A Review of Developer's Profile Across Repositories

Nakul Sharma[1] *, Balasubramnium. Vivekanandam* [2] *, Eugenio Vocaturo*[3]

[1] Linclon University College; [2] Linclon University College; [3] Linclon University College

Email ID

[1] pdf.nakul@lincoln.edu.my ,

[2] vivekanandam@lincoln.edu.my

[3] eugenio.vocaturo@cnt.it

Abstract: Developer's profile exists on different online platforms. This survey studies existing literature on the developer's profile. The databases chosen for conducting research were Sciencedirect and IEEEXplore. The developer networks exist for different purposes across different platform including social media platforms. These are mainly for accomplishing collaboration, mentoring, bug assignment, etc. The existing work can be extended to conduct systematic literature review as well.

Keywords: Developer; profile; software; IT; industry

Introduction

This paper conducts a survey of research paper published in the area of developer's profile. This involves mining keywords of developer's profile related to software in Sciencedirect and IEEEXplore databases.

Software mining of the repositories has progressed to introduce new innovative ideas of repository retrieval and studying the relevance of the extracted information from multiple online sources. There is a growing interest in the mining of large scale software repositories. These repositories are hosted by open source platforms along with various resources which enable easy discovery, retrieval and mining tasks on the software projects. The developer information is very critical which can be tracked in the version control systems. The Version Control System such as GitHub, BigBucket etc. are hosted in order to make distributed development easier and to seek better innovation on the project being developed. The developers make local changes to these online repositories and then initiate commit for reflecting those changes. World of Code (WoC) provides a facility for studying the supply chains of software along with various other facilities for the developer's and researchers [1].

The developer's profile at online repositories is stored as meta-data information and is updated whenever the developer executes commits or undertakes any changes. The developer's profile and the corresponding work potential are also reflected from the other ID's [2].

The developer profile can be categorized according to various different levels of experience. The fresher level experience can be between 0-5 years. The mid-range experience can be between 5-15 years. The senior level experience is 15+ years of experience. The developer's participation across projects vary according to the seniority and level of experience gained in these projects.

The whole paper is organized into different sections. Section-I gives introduction to the topic. Section-II provides the related work in form of literature review. The results found from the literature surveyed are given in section-III. The future scope related to developer's profile is given in section-IV. Section-V gives the future scope.

Related work

This section provides overview of research paper's searched in IEEEXplore and Sciencedirect. The survey method is also described in this section along with inclusion and exclusion criteria's.

A free account was created on IEEEXplore and Sciencedirect using email ID. After account creation, a search query was passed. The search query was passed between 2024-2025. The search query was "developer" "profile" "software". Table-1 provides the query string passed on both the online portals and the results obtained.

Table 1. Search Strings and Their Results

Name of Database	Search String Passed	Results Obtained
IEEEXplore	"developer" "profile" "software"	124
ScienceDirect	"developer" "profile" "software"	16810

Justification of search query

The search query passed was "developer" "profile" "software". These three words contained all the necessary input for getting the relevant research papers. Software keyword in search string was chosen to exclude any civil engineering related housing developer profiles. These papers related to civil engineering could have appeared in Sciencedirect, but not in IEEEXplore.

The count of Sciencdirect papers found after search string was 16810. This count was difficult to analyze hence only first 500 papers were selected for initial study and years of analysis was also restricted to 2020-2025. The most relevant papers were further segregated for the review.

The literature consisted of various categories as provided in table-II. The methodology adopted and the research gaps identified are also provided in table-II.

Table 2. Literature Review and Analysis

Ref. No.	Title	Methodology Adopted	Type of Paper	Research Gaps Identified/ Overcome
4	A systematic mapping study of developer social network research	Kai Peterson Et. al.	SMS	Research Topics, synthesis of literature
5	Profile Mining in CVS-Logs and Face-to-Face Contacts for Recommending Software Developers	Mining logs, Page rank, Mining developer profile	Research	1.Page rank algorithm for extracting developer's profiles. 2. RFID tags for reducing time.
6	A preliminary investigation of developer profiles based on their activities and code quality: Who does what ?	Mining github commits to determine the developer's profile and its activities.	Research	Harmful code practices by the developer can be identified.
7	CVExplorer: Identifying candidate developers by mining and exploring their open source contributions	1. GitHub skill evaluation and mining developer's contribution. 2. Aggregating developer's contributions. 3. Presenting developer's contribution.	Research	Integrating multiple job-portals information for validating credentials.
8	Visual Resume: Exploring developers' online contributions for hiring	Scenario-Based evaluation	Research	Aggregation of developer's profile at GitHub and Stack overflow.
9	Work and career-related features of technology: A grounded theory study of software professionals	Exploratory approach	Research	1. Perception and career expectations influence technology evaluation. 2. Technology can be viewed in lieu of interest of a software professional.

