Computer Science and Computer Engineering

Verification of a labor market domain using an academic crowdsourcing system

Esteban J. Azofeifa, Galina M. Novikova

Department of Information Technologies
Peoples' Friendship University of Russia (RUDN University)
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

Students desiring to become a valuable good in the labor market are willing to pay a considerable monetary cost to obtain knowledge about their prospective job opportunities, nowadays with a diminishing interest in the obtainment of a diploma. Considering the behavior of the labor market as a domain theory under uncertainty, it is straightforward to expect the presence of contradictions, in the form of salaries unable to be classified due to high inconsistency and variation. We provide an algorithm to verify a labor market domain theory based on a crowdsourcing academic system, in which feedback about possible contradictions is generated as a result of consultations with experts inside of the market and clustered into different contexts. We found that the verification process can be repeated iteratively as long as the students' overall tuition is equal or greater than a quantity partially defined by the number of different profiles of the students.

Key words and phrases: labor, market, crowdsourcing, curriculum, multiagent, verification, contradiction

1. Introduction

Assessment of the job market is an important practice in the field of academic curriculum development and in the election of a career path for prospective professionals before becoming part of the educational system [1]. Nowadays, however, students are more than ever concerned about correctly selecting a professional path in which they won't lose time in learning skills that they are not likely going to apply, even if it means the possibility of not obtaining a diploma [2].

As one of the popular topics in today's sharing economy, crowdsourcing [3], [4] has become a feasible tool for people to obtain knowledge at reasonable

costs from a set of experts in a field. Such experts, with varying levels of expertise, can be found in platforms like Amazon's Mechanical Turk [5]. In an academic environment, the roles of teacher and student (expert and requester) can also be integrated in a crowdsourcing system, and if a single agent is allowed to perform both roles interchangeably, then the agent can be further benefitted by an academic platform such as VUZ [6].

Entering an environment like VUZ requires a certain level of previous knowledge of the labor market [7], either compiled by the student or requested to experts. A platform like this, however, can also aid in gathering feedback [8] from the job market in order to actualize itself and optimize the benefits for the requesters and workers. Such feedback provides a way to verify [9] the previous knowledge about the market and points to possible improvements to the system.

In particular, the verification of a domain theory [10], in this case the labor market, allows to tackle all the possible issues that could cause a failure of a forecast, which in this work we take as the projected salary of a new professional in its field of expertise in the job market [11].

We structure this paper as follows: the basic concepts are detailed in the second section, before we explain the concept of labor market in the next part. The fourth section deepens into the concepts of verification and contradiction, while the fifth section focuses on the crowdsourcing system's ability to generate feedback and sustain itself in time.

1.1.   Labor market

The labor market can be thought of as an open multiagent system (MAS) with limited resources [12] where agents seek to maximize their revenues by performing a balance between exploration [13] and working in exchange for a retribution. In the simplest case, requesters can be thought of as being separate from the workers [14] in the sense that their interrelations are not necessary evident to the workers.

1.2.   Academic crowdsourcing

Agents can enter the open MAS freely, but often the process is channeled by an academic institution. In such cases, students agree to pay the institution more depending on its level of prestige in certain field of study [15]. Nowadays, however, prospective professionals are growingly looking for ways to land in the desired position in the market in a simplified way, often bypassing the obtainment of a diploma [16]. Personalized courses designed by a set of experts conform the core functioning of the academic platform VUZ and they correlate and complement the skills already mastered by the student, so that all excessive costs are avoided.

1.3.   Verification

We take into account a formal system $G = <T, F, X, I>$ [17], where T is a finite set of basic characters known as the alphabet of the theory, F corresponds to a collection of formulas, X is referred to as a set of axioms and I is a set of inference rules. Then, the objective of the verification process is

to find all the discrepancies [18] among the elements of the formal system, starting with the set $T' \subseteq T$ that contains the object types, the properties and values of the types and the logical and functional relations between them.

1.4. Contradictions

Contradictions [19] arise in the formal system when the process of recognition of objects in the real environment [20] results in discrepancies with the entities and relations in the respective domain theory under which the recognition is performed. In particular, contradictions are reflected by the presence of 'impossible' elements of the alphabet [21] in the real environment as well as the absence of expected elements. On the other hand, they point to conflicts in the axioms [22] or contrasting states of the system caused by defects in the inference rules.

