Connecting GitHub Issues with Commits in Open Source Software Projects

Georgi Dungarov
Department of Computer Science and Engineering
University of Gothenburg
Gothenburg, Sweden
dungarov@gmail.com

***Abstract*— In the current state of software development a common way to manage and contribute to an Open Source Software Project is to use Version Control Systems. GitHub, one of the largest hosting services for Open Source projects, provides an issue-tracking system allowing users and developers to report issues and offer solutions. Further, developers can assign different labels to an issue, which helps categorize it, as well as, provide basic characteristics. This method could be time and cost inefficient. Lack of connection between issues and commits could lead to Technical Debt by causing developers to return to issues to resolve them. In GitHub, there is no semantic connection expressing an issue to a commit that solves and eventually closes the issue. This lack of connection has a major drawback as tools analyzing meta-data for project measures related to issues and commits cannot be processed. For example, if we want to check whether there is a significant difference among the size of fixes of issues of different types; it is not possible to determine until we have established a connection between issues and commits. This study aims to explore connection between issues and commits, in order to make them traceable. A theoretical framework is developed to target the RQ. The theoretical framework for the research will establish the factors for a possible relation between issues and commits.**

I. INTRODUCTION

Developing software is becoming more rapid and fast paced. The demand for quick delivery of working solutions could cause problems in the projects such as: bugs, miscommunication between the team, difficulty tracking implemented changes and rollbacks to a previous version of the software. In order to reduce those risks, Version Control Systems (VCS) emerged. Version control software allows the storage of previous version of the project, easy collaboration between developers and regular backups. There are multiple VCS on the market that offer both centralized and distributed network. Some of the most popular ones are GitHub, BitBucket, Google Code and SourceForge.

Version Control System, such as GitHub, makes projects publicly available, which turns them into a great source for research. The advantage of publicly available data is the diverse background and access to large amount of software artifacts. GitHub was selected for this research study due to the fact that it is the largest code hosting platform in the world, with more than 10 million repositories. Another benefit to choosing GitHub is the fact that most of the projects hosted on the platform are Open Source.

GitHub provides an issue-tracking system which allows for users and developers to report issues and categorize them by labels. However, in some cases the developers do not submit direct commits that correspond to the issues. Not recognizing and underestimating the importance of the issues and therefore not relating commits to them, could result in Technical Debt. One scenario could be an occurrence of a bug, which due to certain restrictions is not immediately resolved. This means that in a later stage of the project development unpredicted work could arise, due to escalation of the unresolved bug, leading to additional resources cost and time.

So far, researchers that have started exploring GitHub as a mining source have focused on topics like benefits of labels to classify issues and the effects it has on projects [1]. Other topics include the working habits and role of integrators to manage and integrate commits [2] as well as finding patterns for software development by mining source code repositories [3].

This research focuses on mining data from publicly available repositories on GitHub. The main idea is to find a connection between commits and issues of GitHub repositories. To ensure the success of the research a theoretical framework will be developed. The framework will attempt to discover the relationship between commits and issues that have both direct and indirect link. An example of a direct link is when an issue is resolved with a commit. Indirect links between commits and issues could be a commit that makes changes to a file in the project that fixes issues that are not directly linked. To facilitate the study and build the theoretical framework, three Open Source Projects have been selected for the data collection. The data collected is as detailed as possible. That will allow determining a better and stronger connection between the issues and commits. Such connection will be based on one or more factors such as time, tags or committers. Furthermore, a project called GHTorrent mines GitHub repositories and retrieves its contents and their dependencies. The data is stored in MongoDB database. GHTorrent is a data mirroring solution which offers researchers Source tracking, Network analysis and Single developer identities. However, the project lacks relating issues and commits. This research offers a possible solution to that problem that could be realized in the GHTorrent project.

Additionally, by establishing connection between issues and commits, this research could contribute by suggesting candidate commits for issues. The level of accuracy will be achieved through the theoretical framework and will be based on the analyzed metadata. Therefore, another benefit could be lowering the human error in prioritizing issues thus creating technical debt.

This study could be used as analytical framework for multiple projects on different version control systems to find and define connection specific to those projects. It can also be used for analytics of a singular project. Results could be beneficial to developers who maintain projects, as well as, users who would like to contribute to a project.

In this paper, using data from GitHub, the following questions will be explored:

RQ (1) How does the issue landscape look?

RQ (2) How to connect issues and commits in a version control system, so that commits can be traced from their relevant issues or vice versa?

RQ (3) What are the important factors to establish a connection between issues and commits?

II. BACKGROUND

A. Related Work

A study by Cabot (2015) focused on the use of labels to categorize issues in Open Source Projects. The focus of the researchers was the labels/tags used to classify the issues and

more particularly if labelling issues has a beneficial effect to the project and the commits [1]. In this paper, the team has analysed GitHub projects in order to get insight on the labelling system. To achieve that, the researchers have developed GiLA Label Analyzer to work along with GHTorrent in order to perform the analysis as dataset [1]. GHTorrent serves as the source for the study. The team found that even if issues are described poorly, that still favors a commit with a resolution of that issue [1].

Another study focused on the role of the integrator to manage and integrate commits. The study investigated the working habits and challenges of the integrator via a large-scale survey [2]. The results showed some of the factors the integrator takes under account in the decision making if a certain commit that fixes an issue or adds a feature, should be accepted or not [2]. Additionally, the study displayed the struggle of the integrator to prioritize a related commit for an issue and therefore maintain quality [2].

A research by Allamanis (2013) aimed at mining source code repositories in order to find patterns for software development [3]. The developers introduced new metrics of measurement with the help of a probabilistic language model. By using a giga-token corpus of Java code they managed to successfully predict identifiers across different projects. Later, the researchers explored the identifiers - class of tokens - that allowed them to better understand theoretic tools and metrics in source code repositories. That helped them identify different aspects of projects [3].

A study by Gousios (2012) has created a GitHub project called GHTorrent which offers a scalable event stream and persistent data that focuses on users, pull requests and all issues surrounding social coding [4]. In their paper the researchers have demonstrated the initial design of the project as well as presenting datasets that can be requested. The data collected by the project can provide insights on different aspects of Open Source projects such as community dynamics, code authorship and attributions [4].

In a research by Van Der Veen (2015), a tool called *Prioritizer* was developed in order to face the challenges that come when prioritizing pull requests in GitHub [5]. The study investigates the priority criteria the developers have in order to create a tool that provides visualization and suggestions. The research categorizes and examines the pull requests for a certain project, then *Prioritizer* presents the top pull requests that need attention [5]. Further, the developers could sort the pull requests based on criteria [5]. The researchers found that users preferred and appreciated the overview of the project pull requests. But while users found such a priority to be beneficial, participants also requested more insight on how the tool works [5].

The studies exemplified in this section offer different benefits in relation to issues and commits. Few of them are related to the way certain factors affect issues [1][3][5]. The rest of studies are examining patterns in order to explain a phenomenon. However, those studies explore either a single aspect, such as labels [1], or focus on describing patterns in regards to certain events. None of them investigate the relationship between commits and issues to the full extent. Therefore, it is necessary to build a theoretical framework that is capable of collecting and analysing all elements associated with issues and commits.

B. Theoretical Framework

The concept of this research is surrounded by the idea of discovering and defining a relationship between issues and commits in version control systems, such as GitHub. A simple definition of what "connection" is, in relation to version control systems would be:

> *The perception of connection as a result of consciously comparing a variety of factors with solutions through their corresponding issues.*

In other words, a connection between issues and commits could be determined through observation of the factors related to those issues. Then those factors shall be compared with the commits. The degree to which a connection is satisfying or not is determined by the observed commits in relation to the factors.

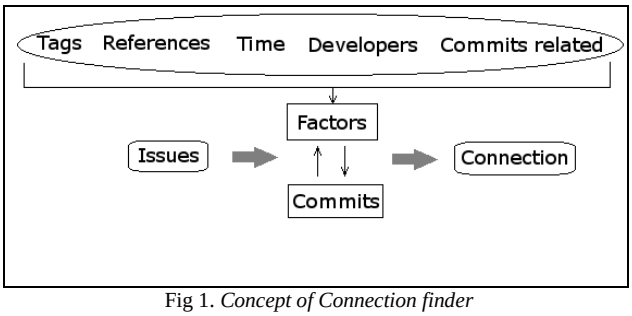

Fig 1. *Concept of Connection finder*

The theoretical framework concept is illustrated in Figure 1. It shows that tags, references, time, developers and commits related data determining the factors. Those factors are compared to the commits (or the lack of them) of corresponding issues. This comparison between the factors and commits is important because it determines the essence of the connection. This model is important for the study, as it can reveal to what extent factors matter and how it affects a software project.

The comparison with commits depends on issue to issue basis. This is due to the fact that not all issues are resolved with a commit. Issues that have direct commits are referenced as *direct commit* or *direct link*. On the other hand, issues that lack direct commits are referenced as *indirect commit* or *indirect link*. This study is looking into both direct and indirect links, because that could be a basis to find direct and indirect connection between issues and commits. The lack of direct commit does not lead to lack of connection. It could mean that the commit resolving the issue is not directly connected, but the relationship exists. The relationship between one issue with a direct commit could lead to another issue being resolved. That type of connection discovery is a way to answer RQ (2). As shown in Figure 1, the connection is based on a comparison between factors and commits. A factor is an attribute related to an issue or a commit. Such attributes are references, labels, timestamp, developers, pull requests,

hashtags and files changed. Further, the factors are compared to the issues that include and exclude commits. That way, there will be strong evidence of which factor defined the connection between a commit and issue. By knowing which factors are more likely to exist between issues and commits, it will be apparent how to connect them and therefore answer RQ (2). Also, by weighting the factors, an there can be a prediction which factors are more likely to be relevant and beneficial to the project and therefore give an answer to RQ (3).

C. Terms

In this research paper, there are several terms that are going to be used. For the purpose of clarity these terms are as follows:

- *Issue* – submitted issue in GitHub Issue Tracker
- *Commit* – represents individual change to a file in a project. Every time a commits is submitted it creates an unique ID (hashtag, hash), which allows tracking the made changes
- *Connection/Link* – the relationship between an issue and a commit investigated by this research paper.
- *Direct commit (direct link)* – commit submitted to a specific issue. Also, the issue contains commits.
- *Indirect commit (indirect link)* – submitted commits for issue A it might also fix indirectly issue B
- *Hashtag* – is the unique ID generated for every commit. *Hashtag* is typically only one, representing a single commit.

III. METHODOLOGY

A. Research strategy

The research strategy for this study will be based on the Design Science Research (DSR) methodology. This methodology focuses on the development and application of designed artifacts for the purpose of understanding a problem and gaining knowledge [6]. One of the reasons to choose DSR is because this method can produce meaningful results in the absence of existing theory base. The mission of this methodology is to provide an innovative solution to a problem in order to not only describe and explain it, but also predict [4]. This research follows the 7 guidelines for a Design Science Research described by Hevner [6]:

1) Design as an artifact - DSR must provide an artifact. In the case of this research - theoretical framework

2) Problem relevance - The developed solution must be relevant to the specified problem/research

3) Design evaluation - The quality of the designed artifact must be ensured by evaluation methods. Observational evaluation will be used for the developed theoretical framework

4) Research contributions - The DSR methodology has to show a clear contribution of the design artifact or methodology. This will be presented in the Result section of the study

5) Research rigor - DSR depends on the correct application of the construction and evaluation of the designed solution

6) Design as a search process - Reach desired results by employing the available means according to the problematic environment it will apply

7) Communication of research - DSR methodology must be presented in an adequate form for variable audiences

The main objective will be to accurately translate issues and commits with all possible relations between them. The artifact developed will be a theoretical framework that includes all commits and issues from a project with all metadata that surrounds it - developers, issue submitter, keywords, date, references and tags. The data that was collected is translated into quantitative. This is because the research is primarily objective and outcome-oriented. The objective nature focuses on testing concepts. With the development of a framework, the relationship between issues and commits is manually analysed. The analysis is focused on a descriptive way which will help determine the weights of the different factors. Such descriptive statistics give a good idea of which factors are more important than others. Beyond that, factors could be combined, in order to illustrate a better defined relationship between issues and commits. For example, if tags do not present significant value, a combination with other factors such as developers, days to get resolved or reference, could give better results.