				3. Mapping of professional's interest and technological fetures.
10	Understanding the social evolution of the Java community in Stack Overflow: A 10-year study of developer interactions	1. Machine learning and Graph mining are integrated for analysis of quality of service and user's reputation across platforms.	Research	1. Information flow across the systems can be undersood from textual nature of question-answers. 2. user's reputation can be linked to quality of their social quality.
11	TDMatcher: A topic-based approach to task-developer matching with predictive intelligence for recommendation	Predictive intelligence used for developer profile prediction.	Research	Developer's prediction and recommendation for crowdsourcing software.
12	Guiding the way: A systematic literature review on mentoring practices in open source software projects	Kai Peterson Et. al.	SMS	Research Topics, synthesis of literature
13	Kieker: A monitoring framework for software engineering research	Logs used for dynamic visualizations	Research	Monitoring at application-level and run-time analysis of software system
14	Learning to rank developers for bug report assignment	1. Commit based developer's profiling. 2. Ranking score as weighted score of features.	Research	Developer's assignment for bug fixing.
15	Mining the Technical Roles of GitHub Users	1. label roles specific to development. 2. Machine-learning models developed for speific job roles.	Research	L-based developer's role identification is necessary.
16	Evaluating and strategizing the onboarding of software	Strategizing on-boarding process by addressing the specific issues.	Research	Issues related to onboarding process for developers addressed.

	developers in large-scale globally distributed projects			
17	We do not appreciate being experimented on": Developer and researcher views on the ethics of experiments on open-source projects	Survey conducted on open-source developers and empiricial software researchers.	Survey	1. Ethics in conducting open-source experimentations. 2. Developers say in conducting open-source experimentation is must.
18	Impact of individualism and collectivism cultural profiles on the behaviour of software developers: A study of stack overflow	1. Data collected via survey 2. Data mining techniques used for analysis.	Research	1. Culture and individualism from regions affect developers behavior. 2. Team work requires learning about developer's cultural and individual background.

Results and Discussion

The selected studies for conducting study gave the following recommendations:-

1. Developer's profile as it is spread across different years [10].
2. Mining tasks conducted on developer's profile for fulfilling different objectives such as bug resolution assignment [14], job profiles mapping [5]
3. Evolution of developer's profile across different years [10].
4. Converting developer's online profile into resume [8].
5. Study of community dynamics along with developer's profile [10].
6. Cultural issues are connected to developer's interests [18].
7. Humanity related issues such as privacy, ethics [17].

Tools Related to Developer's Profile

The literature contained following different software's or online tools for monitoring developer's profile:-

1. CVExplorer
2. Staticat
3. MasterBranch
4. GitHub [8]

These tools provide varying degree of support in making use of developer's profile.

Future Scope

The current focussed on surveying IEEEXplore and sciencedirect databases for the developer's profile. There can be following different research directions:-

1. A Systematic Literature Review (SLR) can be carried using developer's profile for undertaking more exhaustive search and meta-anaysis of the topic
2. The current topic can also be investigated from a humanity related issues also such as ethics, behaviour, atitutes etc.
3. The developer's scope of entrying into different companies along with the probability of joining a company can also be investigated.
4. There can be future scope of addressing different problems faced by developer's by corresponding tools.

Conclusions

This survey consisted of passing developer profile related queries on IEEEXplore and Sciencedirect platforms. The Sciencedirect platform provided 16810 results while IEEEXplore provided 128 results. In case of ScienceDirect only first 500 papers were selected for review, sorted according to the relevance of the papers. The IEEEXplore provided 128 results out of which 24 papers were selected for review. Developer's profile exists on multiple websites/portals in order to accomplish different tasks related to development. These profiles are maintained by the developer in most of the cases.

References

[1] Ma, Yuxing, Chris Bogart, Sadika Amreen, Russell Zaretzki, and Audris Mockus. "World of code: an infrastructure for mining the universe of open source VCS data." In *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)*, pp. 143-154. IEEE, 2019.

