Gender gap in electrical engineering at the University of Belgrade (1923-2010): Analysis of graduates' structure using R

Nadica Miljković[1], Biljana Stanković[2]
[1]*University of Belgrade – School of Electrical Engineering, Belgrade, Serbia*
[2]*Department of Psychology, Faculty of Philosophy, University of Belgrade, Belgrade, Serbia*
nadica.miljkovic@etf.bg.ac.rs, biljana.stankovic@f.bg.ac.rs

Abstract: This paper introduces exploratory analysis of graduates at the University of Belgrade from 1923 to 2010, with the focus on their gender structure. The underrepresentation of women in electrical engineering and computer science is a global issue, and precise data on the local situation is mostly absent from the literature. Preliminary results of objective assessment of gender gap by analysing graduates' gender, graduation age, and module preferences for 87 years are presented. This analysis might be valuable for decision makers in the area of electrical engineering and computer science for designing gender-motivated institutional changes in order to alter male-dominated culture.

Keywords: electrical engineering; computer science; gender equality; gender disparities; women in STEM.

I. Introduction and motivation

A recent StartIT article published in December 2018 promoted a positive trend of number of females applied for entrance exam at the University of Belgrade – School of Electrical Engineering (ETF) [1]. This percent was ~33% in 2018 and 2019 as it increased for ~13% since 2006 [2]. There is a relatively high correlation of these percents of interested females for the studies at ETF and the number of those who passed the exam and became freshwoman (personal communication with the Vicedean for academic affairs at the ETF, [1]). These positive trends are to some extent in line with the tendancy of women choosing more major in computer science more often as seen at the Stanford University in 2015 [1][3]. Gender gap is definitely decreasing, but the process is rather slow (participation of women in engineering is 20% worldwide [4]). It seems that the climate in the tech world did not foster gender gap closedown. At least, not yet [4]. Despite world-wide efforts towards more diversity, electrical engineering is still male-dominant, and adequate efforts are needed to narrow this gap further [5]. In order to direct and plan appropriate actions, assessment of gender gap is required. We believe that a historical perspective and an insight into the number of female graduates in electrical engineering that are presented in this paper by analysis of available data from the address book of graduates from the University of Belgrade (UB) for a period 1923-2010 [6] will be valuable for decision makers to direct more diversity at UB and ETF.

We aimed to answer following research questions: (1) how did gender gap change over the years at the UB, (2) are there any gender differences in age at the time of graduation, and (3) are there "masculine" and "feminine" modules at ETF and UB. Possible origins of the obtained results were discussed.

II. Methods and materials

All processing steps were performed in R environment and programming language [7] using R Studio IDE (Rstudio, Inc., Boston, MA, USA), and LibreOffice (The Document Foundation, Berlin, Germany) was used for storing and handling data. We used dplyr [8], ggplot2 [9], and readODS [10] R packages from an official repository of R packages CRAN (The Comprehensive R Archive Network).

A. Available data and historical background

In 1905 Technical faculty was founded at the UB with three modules (civil engineering, architecture, and mechanical engineering). Lectures in electrical engineering were held at the module of mechanical engineering. University regulation adopted in 1922, envisioned that students could apply for their diploma work in electrical engineering and the first graduates in mechanical and electrical engineering appeared in 1923.

In 1946, mechanical and electrical engineering were separated and the first graduates in electrical engineering appeared in the same year. ETF was formed shortly after, in 1948, under UB as an umbrella institution with two modules. In following years the number of modules and their names changed. The last accreditation included in the available dataset dates back from 2003/04 academic year when Bologna process and educational reforms were implemented [11]. In 2003/04 6 modules were introduced: OG (Power Engineering), OE (Electronics), OT (Telecommunications and Information Technology), OS (Signals and Systems), IR (Computer Engineering and Information Theory), and OF (Physical Electronics), and in 2004/05 the 7th module SI (Software Engineering) was introduced. [6]