B. Research process

The research process started by selecting three software projects from GitHub. The projects were selected with intentional difference in size of total issues in order to show how and if factors differentiate between different projects. The nature of the projects did not play a role in the selection. In order to keep the bias selection to minimum, the criteria was that all projects need to have a significant amount of issues and all projects are maintained and updated regularly. The first project that was selected is called “TestPilot-Containers” [7]. It contains 194 closed issues and 100 Open, which totals in 294 issues. The second project, named “TabCenter” [8], contains 699 closed issues and 148 open issues, resulting in 847 in total. The third project is called “SSH Scan” and it is consisted of 99 closed issues and 32 open [9]. From the three projects, a total of 300 issues (100 per project) both opened and closed was collected.

Issue	No commits	Commits	Closed issues	Total collected	Total
Project 1	36	61	61	100	294
Project 2	42	72	72	100	847
Project 3	25	81	79	100	131
Total	103	109	212	300	1272

Table 1. *Projects summary*

Based on the collected data, a descriptive analysis could be conducted. That will allow the determination of how complex the problem is, as well as, which factors are more important than others. This type of data can be expanded and will allow understanding of how to solve the tractability problem mentioned in RQ (2).
A spreadsheet was created for the data extraction to include all essential information about the issues and the commits. The data has been collected manually. Factors range from tags, dates, keywords, files changes, opened/closed reference and developers. Table 2 explains the different factors that were chosen to collect:

Factor	Description
Issue ID	The unique number of an issue
Status	Current status of an issue. Possible states – Open/Closed
Direct link	The issue has (not) a commit. Possible states – Yes/No
Commit ID	The unique number of a commit
Commit name	The name a commit is submit with
Changed files	Specific files that are added/changed/removed with the commit
Date submitted	When was the issue created
Date resolved	When was the issue closed
Days taken	Time taken for the issue to be resolved
Keywords	Keywords contained in the commit name
Tags	Tags/Labels added to the issue
Issue submitted by	Developer who created the issue
Developers	Developers which worked/resolved the issue
References	References that are related to the issue. Possible states – Open/Closed
Milestone name	Checks if issue is part of a milestone

Table 2. *Selected factors for data collection*

Among the 300 collected issues, 212 are with a status "Closed" issues. The rest are in an "Open" state.
The closed issues that have a direct commit are 107, which make up 50.5% of all closed issues. That makes them relatively 1/3 of all extracted issues.
Additionally, 19 issues without commits from all collected issues have been resolved because they are a duplicate of another issue.
Issue that are open were collected together with the closed issues as they come. Open issue, still in progress, did not have commits. However, they are part of the data collection and while they might not have a decisive factor for the research questions, open issues contain some of the analyzed factors. That could make an ideal test for predictability. However, it is hard to ensure if the prediction is correct due to the time restrains.

C. *Data collection*

The data collection procedure consisted of selecting data that would serve for the purpose of this study, as well as choosing what type of data will be required. The technique for data collection was done through observations and examination of issue record. Observations are a good way of gaining knowledge about a particular situation and the frequency of a certain behavior or phenomenon. Each issue has its own dedicated webpage on GitHub which includes various data. Since the study is developing a theoretical framework based on a new concept a larger amount of data needed to be collected. This allows adequate filtering to determine importance and the answering of the established research questions. Table 2 describes all data that was included in the webpage and that could be useful.
The following figures show a web page of an issue with direct commit. All issue follow similar patterns, with some missing certain metadata. The following example includes all information as represented in Table 2.

Fig 2. *Issue title and ID includes also status and developer who created the issue*

Figure 2 illustrates information which all issues have by default – titles, issue ID, status and user who submitted the issue. Titles vary from simple and basic to detailed and descriptive. At the end of the title, the unique ID of the issue is represented with the symbol #. The developer who created the issue is displayed under the title, together with the date of submission. The status of an issue can be either closed or open. In this example, the status is *Closed*.

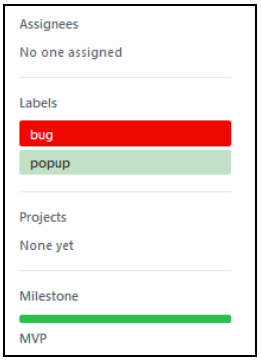

Fig 3. *Tags and Milestone*

Figure 3 presents data which is not included for all reported issues. That statement is true for both *Labels* and *Milestone*. Labels, called also *tags* in this study, have assigned values, which are used as descriptive information

about the reported issue. GitHub offers some default label names, but developers could use custom ones. Milestone is also a customly assigned value, which varies from project to project.

Fig 4 (a). *Commit fixing the issue and developer who closed it*

Fig 4 (b). *Pull request consited of conversation, commits and files changed information*

Figure 4(a) represents a case when the reported issue is closed and it includes a commits that resolves it. The commit is considered a reference to the issue and consisted of a commit name and commit ID. Figure 4(b) represents a Pull requests that consists of one or more commits and provides information which files are changed. Further, GitHub also reports when the issue was closed and by whom. In case a direct commit is not present, the developer could close the issue based on other reasons. GitHub allows for developers to reference issues with issues. This is done for traceability and redundancy purposes.

D. Issue types

This research outlines three main classes of issues. They are based on each possible status of an issue. An issue could be *Closed with commits*, *Closed without commits* or *Open*. Those are the names of the classes as illustrated in Fig 5.

- *Closed with commits* contains a couple of subclasses, which further defines the main class. The commit, included in a closed issue, could be either a part of Pull Request or be directly link to the issue. When the commit is directly resolving an issue it is characterized by a unique ID, also defined as *hashtag*. The Pull Request can contain one or more *hashtags* and it is preferred in cases when one wants to inform other developers about the changes. Despite their similarities, it is important to establish difference between those two subclasses. In the case of a closed issue with a hashtag, it is safe to say that the issue is fully resolved. When Pull Request is discussed, it is important to notice that developers could add follow-up commits as well as review the potential changes. That means that even though in most cases issues are fully resolved through Pull Requests, it is still possible for an issue to be partly resolved while expecting follow-up changes.
- *Closed without commits* is the second class of issues. It is defined by not having commits attached to the issue. This class is divided by several subclasses – *Reference*, *No Reference*, *Questions* and *Duplicate*. The subclass *Reference* is exploring the issues that contains references to other issues with solutions (or commits) or in case the issues have been resolved by a newer version of the software. Subclass *Duplicate*, as the name suggests, refers to those issues that have already been resolved. The duplicates do not contain commits on their own, but the issues they represent could have been resolved by one. In a small amount of cases, people submit questions in the form of an issue. Those issues are resolved by comments from the community and/or developers. The fourth subclass *No reference* represents those issues that do not contain any commits and do not have any connections to other resolved issues. Such cases could be when an issue could not be reproduced or there is no immediate plan to resolve it. There is a scenario of an issue in the *No reference* subclass, which refers to issues that are closed with no explanation. In some cases, the issue in this subcategory might contain comments from the developers, but with no solutions or references offered. In other cases, the issues could be closed if no one offers a solution for a long time.
- *Open* is the third type of issue, which at the time of the data collection those issues have not been resolved. Their status is Open and they could be resolved by any of the ways mentioned above.

.

E. Factor types

Factors need to be established, in order to define the connection between commits and issues. A single issue could contain multiple commits which could relate to it. Initially, the issue has a single commit after it is created. It could have more commits submitted at any point. The final commit is the one at which point the issue is closed. Factors will help filter out the relevant possibilities. The link could be described by a single or multiple factors at once. More factors mean that the link between commits and issues is more firm. That is because factors could narrow down the list of potential relevant commits. The factors are chosen in a way, so that together they make a meaningful mapping between issues and commits. This leads to a commonality between issues and commits. Most common factors are d*eveloper*, *time*, *reference*, *files changed* or *labels/tags.*

- *Developer:* is a factor referring to the name of the developer who submitted the issue and the developer who resolved the issue. In some cases that could be the same person. This could be a significant factor, in a case when it is known that the person submitting the issues is also a developer and not only a user. Therefore, it is important to be able to know if an issue is resolved by the same person or not.
- *Time stamp:* creates a filter which creates a candidate list of commits submitted between the Open and the Closed status of an issue. Naturally, the closer the commit is to the resolved date of an issue, the more likely is to be relevant. A combination with the *developer*, *files changed*, *labels* and *references* could help narrow down the list of possible commit

candidates and therefore, make a close call to which issue it belongs.

- *References:* is a link that points at different issues which may or may not contain commits. An issue could have one or more issues as references. Exploring the referenced issues could lead to the discovery of a commit that is applicable to the original issue. This might be possible due to the fact that referenced issues could be referenced because they already contain the desirable commit which will solve the issue.
- *Files changed:* stands for all the files that are affected by submitted commits. That factor could be helpful in a case of trying to confirm relevance between the commit and the issue. Further it could contribute to linking an issue with commits to other issues with commits.
- *Labels/Tags:* is a descriptive factor that gives an idea about the nature of the issue. For example, issues that are labeled *Questions* can be excluded from the investigation in order to save time and resources.
- *Hashtag:* is a commit directly submitted to the issue and mentioned when the issue is closed. It is the clearest connection between and issue and a commit.
- *Pull request:* could contain one or more commits. It is a formal way of discussing and reviewing changes as well as allowing other developers to add follow-up commits.

F. Data analysis

Once the data extraction was complete, the data was processed and organized for analysis. It was placed into tables where the columns represented each factor as shown in Table 2. Each project has separate tables. In order to get insight and understand the data, a variety of techniques were applied. Descriptive statistics are used to understand the nature of the issues and exemplify how connected they are to the commits. That also gives an indication to how complex the problem is. The overall approach to analyze the data is exploratory, which seeks to summarize the main characteristics and encourage exploration of data that has not been explored in such details before. Further, a conceptual algorithm was created to represent the theoretical solution for RQ (2). The conceptual algorithm establishes a connection between an issue and a commit through the factor filters established beforehand. Single factors can be analyzed individually to see if there is strong evidence that the factor affects the relationship between issues and commits. The state of other factors is also considered. The idea behind is that, two factors by themselves might not cause a difference, but when combined might show significance and therefore help address the RQ (2).

1) Descriptive statistics

a) Labels/Tags

The findings show that the total issues which contain tags/labels are 157 and 104 of all tagged issues are closed. This represents 66.2% of all issues containing tags. Further, it is important to separate the issues based on direct commit (DC) and not direct commit (NC). The DC issues with a tag make up 49% from all closed issues. Further, by dividing in the closed issues into ones with DC and the NC ones, it is determined that 59.8% (64 closed issues) do have commits. That leaves 37.1% for tags with NC.

Issue type	Project 1 (%)	Project 2 (%)	Project 3 (%)	Total (%)
Total issues with tags	74	19	64	157
Total closed issues	61	72	79	212
Closed with tags – from all tagged issues	41 *(55.4%)*	12 *(63.2%)*	51 (79.7%)	104 *(66.2%)*
Closed with tags – from total issues	41 *(67%)*	12 *(17%)*	51 *(64.6%)*	104 *(49%)*

Table 3. *Classification of tagged issues*

Looking further into the issues with tags, it is discovered that 64 (60%) DC issues and 39 (37.1%) NC issues contain a tag. This result shows that it is ~20% more likely for an issue with commit to contain a tag. The closed issues that do not contain any tags for DC and NC are respectively - 43 (40%) and 68 (64.8%) issues.