[2] Fekete, Anett, Máté Cserép, and Zoltán Porkoláb. "Measuring developers' expertise based on version control data." 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2021.

[3] Lamothe, Maxime. "Bridging the divide between API users and API developers by mining public code repositories." Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings. 2020.

[4] Steffen Herbold, Aynur Amirfallah, Fabian Trautsch, Jens Grabowski, "A systematic mapping study of developer social network research", Journal of Systems and Software, Volume 171, 2021,110802, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2020.110802.

[5] B. -E. Macek, M. Atzmueller and G. Stumme, "Profile Mining in CVS-Logs and Face-to-Face Contacts for Recommending Software Developers," 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, Boston, MA, USA, 2011, pp. 250-257, doi: 10.1109/PASSAT/SocialCom.2011.40.

[6] C. Aguilera González, L. Albors Zumel, J. Antoñanzas Acero, V. Lenarduzzi, S. Martínez-Fernández and S. Rabanaque Rodríguez, "A preliminary investigation of developer profiles based on their activities and code quality: Who does what?," 2021 IEEE 21st International Conference on Software Quality, Reliability and Security (QRS), Hainan, China, 2021, pp. 938-945, doi: 10.1109/QRS54544.2021.00103.

[7] G. J. Greene and B. Fischer, "CVExplorer: Identifying candidate developers by mining and exploring their open source contributions," 2016 31st IEEE/ACM International Conference on Automated Software Engineering (ASE), Singapore, 2016, pp. 804-809.

[8] Sandeep Kaur Kuttal, Xiaofan Chen, Zhendong Wang, Sogol Balali, Anita Sarma, "Visual Resume: Exploring developers' online contributions for hiring", Information and Software Technology, Volume 138, 2021, 106633, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2021.106633.

[9] Gunjan Tomer, Sushanta Kumar Mishra, "Work and career-related features of technology: A grounded theory study of software professionals", Information and Software Technology, Volume 164, 2023, 107301, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2023.107301.

[10] Guillermo Blanco, Roi Pérez-López, Florentino Fdez-Riverola, Anália Maria Garcia Lourenço, "Understanding the social evolution of the Java community in Stack Overflow: A 10-year study of developer interactions", Future Generation Computer Systems, Volume 105, 2020, Pages 446-454, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.12.021.

[11] Yiyang Fu, Benjun Shen, Yuting Chen, Linpeng Huang, TDMatcher: A topic-based approach to task-developer matching with predictive intelligence for recommendation, Applied Soft Computing, Volume 110, 2021, 107720, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107720.

[12] Zixuan Feng, Katie Kimura, Bianca Trinkenreich, Anita Sarma, Igor Steinmacher, Guiding the way: A systematic literature review on mentoring practices in open source software projects, Information and Software Technology, Volume 171, 2024, 107470, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2024.107470.

[13] Wilhelm Hasselbring, André van Hoorn, Kieker: A monitoring framework for software engineering research, Software Impacts, Volume 5, 2020, 100019, ISSN 2665-9638, https://doi.org/10.1016/j.simpa.2020.100019.

[14] Bader Alkhazi, Andrew DiStasi, Wajdi Aljedaani, Hussein Alrubaye, Xin Ye, Mohamed Wiem Mkaouer, Learning to rank developers for bug report assignment, Applied Soft Computing, Volume 95, 2020, 106667, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2020.106667.

[15] João Eduardo Montandon, Marco Tulio Valente, Luciana L. Silva, Mining the Technical Roles of GitHub Users, Information and Software Technology, Volume 131, 2021, 106485, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2020.106485.

[16] Ricardo Britto, Darja Smite, Lars-Ola Damm, Jürgen Börstler, Evaluating and strategizing the onboarding of software developers in large-scale globally distributed projects, Journal of Systems and Software, Volume 169, 2020, 110699, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2020.110699.

[17] Dror G. Feitelson, “We do not appreciate being experimented on”: Developer and researcher views on the ethics of experiments on open-source projects, Journal of Systems and Software, Volume 204, 2023, 111774, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2023.111774.

[18] Elijah Zolduoarrati, Sherlock A. Licorish, Nigel Stanger, Impact of individualism and collectivism cultural profiles on the behaviour of software developers: A study of stack overflow, Journal of Systems and Software, Volume 192, 2022, 111427, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2022.111427