In 2010, the address book titled "Address book of graduates in electrical engineering at the University of Belgrade 1923-2010" (original title in Serbian "Imenik inženjera elektrotehnike koji su diplomirali na Univerzitetu u Beogradu 1923-2010") was edited by prof. Miodrag Popović (ETF Dean 2007-2011) and late prof. Dimitrije Tjapkin (ETF Dean 1975-1977) [6]. They merged forgotten address book from 1956 with graduates for a period 1923-1956, an adress book published for the

50th ETF anniversary, and ETF archive. The [6] book contains the following information: graduation year, name, abbreviated middle name, surname, module, and year, place and state of birth for 19596 graduates. For the period 1923-1948 states are available either in abbreviated format (S for Serbia, Sl for Slovenia, M for Macedonia, etc.) or with full names (Russian, Polland, Germany, etc.). For a period from 1948 to 2010, in some cases states were added and in some not (the majority already existed for foreign[1] countries). Authors of this paper had the address book [6] available in an open pdf format (portable document format, ISO 32000-2 standard) stored at the digital optical disc data storage *i.e.* CD (Rainbow books standard) prepared for the 70th ETF anniversary in 2018.

The available data were not in machine-readable format and the graduates' gender[2] was missing. Therefore, the main idea behind the transformation of the available pdf from [6] was to add information on gender based on the graduates' names in a machine-readable format in order to enable analysis of the gender gap in electrical engineering and computer science in UB and ETF.

B. Data formatting and cleaning

In order to analyse given data, we tried to read pdf file directly in RStudio with available R package pdftools [13]. This was not possible without additional and extensive workaround due to a non-consistent formatting (available pdf presents a combination of other documents). Hence, manual manipulation was required. Firstly, we copied text manually (page by page) by Select tool in Adobe Acrobat (Adobe Systems, Inc., San Jose, CA, USA) from [6] into Lexilogos [14] for conversion from Cyrillic to Latin Serbian letters for straightforward analysis in R. Secondly, we copied data in a LibreOffice Writter document with odt (open document text) extension. Thirdly, we converted text to table in LibreOffice Writter with one space as separator and then copied table in LibreOffice Calc with ods (open document spreadsheet) extension. This was done for each graduation year. Column with graduation year was added manually in odt file and excess columns were deleted manually (e.g. additional column was created if the place of birth consisted of more than one word, such as Smederevska Palanka). Also, NAs (Not Available) were added manually for the missing columns. When neccessary, additional corrections due to variety of formatting styles in the original pdf were applied including typos correction (Nikla into Nikola, Nend into Nenad, etc.). In some cases, we added states into the final ods file. However, this cleaning wasn't completed, as this paper is focused on gender gap. N. M. was responsible for filling in the gender column in the dataset, based on the graduates' names, and B. S. performed an additional inspection. The gender was determined for 98.8% of graduates. For inconclusive cases, the raws were filled with "?". This was the case with following unisex names (number of names is given in brackets): Saša (156), Sava (12), Vanja (11), Dobrica (8), Vladica (8), and other unisex and foreign names (33 in total with one replica). We placed NAs for birth year in all graduates older than 80 and younger than 20 at the time of graduation (15 graduates). For Serbian Latin letters (š, č, ć, ž, and đ) representation in R, we used UTF-8 encoding. Then, we added the 10th column of age at the time of graduation by subtracting birth year from graduation year in R. Finally, we recorded data in a text file for further analysis. To summarize, the final dataset consisted of 284 (1.5%) missing abbreviated middle names, one missing module (<0.1%), 70 missing birth years (0.4%), 74 missing birth cities (0.4%), and 14688 missing countries (75.0%).

C. Exploratory data analysis

In order to get an insight into historical trends of gender gap in electrical engineering and computer science at UB and ETF, we visualized data and performed simple descriptive statistics (calcuated mean, standard deviation with Bessel's correction (SD), and the appropriate proportions of females and males, and their age). For comparison of graduation age, we used Welch two sample t-test (significant difference was set at $p<0.01$). We used bar graphs and mirror bar chart for visualization.