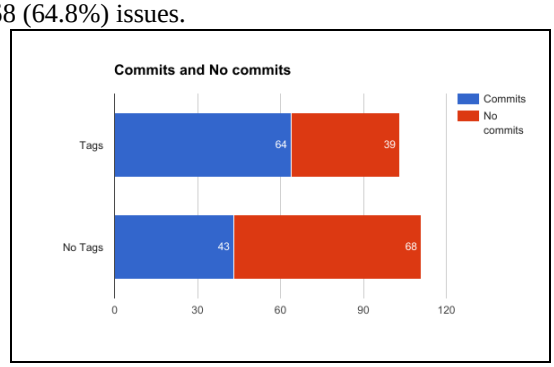

Fig.5. *Tags distribution within closed issues*

GitHub allows tags to be custom, but it also contains some that are default – such as *bug, enhancement, UI.* Many projects use the custom labels for better description of the issues. However, there are labels that are common across projects. The most popular ones that were found in the three projects selected in this study were *bug* and *enhancement.*

Closed issues (DC)	Project 1	Project 2	Project 3	Total
Closed - tag "bug"	12	2	12	26
Closed – tag “enhancement”	1	0	12	13

Table 4. *Issues with commits and specific tags*

Closed issues (NC)	Project 1	Project 2	Project 3	Total
Closed - tag "bug"	9	1	5	15

Closed – tag "enhancement"	2	0	6	8

Table 5. *Issues without commits and specific tags*

b) References

Another factor that was observed are the references. They could be an important link when it comes to indirect commits (NC). A reference for NC could mean that there is a commit that fixes the particular issues and vice versa, a reference in DC could mean that the commit resolves a NC. In the data collection the references that were collected had "Closed" and "Open" state – representing the corresponding issue.

Total issues that include reference are 93. Further, 70 issues with reference are closed. That is 74% of all referenced issues and 33% of the total closed issues. To look further into the reference as a factor, the closed issues were divided into the two sub-states they have. The DC issues represent 54.3% of all references issues, where NC issues represent 45.7%.

Issue type	Project 1 (%)	Project 2 (%)	Project 3 (%)	Total (%)
Total issues with ref	29	37	27	93
Total closed issues	61	72	79	212
Closed with tags – from all ref issues	24 *(83%)*	26 *(70%)*	20 (69%)	70 *(74%)*
Closed with tags – from total issues	24 *(39%)*	26 *(36%)*	20 *(28%)*	70 *(33%)*

Table 6. *Classification of referenced issues*

Figure 6 represents the strength of a reference when compared with closed issues which do not contain any references. The comparison shown below visualizes DC and NC with both possibilities of containing or missing a reference:

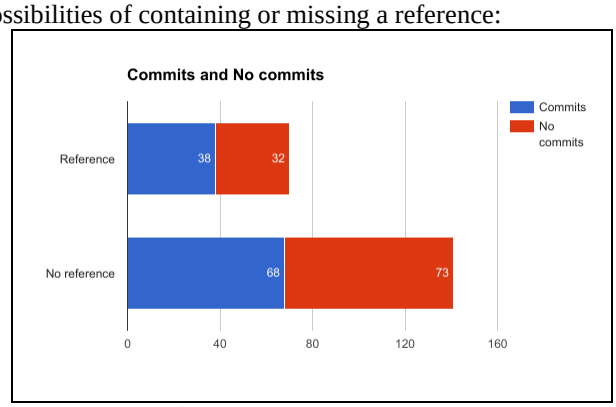

Fig 6. *References distribution within closed issues*

c) Time

Another factor that was examined was time. The spreadsheet includes both date of submission and day of resolve. Based on that it was calculated the days it took for a certain type of issues to be resolved. The average days were calculated for both DC and NC. It turns out that closed issues with commits are closed in about 22 days on average. For the issues that do not have commits, that time range is 23 days.

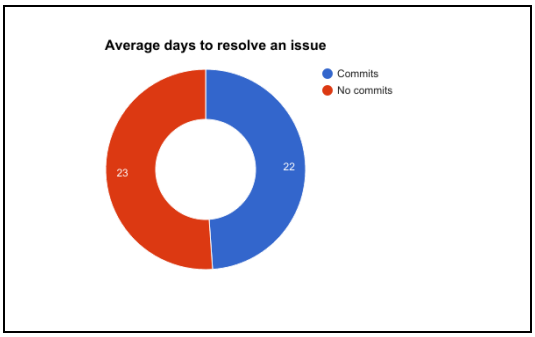

Fig 7. *Average days to resolve an issue with DC or NC*

Figure 7 shows that issues with commits are resolved up to a day faster than the issues with no direct commits.

Closed - commits	Project 1	Project 2	Project 3	Total
Average days to resolve	17	31	18	22

Table 7. *Classification of average days for DC*

Closed – No commits	Project 1	Project 2	Project 3	Total
Average days to resolve	14	25	31	23

Table 8. *Classification of average days for NC*

Table 6 and 7 show that for certain projects the time to resolve issues without commits is less. However, in total that is not the case.

d) Developers

Developers who open issues can also close them. In certain cases this is due to the fact that the project is relatively small and not many developers are maintaining it. In other cases, there are restrictions to who has access to close issues. According to the results from the data extraction, 27.1% of closed issues with commits are also closed by the same developer who has opened the issue. In the case of NC issues, that percentage is 36.2%. That leaves issues closed by a different developer with 72.9% and 63.8% for DC and NC, respectively.

Issues – with commits	Project 1 (%)	Project 2 (%)	Project 3 (%)	Total (%)
Same developer	4 (16%)	0	25 (48.1%)	29 (27.1%)
Different developer	21 (84%)	30 (100%)	27 (51.9%)	78 (72.9%)

Table 9. *Classification of issues with commits closed by developers*

Issues – No commits	Project 1 (%)	Project 2 (%)	Project 3 (%)	Total (%)
Same developer	10 (27.8%)	7 (16.7%)	21 (77.8%)	38 (36.2%)
Different developer	26 (72.2%)	35 (83.3%)	6 (22.2%)	67 (63.8%)

Table 10. *Classification of issues without commits closed by developers*

e) Combination – references and labels

There are multiple factors that could predict and determine the connection between issues and commits. The factors included in the data collection have been chosen based on the information on issues provided by the GitHub issue tracker. The justification is that every factor, even insignificant by itself, could be determining when combined with another one. Therefore, the table includes a combination between labels and references.

G. Limitations and Risks

One important limitation for this design research methodology is that the theoretical framework, as an artifact, cannot be compared to similar tools. The concept of relationship between issues and commits is new and therefore, tools or other theoretical frameworks are not available for comparison. To minimize the risk of not having meaningful results, most factors gathered with the data extraction were analyzed with descriptive statistics were used to determine their significance. Another limitation is the small pool of projects that were explored, which could lead to non-accurate representation of the rest of open source projects. Further, that limitation hides a risk of a biased choice. To limit the risk of biased results the projects were selected at random. No previous knowledge about those projects was known previously. To make sure that those project represent the rest of the open source projects on GitHub, the only selection criteria was the projects to have significant difference in their sizes. That created a filter with three pools – big projects, medium projects and small projects. From those three pools the projects were selected randomly.

IV. RESULTS

The detailed data extraction will help determine a stronger relationship between issues and commits. Even in the case of an issue that has been closed without a direct commits (also named indirect link in the spreadsheet), the issues are reported due to the possibility of an indirect link. An issue could be closed as a duplicate of another issue, or it could be closed because of a commit from another issue. In that case, that issue has an indirect link.

Closed issues	Project 1	Project 2	Project 3	Total
Tags	41 (67%)	12 (17%)	51 (71%)	104 (49%)
References	24 (39%)	26 (36%)	20 (28%)	70 (33%)
Tags and Refer	25 (41%)	8 (11%)	14 (19%)	47 (22%)
No tags or reference	15 (25%)	42 (58%)	23 (32%)	80 (38%)

Table 11. *Classification of combined factors*

In order to determine which factors can be beneficial to tracing issues to commits or vice versa, factors from Table 2 were selected to evaluate.

A. Conceptual algorithm

The conceptual algorithm, illustrated in Fig 8, is the conceptual approach of collection and analysis through different factors. The factors act as a filtering system, in order to evaluate the commits and define a relationship with the issue. In case a factor is missing for a specific issue, the algorithm skips to the next one. The conceptual algorithm checks the status of an issue. The possible outcomes are *Closed* or *Open.* This is the first try to determine the issue type as described in section III (D). If the status is *Open* there are the possibilities of exiting the algorithm or running through the main factors in order to suggest possible commit connections for the issue. If the status comes back as *Closed* then immediately the next iteration is to determine if the issue has commits or not. That is the final step of establishing the issue type as represented in Figure 10. Lack of commits with trigger the factor checking, if confirmed with *Yes*. That will help determine possible connections that fit the specific issue. After the commits are narrowed down, the user has an option to filter through the factors again, potentially discovering other candidates. The system will end if there is no necessity for another filtering. In case, the issue contains a commit, the algorithm will try to establish if the commit is part of a Pull Request. If the commit is not part of a Pull request, the algorithm will proceed with collecting the commit ID (hashtag). If the commits is part of a Pull request, the conceptual approach will collect the Pull Request ID, together with the commits related to the Pull Request. Further, the commits will be filtered through the factor algorithm goes through extra iterations.

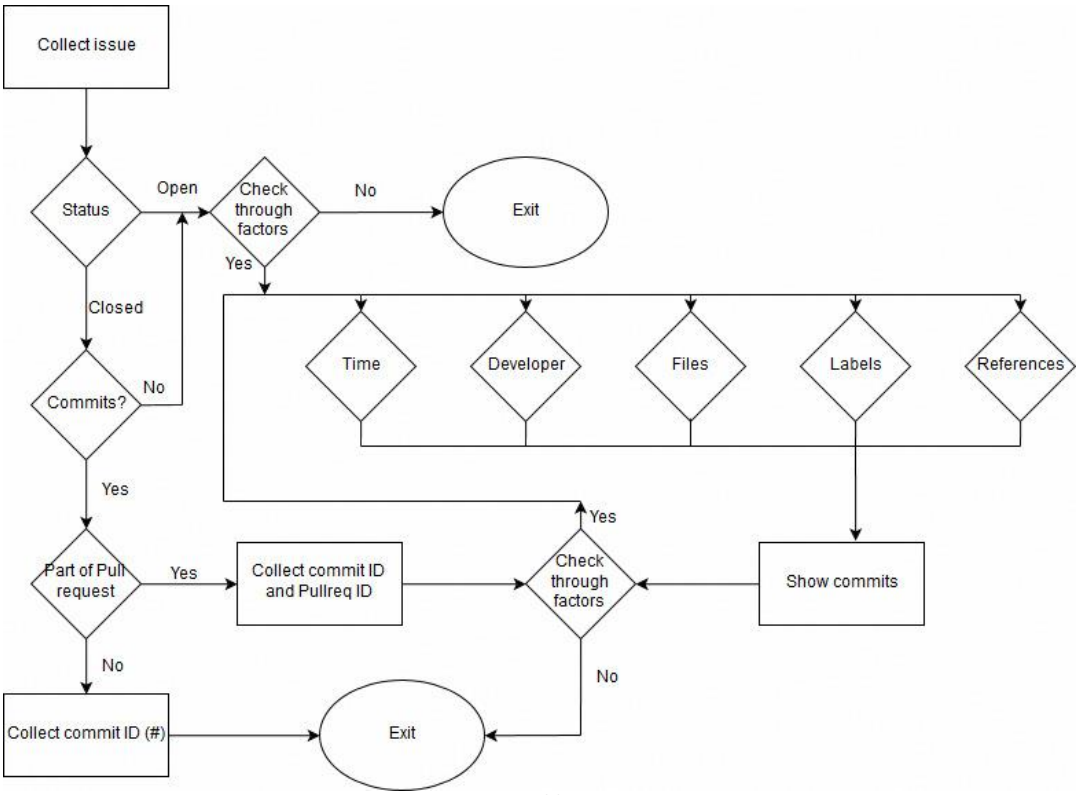

Fig 8.*Concept of factor analysis*

B. *State of factors*

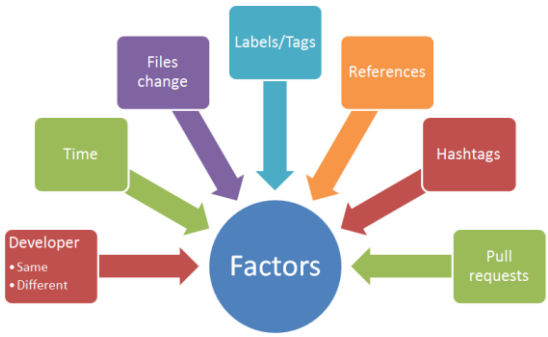

Fig 9. *Represents the state of the factors*

An example of multiple factors could be a case where, we assume there is an issue A that has a commit B which has a time stamp as well as, names of the files changed. Later, if commit C is found with a time stamp close to commit B and changes the same files commit B does, then it could be safe to say that commit C could have a link to issue A. However, that example could work with one of the factors, as long as it gives satisfactory confirmation of a connection between the issue and the commit.