1.5. Retribution

A key point in the VUZ framework is the treatment of students as experts, along with the teachers [23]. This means that every level of expertise from the agents of the system is taken into account, and is utilized in order to teach other students with a lower level of knowledge in a determined field of study. As such, a student is required to pay a certain cost to obtain a crowdsourced assessment of his/her skills, as well as a personalized curriculum [24] for the obtainment of the desired new skills. The cost, however, is mitigated by giving the student the opportunity to share his/her knowledge with others in exchange for a retribution [25], [26]. This is an internal benefit from the academic crowdsourcing system, obtained prior to entering the labor market. Once in the market, the new professionals will obtain a real salary which will become the source of feedback for the system.

1.6. Profiles

The profiles of the agents in the crowdsourced system consist in a set of skills that gradually grows while the agents develops themselves as workers in the system [27]. The description of a skill is maximally general in order for the skill to be used as a variable in different kinds of tasks. In particular, the agent performs a task by adopting a certain role where a set of key skills are necessary and actively used.

On the other hand, roles and task performing are valid and can be evaluated and priced only when observed from the point of view of a certain context [28]. This implies that the same skills can have a different monetary value in different situations: contexts induce different partitions of the set of skills [29]. Like roles, contexts are described by means of a profile, albeit not of skills but of locators: they function as dimensions and can take different values ordered in a scale, giving the context a defined structure to answer to questions involving the place, time and other situational aspects of task performing.

1.7.   Feedback

Gaining direct knowledge from the labor market is crucial in the development and functioning of the academic crowdsourcing system. Once a student leaves the system and joins the job market, he/she is nevertheless asked to communicate his/her salary and to provide a picture of the context in which he/she works, which can include a set of indicators such as weighted coefficients associated to quantitative and qualitative attributes [30]. Such picture is then introduced into the system and helps determine possible contradictions.

2.   Model description

Let $(M, <)$ be a set of salaries ordered in ascending order, R a set of roles, C a set of contexts and $rt : R \times C \to M' \subseteq M$ a function assigning a collection of retributions to agent performing a certain role. Then the system is in presence of a contradiction when $\max(M') - \min(M') > \xi$, where ξ is an arbitrarily defined constant that we call the classification threshold [31].

It is assumed that further refinement of the set of skills and locators for the corresponding role and context will lead to a reduction in the constant ξ, ultimately pointing to the skills and location of a unique agent. This, however, is implausible in real environments, and for this reason we introduced the arbitrary constant into the model.

Such contradictions can be traced back to a single source, provided sufficient resources for analysis.

2.1.   Incomplete role profile, correct context profile

The set of skills or locators associated to a role or context corresponds to a subset of the total collection of skills or locators known to the system, which in turn makes possible the overall classification of roles and contexts. Contradictions can be mitigated and the classification threshold attained by adding skills to a role known or suspected to be incomplete. This additional dimension can be partitioned into different values according to its scale, and a reduced role can be obtained by selecting the corresponding element from the partition.

2.2.   Complete role profile, incorrect context profile

The same procedure can be applied in the case of incorrectly constructed context profiles. In particular, dimensions can be added to a locator in the context in order to refine the description of the time, place and other situational aspects.

2.3.   Complete role profile, correct context profile

Contradictions can also arise when both the role profile and the context profile are guaranteed to be correctly constructed.

Assuming that every possible skill and locator is known to the system and available to be used, this scenario points to issues in the definition of the

parameters of the system. In other words, the assessment of the agent's skill levels and/or the locator value in the scale is not accurate and needs to be revised. On the other hand, there is the possibility that not all the skills or locators are known to the system, case in which it is necessary to engage in exploration of the entirety of the state space [32] in order to add the missing elements to the system.

