COMPUTERS & SECURITY 80 (2019) 238–256

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Design recommendations for online cybersecurity courses

*Lorena González-Manzano**, *Jose M. de Fuentes*

Computer Science Department, Universidad Carlos III de Madrid, Spain

ARTICLE INFO

Article history:
Received 8 March 2018
Accepted 26 September 2018
Available online 10 October 2018

Keywords:
Cybersecurity teaching/learning strategies
Cybersecurity distance education and telelearning
Media in cybersecurity education
Cybersecurity course guidance
NICE framework

ABSTRACT

Nowadays, a significant amount of free online cybersecurity training courses are offered. When preparing further courses, the designer has to decide *what* to teach and *how* to do it. In this paper, we provide with a set of recommendations for both issues. Concerning topic selection, 35 free online courses are analysed using NIST's NICE reference framework. Thus, several training gaps are discovered. Concerning the way of preparing the course (or refining it after the first edition), a set of good practices is proposed based on students' performance and commitment in a cybersecurity MOOC with +2,000 initially active students. To foster further research in this area, an open-source framework is released to enable the analysis of students' performance in EdX MOOCs.

1. Introduction

Cybersecurity needs are becoming more and more widespread. According to the 2015 (ISC)2 Global Information Security Workforce Study, a shortfall of 1.8 million cybersecurity jobs will happen by 2022 (Sullivan, 2017). Allegedly, the most common reason for this fact is the lack of qualified personnel.

As a consequence of this trend, a plethora of training materials and courses have been proposed. For instance, many well-known universities have cybersecurity master programs (Cabaj et al., 2018). However, this paper focuses on online alternatives since their availability make them to become attractive for large amounts of learners. Among them, two types of training actions can be found. On the one hand, virtual security labs are intended to provide with experimental training (Salah, 2014). On the other hand, regular courses are usually considered to offer a theoretical and practical background on a topic. In the last years, a significant amount of platforms (e.g. EdX[1], Coursera[2], etc.) are offering cybersecurity courses as part of their catalogue. Due to their success and general adoption, in this paper we focus on these initatives.

One key aspect is how to determine which Knowledge, Skills and Abilities (KSAs) are needed to become a cybersecurity professional. This specialty is not monolithic – indeed, a huge amount of profiles can be devised. For example, it is not the same to be the network manager in charge of defending a network than being the person responsible for determining the source of an attack. In order to address these needs, in 2013 NIST proposed NICE (National Initiative for Cybersecurity Education) framework to systematically structure this field of knowledge (NIST). In short, NICE proposes a set of work roles, associated with a set of specialty areas. On the other hand, a list of KSAs is also proposed.

When preparing a new cybersecurity course, one important issue is determining which topics are not properly addressed. Paulsen et al. already pointed out early in 2012 that

* Corresponding author.
E-mail address: lgmanzan@inf.uc3m.es (L. González-Manzano).
https://doi.org/10.1016/j.cose.2018.09.009
0167-4048/

[1] https://www.edx.org , last access March 2018.
[2] https://www.coursera.org , last access March 2018.

cybersecurity educators have difficulty gaining a holistic view of the available resources and determining exactly what to teach (Paulsen et al., 2012). However, to the best of authors' knowledge, this issue is still open.

Apart from the topics to be covered, cybersecurity trainers have to pay attention to how these concepts are addressed. For this purpose, characterizing the target audience is at stake. As it happens in many disciplines, the background of the students is relevant to decide the starting point and learning pace. Moreover, student culture, i.e. the intrinsic habits (i.e. cooperation, commitment) that flourish when facing a training activity, are also relevant to design a suitable course. These aspects should also be taken into account when refining the course after each edition.

This paper addresses both topic choice and course preparation issues. Concerning the first aspect, the current coverage of the NICE framework is studied. For this purpose, 35 free online courses are surveyed. This leads us to detect gaps between what is taught and what is required in cybersecurity work roles – those parts of NICE that have received less attention from training designers. With respect to course preparation, we analyse students performance and commitment of a Massive Online Open Course (MOOC) carried out in edX platform in 2017. This MOOC counted on +10,000 worldwide enrolled students among which +2,000 were initially active. This analysis leads us to propose a set of recommendations for course preparation. Moreover, in order to foster further research in this area, an open-source framework to analyse student results in edX courses is released.

Paper organization. The remainder of this paper is organized as follows. Section 2 introduces the NICE framework as background. Afterwards, Section 3 describes the coverage of NICE of existing online courses. Section 4 analyses different aspects of students interaction within our cybersecurity MOOC, which lead us to the proposed recommendations (Section 5). Section 6 introduces the related work. Finally, Section 7 concludes the paper and points out future work directions. For the sake of clarity, a set of comprehensive tables have been placed in appendices, as well as the description of our open-source analysis framework for edX courses.

2. Background. NICE framework

NICE Framework, developed by NIST (NIST), establishes a taxonomy to describe cybersecurity work roles with the intention to be applied in any sector, public, private or academic. In particular, the framework includes three different components: 7 categories which present cybersecurity functions at high level; 33 specialty areas to distinguish cybersecurity areas; and 52 work roles which define cybersecurity work in detailed according to specific KSAs required by each work role. To get a bit deeper in this, here the categories and specialty areas within each of them are briefly introduced (NIST):

- Securely Provision (SP): conceptualize, design, procure and/or build information technology systems. It also involves responsibility for aspects of system and/ network development. The following specialty areas are included:
  - Risk Management (RSK): work in the evaluation, validation, assessment and authorization to ensure the satisfaction of cybersecurity and risk requirements.
  - Software Development (DEV): develop and write or update applications, software or specialized programs applying software assurance best practices.
  - Systems Architecture (ARC): develop system concepts and work on the systems development life cycle; translate technology and environmental conditions into system and security designs and processes.
  - Technology R&D (TRD): conduct technology assessment and integration processes and provide and support prototyping.
  - Systems Requirements Planning (SRP): evaluate and gather functional requirements to be translated to technical solutions.
  - Test and Evaluation (TST): test systems to evaluate compliance.
  - Systems Development (SYS): work on phases of the systems development life cycle.
- Operate and Maintain (OM): provision of support, administration and maintenance to ensure efficient and effective performance and security of information technology systems. The following specialty areas are included:
  - Data Administration (DTA): develop and administer databases and/or data management systems.
  - Knowledge Management (KMG): manage processes and tools towards intellectual capital and information content.
  - Customer Service and Technical Support (STS): address problems, maintenance and training in customer requirements or inquiries.
  - Network Services (NET): network management, at all levels, including hardware and software.
  - Systems Administration (ADM): configure and maintain hardware and software, as well as access control and account management.
  - Systems Analysis (ANA): improve procedures to help the organization operate more securely, efficiently and effectively.
- Oversee and Govern (OV): provision of leadership, management, direction or development and advocacy to carry out an effective conduction of the cybersecurity work. The following specialty areas are included:
  - Legal Advice and Advocacy (LGA): legal advisor and recommender.
  - Training, Education, and Awareness (TEA): personal training on demanding subjects.
  - Cybersecurity Management (MGT): oversee information systems and network cybersecurity programs to put the right resources in the right place.
  - Strategic Planning and Policy (SPP): develop and update or enhance policies.
  - Executive Cyber Leadership (EXL): supervise and manage people working in cybersecurity topics.
  - Program/Project Management (PMA) and Acquisition: manage acquisition programs having the appropriate knowledge and also support by information technologies.

- Protect and Defend (PR): identify, analyse, mitigate threats caused in information technology systems and networks. The following specialty areas are included:
  - Cybersecurity Defense Analysis (CDA): use of collected information to report threats events.
  - Cybersecurity Defense Infrastructure Support (INF): manage the computer network defense service provider network and resources.
  - Incident Response (CIR): response in an urgent situation to mitigate potential threats. Investigation of all relevant response activities.
  - Vulnerability Assessment and Management (VAM): assess vulnerabilities and threats, considering risks, policies and suggest countermeasures.
- Analyze (AN): performance of highly-specialized reviews and evaluations of incoming cybersecurity information to consider its usefulness for intelligence purposes. The following specialty areas are included:
  - Threat Analysis (TWA): assess and produce findings of capabilities of cybersecurity criminals and intelligence entities.
  - Exploitation Analysis (EXP): identification of potentially exploitable vulnerabilities.
  - All-Source Analysis (ASA): analyse threats, synthesize and place intelligence information in context.
  - Targets (TGT): apply knowledge of regions, countries, non-state entities, and/or technologies.
  - Language Analysis (LNG): use of expertise to support information collection, analysis and other cybersecurity activities.
- Collect and Operate (CO): provision of specialized denial and deception operations and collection of cybersecurity information which can be used for intelligence purposes. The following specialty areas are included:
  - Collection Operations (CLO): execute collection with the right strategies and priorities.
  - Cyber Operational Planning (OPL): develop operational plans and orders, together with strategic and operational-level planning.
  - Cyber Operations (OPS): evidence gathering on criminal or foreign intelligence entities.
- Investigate (IV): investigation of events or crimes related to information technology systems, networks and digital evidences. The following specialty areas are included:
  - Cyber Investigation (INV): apply tactics, techniques and procedures for a full range of investigative tools and processes.
  - Digital Forensics (FOR): collect, process, preserve, analyse and present computer-related evidence.

3. Analysis on current NICE coverage

The existence of the NICE framework makes it easier to spot which cybersecurity aspects are needed for every work role. However, to date there is no comprehensive analysis on which matters have received attention from the online training community and which others are lacking.

In the following, an analysis of existing free online cybersecurity courses is presented. The catalogue of courses is introduced in Section 3.1. The coverage of NICE is presented in Section 3.2. Note that work roles have not been considered because most courses are too general to be linked to a particular role.

3.1. *Course catalogue*

In this work, a total of 35 free courses have been identified (Table 1). These courses are mainly taught in four platforms, namely Coursera, Cybrary.it, edX and Udacity. Additionally, a subset of courses belong to other initiatives (Khan Academy, MIT OCW and Future learn). Most of these courses (25 out of 35) are developed by educational institutions, mainly belonging to the United States.

It must be noted that in this work, only courses with more than 10 teaching hours (or 1 week of work, if nothing more concrete is specified) have been selected. Therefore, micro-courses are left out of the scope. Only two courses have been selected that do not specify any length but contain enough material to convey a significant workload.

Regarding their level, 12 courses are for beginners, 12 intermediate and 7 advanced. Apart from this general classification, 2 courses did not specify their level and 2 were for graduates – this latter type being not classifiable as any of the general levels.

3.2. *Coverage of categories and specialty areas*

Fig. 1 shows the coverage of NICE framework categories, though the detailed coverage per courses and Specialty Areas (SAs) is depicted in Appendix A. Note that each course typically covers more than one category. One important issue is that there is a huge difference between categories in what comes to their coverage. Thus, CO and IV are covered by just 3 courses, whereas SP or OV are the most offered aspects with 19 and 18 courses, respectively.