Generally, *hashtags* represent the strongest bond between issues and commits. This is due the fact that a commit specifically targets to resolve an issue. That could be confirmed after the commit is submitted; the issue is marked as *Closed*. The second strongest factor is a *Pull Request* which could accommodate single or multiple commits. Due to the nature of a Pull request, it may or may not fully resolve an issue. That could be confirmed by: combining factors with the Pull request and drawing a confidence level.

C. *Issue types*

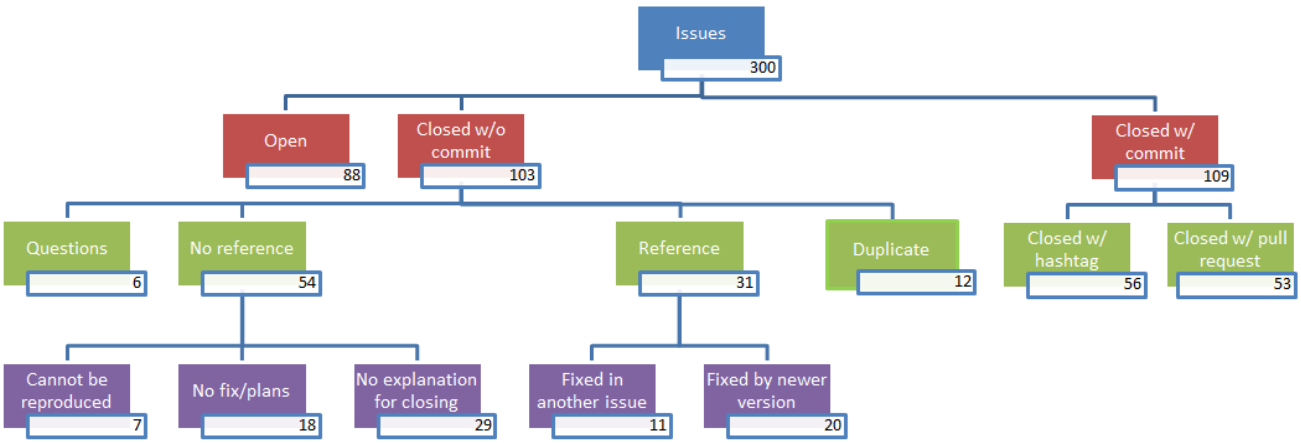

Fig 10. Represents the issue types

Figure 10 represents classification of issue types. The issue types were determined by the performed data collection and relying on the definitions from section III (D). Based on the entire issue collection, each issue type was calculated.

D. Data samples

Each issue type can be investigated by following the algorithm concept described in section III (F) and represented in Figure 8. The issues exemplified have different statuses and factors involved. That is, in order to create diversity and show multiple cases with different variables.

- Issue #74 from the project SSH_Scan is a closed issue that contains a commit. The name of the issue is *Add support for IPv4 fallback when IPv6 cannot be established*. The selected issue goes through the first iteration of the algorithm as shown in Figure 8. Since, it is closed, it proceeds to the next iterations. It has a commit and it is not a part of Pull Request, therefore it reaches the state of *Collecting commit ID* and exiting the program. The issue contains factors that are only secondary since it has a hashtag that resolves the issue. At this point, the additional information in the form of factors could be collected for future references. This example is used to show ideal connection between an issue and a commit, since the latter is directly attached to the first and resolves it.

Fig 11. *Shows the issue name and date of creation*

Fig 12. *Shows hashtag number when closing the issue*

- Issue #410 from the project TabCenter is a closed issue that is resolved by Pull Request. The name of the issue is *Option to open links at top or bottom*. The issue goes through the first two checks that aim to establish the issue type. Next, the algorithm is at the iteration of *Part of Pull Request*. Since the issue is resolved by Pull request, the algorithm proceeds to collect the commits ID and Pull request ID. Since the issue is already resolved by commits, it is safe to say that the issue has a definite connection between the commits. However, because of the nature of the Pull Request and the time difference between the issue being resolved (2016-09-07) and the Pull Request submitted (2016-08-22), a check on other potential commit candidates might be necessary.

Fig 13. *Shows the issue name and date of creation*

Fig 14. *Shows reference, pull request ID, closing date*

If that is the case, then the algorithm proceeds to the next step *Check through factors*. In that stage, all factors mentioned in section III (E) are analyzed. After going through the factors, the conceptual algorithm suggests candidate commits based on the factors it has investigated. Based on the time frame, some of the suggested commits are listed in Figure 15:

Fig 15. Candidate commits based on the time factor

Issue #410 is missing Labels/Tags, so that factor does not apply. Looking at the developers, it seems it has been resolved by a different developer than the developer posting the issue. The developer closing the issue has submitted other issues the same day of the resolvement. For example, looking into the commit *fix:Scroll to the correct tab when Tab Center is expanded* seems that is not only submitted on the same day by the same developer who closed the original issue, but also it pushes changed to the same file *vericaltabs.js*. The investigated issue includes multiple references. They could point to similar issues and offering different solutions.

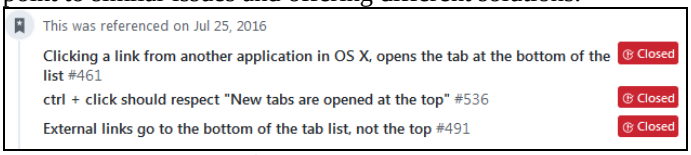

Fig 16. *References attached to issue #410*

Investigating those issues implies that they are either duplicates or have been resolved by other commits submitted for issue #410.

- Issue #71 from the project TestPilot-Containers is an issue without any commits. The name of the issue is *Replace "No Container" copy with something else*.

Fig 17.*Shows issue name and date of creation*

Fig 18. Shows date of resolving the issue

The issue does not have any references, but it contains labels/tags.

Fig 19. *Shows the labels assigned to the issue*

Following the logical direction of the conceptual algorithm, the iteration is at the step where it checks for submitted commits. When it does not find any, it starts checking for candidate list based on the defined factors. After using the time range between opening and closing the issue, it provides multiple commits. The ones presented below are the closest to the closure date:

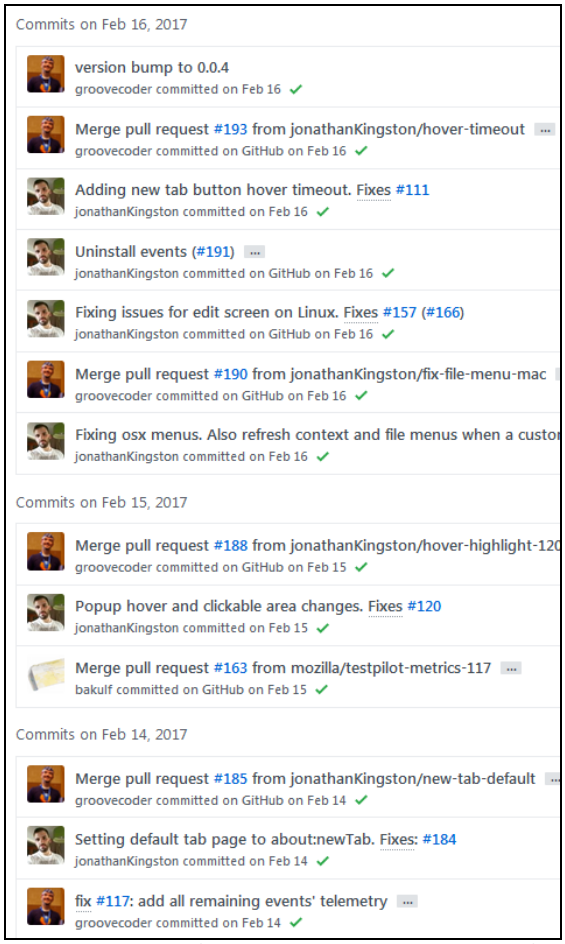

Fig 20. *Presents some of the candidate commits based on time frame*

The next factor is the developer. It is clear from Figure 20 that the developer who has closed the issue, has not submitted any commits prior to the closure of the issue. Issue #71 does not contain any references to other issues and does not change any

files. However, it contains labels which might narrow down the candidate list of commits. Combining the time frame factor and the labels factor return no commits that match the filters. According to the algorithm, the issue could run again through the factors and show the commits based on the time frame. The other option is to end the algorithm.

Some of the investigated issues cannot have possible connection with commits. This is due to the fact that they are either questions which need to be addressed by the developers or the issue itself is open. There are issues that are closed without an explanation.

- Issue #367 from the project TestPilot-Containers is a closed issue marked by a label as Question. Based on the diagram in Figure 5, the issues marked as Questions are issues closed without commits. According to the algorithm diagram, an issue without a commit can run through the factors or it can skip the process and exit it. Considering the fact, the issue is labeled as Question and the issue is closed, it is safe to say that the issue is fully resolve without the need of commits.

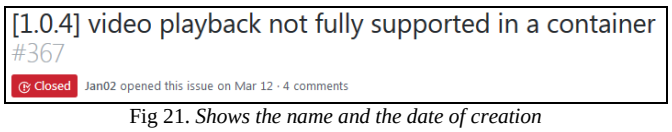

Fig 21. *Shows the name and the date of creation*

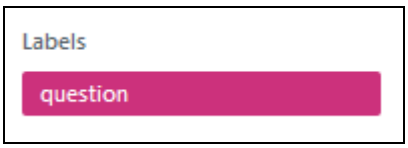

Fig 22. *Issue is marked as Question*

- Issue #984 from the project TabCenter is a closed issue without any factors or commits attached to it. Further, the issue is resolved in the same day.

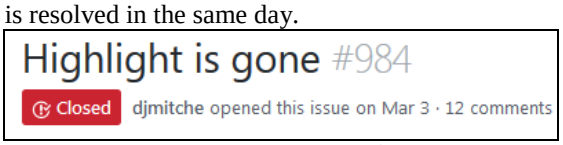

Fig 23. *Shows name and date of creation*

Fig 24. *Issue resolvement date*

In some issue cases, the algorithm is not applicable and candidate list cannot be determine.

- Issue #49 is an open issue from the project SSH_Scan. The issue contains labels and references to other issues. In this case, the algorithm concept could be beneficial to Open issues with suggestions and list of options. Depending on the issue and when it was opened, it might not be a good idea to rely on the time frame factor. This is because the range of possible commits could be too big to analyze. However, when combined with other factors, it could be possible to rely on the start date of the issue.