3. Algorithmic implementation

3.1. A verification algorithm

We now propose an algorithm destined to fix the profiles after the detection of a contradiction. Let P be a set of (role and context) profiles and $d : P \times P \to [0, +\infty[$ a semimetric [33] among the profiles. On the other hand, let $S \subseteq E$ and $L \subseteq E$ be a set of skills and locators respectively, both belonging to a set of elements and yielding $P = \{(S', L')\}_{S' \subseteq S, L' \subseteq L}$. Then a set can be constructed such that each skill or locator will be labelled with the distance of its profile to a determined central profile.

$$D(p) = \{(e, d(p, q)) | e \in E, \ e \notin S', \ e \notin L' \forall p = (S', L') \land p \neq q\}. \quad (1)$$

Equation (1) allows to build an algorithm for gradually adding elements to profiles.

Algorithm 1 The addElements function

Input: a profile set P, a set E of profile elements, and a chosen $p \in P$
Output: a modified profile p
  $O_r \leftarrow D(p)$; a set of tuples of elements with distances
  $O_r^* \leftarrow order(O_r)$; the set O_r in ascending order
  $S' \leftarrow skills(p)$; the set of skills of profile p
  $L' \leftarrow locators(p)$; the set of locators of profile p
  for all $o \in O_r$ **do**
    if $o \notin S'$ **and** $o \notin L'$ **then**
      $p.add(o)$
      if checkVerify(p) = **true then**
        return p
      end if
    end if
  end for

The importance of the function `addElements` (Algorithm 1) is due to the property of the academic crowdsourcing system of continually receiving feedback from new professionals. In particular, each added element in the function corresponds to a new skill or locator which needs to be asked to experts in the labor market in order to refine the system. Their continual contribution will allow to perform the algorithm repeatedly, not forgetting to check for the validity of the profile on each iteration by means of the function

`checkVerify` (which yields $True$ if the desired classification threshold is attained).

3.2. An optimization algorithm

An optional but recommended continuation of `addElements` in Algorithm 1 corresponds to a function to clean (optimize) the resulting profiles. Due to the continual addition of elements to a specific profile in order to mitigate a contradiction, an excess of skills or locators is generated, which in turn calls for an optimization.

Algorithm 2 The cleanOptimize function

Input: a profile set P, a set E of profile elements, and a chosen $p \in P$
Output: a modified profile p

  $E' \leftarrow elements(p)$; the set of elements of profile p
  $C^* \leftarrow combinations(E')$; the set of combinations of elements of profile p
  $C \leftarrow combinations(E)$; the set of combinations of elements of all profiles
  for all $c \in C^*$ **do**
    if $c \notin C - C'$ **then**
      $k = 0$
      for all $e \in C$ **do**
        $p.remove(e)$
        if checkVerify(p) = **false then**
          $p.add(e)$
          $k = k + 1$
        end if
      end for
      if $k = |c|$ **then**
        set $c.isKey = $ **true**
      end if
    end if
  end for
  return p

Concretely, the proposed function `cleanOptimize` (Algorithm 2) takes all the skills and/or locators and subdivides them in groups different from the element sets of other profiles. Afterwards, the function checks by subtraction (by means of the function `checkVerify`) if the eliminated elements are necessary skills or locators, leaving only the key elements in the profile.

4. Discussion

4.1. A cost balance equation

We calculate the cost of performing the verification algorithm for a determined number of profiles in a time lapse, and compare it to the expected revenue

from students' tuition. Let $D_c = |S'| + |L'| + (|E| - (|S'| + |L'|)) * |P - 1|$ the cost of obtaining the set $D(p)$ and $O_c = \gamma(|D(p)| \log(|D(p)|))$ the cost of ordering the set

$$H_{\text{stud}} >= k * \beta * (D_c + O_c + 4 * |D(p)| + 2|E'| * 2|E'| - |E| * (2 + 2|c|)). \quad (2)$$

Then the verification algorithm and its subsequent optimization can be performed along k iterations as long as students' tuition (H_{stud}) is equal or greater than the expression in Equation (2), where γ corresponds to a cost constant for ordering and β is a defined constant for performing common operations.

4.2. Implementation outlook

This paper is incorporated into a body of work related to the processes of recognition and verification under uncertainty. We plan to integrate such concepts into an expanded notion of academic life cycle management, where crowdsourcing will be utilized not only to individually assist students in achieving a certain goal, but also coordinate them collectively so that a certain institution can attain its own objectives with their help.