Beyond categories, NICE contains a refined set of SAs, recall Section 2, whose coverage deserves attention. This issue can be addressed in two directions. On the one hand, how many

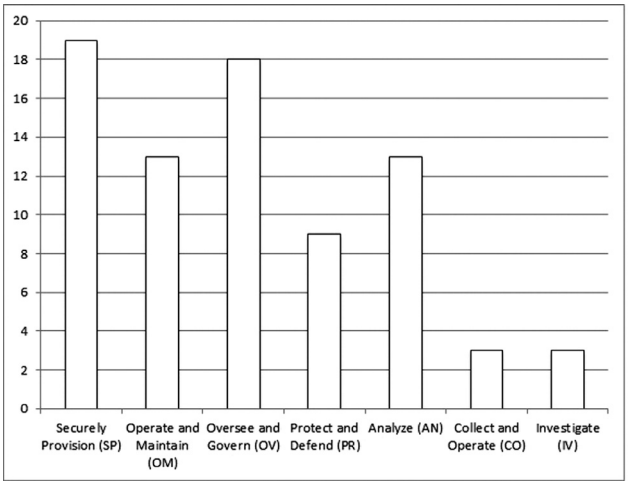

Fig. 1 – Coverage of cybersecurity domains from training courses.

Table 1 – Catalogue of cybersecurity free online courses.

Title	Institution	Platform	Level	Length: week (w) / month (m)
Hardware Security U. Maryland (2017b)	Univ. Maryland	Coursera	Intermediate	7w
Software Security U. Maryland (2017c)			Intermediate	1.5m
Usable Security U. Maryland (2017d)			Intermediate	7w
Cryptography U. Maryland (2017a)			Intermediate	7w
Information Security: Context and Introduction U. of London (2017)	Univ. of London		Beginner	5w
Malicious Software and its Underground Economy: Two Sides to Every Story U. of London (2018)			Beginner	Undefined
Cybersecurity and Its Ten Domains U. S. of Georgia (2018)	Univ. System of Georgia		Beginner	7w
Cryptography I University (2017)	Stanford University		Intermediate	7w
Internet History, Technology, and Security U. of Michigan (2017)	Univ. of Michigan		Beginner	2.5m
Networking and security in iOS applications Irvine (2017)	University California Irvine		Undefined	Undefined
Homeland security and cybersecurity U. of Colorado (2017)	University of Colorado		Beginner	1w
Ethical hacking and penetration testing Cybrary (N/Ae)	No institution	Cybrary.it	Intermediate	13.5h
Certified Information Systems Security Professional Cybrary (N/Ab) (CISSP)			Advanced	13h
Python for Security Professionals Cybrary (N/Af)			Advanced	10.5h
Advanced Penetration Testing Cybrary (N/Aa)			Advanced	14.5h
CompTIA CASP Cybrary (N/Ac)			Advanced	11h
CompTIA Network+ Cybrary (N/Ad)			Beginner	32h
Secure coding Cybrary (N/Ag)			Intermediate	10h
Intro to Cryptography Academy (2011)	Khan Academy	Khan Academy	Undefined	Undefined
Network and Computer Security MIT (2014b)	MIT	MIT OCW	Graduate	3m approx
Computer Systems Security MIT (2014a)			Graduate	3m approx
Introduction to cybersecurity University (N/A)	The open University	Future learn	Beginner	2m
Cybersecurity Fundamentals R. Institute (N/Aa)	Rochester Inst.	edX	Advanced	2m.
Network security Institute (N/Ab)			Advanced	2m
Introduction to cybersecurity U. of Washington (N/Ac)	Univ.Washington		Beginner	4w
Cybersecurity - the CISO's view U. of Washington (N/Ab)			Beginner	4w
Building a cybersecurity toolkit U. of Washington (N/Aa)			Beginner	4w
Cyber security economics U. of Delft (N/A)	University of Delft		Intermediate	10w
Cyberwar, Surveillance and Security U. of Adelaide (N/A)	University of Adelaide		Beginner	1.5m
Cyber Security Basics: A Hands-on Approach C. I. U. of Madrid (2017)	University Carlos III of Madrid		Intermediate	6w
Intro to information security Tech (N/Ac)	Georgia Tech	Udacity	Intermediate	>3m
Cyber-Physical Systems Security Tech (N/Ab)			Intermediate	4m
Network security Tech (N/Ad)			Intermediate	4m
Computer networking Tech (N/Aa)			Intermediate	3m aprox
Applied cryptography U. of Virginia (N/A)	University of Virginia		Advanced	2m

SAs are addressed by each course. This illustrates whether courses are focused on a particular task or have a broader scope. On the other hand, how many courses address each SA. This analysis reveals which areas have been taught in online courses. Both issues are introduced below.

Concerning the first issue, on average 2.94 SAs are addressed in each course. Thus, the vast majority of courses, 74%, are quite targeted, as they cover up to 3 SAs. Only a small portion covers a bigger amount, 17% of courses cover between 4 and 7 SAs and the remaining 9% between 7 and 10 SAs. Among them, our proposed MOOC, *Cyber Security Basics: A Hands-on Approach* (see Section 4), is the most general one, addressing 10 SAs.

With respect to the second aspect, Fig. 2 shows the amount of courses addressing each SA. Particularly, network secure management (NET) is strongly covered (12 courses). Behind this SA, cybersecurity management (MGT) and secure system development (SYS) also receive significant attention, they are covered in 10 and 9 courses, respectively.

On the other side, it is noteworthy that 6 SAs have only been covered by one course. Therefore, issues such as systems architecture (ARE), data administration (DAT) or even training, education and awareness (TEA) count on scarce educational resources.

The most critical situation is found in 8 SAs which have not been covered by any course (TRD, SRP, KMG, STS, ANA,

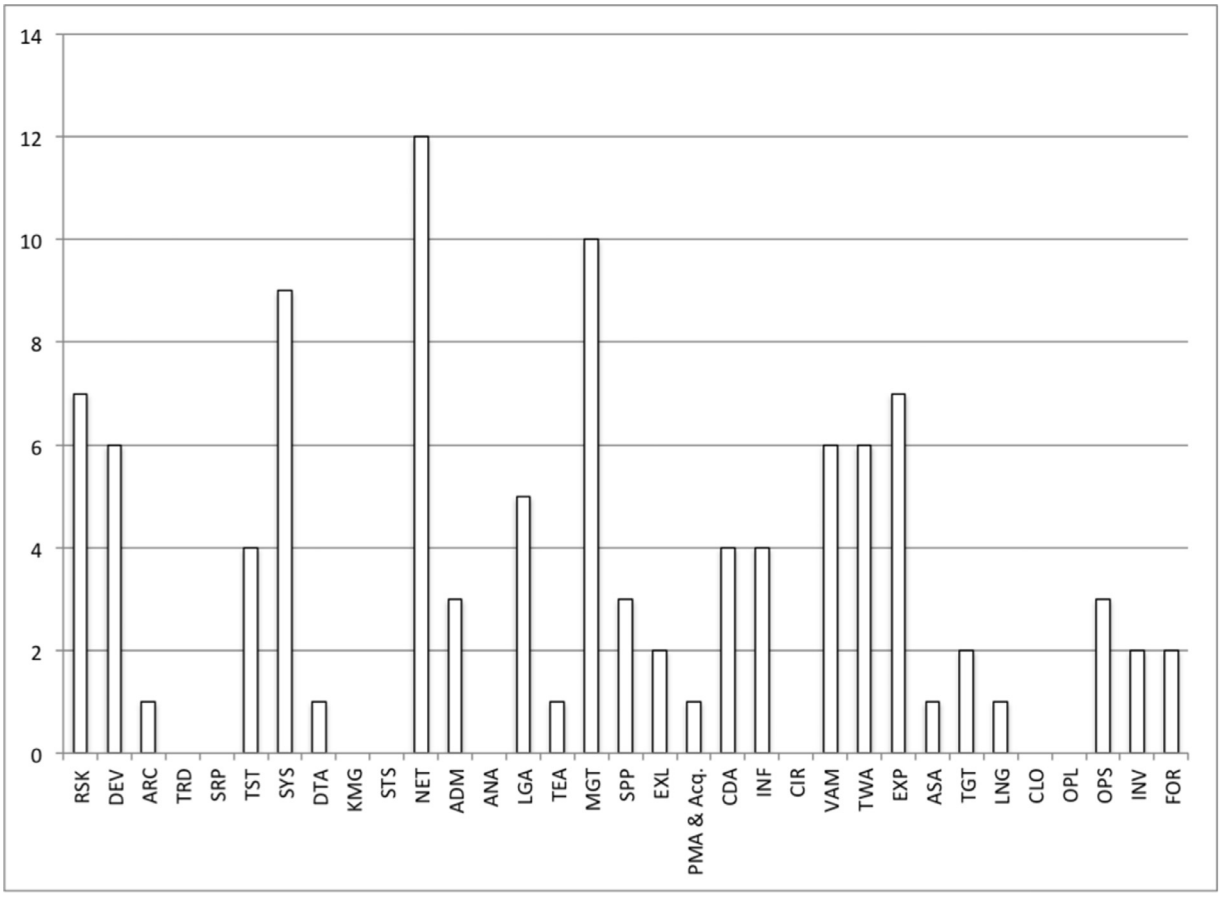

Fig. 2 – Coverage of specialty areas from training courses.

CIR, CLO and OPL). This means that there is no single free course that covers these topics. In some cases, it may be due to the fact that it covers matters which belong to a company know-how. It could be the case of customer service and technical support (STS). In other cases, it may happen because the current training trend is applying other mechanisms, e.g. incident response (CIR) training is typically done in capture-the-flag contests. In any case, the current contents and KSAs pointed out by the NICE framework for all SAs include theoretical issues that could be addressed by an online course. Therefore, the lack of courses covering some SAs should be corrected.

Related to the coverage of SAs, one important consideration is related not only to the amount of courses that are addressing them, but also to which extent they do it. The most convenient situation would be to have every SA taught at beginner level at first, and afterwards increasing the difficulty in intermediate or advanced levels. Table 2 summarizes this issue. Several gaps can be found in many SAs, since they contain intermediate or advanced courses but no beginner ones. This happens, for example, in strategic planning and policy (SPP) or in secure software development (DEV). On the other hand, some areas (such as OPS or INV) are already presented for beginners, but more advanced materials are currently lacking.

4. Cyber security basics: Ahands-on approach MOOC analysis

In order to propose the set of recommendations for cybersecurity course preparation, we leverage the results of students in *Cyber Security Basics: A Hands-on Approach* MOOC. For this purpose, this MOOC is firstly described to be later analysed based on three different aspects: cybersecurity audience composition, students' behavioral trends and students' performance. All these aspects are studied in terms of countries from which students came from, students' gender, age and educational level. Appendix B presents acronyms used for countries and educational level.

In order to carry out this analysis, we have developed a framework to process user interaction data. In its current version, it is suitable for EdX data and it is available in GitHub[3]. Appendix C details the use of this framework which is extensible and can be used to get different views on the data provided by the learning platform. Note that edX does not automatically provide videos and comments data. Then, video-related data, for instance users who have watched videos and for how much time, has been manually downloaded.

[3] https://github.com/lgmanzan/Toolset_EDX_dataProcessing, until being accepted.

Table 2 – Courses covering each SA per level.

