Towards Reduction in MOOCs Dropouts: An Agent-Based Model for Social Network Based Collaborative Learning

Lakshmi Sunil Prakash, Kashif Zia and Ismail Khalil

February 27, 2020

The 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS2019)
18-21 September, 2019, Metz, France

Towards Reduction in MOOCs Dropouts: An Agent-Based Model for Social Network Based Collaborative Learning

Lakshmi Sunil Prakash[1] , Dr.Kashif Zia [1,2] , Prof.Ismail Khalil[2]
[1] Sohar University,Oman,lakshmi@su.edu.om,kzia@su.edu.om,www.soharuni.edu.om,
[2] Johannes Kepler University, Linz, Austria, ismail.khalil@jku.at,www.jku.at

***Abstract* — Universities, hosting Massive Open Online Courses (MOOC) are facing a major challenge of immature dropout of a student. Major research considerations to address this challenge are only able to, identify a student at the verge of a dropout using learning and learner analytics based on different sources of data (MOOCs data, social networking data). There is no significant research done on how to avert this particular state of the learner. An agent-based model is proposed to recreate different scenarios of learners' interactions and social network evolution. The purpose of this study is to identify important phenomena and structures of social networking using a simulated MOOC setting so that the social network interaction can be channeled to avert the possibility of dropout.**

Keywords* — *collaborative learning; MOOCs; dropouts; Agent Based Modelling

I. Introduction

The Massive Open Online Courses (MOOCs) concept came into existence due to the *open learning* vision of Downes and Siemens [1], [2]. They have pioneered educators' attempts to conceptualize and offer non-credit courses to a small group of students over a network. This is widely regarded as the first known, successful cMOOCs (connectivist MOOCs). cMOOCs advocated fewer activities and more content creation by learners connected through social networks. This attempt emboldened like-minded educators to launch similar educational variations. Years later a second and equally popular variant of the MOOCs paradigm was launched by Daniel.J [3], who made a conscious effort to provide these courses globally to a larger number of connected learners through extended MOOCs or xMOOCs – which now prevails as the common definition of MOOCs.

The MOOCs model of online learning was conceptualized with the intention to be make learning accessible to many more learners than would be possible through conventional teaching. They are often free of charge and their participation is not limited by the geographical location of the learners. However, in spite of the availability of resources and self-paced assessments, the number of enrolled learners does not sustain and often it has been observed to report a very large number of dropouts, due to several reasons [4]–[6]. The most important reasons reported in the literature are:

1) The very factors that caused MOOCs to become popular (space and time independence) may have caused lethargy and has made learners less attentive [7], [8]. This behavior of the learner, in an isolated learning space [9] may have resulted or attributed to his decision to drop out of the course.
2) MOOCs are mostly instructor-led courses, however, the feedback and assessments are automated. Online Learners find these feedback systems "limited, insufficient" [10], as they may not be able to know first hand about their chances to succeed or fail in the course.

From the above, it is obvious that the main motivation for completion of a course is self-regulation. This calls for online learners to develop attributes of self-regulated learning (SRL) [11] particularly to have the intention and continuous focus (an active learner) in order to complete a course successfully. An active learner's behavioral commitment provides requisite momentum to course completion [12]. However, the challenge of maintaining SRL is a difficult task due to the very nature of MOOCs environment (as discussed above) and due to the dynamics of the heterogeneous population of learners.

With the popularity of social networks, an opening to address this challenge is widening – the concept of collaborative learning (CL). Therefore, the aim of this research is to explore how the high dropout problem can be addressed through Collaborative learning (CL) mechanisms [13], [14] and how the possibility of a premature dropout can be reduced through a conversion from passive to active learners. However, a theoretically sound (considering social and environmental dynamics) model pacifying this conversion is necessary and proposed in this paper.

The rest of the paper is organized as follows. Section II presents details of proposed collaborative learning framework. Section III is about the target case study and modeling concepts which is adopted to measure the relevance of the framework. Section IV deals with simulation and results analysis and section V concludes the paper.