III. Results and discusion

The results are presented for 19588[3] graduates in a period from 1923 to 2010 (except for the year 1945[4]). During this 87 years range there were 3518 females (~18%) and 15840 males (~82%) graduating in Electrical Engineering at UB. Percentage of females for each year is presented in Figure 1.

The first woman graduated in mechanical and electrical engineering at UB in 1931, only 8 years after the inception of this program. However, the presence of women graduates at UB remained a rare occurrence for the next three decades – women made under or around 10% of graduates until 1970[5]), when women made almost quarter of graduates, although that was an isolated occurrence until the end of the 1980s.

The maximal percent of female graduates (30.6%) was in 1993, which might be the consequence of the breakup

1 The term foreign should be taken with precaution, since Belgrade and UB changed countries during the turbulent history. Following states (with corresponding territories and authority) had Belgrade as capital from 1882: Kingdom of Serbia; Kingdom of Serbs, Croats and Slovenes; Kingdom of Yugoslavia; German-occupied Serbia; Federal People's Republic of Yugoslavia; Socialist Federal Republic of Yugoslavia; Federal Republic of Yugoslavia; Serbia and Montenegro; and from 2006 Republic of Serbia. For more see, https://en.wikipedia.org/wiki/History_of_Belgrade, Assessed in Aug. 2019.

2 Here, we use gender as a common term for both sex and gender. However, sex is commonly used to point out to the person's biological characteristics and gender is frequently used to describe person's internal awareness (https://en.wikipedia.org/wiki/Sex_and_gender_distinction, Assessed in Aug. 2019). For more appropriate and extensive explanation on gender and sex, please see [12].

3 There is a slight difference of reported number of graduates in [6] and this number (8 graduates) and an additional crosscheck is adviced.

4 There were no graduates in 1945. This pause was probably caused by the Belgrade Offensive at the end of the World War II. It resulted in goods shortage, famine, and other unfortunate events which deteriorated significantly during bombardment. Technical faculty of the UB was also damaged during bombing conducted by Serbian allies in a period April-September 1944. Offensive brought also street fights and explosions [15].

5 Except for 11 years when there were no women graduates – 1932-1934, 1939-1940, 1943-1944, 1946-1948, 1950.

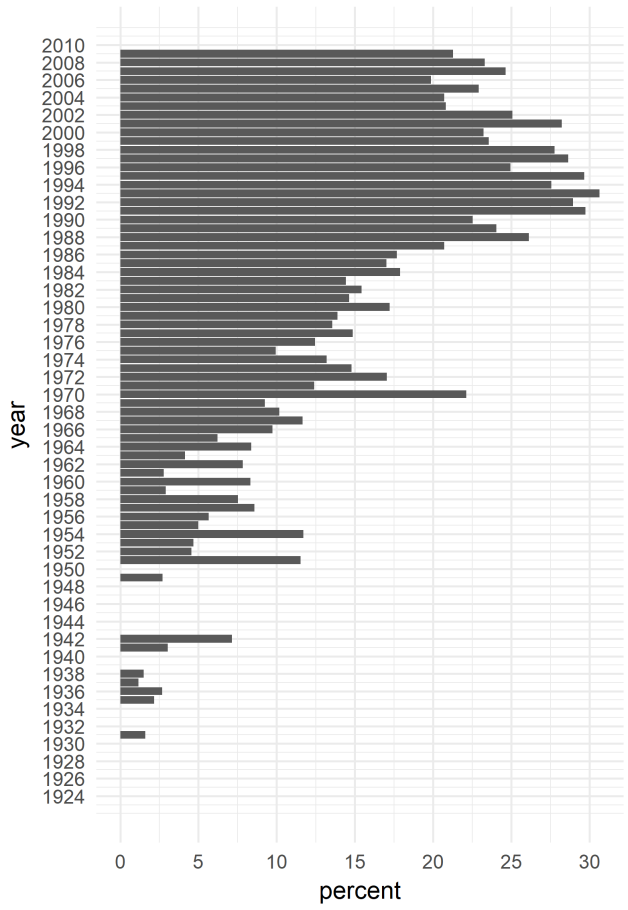

Figure 1: Percents of female graduates per year for a period 1923-2010.

war in Yugoslavia when large refugee migrations (including so-called brain drain migration [16]) and military recruitment took place.