Specialty area (SA)	Beginner	Intermediate	Advanced	Undefined or other
Risk Management (RSK)	3	1	3	0
Software Development (DEV)	0	4	1	1
Systems Architecture (ARC)	0	0	1	0
Technology R&D (TRD)	0	0	0	0
Systems Requirements Planning (SRP)	0	0	0	0
Test and Evaluation (TST)	1	2	0	1
Systems Development (SYS)	2	3	3	1
Data Administration (DTA)	1	0	0	0
Knowledge Management (KMG)	0	0	0	0
Customer Service and Technical Support (STS)	0	0	0	0
Network Services (NET)	5	3	3	1
Systems Administration (ADM)	1	1	1	0
Systems Analysis (ANA)	0	0	0	0
Legal Advice and Advocacy (LGA)	3	2	0	0
Training, Education, and Awareness (TEA)	1	0	0	0
Cybersecurity Management (MGT)	1	4	3	2
Strategic Planning and Policy (SPP)	0	1	2	0
Executive Cyber Leadership (EXL)	1	0	1	0
Program/Project Management (PMA) and Acquisition	0	0	1	0
Cyber Defense Analysis (CDA)	1	2	1	0
Cyber Defense Infrastructure Support (INF)	1	2	1	0
Incident Response (CIR)	0	0	0	0
Vulnerability Assessment and Management (VAM)	1	2	2	1
Threat Analysis (TWA)	3	2	1	0
Exploitation Analysis (EXP)	2	3	2	0
All-Source Analysis (ASA)	0	1	0	0
Targets (TGT)	2	0	0	0
Language Analysis (LNG)	1	0	0	0
Collection Operations (CLO)	0	0	0	0
Cyber Operational Planning (OPL)	0	0	0	0
Cyber Operations (OPS)	3	0	0	0
Cyber Investigation (INV)	2	0	0	0
Digital Forensics (FOR)	1	0	1	0

Similarly, the amount of comments per user has been manually gathered.

4.1. *MOOC description and students' activity overview*

This MOOC is an initiative to learn cybersecurity from a practical point of view. Theoretical explanations, essential for an appropriate comprehension of all concepts, are supported by examples and tools to guarantee a comprehensive learning process. After an introduction to cybersecurity to contextualize its relevance and significance, the main well-known techniques, concepts and tools for a cybersecurity beginner are presented. This course level is intermediate as it requires having a background in computer science.

This course is composed of 6 lectures: 1. Introduction to cybersecurity; 2.Computer forensics; 3.Assembly programming; 4. Cyberdefense; 5. Malware and advanced persistent threats; and 6. Vulnerabilities and exposures. Each lesson contains a wide range of videos to enhance the students learning experience. There are a total of 93 videos with a duration of 6.9 hrs, together with different activities like forums, homeworks or self-assignments, to motivate and guide students. In the case of graded activities they correspond to an exam at the end of each lecture, which is graded 10% of the total grade (i.e. a total of 60%), and 10 readings located along all lectures, which are graded 15% of the total grade. A final exam is performed at the end of the course, which contains contents from all lectures and it is graded the remaining 25%. Table 3 presents a summary of the MOOC's content.

By mid March 2017 this course was opened, having active participation of teachers for the initial 8 weeks. After that, the course remained open without teacher intervention. In this study, we consider results obtained until 11 September 2017. To this date, 10,802 students were enrolled. However, only a fraction (2387 students) were active, meaning that they have watched videos, done exams/readings or written comments. Fig. 3 presents the evolution of active students per lecture. As expected, the number of involved students decreases every week which is in line with the fact that the completion rate is between 2% and 10%[4], 216 in this case. However, the number of students from lecture 4 on can be considered constant in terms of doing exams and readings. Note that in this MOOC Lectures 3, 4 and 6 are the most difficult ones and thus, in line with achieved results.

When the MOOC was opened teachers assisted students for 2 months and in this period of time the amount of generated comments is depicted in Fig. 4. Regarding exams, most

[4] http://www.katyjordan.com/MOOCproject.html, last access March 2018.

Table 3 – MOOC description summary.

Lecture	Lecture length (weeks)	# of videos	Video length (min)	# of readings
1	0.5	6	29.0	0
2	0.5	11	29.7	1
3	1	20	88.6	1
4	1	17	94.8	3
5	1	13	42.8	2
6	2	26	133.7	3
Total	6	93	418.6	10

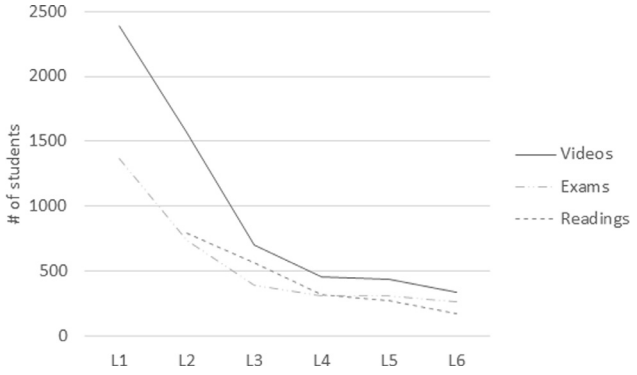

Fig. 3 – Number of students per lecture for exams, readings and videos.

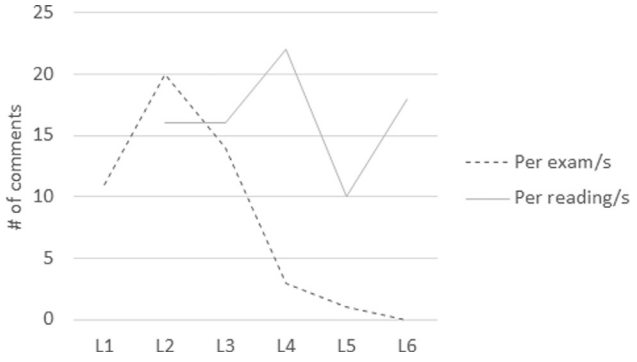

Fig. 4 – Number of comments per lecture per exams and readings.

comments were written at the very beginning of the course. Most of them were related to errors in several questions. As the course went on, teachers improved exams and the amount of comments decreased in the last lectures. A different situation is identified in readings, several comments were performed in readings of all lectures and though teachers did their best to present clear questions, they should still be improved for future editions of the MOOC.

4.2. *Analysing cybersecurity audience*

Cybersecurity interest is measured in terms of the amount of enrolled students. It is different depending on the targeted feature, namely country, age, gender and educational level.

Concerning countries, Fig. 5a, IN is the leader as there are 2,917 students enrolled, followed by US with 1,418 users and SP, GB, PK and NG. This is in line with previous results, which suggest that US and some European students enroll in MOOCs more than other nationalities, with the increasing participation of CN and IN (EDUCATION et al., 2016). Our results are also consistent with findings from the World Economic Forum, which states that US and IN are among the top 5 countries with more science, technology, engineering and mathematics graduates.

Gender speaks for itself, as shown in Fig. 5b. The amount of males is significantly larger then females, 56.5% more. This is a common pattern in the computer science world (Reuben et al., 2014) and in online courses (EDUCATION et al., 2016), and this MOOC is not an exception.

Number of students per age is also in line with expectations, Fig. 5c. The majority of the audience is between 21–40 years old which is the most common age for learning. (OECD) points out that around 20 is the common age to go to university and around 24 to be graduated. After been graduated, people usually look for a job, where the average age of students in the labor force is between 23 and 38 (OECD). Additionally, EDUCATION et al. (2016) highlights that MOOC students are around 30 years old.

The educational level highlights that Bachelor (B) and high school (HS) students, followed by Master (M) ones, correspond to the majority of students enrolled in this course, Fig. 5d. Based on the intermediate level of the MOOC the expected audience is B and M students but HS ones, which are the second larger group, though they should work harder, can also learn interesting things. Indeed, as it is the case, 60% of MOOC students have completed a bachelor (EDUCATION et al., 2016). The point is that contents of the MOOC can be easily understood but a really deep understanding can only be achieved by those with some computer science knowledge, presumably students with bachelor or higher educational level.

4.3. *Analysis of students' behavioral trends*

Performed comments and the percentage of time during which videos have been watched help to estimate the students' participation and to what extent they are really interested in learning cybersecurity.

4.3.1. *Comments analysis*

Students have written 514 comments and 242 comments more have been written by teachers in response to students' ones.

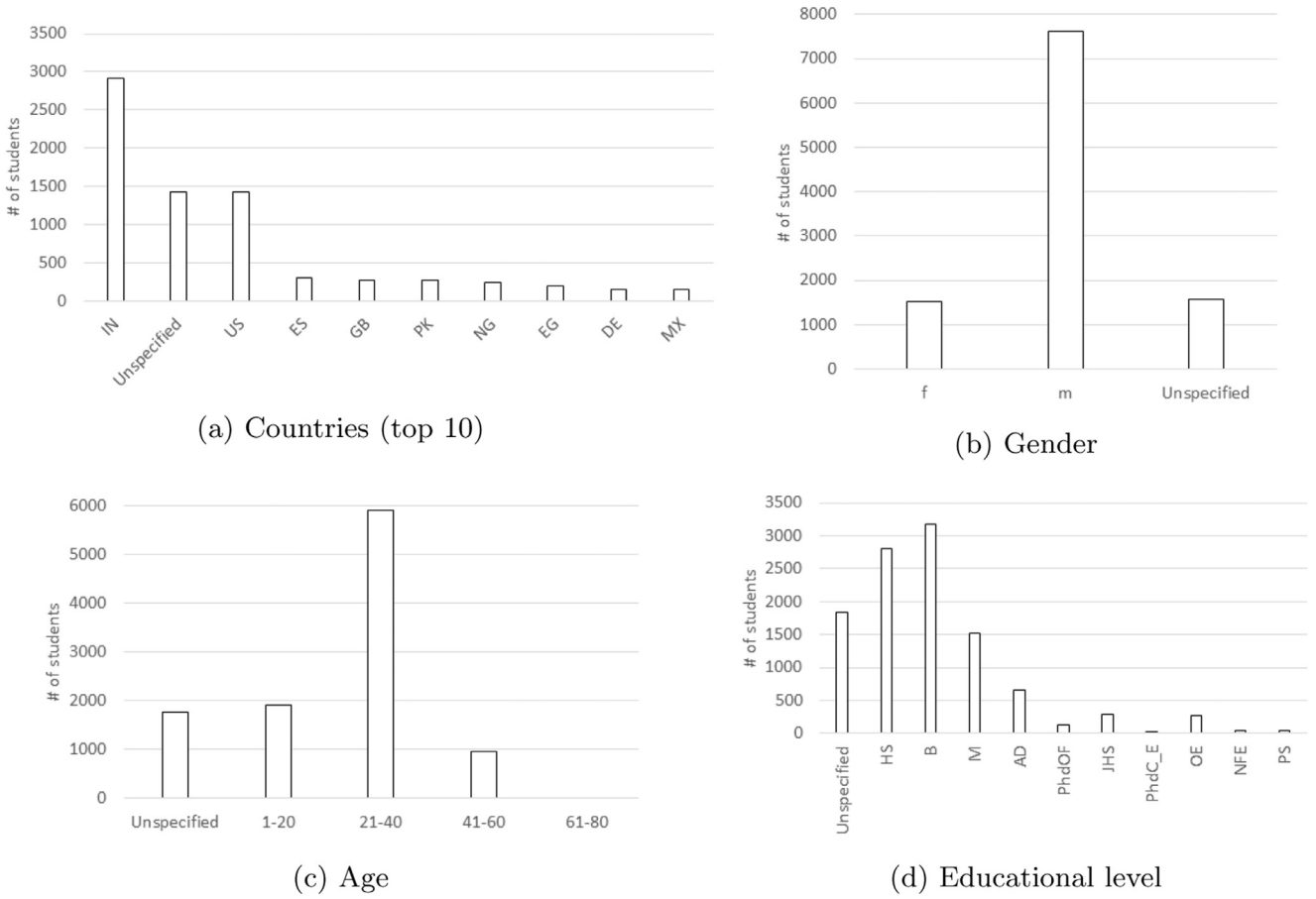

Fig. 5 – Cybersecurity audience characterization.