V. Discussion

After examining the results it is clear that only part of the factors as described in Table 2 are relevant to the current study. While, some such as *Changed files* are not currently applicable to this research, the theoretical framework requires all metadata to be collected. This is due to the fact that in the future different projects might need additional descriptive factors in order to define a connection. The current factors that contribute to the research are – *tags, time range, references, hashtags, pull requests and developers.* To define a more accurate relationship and obtain a better result, a combination of factors were used. Since, most issues rarely contain all factors it is important to be careful to which factors are included. When multiple factors are available in an issue, it is important to be cautious about the choice. This is well illustrated in issue #71, where combining multiple factors does not help narrow down the list with possible candidate commits. However, in most cases, multiple factors point at more relevant commits, as described in issue #410.

The descriptive statistics in section III (F) are important for this study as it shows the frequent occurrence of certain factors. It affirms the selection of factors as described in section III (E). Looking into the statistical results, it is obvious that *Labels* have a significant application when it comes to issues. That is evident from the fact that 49% of all issues included a label. Even more, 60% of issues with direct commits are likely to be resolved, as pointed in Figure 8. Labels do not have a pattern for repeatability, since two of the most popular labels – *bugs* and *enhancement*, represent only 16.2% and 8.3% respectively. The fact that labels tend to be different increases the chance to have more accurate match when looking for a connection between an issue and a commit. However, it might also be more difficult to suggest even broader candidate commits.

Next factor that has significant values is *References*. They are included in 33% of all issues and 74% of all references issues are closed. When looking further into the data, it seems that 40.8% of all referenced issue are issues that contain commits. While it is less than half, it does seem like an important factor. Most references are used to close duplicates or to point to an issue that has been resolved. When an issue with a commit contains a reference, it is rather safe to say that this commit resolves at least one more issue. That is one way of defining indirect relationship between commits and issues. Further, references could be used as a fast and easy way to trace issues to commits or vice versa.

When it comes to the time it took to resolve issues, it is evident in Figure 7, that time by itself does not make a big difference. Time is most useful as a filter that sets range between the opening and the closing of an issue. That way it significantly limits the possibilities of fittings commits. Time range should be consider as one of the most useful factors when it comes to linking issues to commits and vice versa. Table 8 and 9 do not show significant difference in the time to resolve issues of different types.

Developers that open issues tend to not be the ones who are closing them. Only 27.1% of the users that have opened

issues have closed it with a commit. Most of that could be the people who maintain the projects or people who are aware of the issue and want to fix it officially. When it comes to issues without commits, 36.2% of the users that started it have resolved it. That tendency could be to the fact that issues without commits could be duplicates or questions. That statement is supported in the Results section by issue #367.

The statistic results point out that Labels and References are a significant part when it comes to defining connection between issues and commits. Therefore, a combination of the two was made, as shown in Table 11, to explore if the factors could be strengthened when combined. Two different approaches were taken. One was to introduce the combination of factors and the other was to resemble the issues that contain one or the other. When factors are combined, the closed issues represent only 22% of all issues. However, when the lack of combination is introduced the percentage raises to 62%. That means that more than half of the issues contain either a label or a reference. That sort of combination can be beneficial to the research as it provides evidence that issues that are fixed are more likely to have one or the other and therefore helping with addressing RQ (1). Further it also, helps with presenting which factors are important and therefore addressing RQ (2).

The hashtag is the factor that provides robust connection between issues and commits. This is due to the fact that an issue that is closed with a direct commit (hashtag), is with full certainty, resolved. Based on Figure 10, it is evident that hashtags are in 51.4% of all issues with commits. This means that more than half of the issues that contain commits already have a firm connection with those commits.

Pull request is the second most defining factor after hashtags. Pull requests contain one or more commits which close the issue. However, that might now always be the case. Sometimes, Pull requests are a work in progress. Multiple developers could submit commits, but that might not mean the issue is solved. On the other hand, it is very likely for a Pull request to be able to build a bridge between issues and commits. Especially after a Pull request is merged with the issue and that issue is closed.

The statistical data shows solid evidence that the chosen factors are not only relevant, but also important. It is safe to assume that the examples given in section IV(D) are representatives of the pool of data collection. All examples are chosen at random, based only on the criteria of representing each issue type as shown in Figure 10. The selected issues were put through the conceptual algorithm defined in Figure 8. The randomization lowered the bias choice as well as it gave more natural representation of the limitations of each issue type.

Issue #49 represents the Open status of issues. Typically, Open issues, depending on their age, do not have many descriptive characteristics. Therefore, using factors to connect or even suggest a commit is very limiting. The time frame factor could be beneficial in combination with another factor. It narrows down the possibilities of commits.

Issue #984 is representing the issues closed without explanation. That is a subclass of issues *closed without commits* as seen in Figure 10. #984 represents 28% of all issues without commits. That means that there is a chance 28% of the issues to not have connection with commits. This is not only due to the fact that those issues types do not have any of the other factors. It is possible that those issues are duplicates, that were never marked or issues that have been opened for a long time that nobody attempted to resolve.

Only 6% are issues that are questions. Even though, it represents a small amount of issues, it will not be possible to make a link between those issues and a commit. Further analyzing the issue types from Figure 10 shows that 52.4% of the issues have no references. However, those issues still have time and developers as descriptive factors. Some of them might include labels as well. That will allow for the conceptual algorithm to try and determine some possible links with commits. The commits shown by the conceptual algorithm have no definite character and are mostly suggestive. That is especially good in situations where there is no fix, a plan for an issue to be fixed or the issue cannot be reproduced. That way, even if those particular issues were not fixed, some similar and resolved issues might be available in the future.

Those results could not only be beneficial to future research at multiple projects at once, but also could be helpful to researchers working on a single project. Furthermore, it summarizes how the different factors affect the projects in GitHub.

VI. Conclusion and Future work

This research paper identified several issue types. Each issue type has specific characteristics. Knowing those characteristics allows for different approaches when establishing connection with commits. The approach is defined in the conceptual algorithm. It handles the different types of issues while trying to make a connection with possible commits.

After examining the descriptive statistics, it is evident that descriptive data such as labels, references, developers and timestamp, are beneficial for the projects. The statistics show that those factors are relevant and important for establishing connection between issues and commits. If issues contain any combination of the factors described in this study it is favorable that there will be a way to define the relationship of the issues and commits for that particular project. Beneficial outcome of this research is the theoretical framework and the landscape of issues which points to way of connecting issues with commits, as well as, which factors are involved and important. This is cost and time efficient for projects. That outcome is possible with the conceptual algorithm describe in the thesis. Identifying different types of issues and analyzing the issue through the factors are key contributions of this research, answering research questions RQ (1), RQ (2) and RQ (3).

For future work, more studies could be added in order to have better resemblance of the GitHub project base. Also, following the theoretical framework and the conceptual

algorithm, an automatic tool can be created gathering issues from the GitHub API and then analyzing the issues and their corresponding factors. Such a tool can be used by project developers and users contributing to the project.

ACKNOWLEDGEMENT

I would like to thank: my supervisor, Abdullah Mamun, for his valuable feedback and guidance, my wife, Tierra Dungarova, for her insightful advice and firm support, and my parents for their constant patience and love.

REFERENCES

[1] Cabot, J., Izquierdo, J. L. C., Cosentino, V., & Rolandi, B. (2015, March). Exploring the use of labels to categorize issues in Open-Source Software projects. In *2015 IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)* (pp. 550-554). IEEE.

[2] Gousios, G., Zaidman, A., Storey, M. A., & Van Deursen, A. (2015, May). Work practices and challenges in pull-based development: the integrator's perspective. In *Proceedings of the 37th International Conference on Software Engineering-Volume 1* (pp. 358-368). IEEE Press.

[3] M. Allamanis and C. Sutton. Mining source code repositories at massive scale using language modeling. In MSR, pages 207–216. IEEE, 2013

[4] G. Gousios and D. Spinellis, "GHTorrent: Github's data from a firehose," in Mining Software Repositories (MSR), 2012 9th IEEE Working Conference on. IEEE, 2012, pp. 12–21

[5] Van Der Veen, E., Gousios, G., & Zaidman, A. (2015, May). Automatically prioritizing pull requests. In *Proceedings of the 12th Working Conference on Mining Software Repositories* (pp. 357-361). IEEE Press.

[6] Van Aken, J. E. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. Journal of Management Studies

[7] https://github.com/mozilla/testpilot-containers/issues?utf8=%E2%9C%93&q=is%3Aissue

[8] https://github.com/bwinton/TabCenter/issues?utf8=%E2%9C%93&q=is%3Aissue

[9] https://github.com/mozilla/ssh_scan/issues?page=5&q=is%3Aissue+sort%3Acomments-des