II. Towards Modeling Collaborative Learning in MOOCs

Major research considerations to address the challenge of dropouts in MOOCs are only able to identify a student at the verge of a dropout using learning and learner analytics based on different sources of data (MOOCs data, social networking data). There is no significant research done on how to avert this particular state of the learner. We propose to use social network data and interaction patterns of a learner and his contacts to identify important structures of social networking so that the social network interaction can be channeled to avert the possibility of dropout.

Quite a few researchers have classified different kind of learners based on themes derived from features of SRL. Authors in [12] presents a framework for measuring dropout ratio based on the self-directed learning process, proposing a causal relationship from intention to the commitment and identifying behavioral commitment a prerequisite for an "active learner". Authors in [15] have proposed an extension of the concept and proposed a typology of learners being "inclined actors", "inclined abstainers", "disinclined actors", and "disinclined abstainers", where an actor is a learner who acts (or shows behavioral commitment) and abstainer being opposite of it, and the word inclined represents the original intention of a learner and disinclined being opposite of it.

The question now is what to do with the type of learner identified from the above process? Definitely, if the learner has high chances of getting dropped from the course, we will try to prevent it. This can be achieved by using his social network as a medium of influence. CL generally and social network analytics, in particular, provides a significant opportunity to get a learner influenced by its connections in a positive manner.

Collaborative learning (CL) is the learning achieved when a group attempts to learn together, [16]. The learning process within collaborative learning environments encourages the creation of support groups, [17]. The creation of support groups encourages better learning engagements, based on feedback that each learner receives from within the learner group [18]. In this work, we have evidenced a case study that signifies the importance of CL towards a reduction of dropout rate in MOOCs.

Another aspect of MOOCs research is its overwhelming reliance on analytical modeling. However, a complex system like this cannot be exhaustively modeled in this way. We propose a solution, a "bottom-up" approach using agent-based modeling. Our system can be categorized as a computational social system. Computational Social Science (CSS) uses computationally intensive methods to analyze and model social phenomena. Using computer simulations, artificial intelligence, complex statistical methods, and analytic approaches like social network analysis, computational sociology develops and tests theories of complex social processes through bottom-up modeling of social interactions.

III. Modeling Dynamics of Social Networking and Collaborative Learning

A. Model Overview

Contemporary research by [19], [20] have been inspired by the social networks dynamics of clustering, to explore the plausibility of simulating an optimal network and to propose a new tool for modeling message distribution and opinion formation in societies, respectively. In [21], the model describes how informal social networks play a strategic role in generating relationships among agents exhibiting homophilic nature in certain types of populations. The main ingredient of collaborative learning is the identification of learners' types. Based on interactions between different types of learners, an agent (representing one type of learner) can influence others and get itself influenced by others. However, the model representing this influence is abstracted from both the (i) cognitive dimensions of human learning mechanisms, and the (ii) social dimensions emerging due to the evolution of networking dynamics. The model only focuses on a *presumed* influence of agents on each other due to homophily.

In our research, we will employ a simplified form of homophily, which is an empirically observed sociological fact that people tend to connect to those who are similar to themselves. Homophily is a term [22], used to explain the natural tendency of people to connect with others in their vicinity, depending on the similarity in their beliefs and this is more to confirm their own beliefs. Nevertheless, the model conceptualizes the effect of one type of agents onto others (in homophilic terms) and provides means to ask interesting *what-if* questions by mimicking (i) distribution of learners in a realistic landscape and (ii) possible changes in the neighborhood of the learners due to changes in social networking.

The model demonstrates the relationship between social networking dynamics and a potential decrease in dropouts, in a MOOCs space.

