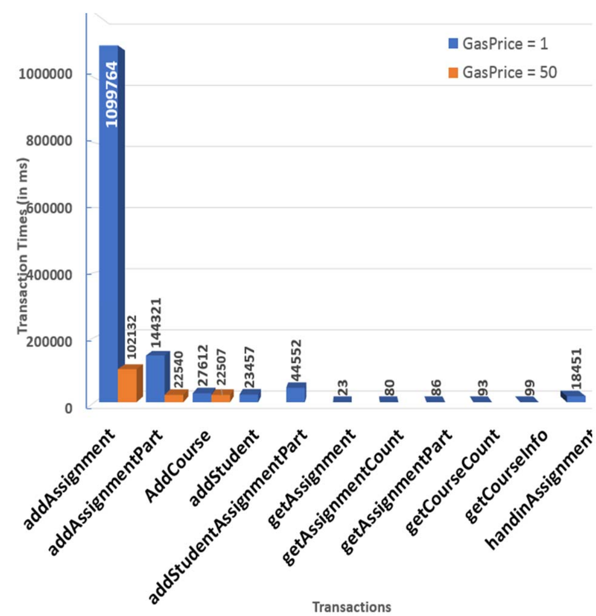

Fig. 7 Transaction Times for Student Transactions – These are transactions to and for Students and their respective cost times in milliseconds to transact. Similar to (Fig 6) for Gas Prices, when the Price was set higher, the miners were much more eager to do the work, and thus times were significantly faster. it does the transaction in a reasonable time. More work is needed on transaction times overall.