Appendix

	Issues		Connection	Commits						Dates			Keywords			Tags		Developers			References - issues ID		
№	Issue	Issue Name	Status	Direct link	Commit ID	Commit Name	Changed file 1	Changed file 2	Change file 3	Change file 4	Change file 5	Date submitted	Date resolved	Days taken	Keyword 1	Keyword 2	Keyword 3	Tag 1	Tag 2	Issue submitted by	Developer 1	Developer 2	Closed references	Opened references	Milestone name
1	311	**Consider a tab context menu option**	Open	No								03/02						enhancement	needs:Product	pmac			325, 326, 329, 377, 3	2281, 317, 336, 387,	Stretch
2	113	**Measurements Questions Draft**	Closed	No								01/26	02/16	21						johngruen	groovecoder	TanviHacks	99, 117	118	MVP
3	336	**Embrace the use of containers for t**	Open	No								03/04		#NUM!				enhancement	vote for me	aleth			328	311, 535, 428, 430, 4	Stretch
4	346	**Custom containers reset**	Closed	No								03/08	03/15	7				bug		attis	groovecoder		349, 361, 388		Sprint 2
5	242	**Moving Container tabs to a new win**	Closed	Yes	3fb6976401	for #242: disable me	index.js	webextension/cs	webextension/js/popup.js			02/22	02/27	5	disable	incompatible	detected			SoftVision-CiprianMuresan	groovecoder				MVP
6	59	**Learn more link in onboarding does**	Closed	Yes	7f6b37db94	WIP: clean up for lau	popup.css	popup.js	popup.html			01/10	02/20	41	clean up	panel	WIP	popup	needs:Product	pdehaan	TanviHacks	groovecoder	204		Sprint 2
7	306	**Allow classifying of sites or urls as**	Closed	Yes	3e657a2e8d	Auto send users to s	index.js	package.json	webextension/ba	webextension/cs	webextension/js/	03/02	04/04	33	assigned	container	fixes	enhancement	needs:Product	crabel	jonathanKingston	groovecoder	418, 420	429, 430	Stretch
8	263	**First log in attempt fails when open**	Open	No								02/24		#NUM!				bug		SoftVision-CiprianMuresan				5528	Stretch
9	245	**Opening a new tab from a container**	Open	No								02/22		#NUM!				enhancement	needs:Product	SoftVision-CiprianMuresan				311, 356	Stretch
10	237	**original icons seem like they're ove**	Closed	Yes	d13122a26a	Fix for icons overlayi	data/usercontext.css					02/22	02/22	0	Fix	icons	overlaying	bug		kjozwiak	jonathanKingston				Stretch
11	441	**Extension opens a new tab when ar**	Open	No								04/10		#NUM!				enhancement	question	cpeterso					
12	99	**What survey question do we want t**	Closed	No								01/19	02/22	34				needs:Product		johngruen	johngruen			113	MVP
13	315	**Container menu slow to show up**	Open	No								03/02		#NUM!				bug		dCosminn					Stretch
14	172	**TypeError: ContextualIdentityServic**	Closed	Yes	4a5129a999	Fix the CookieMana	index.js					02/10	02/21	11	Fix	CookieMana	issue	bug		kjozwiak	jonathanKingston	bakulf			MVP
15	118	**Data dashboard**	Open	No								01/26		#NUM!						TanviHacks	groovecoder	clouserw			Sprint 4
16	390	**NewTab page middle click opens in**	Open	No								03/17		#NUM!						Drouchka	bakulf				Sprint 5
17	327	**In a contextualized "new tab" panel**	Closed	Yes	82cad13630	Fix the opening of a	index.js					03/03	03/23	20	Fix	menu	issue	bug		shavounet	bakulf	groovecoder		390	Sprint 2
18	119	**Add keyboard shortcuts for the plu**	Open	No								01/26		#NUM!				needs:Product		TanviHacks				393	323, 342, 311	Stretch
19	331	**Container options flashes up when**	Closed	No								03/03	03/09	6						jnatchtigall	groovecoder			338	
20	97	**Make sure experiments is E10S con**	Closed	No								01/19	02/14	26						johngruen	bakulf	kjozwiak			MVP
21	302	**Containers configuration is empty**	Closed	Yes	08fe3551010	If localStorage is dis	webextension/js/popup.js					03/02	03/07	5	localStorage	disabled	skip	bug		crabel	bakulf			395	Stretch
22	374	**Containers disappeared while I was**	Closed	Yes	f6d79cd6c2b	fix #374 for 2.0+ - dc	index.js	package.json	webextension/manifest.json			03/13	03/15	2	Fix	remove	containers	bug		ruihildt	groovecoder		388, 403, 399	397	Sprint 2
23	108	**Pass a security review**	Closed	No								01/25	02/15	21						wresuolc	wresuolc				MVP
24	8	**Collect Final Icon Assets**	Closed	No								12/15/2016	01/26	42				needs:UX		johngruen	fangshih				MVP
25	71	**Replace "No Container" copy with c**	Closed	No								01/11	02/17	37				needs:Product	tab bar	groovecoder	pdehaan				MVP
26	396	**Secure connection error when conn**	Open	No								03/19		#NUM!				bug	needs:Testing	lisajx					
27	244	**After Container tabs are moved to a**	Closed	No								02/22	03/10	16				bug		SoftVision-CiprianMuresan	bakulf		327		Sprint 1
28	328	**Show only tabs from the current co**	Closed	No								03/03	03/15	12						maverick74	groovecoder			336	
29	11	**Implement hide containers API func**	Closed	Yes	f87eece6cc0	Issue #11 - Showing	webextension/js/popup.js					12/15/2016	01/09	25	Issue	Showing	tabs	needs:Product		jonathanKingston	bakulf	jonathanKings	12, 40		
30	33	**51 -> 52 suport issues**	Closed	No								01/05	01/06	1						jonathanKingston	bakulf				
31	14	**Implement select-on-hover containe**	Closed	No								12/15/2016	01/09	25				needs:Product		johngruen	bakulf				
32	20	**Impelement Containers Onboarding**	Closed	Yes	679d8b9fbb	for #20: start onboar	webextension/cs	webextension/js/	web/extension/pc	webextension/im	webextension/im	12/22/2016	02/01	41	fix	additional	onboarding	popup		TanviHacks	groovecoder				MVP
33	62	**Plus' icon missing from assets**	Closed	No								01/10	02/03	24						bakulf	TanviHacks				
34	75	**"Error: this._identitiesState[identity**	Closed	Yes	2238775d0c	Remap tabs in case	index.js					01/11	02/01	21	Remap	tabs	unknown	bug		pdehaan	bakulf				MVP
35	7	**Default Containers: change "Bankir**	Closed	Yes	bdc81fb4a6	Banking to Finance	index.js					12/15/2016	02/19	66	Banking	Finance		has PR		johngruen	bakulf				MVP
36	321	**Combine addon telemetry data with**	Open	Yes	6282201b2c	for #321: only send a	webextension/background.js					03/03		#NUM!	send	counts	activated			bakulf	groovecoder				Sprint 4
37	156	**Restoring hidden containers after re**	Closed	No								02/07	02/11	4				needs:Product	needs:UX	TanviHacks	bakulf			165	417	
38	116	**Test Addon with FF 52 and report b**	Closed	No								01/26	02/16	21						TanviHacks	kjozwiak	groovecoder	127, 135, 150		MVP
39	454	**Browser is crashing when "open in**	Open	No								04/14		#NUM!				bug		trevan	kjozwiak				
40	398	**Customized containers lose their co**	Open	No								03/20		#NUM!				bug		nkestrel	bakulf				Sprint 4
41	397	**Permanent cookies set on same site**	Open	No								03/20		#NUM!				bug	needs:Testing	nkestrel	kjozwiak			449	417	Stretch
42	408	**Open tab in new window does not p**	Open	No								03/30		#NUM!						hwine	kjozwiak				Sprint 4
43	46	**hide tabs != close tabs**	Closed	No								01/08	02/01	24						bakulf	kjozwiak	bakulf			Stretch
44	131	**adding two new colors under "choo**	Closed	No								01/31	02/24	24				enhancement	needs:UX	kjozwiak	TanviHacks			7	Stretch
45	434	**Consider making search not abide t**	Open	No								04/06		#NUM!				enhancement	vote for me	jonathanKingston					Stretch
46	347	**List Container as experiment on tes**	Closed	No								03/08	03/09	1						D219	D219				
47	168	**Add ability to close a tab without sh**	Open	No								02/10		#NUM!				enhancement	good first bug	bakulf	kjozwiak				308 Stretch
48	339	**Containers not synchronised**	Open	No								03/05		#NUM!				enhancement	question	aplatypus					Stretch
49	424	**Add assignment menu on the tab co**	Open	Yes	2ee9682950	Adding in context me	webextension/background.js					04/05		#NUM!				bug	enhancement	CircleCode	jonathanKingston				
50	55	**Implement "Edit Containers" UI**	Closed	Yes	faaa3008a8d	for #55: edit-caontai	index.js	webextension/cs	webextension/js/	webextension/popup.html		01/10	01/26	16	container	panel	update	popup		groovecoder	groovecoder			131	MVP
51	417	**"Restore previous session" should**	Open	No								04/04		#NUM!				enhancement		cbrgalson				156	Stretch
52	318	**Allow per-container privacy settings**	Open	No								03/03		#NUM!				enhancement		amluto					Stretch
53	356	**Default Container**	Open	No								03/09		#NUM!				enhancement	vote for me	comzeradd					419 Stretch
54	125	**Audit files in XPI**	Closed	Yes	06a0e12ecd	Merge pull request #	.jpmignore	package.json	webextension/background.js			01/31	01/31	0	pull	request				pdehaan	pdehaan			128	
55	399	**Container is lost when pinning a tat**	Closed	No								03/20	03/31	11				bug	duplicate	ase1116	groovecoder			374	
56	367	**[1.0.4] video playback not fully supp**	Closed	No								03/12	04/11	30				question		Jan02	jonathanKingston				
57	431	**Tabs and history are lost when disa**	Open	No								04/06		#NUM!				bug		longears					
58	47	**Disable history per-container settin**	Open	No								01/08		#NUM!				enhancement	vote for me	nt1m				419	
59	243	**If Containers add-on is disabled, all**	Closed	Yes	b9cc9ad7dc	wip #243: confirm pr	index.js					02/22	03/06	12	wip	confirm	add-on	bug		SoftVision-CiprianMuresan	groovecoder			380	431
60	317	**Can't add Open Tab to a Container**	Open	No								03/03		#NUM!				enhancement		aplatypus				311, 336	Stretch
61	354	**Tab Center integration**	Closed	Yes	8fd58ffe855	Adding container CSS	for highlight color change					03/09	04/11	33	adding	container	change	addon incompatibility		fvsch	abaez	ChrisW-B			
62	419	**Proposal: Private browsing by defa**	Open	No								04/04		#NUM!				enhancement	vote for me	pmac				429, 356	Stretch
63	395	**Having "Allow cookies from sites" c**	Open	No								03/19		#NUM!				bug	needs:Testing	mertzi					
64	337	**Awkward delay when opening a new**	Open	No								03/04		#NUM!				bug		emmyemi					Stretch
65	32	**How does Unhide behave**	Closed	No								01/05	01/06	1				question	needs:UX	TanviHacks	TanviHacks				
66	349	**User created containers lost after fi**	Closed	No								03/08	03/09	1				bug	duplicate	src-101	groovecoder			346	
67	179	**listing tabs via ">" doesn't always w**	Closed	Yes	76fc91be9a5	Remap tabs when o	index.js					02/14	02/17	3	remap	tabs		bug	popup	kjozwiak	bakulf			202	MVP
68	247	**Context menu persists after openin**	Open	No								02/22		#NUM!				bug		SoftVision-CiprianMuresan					Stretch
69	391	**Let users set custom colors**	Open	No								03/17		#NUM!				enhancement	vote for me	johngruen					Stretch
70	30	**File Menu -> New Container Tab**	Closed	No								01/05	01/19	14				needs:Product	question	TanviHacks	bakulf	johngruen			
71	273	**See if there is a bug in moving tabs**	Open	No								02/25		#NUM!						TanviHacks	groovecoder				Stretch
72	335	**Bug in hover menu display (via "Tat**	Closed	No								03/04	03/10	6				bug	addon incompability	bunnybooboo	groovecoder			334	384 Sprint 1
73	159	**Getting a signed build that works w**	Closed	No								02/08	02/09	1						kjozwiak	kjozwiak				
74	112	**Create tooltips for containers icons**	Open	No								01/26		#NUM!				good first bug	popup	TanviHacks					Stretch
75	324	**Inconsistent terminology ("containe**	Open	No								03/03		#NUM!				bug		attis					Stretch
76	457	**Features**	Closed	No								04/21	04/23	2						raymanB9	jonathanKingston				
77	96	**Focused Container Tab on MacOS t**	Closed	No								01/19	02/17	29				bug	tab bar	johngruen	johngruen		13, 95, 195		MVP
78	183	**Restoring container tabs after the a**	Closed	Yes	6d1c5ff09b5	Fix the restoring con	index.js					02/14	02/19	5	Fix	container	removed	bug		kjozwiak	bakulf		173, 211		MVP
79	13	**Implement tab color bars to match t**	Closed	Yes	f62873fffc19	Initial tab restyle wor	index.js	data/chrome.css				12/15/2016	01/26	42	tab	experiment	bug			johngruen	jonathanKingston			50	MVP
80	271	**Add question to survey about incor**	Closed	No								02/24	03/10	14						groovecoder	johngruen				Sprint 1
81	403	**Custom container disappeared afte**	Closed	No								03/24	03/27	3				bug	duplicate	youwenliang	groovecoder				
82	323	**Open bookmark into Container**	Open	No								03/03		#NUM!				enhancement	vote for me	attis					Stretch
83	54	**Error when running ESLint**	Closed	No								01/09	01/10	1						pdehaan	bakulf				
84	423	**(un)assignment menu not shown**	Open	No								04/05		#NUM!				bug		CircleCode	jonathanKingston				Sprint 4
85	382	**Popup panel is empty/blank, does n**	Open	No								03/15		#NUM!				bug		nkestrel					Stretch
86	94	**Implement new container creation U**	Closed	Yes	b5bde290af7	for #94, for #95: mor	webextension/cs	webextension/im	webextension/js/	webextension/popup.html		01/19	02/10	22	ui	updates		popup		johngruen	groovecoder				MVP
87	295	**container menu's not initially refres**	Closed	Yes	be189ac7ec	Fix the configuring o	index.js					03/01	03/01	0	fix	configuring	issue	bug		kjozwiak	bakulf				MVP
88	290	**All hidden tabs are displayed when**	Closed	No								02/28	02/28	0						SoftVision-CiprianMuresan	groovecoder				
89	404	**Cookies in containers are not isolat**	Closed	No								03/26	04/05	10				bug	question	SCmkQrK1NhwcWoPJF3F	jonathanKingston				
90	353	**Container colors missing from tree**	Closed	No								03/09	03/13	4				bug	addon incompatibly	escapewindow	escapewindow				Stretch
91	440	**If you add containers, then remove**	Open	No								04/10		#NUM!				bug	needs:Testing	Kryze					
92	197	**Customise menu breaks new tab m**	Open	No								02/17		#NUM!				bug		jonathanKingston					Stretch
93	322	**Automatically open a hostname in t**	Closed	No								03/03	03/31	28				enhancement		option-greek	kjozwiak		306, 386		
94	70	**Add Travis or Circle CI for running t**	Closed	No								01/11	02/10	30						pdehaan	pdehaan	groovecoder		83	Stretch
95	449	**Facebook logged in status not kept**	Closed	No								04/13	04/13	0				bug	duplicate	ruihildt	jonathanKingston				
96	438	**"Containers confirm navigation" pa**	Closed	Yes	53c1b07d57	Wrap assignment UI	webextension/css/confirm-page.css					04/08	04/10	2	wrap	prevent	fixes			dholbert	jonathanKingston	groovecoder			
97	410	**Custom container with shopping ca**	Closed	Yes	7d9b4c1561	Fixing shopping icon	data/usercontext.css					04/03	04/03	0				bug		jonathanKingston	jonathanKingston	groovecoder			
98	105	**Plus Panel (New Tab) UI Broken**	Closed	Yes	1e6b684915	Fixing usability of ne	data/usercontext.	index.js				01/24	01/27	3	fixing	usability	popup	bug	tab bar	kjozwiak	jonathanKingston				
99	225	**Linting Errors on Master**	Closed	No								02/21	02/21	0						johngruen	johngruen			226	
100	314	**Container security "personas"**	Open	No								03/02						enhancement		brotherdust					311 Stretch