B. Decision Making under Homophily

The learners' types used in the model are adapted from the learners' typology model presented in [15]. These types are: "inclined actors", "inclined abstainers", "disinclined actor" and "disinclined abstainers". An actor is a learner agent who acts (or shows behavioral commitment) and an abstainer, on the other hand, does not show the required commitment. The word inclined represents the original intention of a learner and disinclined being opposite of it. An agent's current behavior is analogous to its type. An inclined abstainer, as well as a disinclined actor, is a potential dropout and only an inclined actor is

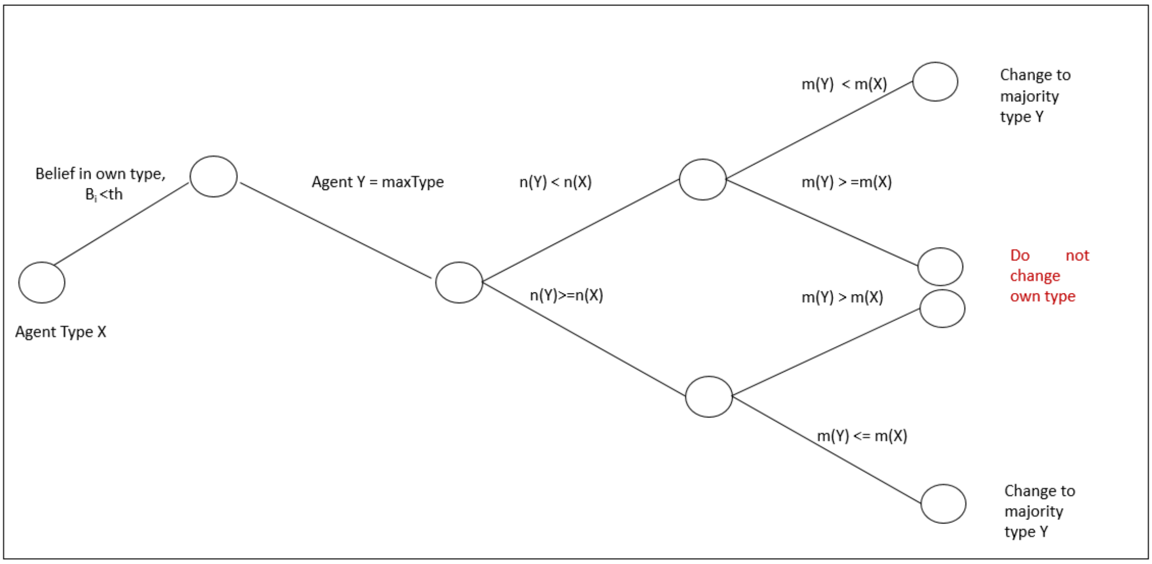

Figure 1. Change Type process of an agent i, where Y represents the type which appear in maximum percentage and X is agent's own type. n represents the accumulation of incidents in which the belief was enforced, and m represents the accumulation of incidents in which the belief was depleted.

expected to complete the course. Disinclined abstainers are obviously redundant and are not considered in the model. However, the model does not differentiate between these types - it just treats an agent's type as *a* type, which is seen in comparison with other agents in the neighborhood.

Since an agent is susceptible to change its type (which is analogous to its behavior) due to homophily, the effect of neighboring agents on the agent is reflected in its state "belief" - how much an agent believes in its own type. The belief of an agent changes a little (a small fraction between 0.01 - 0.10) due to neighborhood effect - a clue. A positive clue - if the majority of agents in the neighborhood of an agent are of its own type - reinforces agent's belief. Contrarily, a negative clue - if the majority of agents in the neighborhood of an agent are NOT of its own type - depletes the agent's belief. Starting from an initial belief equal to 0.5, repetitive encounter to a negative clue forces the agent to change its type - the type of the majority - when its belief (on its own type) is depleted to an extent that it is below a threshold, 0.2. The exact mechanism is as follows.

Let X be the type of agent A which is making the decision. Let Y be the type of the majority of the agents in the neighborhood of A. We use two opposite aspects of historical proceedings driving the model, n and m. n represents the accumulation of incidents in which the belief was enforced, and m represents the accumulation of incidents in which the belief is depleted.

If n(Y) is *greater* than n(X) – belief enforcement of agent's own type is less than enforcement of other types – the agent would change its type from X to Y only if m(Y) is less than m(X) – belief depletion of agent's own type is greater than depletions of other types. On the contrary if n(Y) is *less* than n(X) - belief enforcements of the agent's own type is greater than the belief enforcements of other types, the agent would change its type from X to Y, if m(Y) is less than m(X) – belief depletions of agent's own type is greater than belief deletions of other types.