In line with cybersecurity interest, comments per country show that US, ES and IN are the most active countries, followed very closely by CA and MX, Fig. 6a. As most users came from US, it is not surprising that the highest number of comments are written by US students, 103 comments in total. However, users from ES and IN have written multiple comments as well, that is 46 and 44 respectively. Note that this study includes countries whose students have written more than 5 comments as the remaining ones are not considered representative enough.

Concerning gender and age, Fig. 6b and c, and in line with the reasoning of Section 4.2, males and students between 21–40 are the most participatory ones.

Quite different is the case of education level, Fig. 6d, M students are those more willing to participate. These students are older that those in HS or doing a B and thus, more mature and commonly more demanding.

4.3.2. *Watched videos analysis*

The average duration of videos is 270 sc (4.5 min) and, on average, videos have been watched during 37.8% of time, that is 102 sc (1.7 min). Videos of Lecture 1 are watched during more time than the remaining ones. This is a sensible result because this lecture is the initial contact of students with the MOOC and in the remaining lectures, apart for having more videos, the difficulty level increases.

In terms of countries, Fig. 7a, just those in which the average exceeds 30% in every lecture are considered. The first interesting issue is that among the top ten countries with more enrolled students (recall Fig. 5a), IN, PK, NG and MX (in bold) do not appear in this Figure. CD is the only country in which students have watched videos almost to the end, 94.3% on average and only 3 countries, NL, FI and SE have watched videos until 50% and 70% in all lectures. In the remaining countries all videos are watched, approximately, the same amount of time in every lecture, being surprising the case of US, ES and GB. This three last countries, in spite of being in the top ten list of most enrolled students, videos are watched 53.7% of time.

In terms of gender it is clear that males, Fig. 7b, apart from those whose gender is not specified, watch videos during 48.7% of videos' length, in contrast to 41.3% in case of females. Besides, videos of all lectures except for the former one are watched a similar amount of time.

More differences are noticed in the analysis of age, Fig. 7c. All students, with the exception of those between 1 and 20 years old, watch videos during a similar amount of time in all lectures.

Studying educational level, Fig. 7d, Phd and M students followed by HS ones, watch videos during more time than the remaining students, 52.3%, 51.8% and 49.6% on average, respectively. As in gender, videos are watched during similar time regardless of the lesson.

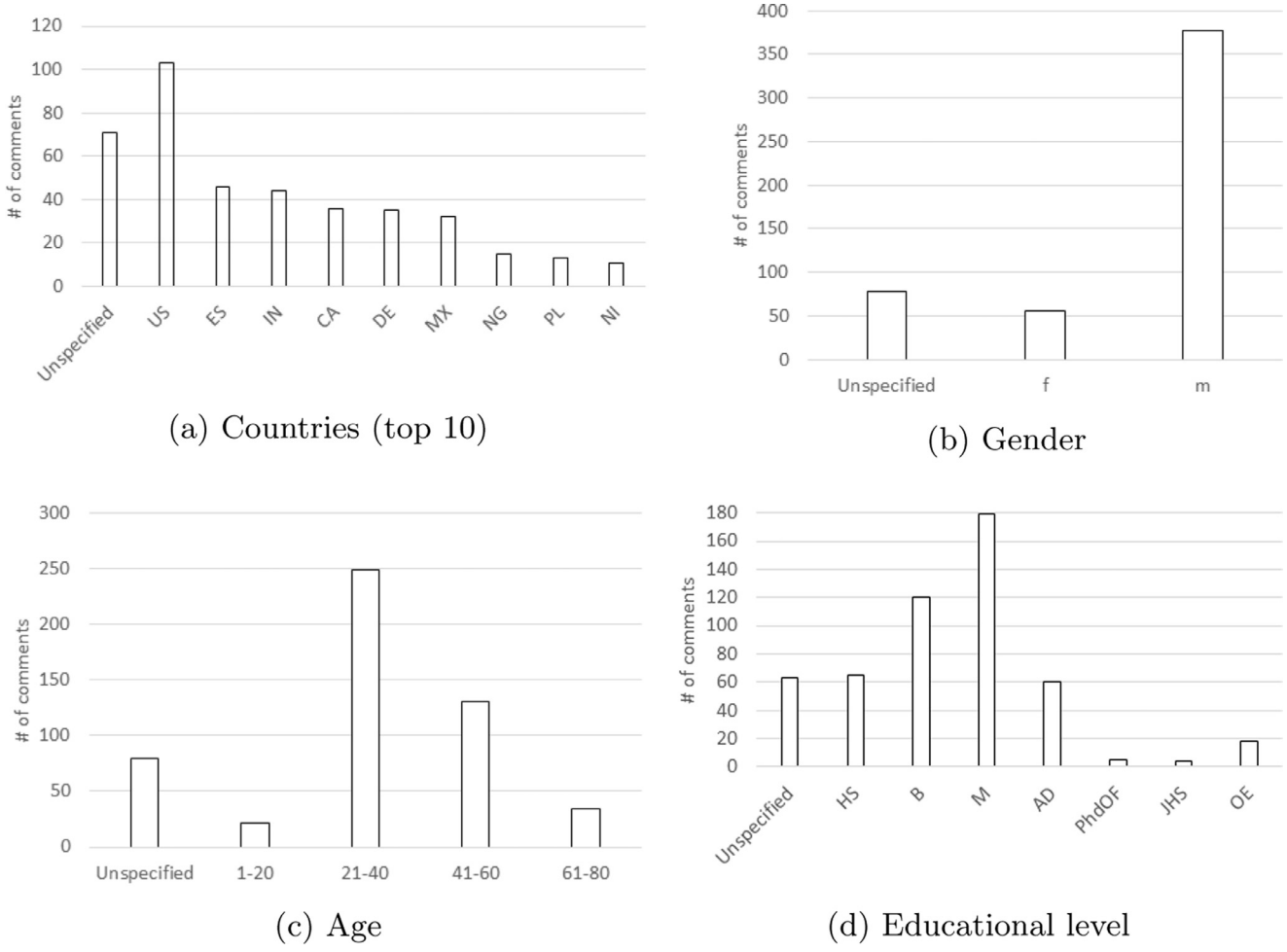

Fig. 6 – Students' behavioral trends: number of comments.

4.4. *Analysis of students' performance*

The analysis of marks leads us to study students' performance. Firstly, average marks obtained in exams and readings of each lessons are studied, to analyse average marks of the final exam afterwards.

4.4.1. *Marks of exams and readings analysis*

Marks in graded activities (readings and exams) per lecture are reasonably good. The average marks are 6.6 and 6.62 (out of 10) for readings and exams respectively. Marks of readings are similar in all lectures. However, they are different for exams, being significantly high in the first exam, 8 on average, and surprisingly low in the last one, 3 on average. One possible reason behind this is that the first Lecture is the easiest one, opposite to Lecture 6 which without adequate time and effort passing Exam6 is not that easy.

Most countries get a very good average mark in exams, Fig. 8a, 56.1% of them between 7 and 9 and 31.4% between 5 and 7. Students from 105 different countries have done exams and just ZA, UY, NZ, RS and SI have failed them. A potential reason is that ZA, UY, SI and NZ are not within those countries which have watched videos during a significant amount of time. RS presents surprising results because students in this country have watched videos during 80% of the time, thus the level of the course seems to be unsuitable for them. It is also noteworthy that countries with more number of students (in bold) have marks between 7 and 9 which means that they are more committed to this course.

Readings present worse results, Fig. 9a, in this case considering students from 97 different countries. Students from 32.9% of the countries have failed exams and 50.5% have passed with marks between 5 and 7. In light of countries with more students (in bold), results are similar to exams, except for EG, all students have passed readings and many of them with really good marks, that is between 7 and 9 those from MX, GB and IN. Nonetheless, the number of comments in readings (recall Section 4.3.1) leads to conclude that attached questions should be improved to prevent mistakes or confused answers.

Talking about gender, females seem to surpass marks of males, Fig. 8b. In exams, marks of both genders are similar, though females get, on average, 0.1 marks more. This difference is more acute in readings, Fig. 9b, in which females get, on average, 0.6 marks more than males. This result could be related to self discipline as some studies point out that females have more self-discipline than males, though it is not clear in adulthood (Duckworth and Seligman, 2006).

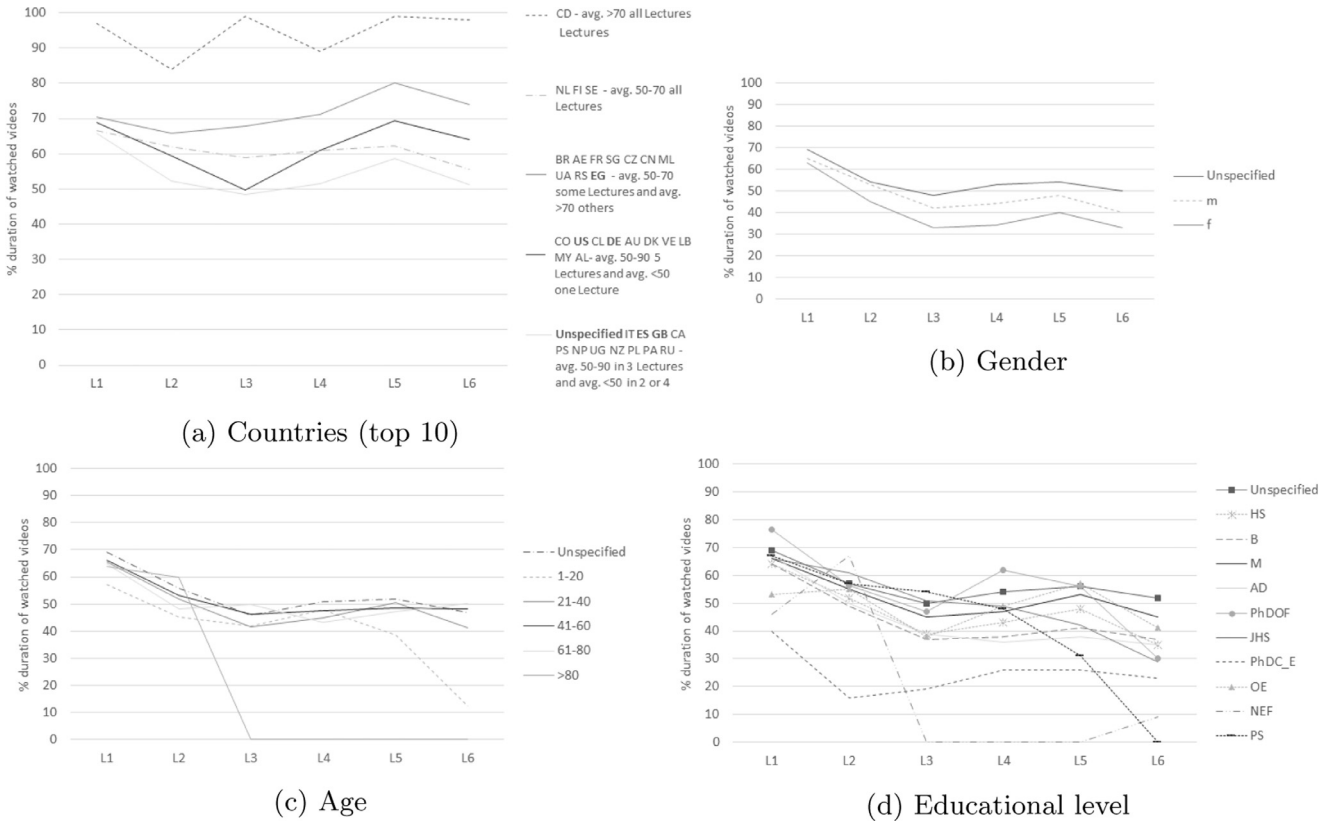

Fig. 7 – Students' behavioral trends: % duration of watched videos.