Overall, the larger percent of females in certain historical periods might be the consequence of historical events, as well as other societal changes: during 23 years (out of 87 years range), the number of female graduates was >20%; and for the period (1988, 1991-1995, 1997, 1998, 2001-2002) it was >25% (see Figure 1). The largest number of students was in 2007 (755 students) with 24.6% of females.

On average, during graduation, females were 25.98±2.59 years old (range from 20 to 48 years) and males were 26.85±3.48 years old (range from 20 to 61). Mirror bar in Figure 2 presents comparison of age for female and male graduates for each graduation year. SDs are not presented Figure 2 as they would not be visible (on average SDs were 9.15±3.06 and 12.51±2.28 for females and males, respectively). T-test showed statistically significant difference of average age during graduation per year between males and females for 1923-2010 ($p = 0.007$), and for 1952-2010 ($p < 10^{-8}$) indicating probably more efficient studying of females. A number of reasons could be at the root of this difference, many of them "gendered" in nature – e.g. girls are under bigger pressure to finish schooling so they can marry and have children (since their "biological clock is ticking") and, on the other side, there was a compulsory military service

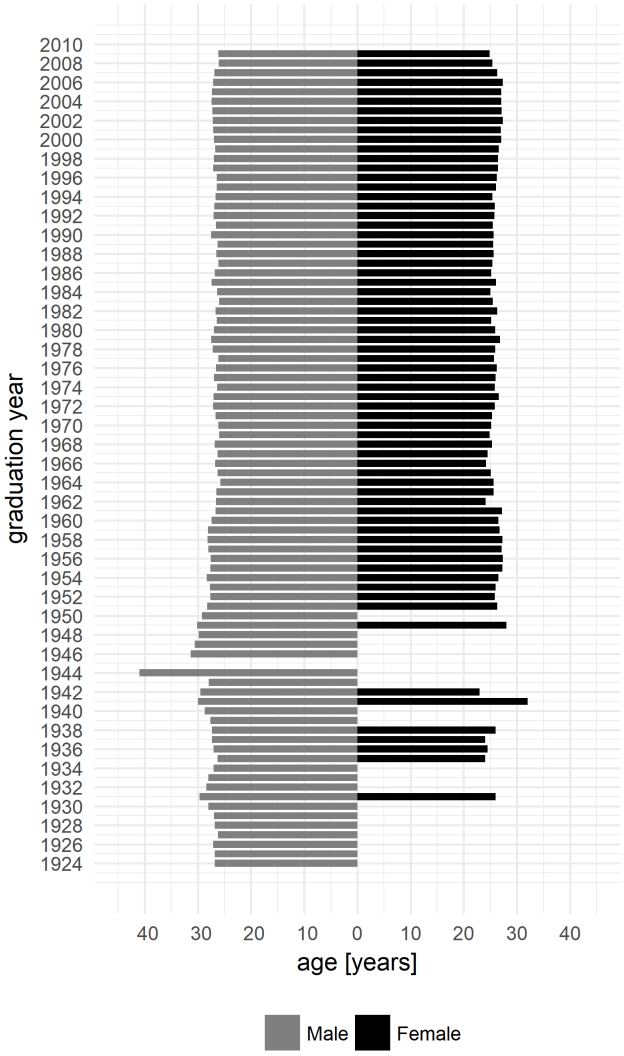

Figure 2: Mirror bar chart of average graduates' age.

for men in Serbia (which was suspended in 2011) tha t could have delayed their graduation.