Table A1. *Data collection from Project "Containers"*

Issues			Connection	Commits							Dates			Keywords			Tags		Developers			References - issues ID		
№	Issue ID	Issue Name	Status	Direct link	Commit ID	Commit Name	Changed file 1	Changed file 2	Change file 3	Change file 4	Change file 5	Date submitted	Date resolved	Days taken	Keyword 1	Keyword 2	Keyword 3	Tag 1	Tag 2	Issue submitted	Developer 1	Developer 2	Closed referenc	Opened reference	Milestone name
1	7	**Add an option t**	Open	No								03/25/2016								bwinton			173,188, 282, 28	782	Post-launch
2	655	**"Open in new ta**	Open	No								11/04/2016								fotz				667, 674, 815	
3	405	**Support for cor**	Open	Yes	8fd58ffe8551272	Adding container	skin/base.css					07/12/2016			adding	container	change			martinthomson	abaez	ChrisW-B	488, 944, 354		
4	583	**Right click men**	Closed	Yes	cad05ca6a2b971	Fix: Restore Place	verticaltabs.js					08/26/2016	02/22	180	fix	restore				Cowpatiz	ericawright		490, 706,704, 710, 715, 723, 760, 800, 965, 978		
5	336	**Tabs not expan**	Closed	No								07/05/2016	07/26/2016	21						alexanderkjeldaa	ericawright		342		
6	439	**Option to turn o**	Closed	Yes	ebd4ff8930c1d43	feat: Option to di	main.js					07/22/2016	08/06/2016	15	feat	option	disable			tobia	Popog	bwinton	486, 430, 444	434	
7	1024	**New link not dis**	Closed	No								03/20	03/24	4						sap7772000zilla	sap7772000zilla				
8	716	**Tabs in inverse**	Closed	Yes	f4e22cf701ecd9(	feature: Reverse	package.json	skin/base.css	vertical-tabbrowse	verticaltabs.js		10/20/2016	11/02/2016	13	feature	display	order			MisterY	ericawright		709, 771	815	
9	75	**hyperlink hover**	Closed	No								04/27/2016	04/29/2016	2						edwinwongsf	bwinton		74		
10	922	**Pinned tabs co**	Open	No								02/10		#NUM!						fvsch					
11	427	**All open tabs ar**	Closed	Yes	43b44fd0e048d4	fix: Use a better v	verticaltabs.jsm					07/21/2016	07/22/2016	1	fix	use	loading			mewman	bwinton	ericawright			
12	866	**Pin button has t**	Closed	No								12/29/2016	12/30/2016	1						Keith94	ericawright				
13	230	**Incognito mode**	Open	No								05/16/2016		#NUM!				bug	ready	meandavejustice			126		
14	984	**Highlight is gon**	Closed	No								03/03	03/03	0						djmitche	ericawright				
15	90	**No visual indica**	Closed	Yes	a1610aeb667ca	Pinned tabs disp	skin/light/light.css					04/30/2016	05/02/2016	2	tabs	display	support			Verdi	bwinton				Testing Pilot Laun
16	164	**Firebug do not t**	Open	No								05/11/2016		#NUM!				Add-on compat	ready	beholdr					
17	156	**Full-screen moc**	Closed	No								05/11/2016	05/16/2016	5				bug		fx4waldi	wresuolc		779		
18	678	**Add a small del**	Closed	No								10/05/2016	01/04	91						dannyhaij	ericawright				
19	188	**An option to hid**	Closed	No								05/12/2016	05/12/2016	0						Nidre	bwinton		87		
20	636	**"Block pop-up v**	Open	No								09/27/2016		#NUM!						merlin67					
21	1043	**Search bug - Lo**	Closed	Yes	9bd4b290afe0b7	do not force new	verticaltabs.js					04/03	04/06	3	not	force	tabs			TDA	ericawright				
22	115	**Tab center redr**	Closed	No								05/05/2016	05/17/2016	12				Add-on compat		rpappalax	bwinton		193, 234		
23	490	**Places context**	Closed	No								08/05/2016	09/14/2016	40						Keith94	Keith94		552, 555, 570, 583		
24	116	**Open new tabs**	Closed	No								05/05/2016	05/12/2016	7						marco-c	johngruen				
25	3	**New tabs shoul**	Closed	Yes	cbe87544c5555(	Open new tabs a	verticaltabs.jsm					03/15/2016	03/30/2016	15	open	tabs	list	p2		philsa	bwinton				
26	918	**New tab appear**	Closed	Yes	6c85a130409615	Fix #918 janky fa	skin/base.css						02/07	20	fix	animation	tabs			fvsch	ericawright	fvsch	944		
27	693	**TabCenter Jum**	Closed	Yes	3742c1c02d1f20	fix: scroll jump to	verticaltabs.js					10/12/2016	11/15/2016	34	fix	tabs	list			ksec	ericawright	bwinton	738		
28	309	**Tabs opened by**	Closed	No								05/30/2016	06/01/2016	2						djmylt	wresuolc		322		
29	594	**Move te close b**	Open	No								08/28/2016		#NUM!						ajagmx					
30	61	**Browser windo**	Closed	Yes	9bd4b7957cab4(	Move the one-tir	bootstrap.js	verticaltabs.jsm				04/19/2016	04/26/2016	7	move	initialization	function	Arcane shit		pdehaan	bwinton				
31	686	**search bar disa**	Open	No								10/07/2016		#NUM!						cinna01					
32	959	**Lock resize?**	Open	No								02/24		#NUM!						dubcanada					
33	1038	**Throbber/spinn**	Closed	Yes	caf419aff625d20(	Fix: show spinne	skin/base.css	vertical-tabbrowse	verticaltabs.js			03/30	04/07	8	fix	restore	versions			hcs64	ericawright				
34	62	**Cannot use Boo**	Closed	Yes	5c7040beceba72	Sidebards!	skin/light/light.css					04/19/2016	05/14/2016	25	sidebars			Feature request	p1	pdehaan	bwinton		36, 108, 127, 781, 187, 841, 844		
35	475	**Rearranging tat**	Closed	Yes	fc815b9f405344(	fix: Disappearing	vertical-tabbrowse	verticaltabs.js				07/30/2016	08/12/2016	13	fix	disappearing				Keith94	bwinton		503, 511		
36	665	**Shrink sidebar**	Open	No								10/04/2016		#NUM!						NaRuBeTo					
37	827	**Since 50 seeing**	Closed	No								12/01/2016	03/16	105						supnah	ericawright				
38	142	**Back button is t**	Open	No								05/10/2016		#NUM!				Feature request	ready	potch				89, 639, 56	
39	410	**Option to open**	Closed	Yes	54260b5d700e7(	Feature:link oper	verticaltabs.js					07/14/2016	09/07/2016	55	feature	tab				kkot	bwinton		567, 461, 536, 491		Shield ready
40	430	**Allow thumbna**	Closed	No								07/21/2016	07/22/2016	1						mewman	bwinton				
41	774	**pontoon can't c**	Closed	No								11/09/2016	11/09/2016	0						wresuolc	bwinton				
42	171	**Active tab chan**	Open	No								05/11/2016		#NUM!				bug	ready	abdulmoeedammar			47		
43	552	**Bookmarks too**	Closed	No								08/21/2016	08/22/2016	1						El112358	ericawright				
44	112	**Feedback:Pinne**	Open	No								05/05/2016		#NUM!				Feature request	ready	chutten			435, 608	434	
45	377	**UX: thinner tabs**	Open	No								06/22/2016	08/22/2016	61						armenzg	bwinton				
46	782	**#titlebar does n**	Open	No								11/15/2016		#NUM!						Shickmo					
47	437	**Search box doe**	Open	No								07/22/2016		#NUM!						TDA					
48	265	**LastPass scree**	Closed	No								05/21/2016	06/08/2016	18						AutomatedTeste	AutomatedTester				
49	502	**Private browsin**	Closed	Yes	4ed42ca0984b8(	skin/base.css						08/08/2016	08/09/2016	1	fix	hide	icon			Keith94	bwinton				
50	629	**[omitted]**	Closed	No								09/21/2016	09/21/2016	0						Red-Squid	bwinton				
51	365	**Downloads butt**	Closed	No								06/17/2016	06/17/2016	0						Zaggy1024	bwinton				
52	1029	**Add options to**	Closed	No								03/24	03/24	0						YPetremann	bwinton				
53	273	**With large num**	Closed	Yes	7fab8fa344a51b(	Fix: rough scroll	verticaltabs.js					05/24/2016	08/25/2016	93	fix	expanded		p2		mystor	bwinton	ericawright	393, 264	545	
54	975	**Change "move**	Closed	Yes	1ef74129b7fbaa	fix: Search stays	main.js	skin/base.css	verticaltabs.js			03/01	03/20	19	fix	search	visible			ShalokShalom	bwinton	ericawright			
55	902	**Tab center**	Closed	No								01/23	02/18	26						Aiswarya-ksr	ericawright				
56	724	**Sidebar stays o**	Closed	No								10/24/2016	10/25/2016	1						macxcool	ericawright				
57	875	**Closing Tabs M**	Open	No								01/04		#NUM!						hot-soup					
58	639	**TabCenter shou**	Open	No								09/29/2016		#NUM!						myBestSoftAndPref			678, 867, 949		
59	675	**Race condition**	Closed	No								10/05/2016	10/24/2016	19						callahad	ericawright		651, 623, 1607, 718		
60	1064	**TabCenter is ov**	Closed	No								04/22	04/26	4				p1		Defman21	Defman21				
61	985	**Tabs can't be cl**	Open	No								03/03		#NUM!						AgentSteel					
62	832	**Provide an opti**	Closed	No								12/05/2016	03/05	90						andrea-capri	bwinton				
63	146	**Tab Center pus**	Closed	No								05/10/2016	05/11/2016	1						pdehaan	bwinton		789, 760		
64	464	**Add Compatibil**	Open	No								07/26/2016		#NUM!						ericawright			458, 454, 453, 433, 424, 469, 390, 370, 517, 499, 624, 6		
65	669	**Tab list scroll p**	Closed	Yes	f46f7e0e8da97a(	fix: On chnage of	skin/base.css	verticaltabs.js				10/05/2016	10/12/2016	7	fix	change	position			Maeyanie	bwinton		694, 693, 730		
66	687	**Sidebar become**	Closed	No								10/09/2016	10/11/2016	2						glandium	ericawright				
67	434	**Option to have**	Open	No								07/22/2016		#NUM!				in progress		nt1m			435, 512, 530		
68	379	**High CPU usage**	Closed	No								06/22/2016	11/08/2016	139						lissyx	lissyx				
69	78	**First tab is alwa**	Closed	No								04/27/2016	05/18/2016	21						wresuolc	wresuolc				
70	169	**Get rid of scroll**	Closed	Yes	09ea2c5c11ed1f(	Style: Windows e	skin/light/light.css					05/11/2016	06/20/2016	40	style	edits		bug	p3	johngruen	bwinton		216, 274, 272, 288, 359		
71	73	**When uninstalli**	Closed	No								04/27/2016	04/27/2016	0						sevaan	bwinton				
72	553	**Control-Shift-K**	Closed	Yes	e471dcaeb80c9(	Don't conflict with	main.js					08/21/2016	08/23/2016	2	don't	conflict	keys			callahad	ericawright		628, 766		
73	691	**Add a toggle to**	Closed	Yes	4af573d5c305b6	feature: Add altog	main.js	skin/base.css	skin/osx.css	verticaltabs.js	util.js	10/10/2016	10/21/2016	11	feature	add	tabs			philsa	bwinton				
74	95	**feature request**	Closed	Yes	a4894c3cf8af8f1	fix: Add a hotkey	main.js	skin.base.css	verticaltabs.js			05/03	08/19/2016	108	fix	add	toggle			lundjordan	ericawright		579		Post-launch
75	438	**Malfunction wh**	Closed	No								07/22/2016	07/26/2016	4						abFx	bwinton				
76	988	**Scrolling tabs u**	Closed	Yes	8299a83a9b021	fix: Enable scroll	vertical-tabbrowser.xml					03/05	03/07	2	fix	enable	scrolling			kkot	bwinton				
77	462	**Opening new ta**	Closed	Yes	54260b5d700e7(	feature:links oper	verticaltabs.js					07/25/2016	09/07/2016	44	feature	link	open			powerfool	bwinton				
78	398	**Tabbar overlaps**	Closed	No								07/05/2016	07/05/2016	0						Celti	bwinton				
79	761	**Option for hidin**	Closed	No								11/03/2016	11/03/2016	0						ostrolucky	bwinton				
80	936	**Tab Center isn't**	Closed	No								02/17	03/20	31						SoftVision-Ciprian	SoftVision-CiprianMuresan				
81	587	**Tabs temporarili**	Open	No								08/26/2016		#NUM!						pikestaff					
82	804	**"Feature Reque**	Open	No								11/21/2016		#NUM!						Vouskopes					
83	234	**Let's add an ind**	Closed	Yes	4b0e17c4e2183(	docs: add incom	README.md					05/16/2016	06/29/2016	44	docs	add	incompatibility	p1		wresuolc	ericawright		222, 220, 213, 201, 166, 115, 137, 259, 292, 325, 339		
84	360	**Make Tab Cente**	Closed	Yes	3101f4f41d6c9c5	feature: first cut a	verticaltabs.js	package.json	main.js	locales/en-US/ac	locales/README.r	06/13/2016	11/08/2016	148	feature					brwolfgang	ericawright				
85	424	**Breaking bug w**	Closed	No								07/21/2016	07/26/2016	5						KingZee	ericawright		443, 375		
86	572	**Pasting a string**	Closed	Yes	df7bd290e18b5ff	fix: Allow pasting	verticaltabs.js					08/24/2016	09/08/2016	15	fix	allow				SoftVision-Paul(	bwinton	ericawright			
87	118	**Make sidebar ar**	Closed	No								05/05/2016	05/18/2016	13				Polish		marco-c	johngruen				
88	1015	**[Remember tab**	Closed	Yes	2264b16fb7636d	fix: Tabs are pern	verticaltabs.js					03/15	03/20	5	fix	tabs	filtered			SoftVision-Paul(	bwinton	ericawright			
89	45	**Can't scroll all t**	Closed	No								04/14/2016	04/28/2016	14						pdehaan	bwinton				
90	313	**new tab order ir**	Closed	No								05/31/2016	10/25/2016	147				p3	ready	meandavejustice	ericawright		931		
91	292	**Problem with P**	Closed	No								05/26/2016	05/26/2016	0						blue101	bwinton				
92	717	**Remote the title**	Open	No								10/20/2016		#NUM!						bwinton			994		
93	258	**Allow creating t**	Open	No								05/19/2016		#NUM!				Feature request	ready	Zaggy1024	ericawright		378, 967		
94	784	**Resizing the tat**	Open	No								11/15/2016		#NUM!						Quicksaver			197		
95	808	**Graphical glitch**	Open	No								11/22/2016		#NUM!						johancz					
96	168	**Opening FF's C**	Closed	Yes	072846ed3fb52f4	Fix/customize	bootstrap.js	verticaltabs.jsm				05/11/2016	05/19/2016	8	fix	customize		bug	p2	riyad	bwinton	ericawright	778, 154, 195, 219, 228, 236		
97	370	**On start lack of**	Closed	No								06/18/2016	08/18/2016	61						pasi1	ericawright				
98	333	**new tab button**	Closed	No								06/03/2016	08/22/2016	80						johngruen	bwinton				
99	119	**Add Test Pilot i**	Closed	Yes	5c109c3dc751fa	install.rds	tab-center.update.rdf					05/06/2016	05/06/2016	0						javaun	bwinton				
100	1056	**Possibly only d**	Open	No								04/12		#NUM!						sidpagariya					