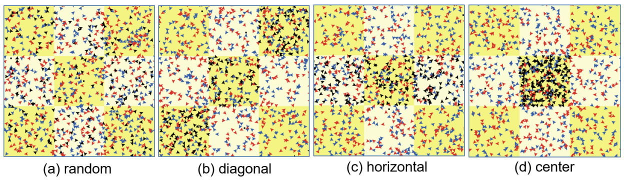

Figure 2. Agents' initial placements; a population of 1000 agents with 20% inclined actor (black), 40% disinclined actors (blue) and 40% inclined abstainers (red).

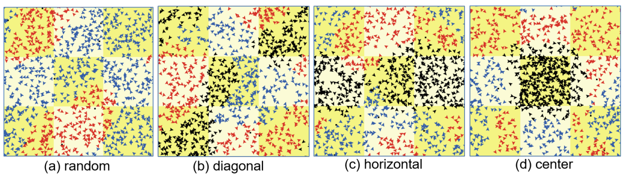

Figure 3. State of agents at iteration 100 in case of four placements for stationary mobility.

These are the only two situations in which the agent changes its type from X to Y. The complete decision-making process is depicted through a decision tree shown in Figure 1. In this way, the model provides a simple, yet

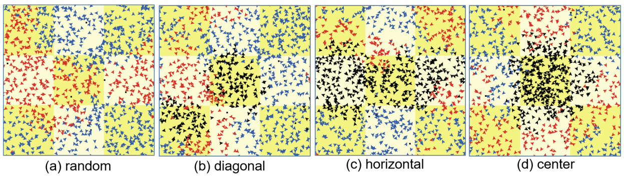

Figure 4. State of agents at iteration 100 in case of four placements for random walk mobility.

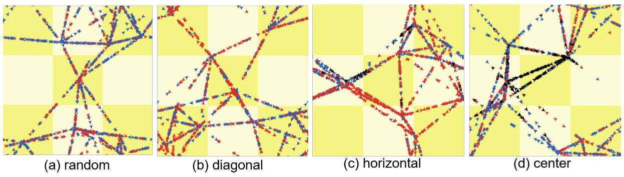

Figure 5. State of agents at iteration 100 in case of four placements for location based mobility.

rational decision support to the agents.

C. Network Dynamics

The network dynamics are incorporated into the model through the placement of agents and mobility. The initial placement of agents defines how agents are knitted together. In the following, we detail considerations related to agents' placement. It is a plausible notion that these learners' placements drive the formation of ties between various types of learners. The type of learner, the number of that particular type within his learning space and the frequency and strength (strong/weak) of these interactions.

A **random** placement has no networking except for some random incidents in which agents become neighbors. There is no concept of *institutions* as well. Figure 2 (a) shows such a situation in which the type of agent we are interested in (the inclined actors) have no spatial correlation with each other. They are just randomly placed. Whereas, all inclined actors are placed at the center of the world in case of **center** placement. As shown in Figure 2 (d), all inclined actors start by occupying the central section (consider it to be a district). In terms of institutions, we can assume that this corresponds to one specialized University of research center; and all the inclined actors reside in the same institution initially. Another situation can be when inclined actors are distributed across more than one institutions and these institutions have no correlated interaction - for example, three universities with no relationship between their students. This situation is represented by only three **diagonal** districts having initial placements of inclined actors as shown in Figure 2 (b).

On the contrary, an entirely opposite situation would be when inclined actors are distributed across more than one institutions and these institutions have a relationship - more specifically in terms of students being friends to each other - for example, three universities with a lot of inter-university activity. This situation is represented by three **horizontal** districts having initial placements of inclined actors as shown in Figure 2 (c).

After initial placements, at each time-stamp, each of the agents would - within its radial range - interact with other agents and make a decision (should I change my type or not). The purpose is to find those conditions in which the percentage of inclined actors increase with time. Remember, that the decision-making model itself, does not differentiate between agents' type and there is no privilege of being inclined actor or penalty of being inclined abstainer, for example.