We calculated percents of females per modules that were introduced by the accreditation and Bologna reform in 2003/2004 academic year for period 2007-2010 for 7 modules. Results are presented in Figure 3. Females' share for 2007-2010 was 33.0% for OT (Telecommunications and Information Technology), 30.2% for OS (Signals and Systems), 26.9% for OF (Physical Electronics), 15.6% for OE (Electronics), 14.4% for IR (Computer Engineering and Information Theory), 13.9 for OG (Power Engineering), and 11.7% for SI (Software Engineering). Data presented in Figure 3 was analyzed for 1636 graduates, which encompasses only 8.4% of the entire dataset. When looking at the more comprehensive analysis of top modules for 87 years, which is presented for 7665 graduates (39.1% percent of the dataset), the results are fairly similar[6]. TE (Electronics and Telecommunications) had 27.4% of females, EL (Electronics with Automaticcontrol) 21.3%, and EN (Power Engineering) with 17.6%. This shows that males'

6 The number of enrolled students at the module should be taken into account for appropriate discussion of these results, as we determined these top modules by the largest proportion of females entering the field.

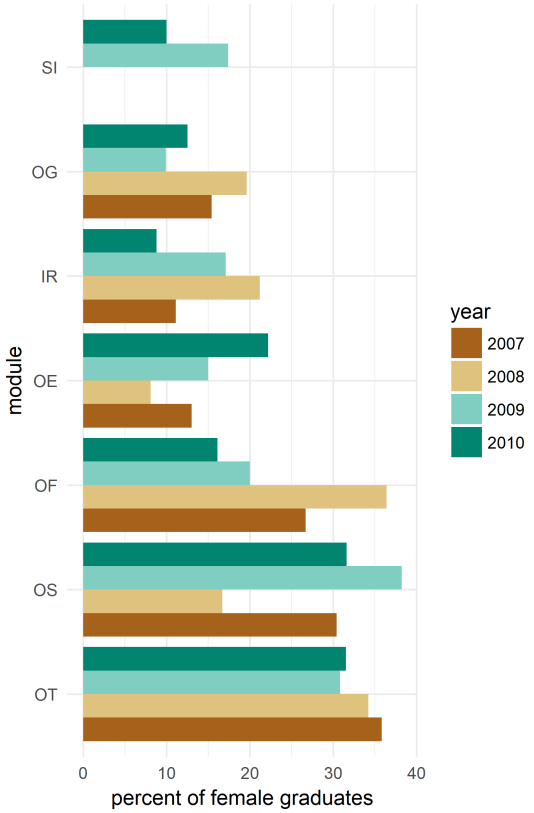

Figure 3: Percents of female graduates per year for modules introduced in 2003/2004 academic year (SI module was introduced in 2004/2005).

preference is Electronics and females' preference is Telecommunications. However, a more comprehensive analysis is necessary, that would include a larger proportion of the sample and also look into separate periods with different modules. But, even this partial insight into the gender structure across modules indicates that it reproduces some typical gender stereotypes when it comes to the professional interests and choice of career.

One interesting result when it comes to historical trends is related to computer science: RT (predecessor of IR) had 29.9% females (1923-2010) and this number decreased over years to 14.4% for IR. Further research is needed to adequately interpret this trend. Another interesting phenomenon is SI. There were no females during the first graduation year (2008), although it should be mentioned that only 7 males graduated in 2008. Higher gender gap in SI can be explained by discouragement of females to enter a newly opened field. Namely, although other modules in Figure 3 were introduced one year before, they already existed under other names. Sadly, but it appears the females might not be as bold as males for entering novel field [17].

IV. Instead of a conclusion

We did not present an exhaustive analysis, rather an insight into gender gap over 87 years long period. Future work will be focused on a more detailed analysis as we will try to acquire information on enrolled students for a period 1923-2019 and on their careers after graduation.