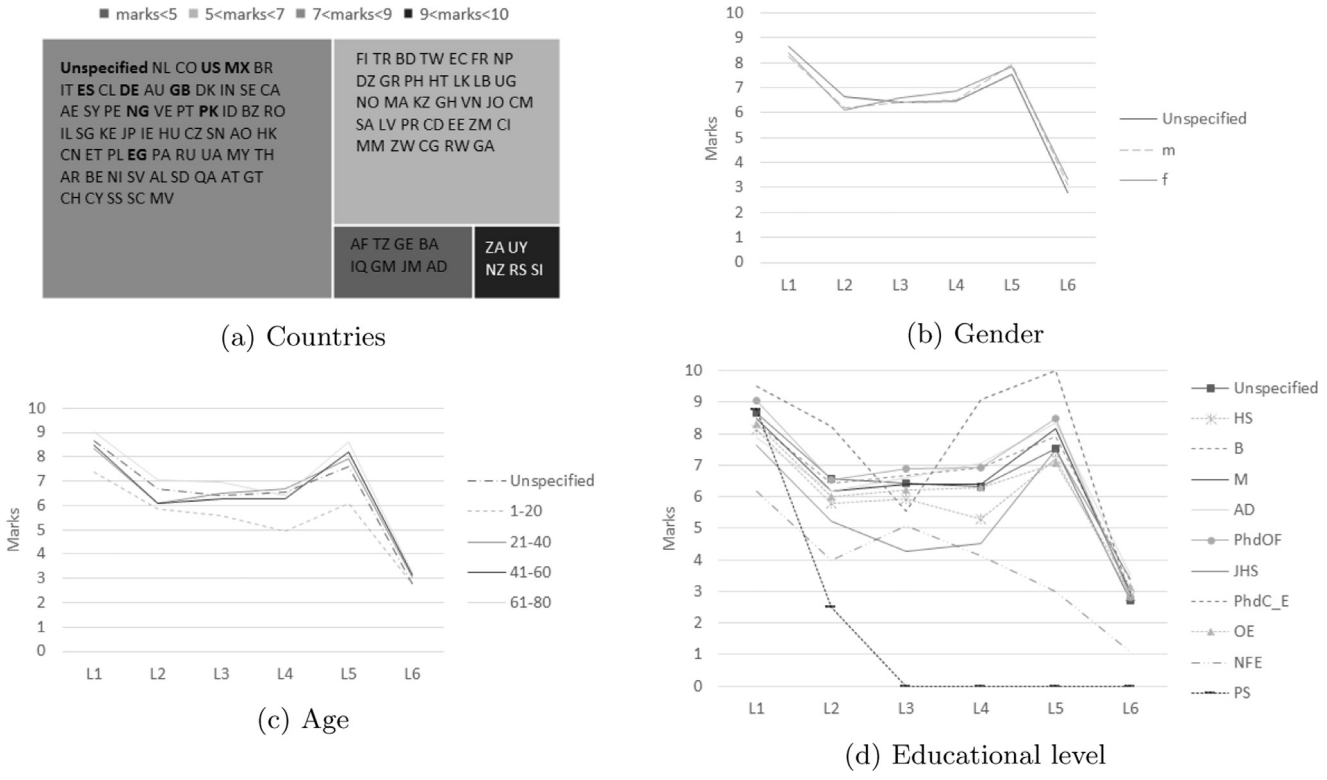

Fig. 8 – Students' commitment: marks per lesson in exams.

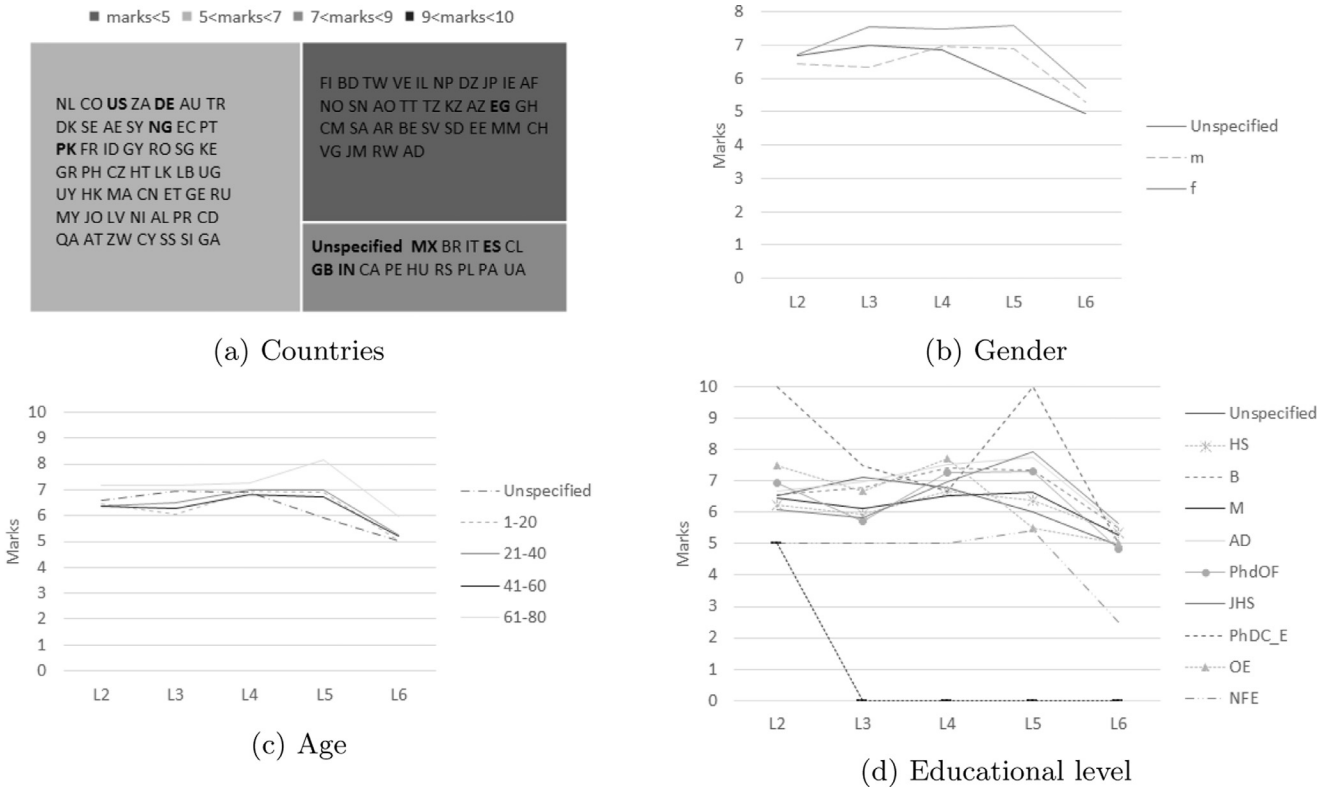

Fig. 9 – Students' commitment: marks per lesson in readings.

Age is not a high differentiating factor, Fig. 8c, neither for exams nor for readings. In exams just those students under 20 years old get worse marks, probably due to the intermediate level of this course. By contrast, marks in readings are extremely similar among all ages, Fig. 9c. The main reason could be that readings have questions tightly close to their content without requiring a deep knowledge of each lesson. Indeed, this is in line with the fact that students with less than 20 years old watch videos during less amount of time than older students. Other noticeable issue is that students over 61 years old present slightly better results in all cases.

Contrary to age, the academic level shows that those with a degree, master or doctorate get better results specially in exams, Fig. 8d, which follows our expectations based on the level of the MOOC. However, HS students also get really good marks and, except for Lecture 4, results are comparable with B students and readings marks with M ones.

4.4.2. *Marks of the final exam analysis*

As a final part of the MOOC, students do a final exam. The average mark is 5.3. This mark is quite low but this exam involves all contents of the MOOC and students should make a great effort to pass it. Results are in most cases comparable to those achieved in lectures' exams.

In terms of countries, Fig. 10a, students from 59 of them have done the final exam. Those with more enrolled students (in bold) have better marks. IN should be highlighted within this group because the average mark is between 7 and 9, pointing out the interest of Indian students in this course.

Gender follows a pattern similar than lectures' exams, Fig. 10b, females surpass males but for 0.78 marks. This is opposite to the fact that females watch videos during less time than males.

A similar situation happens with age, Fig. 10c, all ages present similar results to lectures' exams and just those over 61 years old get a bit higher marks, 1.7 marks more.

Educational level does not significantly affect the final exam and only those with Phd. in science or engineering (PhdC_E) stand out from the rest, Fig. 10d, being surprising marks of junior high school (JHS) which are comparable to B students and even higher than M ones. However, given that most of students have B, M and HS educational level, we have done a deeper analysis and we have identified that only one student with a PhdC_E and another one with JHS did the final exam. Thus, both results are not representative but those from M, B and HS which have got 5.5, 6.3 and 5.1 marks respectively.