We also study the implementation of means to store data and multimedia related to the academic platform, so that every bit of information needed by the students could be one day provided by the system without delay. In addition, we have also studied new ways to incorporate knowledge related to Big Data to correctly manage uncertainty from the recollection of the primary data.

On the other hand, we plan to extend the knowledge gained in the academic crowdsourcing environment to other fields such as the legal one. In particular, we hope to implement a crowdsourcing management system capable of registering normative events and linking them in a network of judicial decisions assessed by a set of experts with varying degrees of expertise.

5. Conclusion

In this work we presented an example of verification using a labor market as a domain theory. We outlined our verification approach based on the application of an academic crowdsourcing system, which serves as an introduction of new professionals into the open MAS of the job market. Students in the crowdsourcing system not only receive knowledge from experts, but also become experts according to their level of experience and knowledge. Once outside of the system, students become professionals and are expected to collaborate with the functioning of the system by providing feedback from the labor market.

In particular, information about the work salary is recollected from professionals in different contexts of the job market, and later processed in the system in order to identify contradictions in the understanding of the domain. Such contradictions arise in the form of salaries unable to be classified due to high inconsistency and variation. In our model, we characterized agent roles by means of skills and context by means of locators. We proposed a method to tackle salarial contradictions, which is based in the gradual addition and/or

removal of skills or locators in order to achieve a reduced salarial gap and attain the desired classification threshold.

Based on our proposed algorithms, we found that the verification process can be repeated iteratively as long as the students' overall tuition is equal or greater than a quantity defined by the sizes of the profile set and the element set, as well as the values of certain operational constants.

Acknowledgments

References

[1] M. G. Staton, "Improving Student Job Placement and Assessment Through the Use of Digital Marketing Certification Programs," *Marketing Education Review*, vol. 26, pp. 20–24, 2016. DOI: `10.1080/10528008.2015.1091665`.

[2] M. Scholten and N. Tieben, "Vocational qualification as safety-net? Education-to-work transitions of higher education dropouts in Germany," *Empirical Research in Vocational Education and Training*, vol. 9, 2017. DOI: `10.1186/s40461-017-0050-7`.

[3] I. Fedorenko and P. Berthon, "Beyond the expected benefits: unpacking value co-creation in crowdsourcing business models," *AMS Review*, vol. 7, pp. 183–194, 2017. DOI: `10.1007/s13162-017-0106-7`.

[4] J. Jiang, B. An, Y. Jiang, D. Lin, Z. Bu, J. Cao, and Z. Hao, "Understanding Crowdsourcing Systems from a Multiagent Perspective and Approach," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 13, 2018. DOI: `10.1145/3226028`.

[5] D. E. Difallah, M. Catasta, G. Demartini, P. G. Ipeirotis, and P. Cudr-Mauroux, "The dynamics of micro-task crowdsourcing," *Proceedings of the 24th International Conference on World Wide Web – WWW'15*, pp. 238–247, 2015. DOI: `10.1145/2736277.2741685`.

[6] E. J. Azofeifa and G. M. Novikova, "VUZ: a crowdsourced framework for scalable interdisciplinary curriculum design," *2018 IV International Conference on Information Technologies in Engineering Education (Inforino)*, pp. 188–193, 2018. DOI: `10.1109/inforino.2018.8581804`.

[7] R. Baker, E. Bettinger, B. Jacob, and I. Marinescu, "The Effect of Labor Market Information on Community College Students' Major Choice," *Economics of Education Review*, vol. 65, pp. 18–30, 2018. DOI: `10.1016/j.econedurev.2018.05.005`.

[8] N. Engbom, "Firm and Worker Dynamics in an Aging Labor Market," Federal Reserve Bank of Minneapolis, Working Papers 756, Apr. 2019. DOI: `10.21034/wp.756`.

[9]   P. Pollok, D. Lüttgens, and F. T. Piller, "How Firms Develop Capabilities for Crowdsourcing to Increase Open Innovation Performance: The Interplay between Organizational Roles and Knowledge Processes," *Journal of Product Innovation Management*, vol. 36, pp. 412–441, 2019. DOI: `10.1111/jpim.12485`.