Further, the network connectivity may change as the simulation progress making the world a dynamically changing network. This is achieved through three mobility modes. These mobility modes represent how people interact with each other. **Stationary mode** is a static network without any dynamics. **Random walk mobility** is a popular interaction mechanism in which people move and influence their neighborhood with a *little* randomness and **location based mobility** is about how people traverse from one location to another independently.

It is important to note that the interactions only have one layer. People having two or more layers (a layer of social network and another layer of people in a market together, for example) is not considered. Another limitation is that the network is volatile and instant in nature, that is agents do not have any history of the contacts and their past behavior or states.

D. The Agent-based Model

Using Agent-Based Modeling (ABM), an emerging and popular computational paradigm, we will prepare a simulation of these adaptive learner agents connectivity dimensions and their placement models. The agents are behaviorally adaptive, in a sense, they influence others in response to the influence they receive. ABM provides a possibility to model these complex local interactions with the rational that controls or influences the behavior of these adaptive agents. These three distributions are realized through spatial arrangements of the inclined actors shown in *center* representing the first case, *horizontal* representing the second case, and *sides* and *diagonal* representing the third case. Mobility of learners generates connectivity dynamics and defines the dynamics of influence from the emerging neighborhood. The purpose of the model is to find the conditions for which dropouts are decreased indicated by more and more inclined actors in the system as the simulation progresses in time. These *learner agents* are distinguished by their intent or motivation to be an active participant in the course and engage in a majority of the assessments that steer them towards course completion and also earn credit. *Abstainers agents* are MOOCs participants who register in courses but who are not showing the behavioral commitment required by learners to be considered as active learners. Hence we have three learner(agent) typologies - inclined actor actor, the disinclined actor and inclined abstainer. One typology, the *inclined abstainer* has not been included in our conceptual model as this type is quite rare and we feel, and do not need representation in MOOCs learning spaces. In order to model, the time-based agent interaction-influence scenarios in social learning spaces, we will create the following connectivity dimensions.

IV. Simulation and Results

The purpose of the model is to find the conditions for which dropouts are decreased indicated by more and more inclined actors in the system as the simulation progresses in time. The type of agent which we focus on is inclined actors. The success of a mechanism is directly proportional to the number of inclined actors in the system after time t.

A. Simulation Setup

A 2D square torus of size 99×99 (each point represented by a unique XY-coordinate) is distributed into 9 square cells of equal size. A total of 9 random points (coordinates) are chosen each of which represents a point of interest (poi), which need not be one poi per cell. In one of the mobility mode (location-based), all the agents created on the torus already have four nearest points as locations to visit. All other mobility modes do not require any initialization or pois. A total of 1000 agents are created and randomly distributed across space with a default belief of 0.5.

A relatively small percentage (20%) of agents are inclined actors. The rest of the population is equally distributed between the agents who are inclined abstainers and disinclined actors, 40% in each case. The inclined actors are then re-stationed according to placement features of the space. Each agent also has a count of its neighbors (within a radius = 10), which does not change as the simulation progresses (although the neighbors count changes). This provides us with an opportunity to look at the progression of simulation from viewpoint of agents' initial perception; they do not have perception capabilities to realize a highly volatile neighborhood (very similar to social media where the ties are often not broken or made as rapidly as the messages are disseminated).

The simulation initializes the positioning of the inclined actors according to four distributions as shown in Figure 2, namely:

- randomly - without any cell-based positioning.
- diagonally - 3 cells at the diagonal of the 9 cells lattice - corresponding to islands of institutions.
- horizontally - 3 cells at the central horizontal row of the 9 cells lattice - corresponding to correlated institutions.
- center - a single cell right at the center of the 9 cells lattice - corresponding to a single institute with a very dense placement of inclined actors.

B. Results

The first mode is **stationary** mode in which agents do not move. Based on initial placement, different pattern emerge. For example, in case of a random placement, the neighborhood of an inclined an actor would be 80% of not its own kind. Therefore, due to the homophilic nature of the decision-making model would let all inclined actors change their type to one of the other two types, as shown in Figure 3 (a). However, for all other cases (diagonal Figure 3 (b), horizontal Figure 3 (c), and center Figure 3 (d) placement), the inclined actors enforce each other. We will explain these results in more detail next.