5. Summary of recommendations

After presenting the coverage of the NICE framework by existing courses and having inspected several issues of the student behavior in the cybersecurity MOOC, a set of recommendations can be proposed. They are intended to guide the training

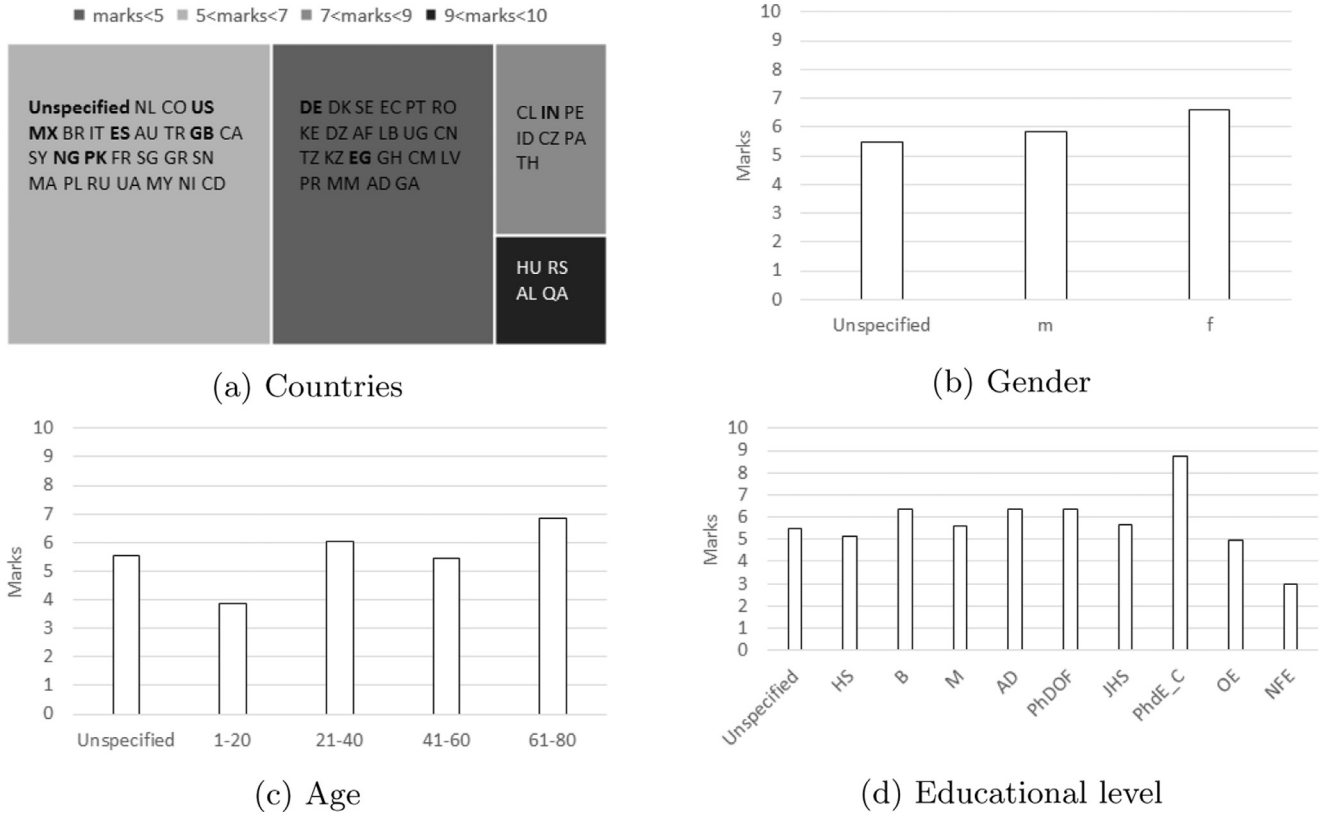

Fig. 10 – Students' performance: marks final exam.

community in designing more effective cybersecurity courses. This catalogue is divided into the following pair of categories:

Recommendations to consider before the development of the course

- *Topic choice*. Supported by our findings in NICE coverage, it is critical to point towards those issues that have received less attention. For example, courses regarding collection operations or cyber operational planning are nowadays lacking.
- *Level choice*. Our analysis on existing courses has revealed that different levels are offered for each specialty area. In particular, some areas (e.g. strategic planning and policy) have been taught for intermediate and advanced levels, but not for beginner ones.
- *Teachers commitment*. MOOC analysis shows that if teachers pay attention to comments when the course is just opened and try to improve it at the very beginning, students have less comments along the remaining parts of the course and this could affect their marks.
- *Gender adaptation*. According to our MOOC indicators, women underrepresentation is not due to any learning obstacle. Therefore, difficulty is not an issue but attractiveness to the topic could be a potential reason instead.
- *Regional adaptation*. Students from IN and US are interesting targets in light of the MOOC analysis and in line with existing works (EDUCATION et al., 2016). If enrolled students do not follow this pattern, it could be the case that marketing campaigns are not focused on the right audience. On the other hand, be aware that some communities such as China may have language-, background- or cultural-related issues that have to be considered. Thus, if the MOOC coverage is intended to include these sectors, MOOC contents have to be designed to be suitable for them.
- *Interaction design*. If the course is intended to have some form of participation or interactivity, the audience has to be considered. Our MOOC findings show that whereas some countries are prone to interaction (e.g. US, IN, ES), others are less likely to take an active role (e.g. PK, NG).
- *Age impact*. Students between 20 and 40 are most committed to the analysed MOOC, as well as those older than 60. This issue depends on the level of the course. The analysed MOOC was intermediate level and then, students younger than 20 are not the main target.

Recommendations to consider after the first edition of the course

- *Suitability of content design*. Analysing results of graded activities, together with students' commitment, helps in making decisions about the appropriateness of the content vs level of the course. Results of our MOOC show that marks are acceptable in regard to performed comments and the amount of time students have watched videos. Then, contents and level seem to be appropriate. However, RS students are an exception because they fail exams even

watching videos towards the end (80%). These students may require some additional material, as well as watching videos completely to ensure an appropriate learning process.

- *Activities review*. Those activities which many students fail or get low marks, should be reviewed. This is the case of readings in the studied MOOC in which multiple comments in readings' questions have come up.
- *Students videos commitment*. If videos are not watched towards the end, it could be required to a) decrease videos' length, as small videos can be easier for students, or b) create more attractive videos to encourage students to watch them. In the studied MOOC, on average, the length of each video is of 4.5min, as this is an appropriate length[5], the development of more attractive videos should be considered.

6. Related work

Many works propose cybersecurity games, such as capture the flag, to learn this discipline (Boopathi et al., 2015; Cone et al., 2007; Trickel et al., 2017). Similarly, Mirkovic and Benzel (2012) presents a laboratory for cybersecurity education purposes and Pusey et al. (2016) analyses a cybersecurity competition together with the benefits of students interactions. Just Ferrer Mico et al. (2016) present a study of an introduction to cybersecurity MOOC. They mainly analyse the success of the course based on their 3320 participants. Our recommendations are focused on helping the designer on choosing a topic that deserves attention and preparing the course in a suitable way based on different aspects of the audience and results from previous editions (if any).

Beuran et al. propose a set of requirements to make a cybersecurity programme be successful (Beuran et al., 2016). In particular, they impose five conditions: (i) must be appropriate for the target audience in terms of knowledge and ability levels, (ii) its content must be in accordance with the intended skills, (iii) it should use hands-on activities, (iv) it should reach as large an audience as possible, and (v) it must be sustainable in the long term (i.e. good cost/performance). In our work these conditions are indirectly considered because we provide guidelines to address some of them, e.g. (i), and some others are implicitly managed in a MOOC, e.g. (iv). Cultural and behavioral issues have been left out of the scope, but Wang et al. have already pointed out that audience's culture has impact on the way of learning in online courses (Wang, 2007). Therefore, in this paper we propose a complementary set of recommendations that contribute to the success of a cybersecurity training action.

Patriciu and Furtuna proposed a set of guidelines to design cyber security exercises (Patriciu and Furtuna, 2009). For this purpose, several steps are identified – defining the objectives, choosing an approach, designing network topology, creating a scenario, establishing a set of rules, choosing appropriate metrics and learning lessons. As opposed to this paper, our set of guidelines are inspired by experimental results of a MOOC. Moreover, we make recommendations that also take into account the audience background and profile. Indeed, Wang et al. have already pointed out that audience's culture has impact on the way of learning in online courses (Wang, 2007). In this sense, our work is similar to Bashir et al. proposal (Bashir et al., 2015), in which they aim to develop a demographic, psychological, cultural and vocational profile of participants in cybersecurity competitions. From the methodological viewpoint, there are significant differences between their approach and ours. Concerning their working data, their approach is based on a 229-question survey, independent from the competition. In our case, data is extracted from students' participation, with no requirements for explicit additional interaction. Moreover, their output is related to understanding whether competitions are effective recruiting tools in cybersecurity. In our case, we aim to improve the design of future training courses.

The analysis on the success of a cybersecurity education activity has been explored by Vasserman et al. (2015). In particular, they focus on measuring students interest and self-efficacy. Our work is orthogonal to theirs in that we focus on how to prepare the course whereas they assess its successfulness. Therefore, their approach could serve as a future way of assessing the effectiveness of courses adopting our recommendations.

7. Conclusion

There is a current demand for cybersecurity professionals and many courses have been developed to address this issue. Online courses are specially interesting as they reach out more people. This paper presents an analysis of 35 cybersecurity online courses concerning NICE framework to help in the selection of topics while preparing a cybersecurity course. Results show that there are gaps to patch, thus place for many more new courses. Additionally, guidelines on the way to prepare a cybersecurity course are presented, all of them based on the analysis of a cybersecurity edX MOOC with +2,000 active users. A framework for the analysis of edX courses have been also released to promote the research in this direction.

Future work will be focused on addressing some limitations of the current work. In particular, one interesting issue is to analyze students' recurrence in different courses. This will illustrate whether prior training activities are enough to prepare for subsequent ones.

Acknowledgments

This work was supported by theMINECO grant TIN2016-79095-C2-2-R (SMOG-DEV) and by the CAM grant S2013/ICE-3095 (CIBERDINE: Cybersecurity, Data, and Risks) co-funded with European FEDER funds.

Appendix A. Summary of SAs per courses

Table A.4 presents SAs addressed per course.

[5] http://blog.edx.org/optimal-video-length-student-engagement , last access March 2018.

Table A.4 – SAs per courses. B, I, A an G refer to level of courses, being X undefined.

	SA	Hardware Security	Software Security	Usable Security	Internet History, Technology, and Security	Malicious Software and its Underground Economy: Two Sides to Every Story	Cybersecurity and Its Ten Domains	Cryptography I	Information Security: Context and Introduction	Cryptography (U. Maryland)	Networking and security in iOS applications	Homeland security and cybersecurity	Ethical hacking and penetration testing	Certified Information Systems Security Professional (CISSP)	Python for Security Professionals	Advanced Penetration Testing	CompTIA CASP	CompTIA Network +	Secure coding	Intro to Cryptography	Network and Computer Security	Computer Systems Security	Introduction to cybersecurity	Cybersecurity Fundamentals	Network security	Introduction to cybersecurity	Cybersecurity - the CISO's view	Building a cybersecurity toolkit	Cyber security economics	Cyberwar, Surveillance and Security	Intro to information security	Cyber-Physical Systems Security	Network security	Applied cryptography	Computer networking	**Cyber Security Basics: A Hands-on Approach**
SP	RSK						B		B					A			A								A				B	I							
	DEV		I	I							X					A			I														I			
	ARC																A																			
	TRD																																			
	SRP																																			
	TST		I											I								G														I
	SYS		I		B									I	A	A		A				G									I					I
OM	DTA																																			I
	KMG																																			
	STS																																			
	NET						B		B								A	B				G	B	A	A						I	I	I			I
	ADM	I							B															A												
	ANA																																			
	LGA			I			B																			B				B	I					
	TEA								B																											
OV	MGT							I		I				A						X	G		B	A							I			A	I	
	SPP													A			A												I							
	EXL													A													B									
	PMA and Acq.													A																						
PR	CDA					B											A												I			I				
	INF	I	I														A	B																		
	CIR																																			
	VAM		I											I		A	A						G													I
	TWA		I			B								I										A		B									I
	EXP		I											I	A		A											B				I				I
AN	ASA																														I					
	TGT												B																	B						
	LNG					B																														
CO	CLO																																			
	OPL																																			
	OPS																							B						B						I
IV	INV																											B								I
	FOR													A																						I

Appendix B. Acronyms: countries and educational level

Tables B.5 and B.6 presents countries and education level notation.

Table B.5 – Countries acronyms.