[10]   A. Sutcliffe, *The Domain Theory: Patterns for Knowledge and Software Reuse*. Mahwah NJ: Lawrence Erlbaum Associates, 2002. DOI: `10.1201/b12455`.

[11]   H. S. Sugiarto, E. Lim, and N. L. Sim, "On analysing supply and demand in labor markets: Framework, model and system," in *Proceedings of the 6th IEEE International Conference on Data Science and Advanced Analytics*, Research Collection School of Information Systems, 2019. DOI: `10.1109/DSAA.2019.00066`.

[12]   A. Artikis, M. Sergot, J. Pitt, D. Busquets, and R. Riveret, "Specifying and Executing Open Multi-agent Systems," in *Social Coordination Frameworks for Social Technical Systems. Law, Governance and Technology Series*. Cham: Springer, 2016, vol. 30. DOI: `10.1007/978-3-319-33570-4_10`.

[13]   T. Wiedemann, C. Manss, and D. Shutin, "Multi-agent exploration of spatial dynamical processes under sparsity constraints," *Autonomous Agents and Multi-Agent Systems*, vol. 32, pp. 134–162, 2018. DOI: `10.1007/s10458-017-9375-7`.

[14]   C. Fieseler, E. Bucher, and C. P. Hoffmann, "Unfairness by Design? The Perceived Fairness of Digital Labor on Crowdworking Platforms," *Journal of Business Ethics*, vol. 156, pp. 987–1005, 2019. DOI: `10.1007/s10551-017-3607-2`.

[15]   R. Casidy and W. Wymer, "A taxonomy of prestige-seeking university students: strategic insights for higher education," *Journal of Strategic Marketing*, vol. 26, pp. 140–155, 2018. DOI: `10.1080/0965254X.2016.1182573`.

[16]   H. Buchs, E. Murphy, and M. Buchmann, "Landing a job, sinking a career? The trade-off between occupational downgrading and quick reemployment according to unemployed jobseekers' career stage and job prospects," *Research in Social Stratification and Mobility*, vol. 52, pp. 26–35, 2017. DOI: `10.1016/j.rssm.2017.10.001`.

[17]   R. M. Smullyan, *Theory of Formal Systems: Annals of Mathematics Studies*. Princeton University Press, 1961.

[18]   P. Cousot, "Verification by Abstract Interpretation," in *Verification: Theory and Practice. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2003, vol. 2772. DOI: `10.1007/978-3-540-39910-0_11`.

[19]   G. M. Novikova, E. J. Azofeifa, and V. R. Milov, "Towards the formalization of contradictions in a legal multiagent system," *Procedia Computer Science*, vol. 150, pp. 488–494, 2019. DOI: `10.1016/j.procs.2019.02.083`.

[20]   J. Eklundh and M. Björkman, "Recognition of Objects in the Real World from a Systems Perspective," *Kuenstliche Intelligenz*, vol. 19, pp. 12–17, 2005.

[21] S. Shapiro, "Incompleteness and Inconsistency," *Mind*, vol. 111, pp. 817–832, 2002. DOI: 10.1093/mind/111.444.817.

[22] P. Cousot, "Conflicts of Ontologies – Classification and Consensus-Based Methods for Resolving," in *Knowledge-Based Intelligent Information and Engineering Systems. KES 2006. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2006, vol. 4252. DOI: 10.1007/11893004_34.

[23] D. H. Wexler, "Integrating computer technology: Blurring the roles of teachers, students, and experts," in *Computing and Educational Studies: A Special Issue of Educational Studies*. Routledge, 2000, vol. 31, pp. 33–43.

[24] C. M. Chen, C. Y. Liu, and M. H. Chang, "Personalized curriculum sequencing utilizing modified item response theory for web-based instruction," *Expert Systems with Applications*, vol. 30, pp. 378–396, 2006. DOI: 10.1016/j.eswa.2005.07.029.