The second mobility mode is **random walk** mode. Similar to the above, based on initial placement, different pattern emerge. The mobility would not help here either and all inclined actors change their type to one of the other two types, as shown in Figure 4 (a). However, for all other cases (diagonal Figure 3 (b), horizontal Figure 3 (c), and center Figure 3 (d) placement), the inclined actors enforce each other, most in case of central placement due to their close vicinity to each other. We will explain these results in more detail next. The third mobility mode is **location-based** mobility mode, in which agents choose some nearest pre-specified locations, and they move from one location to another. Overall, with this mobility, the enforcement of inclined actors for each other depletes due to mobility dynamics itself - the agents form some kind of lines while traversing from one location to another - thus reducing inclined actors (see Figure 5). We will explain these results in more detail next.

C. Discussion

Our simulation represents some of the connectivity dimensions possibilities in the simulated MOOCs world and the three chosen mobility models. The results of the simulation are shown in Figure 6. The simulation generates best results, in terms of inclined actor population at the end of 100 runs for the stationary mode in horizontal, random and diagonal placement models. However, the number of inclined learners drops for randomly based mobility with the center and diagonal placement which has better enforcement than the random and horizontal placements models. The worst results were received for the location-based mobility where the center placement model fared better than the other three placement models. The dense placement of inclined actors (center placement) does have a positive increase in actors of the same type in both the stationary mode, random mode and location-based mode as compared to other placement modes. This reiterates our earlier results that learners who exhibit less mobility due to their affinity within their neighborhood with learners of their own type can be expected to show similar behavior in the social space too.

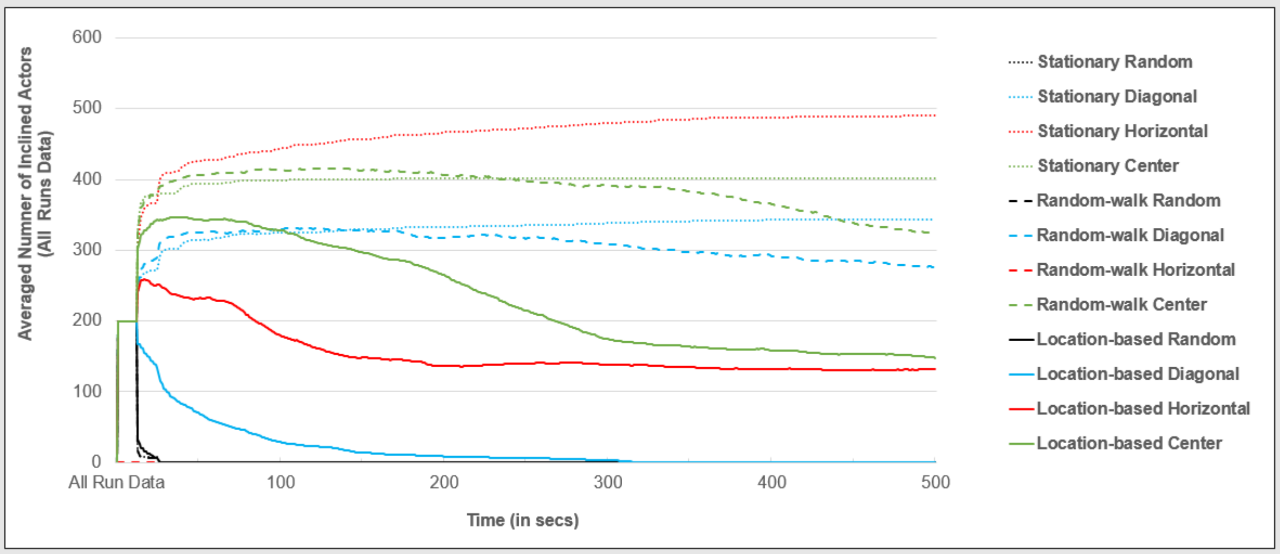

Figure 6. Graph for the set of 12 Cases from the Simulation Experiment

V. Conclusion

Massive Pervasive Learning spaces such as MOOCs face several challenges with regards to learner disengagement and dropouts. In this paper, we have proposed a framework which combines self-regulated learning with collaborative learning. The paper focuses on a case study which manifests the contribution of collaborated learning interactions with social world connectivity, for course completion rates of MOOC courses. An agent-based model simulating simplistic social networking modalities and learners' categories is simulated. It is observed that the mobility (social connectivity) and initial distribution (physical placement) of agents have a deciding role in decreasing the dropout rate. There are some limitations in this model in that the model is static. We have simulated the agent interactions within a physical space of the learner. However, we will consider the social layer of agent interactions and take into consideration the history of contacts and variations in their behavior.