Acronym	Description	Acronym	Description	Acronym	Description	Acronym	Description
AE	United Arab Emirates	EG	Egypt	LB	Lebanon	RS	Serbia
AF	Afghanistan	EH	Western Sahara	LC	Saint Lucia	RU	Russian Federation
AG	Antigua and Barbuda	ER	Eritrea	LI	Liechtenstein	RW	Rwanda
AI	Anguilla	ES	Spain	LK	Sri Lanka	SA	Saudi Arabia
AL	Albania	ET	Ethiopia	LR	Liberia	SB	Solomon Islands
AM	Armenia	FI	Finland	LS	Lesotho	SC	Seychelles
AO	Angola	FJ	Fiji	LT	Lithuania	SD	Sudan
AQ	Antarctica	FK	Falkland Islands (Malvinas)	LU	Luxembourg	SE	Sweden
AR	Argentina	FM	Micronesia, Federated States of	LV	Latvia	SG	Singapore
AS	American Samoa	FO	Faroe Islands	LY	Libya	SH	S. Helena, Ascension and Tristan da Cunha
AT	Austria	FR	France	MA	Morocco	SI	Slovenia
AU	Australia	GA	Gabon	MC	Monaco	SJ	Svalbard and Jan Mayen
AW	Aruba	GB	UK	MD	Moldova, Rep. of	SK	Slovakia
AX	Aland Islands !	GB	United Kingdom	ME	Montenegro	SL	Sierra Leone
AZ	Azerbaijan	GD	Grenada	MF	Saint Martin (French part)	SM	San Marino
BA	Bosnia and Herzegovina	GE	Georgia	MG	Madagascar	SN	Senegal
BB	Barbados	GF	French Guiana	MH	Marshall Islands	SO	Somalia
BD	Bangladesh	GG	Guernsey	MK	Macedonia, the former Yugoslav Rep. of	SR	Suriname
BE	Belgium	GH	Ghana	ML	Mali	SS	South Sudan
BF	Burkina Faso	GI	Gibraltar	MM	Myanmar	ST	Sao Tome and Principe
BG	Bulgaria	GL	Greenland	MN	Mongolia	SV	El Salvador
BH	Bahrain	GM	Gambia	MO	Macao	SX	Sint Maarten (Dutch part)
BI	Burundi	GN	Guinea	MP	Northern Mariana Islands	SY	Syrian Arab Rep.
BJ	Benin	GP	Guadeloupe	MQ	Martinique	SZ	Swaziland
BL	Saint Barthélemy	GQ	Equatorial Guinea	MR	Mauritania	TC	Turks and Caicos Islands
BM	Bermuda	GR	Greece	MS	Montserrat	TD	Chad
BN	Brunei Darussalam	GS	South Georgia and Sandwich Islands	MT	Malta	TF	French Southern Territories
BO	Bolivia, Plurinational State of	GT	Guatemala	MU	Mauritius	TG	Togo
BQ	Bonaire, Sint Eustatius and Saba	GU	Guam	MV	Maldives	TH	Thailand
BR	Brazil	GW	Guinea-Bissau	MW	Malawi	TJ	Tajikistan
BS	Bahamas	GY	Guyana	MX	Mexico	TK	Tokelau
BT	Bhutan	HK	Hong Kong	MY	Malaysia	TL	Timor-Leste
BV	Bouvet Island	HM	Heard Island and McDonald Islands	MZ	Mozambique	TM	Turkmenistan
BW	Botswana	HN	Honduras	NA	Namibia	TN	Tunisia
BY	Belarus	HR	Croatia	NC	New Caledonia	TO	Tonga

(*continued on next page*)

Table B.5 (*continued*)

Acronym	Description	Acronym	Description	Acronym	Description	Acronym	Description
BZ	Belize	HT	Haiti	NE	Niger	TR	Turkey
CA	Canada	HU	Hungary	NF	Norfolk Island	TT	Trinidad and Tobago
CC	Cocos (Keeling) Islands	ID	Indonesia	NG	Nigeria	TV	Tuvalu
CD	Congo, the Democratic Rep. of the	IE	Ireland	NI	Nicaragua	TW	Taiwan, Province of China
CF	Central African Rep.	IL	Israel	NL	Netherlands	TZ	Tanzania, United Rep. of
CG	Congo	IM	Isle of Man	NO	Norway	UA	Ukraine
CH	Switzerland	IN	India	NP	Nepal	UG	Uganda
CI	Cote d'Ivoire !	IO	British Indian Ocean Territory	NR	Nauru	UM	United States Minor Outlying Islands
CK	Cook Islands	IQ	Iraq	NU	Niue	US	United States
CL	Chile	IR	Iran, Islamic Rep. of	NZ	New Zealand	UY	Uruguay
CM	Cameroon	IS	Iceland	OM	Oman	UZ	Uzbekistan
CN	China	IT	Italy	PA	Panama	VA	Holy See (Vatican City State)
CO	Colombia	JE	Jersey	PE	Peru	VC	Saint Vincent and the Grenadines
CR	Costa Rica	JM	Jamaica	PF	French Polynesia	VE	Venezuela, Bolivarian Rep. of
CU	Cuba	JO	Jordan	PG	Papua New Guinea	VG	Virgin Islands, British
CV	Cape Verde	JP	Japan	PH	Philippines	VI	Virgin Islands, U.S.
CW	Curaçao	KE	Kenya	PK	Pakistan	VN	Viet Nam
CX	Christmas Island	KG	Kyrgyzstan	PL	Poland	VU	Vanuatu
CY	Cyprus	KH	Cambodia	PM	Saint Pierre and Miquelon	WF	Wallis and Futuna
CZ	Czech Rep.	KI	Kiribati	PN	Pitcairn	WS	Samoa
DE	Germany	KM	Comoros	PR	Puerto Rico	XK	Kosovo
DJ	Djibouti	KN	Saint Kitts and Nevis	PS	Palestine, State of	YE	Yemen
DK	Denmark	KP	Korea, Democratic People's Rep. of	PT	Portugal	YT	Mayotte
DM	Dominica	KR	Korea, Rep. of	PW	Palau	ZA	South Africa
DO	Dominican Rep.	KW	Kuwait	PY	Paraguay	ZM	Zambia
DZ	Algeria	KY	Cayman Islands	QA	Qatar	ZW	Zimbabwe
EC	Ecuador	KZ	Kazakhstan	RE	Reunion !		
EE	Estonia	LA	Lao People's Democratic Rep.	RO	Romania		

Table B.6 – Educational level acronyms.

Acronym	Description
M	Master
B	Bachelor
AD	Associate degree
HS	Secondary/High school
JHS	Junior secondary/junior high/middle school
PS	Elementary/primary school
NFE	No formal education
PhdOF	PhD other field
PhdC_E	PhD in science or engineering
OE	Other education

Appendix C. Appendix. EdX advanced analytics framework

For the sake of comparison and in order to ease the process of extracting MOOC performance data, we hereby present the proposed framework. It is presented as a Python open-source framework freely available on GitHub[6]

This framework is composed of a pair of folders, *scripts* and *statistics* together with a pair of execution files, *Structure-Development_edX.cmd* and *Statistics_edX.cmd*. On the one hand, *scripts* folder contains all required scripts to create a sqlite3 database called *moocDB.db*, see Fig. C.11, with all required data stored in it. These scripts should be initially executed

[6] https://www.dropbox.com/sh/x6f78doafv8efkt/AADVh-W_3AhtskKs9GvjmR7Sa?dl=0 , once accepted it will be uploaded to GitHub.

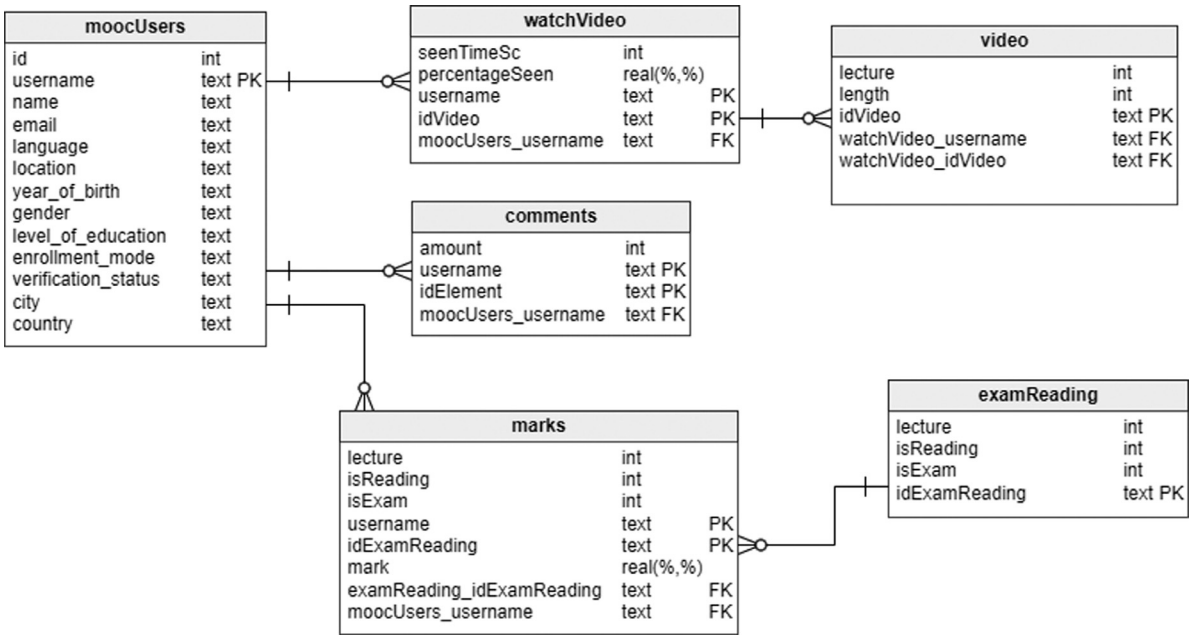

Fig. C.11 – Database *moocDB.db* definition.

running *StructureDevelopment_edX.cmd*. On the other hand, *statistics* folder contains a set of scripts to obtain statistics from data. These scripts can be executed after the creation of the database through the use of *Statistics_edX.cmd*. Note that execution files are prepared for Windows operating systems but as they run Python, they can be easily adapted to any other operating system.

The execution of *StructureDevelopment_edX.cmd* involves five parameters which refer to the location of (1) *scripts* folder; (2) *Data* folder; (3) the number of exams in total in all lectures (not including the final exam); (4) the number of readings in total in all lectures.; and (5) folder of python.exe. Specifically, *Data* folder should contain the following files, some of them can be directly downloaded from edX platform and some others should be pre-processed or created:

- *Data_tables_videos.csv* should be created. It should contain three elements separated by commas, namely an identifier of each video, the lecture linked to each video and the length in seconds of each video. The first line of the file should contain *idVideo,lecture,length*.
- *Data_tables_examsReadings.csv* should be created. It should contain four elements separated by commas, namely the identifier given to each particular exam and/or reading, the lecture linked to any of them, 1 if it is a reading or 0 otherwise and 1 if it is an exam or 0 otherwise. The first line of the file should contain *idExamReading,lecture,isReading,isExam*. In this case, exams and readings should have the following form respectively ExamL+NUM_EXAM, e.g. ExamL4, and Reading+NUM_READING, e.g. Reading8.
- *changeNamesFromedXtoidVideo.csv* should be created. It should contain three elements separated by commas, namely the identifier of each video, its duration in seconds and the edX identifier which refers to a string between 20–32 random letters and numbers, e.g. 1c5fk6i01n7gttl17upof71tskv. The edX identifier can be found in each video and/ or reading clicking on 'STAFF DEBUG INFO'. The first line of this file should contain *idVideo,duration,idedX*.
- *Data_tables_marks.csv* should contain the list of marks delivered by edX through "Instructor > Download Data > Generate Grade Report". The first line should be updated according to identifiers of exams and readings written in **Data_tables_examsReadings.csv**.
- *Data_tables_comments.csv* should contain comments per user and element (namely, video, exam and reading). This data is not provided by edX so it has to be manually collected and located in a csv file in which the first element of each row refers to the username given in edX, followed by the amount of comments per element. The first line should contain the username followed by identifiers of videos, readings and exams according to files **Data_tables_examsReadings.csv** and **Data_tables_videos.csv**.
- *usersProfilePreprocessed.csv* should contain data provided by edX in "Instructor > Download Data > Download Profile Information as a csv". The downloaded file should be processed, e.g. using text editors like Notepad[7] and Apache OpenOffice Calc[8], such that elements *goals* and *mailing_address* are removed. Similarly, commas within a particular element, double quotes or simple quotes, e.g. from usernames, should be also removed.
- *Videos folder* should contain data provided by edX. Data of each video should be individually downloaded, i.e. "Instructor > Download Data > 'location from STAFF

[7] https://notepad-plus-plus.org/ , last access March 2018.
[8] https://www.openoffice.org/es/producto/calc.html , last access March 2018.