[25] M. De Paola, V. Scoppa, and R. Nisticó, "Monetary Incentives and Student Achievement in a Depressed Labor Market: Results from a Randomized Experiment," *Journal of Human Capital*, vol. 6, pp. 56–85, 2012. DOI: 10.1086/664795.

[26] T. Collins, R. Hawkins, and E. M. Flowers, "Peer-Mediated Interventions: a Practical Guide to Utilizing Students as Change Agents," *Contemp School Psychol*, vol. 22, pp. 213–219, 2018. DOI: 10.1007/s40688-017-0120-7.

[27] R. Suzuki, N. Salehi, M. S. Lam, J. C. Marroquin, and M. S. Bernstein, "Atelier: Repurposing Expert Crowdsourcing Tasks as Micro-Internships," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 2016, pp. 2645–2656. DOI: 10.1145/2858036.2858121.

[28] A. Felce, S. Perks, and D. Roberts, "Work-based skills development: a context-engaged approach," *Higher Education, Skills and Work-Based Learning*, vol. 6, pp. 261–276, 2016. DOI: 10.1108/HESWBL-12-2015-0058.

[29] J. Beck, K. Ostrow, and Y. Wang, "Students vs. skills: Partitioning variance explained in learner models," in *Student modeling from different aspects*. Worcester Polytechnic Institute, 2016, pp. 2–9.

[30] G. M. Novikova, "Evaluation of the competence of personnel in the corporate management system [Otsenka kompetentnosti personala v korporativnoy sisteme upravleniya]," *Competencies and expertise of the modern specialist [Kompetentsii i kompetentnost' sovremennogo spetsialista]*, pp. 78–95, 2011, in Russian.

[31] Q. Zou, S. Xie, Z. Lin, M. Wu, and Y. Ju, "Finding the Best Classification Threshold in Imbalanced Classification," *Big Data Research*, vol. 5, pp. 2–8, 2016. DOI: 10.1016/j.bdr.2015.12.001.

[32] S. Russell and P. Norvig, *Artificial intelligence*. Harlow: Pearson Education UK, 2016.

[33]   L. A. Gottlieb, A. Kontorovich, and P. Nisnevitch, "Nearly Optimal
       Classification for Semimetrics," *Journal of Machine Learning Research*,
       vol. 18, pp. 1233–1254, 2017. DOI: 10.5555/3122009.3122046.

For citation:

Information about the authors:

Azofeifa, Esteban J. — PhD student at the Department of
Information Technologies of Peoples' Friendship University of
Russia (RUDN University) (e-mail: azofeyfa-gomez-e@rudn.ru,
phone: +7(495)9522665, ORCID: https://orcid.org/0000-0003-0604-4234,
Scopus Author ID: 57193714543)

Novikova, Galina M. — Candidate of Engineering Sciences, di-
rector of the Center of Development of Digital Technologies for
Educational Processes and assistant professor at the Department
of Information Technologies of Peoples' Friendship University
of Russia (RUDN University) (e-mail: novikova-gm@rudn.ru,
phone: +7(495)9522665, ORCID: https://orcid.org/0000-0002-9292-0597,
ResearcherID: AAD-2296-2019, Scopus Author ID: 8971100500)

Верификация рынка труда с помощью академической краудсорсинговой системы

Э. Х. Азофейфа, Г. М. Новикова

Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

В настоящее время студенты, желающие иметь преимущество на рынке труда, готовы платить значительные деньги за информацию о потенциальных возможностях трудоустройства, тогда как получение диплома волнует их все в меньшей степени. Рассматривая поведение этого рынка труда в качестве теории предметной области в условиях неопределённости, ожидаются некоторые противоречия в виде уровней заработной платы, которые невозможно классифицировать из-за высокой противоречивости и изменчивости. Нами представлен алгоритм верификации теории предметной области рынка труда на основе краудсорсинговой академической системы, в которой обратная связь о возможных противоречиях формируется в результате консультаций с экспертами на рынке и группируется в различных контекстах. Нами обнаружено, что процесс проверки может повторяться итеративно, если общая стоимость обучения студентов равна или превышает количество, частично определяемое числом различных профилей студентов.

Ключевые слова: рынок труда, краудсорсинг, учебный план, мультиагент, верификация, противоречие