Even with this scratch of the gender gap surface, there are strong evidences that pervasive gender gap exists. Institutional and planned actions should be performed to close the gap permanently and efficiently [5].

Acknowledgement

The work on this paper was partly supported by the Ministry of education, science, and technological development, Republic of Serbia Grants TR-33020 (N. M.) and 179018 (B. S.).

References

[1] N. Rangelov, "Na ETF-u studira čak jedna trećina devojaka – broj raste iz godine u godinu," *StartIT, Belgrade, Serbia*, 2018, [Online] https://startit.rs/na-etf-u-studira-cak-jedna-trecina-devojaka-broj-raste-iz-godine-u-godinu/ Assessed in Aug. 2019

[2] https://prijemni.etf.bg.ac.rs/statistika/pol.php, Aug. 2019

[3] S. McBride, "Computer science now top major for women at Stanford University," *Technology news, Reuters*, 2015, [Online] https://www.reuters.com/article/us-women-technology-stanford-idUSKCN0S32F020151009 Assessed in Aug. 2019

[4] J. E. Mills, S. Franzway, J. Gill, and R. Sharp, "Challenging knowledge, sex and power: Gender, work and engineering," Routledge, New York, USA, 2013

[5] S. Ihsen, S. Gebauer, and A. Buschmeyer, "Gender motivated institutional changes at Technische Universität München and its influence on engineering education," In *Book of abstracts: Annual SEFI Conference*. Aalborg, Denmark, 2008

[6] "Imenik inženjera elektrotehnike koji su diplomirali na Univerzitet u Beogradu 1923-2010," Editors: Prof. Miodrag Popović and Prof. Dimitrije Tjapkin, *University of Belgrade – School of Electrical Engineering and Academic Mind, Belgrade, Serbia*, 2010

[7] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, 2017, [Online] https://www.R-project.org Assessed in Aug. 2019

[8] H. Wickham, R. Francois, L. Henry, and K. Miller, "dplyr: A grammar of data manipulation," *R package*, ver. 0.8.0.1, 2019 [Online] https://cran.r-project.org/web/packages/dplyr/index.html Assessed in Aug. 2019

[9] H. Wickham, "ggplot2: Elegant graphics for data analysis," Springer-Verlag, New York, USA, 2016

[10] S. Gerrit-Jan, C. Chung-Hing, T. J. Leeper et al., "readODS: Read and write ODS files," *R package*, ver. 1.6.7, 2018 [Online] https://cran.r-project.org/web/packages/readODS/index.html Assessed in Aug. 2019

[11] Bologna Process, Wikipedia [Online], Accessed in Aug. 2019, https://en.wikipedia.org/wiki/Bologna_Process

[12] A. Oakley, "Sex, gender and society," Routledge, 2016

[13] J. Ooms, "pdftools: Text extraction, rendering and converting of PDF documents," *R package*, ver. 2.0, 2018 [Online] https://cran.r-project.org/web/packages/pdftools/index.html Assessed in Aug. 2019

[14] Lexilogos, Assessed in Aug. 2019 https://www.lexilogos.com/keyboard/serbian_conversion.htm

[15] S. Lazić, S. Mandić, V. Mijatović, J. Mitrović Kocev, M. Obradović, I. Stojanović, "Oslobođenje Beograda – 70 godina posle, kroz fondove i zbirke Istorijskog arhiva Beograda," *Historical Archives of Belgrade*, 2014 [Online] https://www.arhiv-beograda.org/images/publikacije_elektronske_pdf/katalog%2070%20godina%20oslobodjenja%20BGD%20web.pdf Assessed in Aug. 2019

[16] Z. Avramović, "Odliv mozgova iz Srbije – jedno lice globalizacije obrazovanja," *Sociološki pregled*, 46(2): 189-202, 2012.

[17] B. Stanković, "Programirati kao devojčica: Rodne razlike u računarstvu," In *PSSOH Proceedings*. Belgrade, Serbia, 2018, doi: 10.5281/zenodo.1476518.