DEBUG INFO' > Download a csv from problem responses". Files corresponding to all videos of the course, once downloaded, should be delivered in this folder.

After the database is created, *Statistics_edX.cmd* can be executed, passing as parameters the location of 1) *scripts* folder; 2) *Results* folder; and 3) python.exe folder. The second folder is the one in which results from executing all statistics are stored. In particular, the following files will be generated as a result of executing *Statistics_edX.cmd*:

- *WhoHasPerformedComments.txt* contains per country, gender, level of education and year of birth, the amount of performed comments.
- *WhoHasCybersecurityInterest.txt* contains per country, gender, level of education and year of birth, the amount of users enrolled in the course.
- *WhoHasAnsweredBetterToExamsReadings.txt* contains per country, gender, level of education and year of birth, the average marks per exams and readings of each lecture.
- *WhoHasAnsweredBetterFinalExam.txt* contains per country, gender, level of education and year of birth, the average mark of the final exam. Note that this statistics can only be achieved if the course has a final exam and if in file **Data_tables_marks.csv** an element *FinalExam* is included.
- *WhoHasSeenMoreVideosPerLecture.txt* contains per country, gender, level of education and year of birth, the average percentage of time videos have been watched per lecture.
- *StatisticQueries.txt* contains data about the total amount of comments, videos and users; the average percentage of time students have watched videos per lecture; number of students who have done exams per lecture; and the number of students who have done the final exam.

Note that *Statistics_edX.cmd* executes a Python script per statistic and it is extensible, such that considering the structure of the database (recall Fig. C.11), new scripts can be created to get desired statistics.

REFERENCES

Academy K. Intro to cryptography. https://www.khanacademy.org/computing/computer-science/cryptography/crypt/v/intro-to-cryptography; 2011.

U. of Adelaide. Cyberwar, surveillance and security. https://www.edx.org/course/cyberwar-surveillance-security-adelaidex-cyber101x-0#!; N/A.

Bashir M, Lambert A, Wee JMC, Guo B. An examination of the vocational and psychological characteristics of cybersecurity competition participants. 2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15); 2015.

Beuran R, Chinen KI, Tan Y, Shinoda Y. Towards effective cybersecurity education and training; 2016.

Boopathi K, Sreejith S, Bithin A. Learning cyber security through gamification. Indian Journal of Science and Technology 2015;8(7):642–9.

Cabaj K, Domingos D, Kotulski Z, Respício A. Cybersecurity education: evolution of the discipline and analysis of master programs. Computers & Security 2018.

U. of Colorado. Homeland security and cybersecurity. https://www.coursera.org/specializations/homeland-security-cybersecurity; 2017.

Cone BD, Irvine CE, Thompson MF, Nguyen TD. A video game for cyber security training and awareness. Comput Secur 2007;26(1):63–72.

Cybrary. Advanced penetration testing. https://www.cybrary.it/course/advanced-penetration-testing/; N/Aa.

Cybrary. Certified information systems security professional. https://www.cybrary.it/course/cissp/; N/Ab.

Cybrary. Comptia casp. https://www.cybrary.it/course/comptia-casp; N/Ac.

Cybrary. Comptia network+. https://www.cybrary.it/course/comptia-network-plus/; N/Ad.

Cybrary. Ethical hacking and penetration testing. N/Ae.

Cybrary. Python for security professionals. https://www.cybrary.it/course/python/; N/Af.

Cybrary. Secure coding. https://www.cybrary.it/course/secure-coding; N/Ag.

U. of Delft. Cybersecurity economics. https://www.edx.org/course/cyber-security-economics-delftx-secon101x; N/A.

Duckworth AL, Seligman ME. Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. Journal of educational psychology 2006;98(1):198.

O. for Economic Co-operation, D. (OECD). At what age do university students earn their first degree? https://www.oecd.org/education/skills-beyond-school/EDIF_23eng(2014)EN.pdf; 2014a.

O. for Economic Co-operation, D. (OECD). Skills beyond school. synthesis report. https://www.oecd.org/edu/skills-beyond-school/Skills-Beyond-School-Synthesis-Report.pdf; 2014b.

EDUCATION DF, RESEARCH SCFE, BOARD ICG. Massive open online courses (moocs): Trends and future perspectives. EDU/CERI/CD/RD; 2016.

Ferrer Mico, M.T. (2016). Community of Inquiry (COI) and Self-Directed Learning (SDL) in Online Environments: An Exploratory, Correlational and Critical Analysis of MOOCs. Introduction to Cybersecurity MOOC Case Study (Doctoral dissertation, Universitat Ramon Llull); 2016.

R. Institute. Cybersecurity fundamentals. https://www.edx.org/course/cybersecurity-fundamentals-ritx-cyber501x-1; N/Aa.

R. Institute. Network security. https://www.edx.org/course/network-security-ritx-cyber504x; N/Ab.

Irvine UC. Networking and security in ios applications. https://www.coursera.org/learn/security; 2017.

U. of London. Information security: context and introduction. https://es.coursera.org/learn/information-security-data; 2017.

U. of London. Malicious software and its underground economy: Two sides to every story. https://es.coursera.org/learn/malsoftware; 2018.

C. I. U. of Madrid. Cyber security basics: a hands-on approach. https://www.edx.org/course/cyber-security-basics-hands-approach-uc3mx-inf-2x; 2017.

U. Maryland. Cryptography. https://es.coursera.org/learn/cryptography; 2017a.

U. Maryland. Hardware security. https://en.coursera.org/learn/hardware-security; 2017b.

U. Maryland. Software security. https://es.coursera.org/learn/software-security; 2017c.

U. Maryland. Usable security. https://en.coursera.org/learn/usable-security; 2017d.
U. of Michigan. Internet history, technology, and security. https://es.coursera.org/learn/internet-history; 2017.
Mirkovic J, Benzel T. Teaching cybersecurity with deterlab. IEEE Secur Priv 2012;10(1):73–6.
MIT. Computer systems security. https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-858-computer-systems-security-fall-2014/index.htm; 2014a.
MIT. Network and computer security. https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-857-network-and-computer-security-spring-2014; 2014b.
Patriciu VV, Furtuna AC. Guide for designing cyber security exercises. In: Proceedings of the 8th WSEAS international conference on E-activities and information security and privacy. World Scientific and Engineering Academy and Society (WSEAS); 2009. p. 172–7.
Paulsen C, McDuffie E, Newhouse W, Toth P. Nice: Creating a cybersecurity workforce and aware public. IEEE Secur Priv 2012;10(3):76–9.
Pusey P, Gondree M, Peterson Z. The outcomes of cybersecurity competitions and implications for underrepresented populations. IEEE Secur Priv 2016;14(6):90–5.
Reuben E, Sapienza P, Zingales L. How stereotypes impair womens careers in science. Proc Natl Acad Sci 2014;111(12):4403–8.
Salah K. Harnessing the cloud for teaching cybersecurity. In: Proceedings of the 45th ACM technical symposium on computer science education. New York, NY, USA: ACM, SIGCSE '14; 2014. p. 529–34. doi:10.1145/2538862.2538880.
Sullivan F. 2017 global information security workforce study; 2017.
U. S. of Georgia. Cybersecurity and its ten domains. https://en.coursera.org/learn/cyber-security-domain; 2018.
N. I. of Standards, T. (NIST), National initiative for cybersecurity education (nice). cybersecurity workforce framework. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800--181.pdf; 2017.
G. Tech. Computer networking. https://www.udacity.com/course/computer-networking--ud436; N/Aa.
G. Tech. Cyber-physical systems security. https://www.udacity.com/course/cyber-physical-systems-security--ud279; N/Ab.
G. Tech. Intro to information security. https://www.udacity.com/course/intro-to-information-security--ud459; N/Ac.
G. Tech. Network security. https://www.udacity.com/course/network-security--ud199; N/Ad.
Trickel E, Disperati F, Gustafson E, Kalantari F, Mabey M, Tiwari N, Safaei Y, Doupé A, Vigna G. Shell we play a game? ctf-as-a-service for security education. Proceedings of the 2017 {USENIX} workshop on advances in security education ({ASE} 17). {USENIX Association}, 2017.
S. University. Cryptography i. https://es.coursera.org/learn/crypto; 2017.
T. O. University. Introduction to cybersecurity. https://www.futurelearn.com/courses/introduction-to-cyber-security; N/A.
Vasserman EY, Bell RS, Sayre EC. Developing and piloting a quantitative assessment tool for cybersecurity courses. American Society for Engineering Education, 2015.
U. of Virginia. Applied cryptography. https://www.udacity.com/course/applied-cryptography--cs387; N/A.
Wang M. Designing online courses that effectively engage learners from diverse cultural backgrounds. Br J Educ Technol 2007;38(2):294–311.
U. of Washington. Building a cybersecurity toolkit. https://www.edx.org/course/building-cybersecurity-toolkit-uwashingtonx-cyb003x; N/Aa.
U. of Washington. Cybersecurity - the cisos view. https://www.edx.org/course/cybersecurity-cisos-view-uwashingtonx-cyb002x; N/Ab.
of Washington U.. Introduction to cybersecurity. https://www.edx.org/course/introduction-cybersecurity-uwashingtonx-cyb001x; N/Ac.

Lorena González-Manzano is lecturer in the Computer Security Lab at the University Carlos III of Madrid, Spain. She is Computer Scientist Engineer and Ph.D. in Computer Science by the University Carlos III of Madrid. Her Ph.D. focuses on security and privacy in social networks. She is currently focused on Internet of Things and cloud computing security, as well as,on cybersecurity. Indeed, she has published several papers in national and international conferences and journals and she is also involved in national R+D projects.

José María de Fuentes is associate professor in the Computer Science and Engineering Department at University Carlos III of Madrid, Spain. He is Computer Scientist Engineer and Ph.D. in Computer Science by the University Carlos III of Madrid. His main research interests are digital evidences management, non-repudiation in vehicular environments, as well as security and privacy in the internet of things and ad-hoc networks. He has published several articles in international conferences and journals. He is participating in several national R+D projects.