References

[1] G. Siemens, "Connectivism: Learning as network-creation," 2005.

[2] S. Downes and G. Siemens, "Cck08-the distributed course," *The MOOC guide*, 2008.

[3] J. Daniel, "Making sense of moocs: Musings in a maze of myth, paradox and possibility," *Journal of interactive Media in education*, vol. 2012, no. 3, 2012.

[4] D. Clow, "Moocs and the funnel of participation," in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 185–189.

[5] J. Kay, P. Reimann, E. Diebold, and B. Kummerfeld, "Moocs: So many learners, so much potential," *IEEE Intelligent systems*, vol. 28, no. 3, pp. 70–77, 2013.

[6] J. Reich, "Mooc completion and retention in the context of student intent," *EDUCAUSE Review Online*, 2014.

[7] S. M. North, R. Richardson, and M. M. North, "To adapt moocs, or not? that is no longer the question." *Universal Journal of Educational Research*, vol. 2, no. 1, pp. 69–72, 2014.

[8] J. Sinclair, R. Boyatt, C. Rocks, and M. Joy, "Massive open online courses: a review of usage and evaluation," *International Journal of Learning Technology*, vol. 10, no. 1, pp. 71–93, 2015.

[9] L. S. Prakash, K. Zia, and I. Khalil, "Collaborative learning to reduce learners dropouts in moocs: Modeling an agent-based simulation," in *Proceedings of the 20th International Conference on Information Integration and Web-based Applications & Services*. ACM, 2018, pp. 55–64.

[10] T. Daradoumis, R. Bassi, F. Xhafa, and S. Caballé, "A review on massive e-learning (mooc) design, delivery and assessment," in *P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2013 Eighth International Conference on*. IEEE, 2013, pp. 208–213.

[11] C. Milligan, A. Littlejohn, and N. Hood, "Learning in moocs: A comparison study," *Research Track*, p. 15, 2016.

[12] L. Huin, Y. Bergheaud, P. A. Caron, A. Codina, and E. Disson, "Measuring completion and dropout in moocs: A learner-centered model," *Research Track*, p. 55, 2016.

[13] K. Jordan, "Massive open online course completion rates revisited: Assessment, length and attrition," *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 3, 2015.

[14] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: analyzing learner subpopulations in massive open online courses," in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 170–179.

[15] M. A. Henderikx, K. Kreijns, and M. Kalz, "Refining success and dropout in massive open online courses based on the intention–behavior gap," *Distance Education*, vol. 38, no. 3, pp. 353–368, 2017.

[16] P. Dillenbourg, "What do you mean by collaborative learning?" 1999.

[17] C.-H. Tu, *Online collaborative learning communities: Twenty-one designs to building an online collaborative learning community*. Intellect Books, 2004.

[18] C. Crook, "Editorial changing times," *Computer Assisted Learning*, vol. 19, no. 4, pp. 399–400, 2003.

[19] R. E. Hiromoto, "Parallelism and complexity of a small-world network model," *International Journal of Computing*, vol. 15, no. 2, pp. 72–83, 2016.

[20] A. Bomba, N. Kunanets, V. Pasichnyk, and Y. Turbal, "Process modeling of message distribution in social networks based on socio-communicative solitons," *International journal of computing*, vol. 17, no. 4, pp. 250–259, 2018.

[21] M. O. Jackson and A. Wolinsky, "A strategic model of social and

economic networks," *Journal of economic theory*, vol. 71, no. 1, pp. 44–74, 1996.

[22] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, vol. 27, no. 1, pp. 415–444, 2001.