Table A2. *Data collection from Project "Tab Center"*

Issues			Connection		Commits							Dates			Keywords			Tags		Developers			References - Issues ID			Note
Nr	Issue ID	Issue Name	Status	Direct link	Commit ID	Commit Name	Changed file 1	Changed file 2	Change file 3	Change file 4	Change file 5	Date submitted	Date resolved	Days taken	Keyword 1	Keyword 2	Keyword 3	Tag 1	Tag 2	Issue submitted	Developer 1	Developer 2	Closed referenc	Opened referen	Milestone name	
1	140	resolv.rb:164:in	Closed	Yes	02b98d205a168	See if we can fix	lib/string_ext.rb					09/01/2016	09/02/2016	1	fix	timeout	issues	bug		christophprokop	claudijd		147			Pull request
2	27	Add support for	Closed	Yes	3dfe82b11bcb6d	Reading IPs via f	bin/ssh_scan					02/23/2016	08/15/2016	174	reading	ip		enhancement	help wanted	claudijd	claudijd	jinankjain	78, 65			Pull request
3	50	Create a pretty	Open	No								08/04/2016		#NUM!				enhancement	help wanted	claudijd			198, 230			
4	93	negotiate': coul	Closed	No								08/17/2016	08/17/2016	0				bug		dguido	dguido					Fixed in update
5	77	NoMethodError	Closed	No								08/15/2016	08/16/2016	1				enhancement	question	claudijd	claudijd					Question. Fixed i
6	236	Add something	Open	No								10/06/2016		#NUM!						claudijd			242, 246, 247, 249			Too much work
7	28	Add support for	Closed	Yes	bc537529c42a80	Fixes #28 Add su	lib/ssh_scan/scan_engine.rb					02/23/2016	08/09/2016	168	fix	add	support	enhancement	help wanted	claudijd	claudijd					
8	330	current master	Closed	No								02/28	03/24	24						gerhard-tinned	claudijd					Newer version fix
9	118	Add Gitter for a	Closed	Yes	22ad92543fa75c	Added gitter chat	README.md					08/23/2016	09/07/2016	15	add			help wanted	question	jinankjain	jinankjain		174			Pull request
10	238	Add a heartbeat	Closed	Yes	55c3c4c5246e4e	lbheartbeat_rout	.gitignore	lib/ssh_scan/ari.r	spec/ssh_scan/api_spec.rb			10/19/2016	10/25/2016	6	api			help wanted		claudijd	Rajat-Goyal					
11	201	Make ssh_scan	Open	No								09/20/2016		#NUM!				help wanted	feature	claudijd			237, 255			Pull request not a
12	281	False "complia	Closed	Yes	f130fd2aa05b101	fix nit in policy_m	lib/ssh_scan/policy_manager.rb					12/05/2016	02/07	64	fix	nit				gerhard-tinned	gerhard-tinned		330			Pull request
13	47	Add support for	Open	No								06/20/2016		#NUM!				enhancement	help wanted	claudijd						
14	135	Work around SS	Closed	Yes	c0b3fd6fc3b9a18	Use null host key	lib/ssh_scan/scan_engine.rb					08/30/2016	09/19/2016	20	use	null	verifier	bug	help wanted	claudijd	claudijd		136, 189			Pull request
15	228	Create high-lev	Open	No								09/28/2016		#NUM!				documentation	feature	claudijd						
16	43	Install on CentO	Closed	No								03/02/2016	03/03/2016	1						gerhard-tinned	claudijd					Old version of Ru
17	51	Add moar OS g	Closed	Yes	02dfd2d738d852	add ubuntu bann	lib/ssh_scan/ban	lib/ssh_scan/os/u	spec/ssh_scan/banner_spec.rb			08/05/2016	09/02/2016	28	add			enhancement	help wanted	claudijd	claudijd					Pull request
18	115	Set a default m	Closed	No								08/23/2016	08/23/2016	0				bug	help wanted	claudijd	claudijd		121			Fixed in #121
19	24	Refactor policy	Closed	No								02/19/2016	02/23/2016	4						claudijd	claudijd		31			Question
20	130	Potential port b	Closed	Yes	5f90a827ba8a4	fixes #130	bin/ssh_scan					08/29/2016	09/03/2016	5	fix			bug	test	claudijd	claudijd	agaurav77				
21	203	Add full support	Open	No								09/20/2016		#NUM!				fingerprint	help wanted	claudijd			204, 229			Pull request need
22	190	Make a gh-page	Closed	Yes	b2dd7e1fad59f7	fix css/js on gh-p	_layouts/default.html					09/19/2016	09/27/2016	8				documentation		claudijd	agaurav77		219, 220, 221, 222			Pull requests
23	86	Add the ability t	Open	No								08/16/2016		#NUM!				enhancement	help wanted	claudijd						
24	357	Documentation	Open	No								04/17		#NUM!						JoeArcher007						Re-opened
25	262	Need a mechani	Closed	Yes	ec3a2355e482ef	Fix/ignore a bunc	lib/ssh_scan/api.rb					11/17/2016	12/02/2016	15	fix	ignore				claudijd	claudijd					Pull request
26	233	Follow up API ir	Closed	Yes	b65e59aa089e7	add https support						10/04/2016	12/02/2016	59	add	https		enhancement	feature	claudijd	Rajat-Goyal		260			Pull request
27	170	IPv6 fallback bu	Closed	Yes	587133c4c1581	fixes #170	lib/ssh_scan/client.rb					09/06/2016	09/06/2016	0	fix			bug	help wanted	claudijd	jinankjain					
28	224	Mixed content b	Closed	No								09/26/2016	09/27/2016	1				bug	help wanted	claudijd	claudijd					Fixed in newer v
29	200	Write specs for	Closed	Yes	b8c1e33a075ade	specs for string_	spec/ssh_scan/string_ext_spec.rb					09/20/2016	10/19/2016	29	specs			help wanted	test	claudijd	flash1452		232			Pull request
30	133	Update gem ver	Closed	No								08/30/2016	08/30/2016	0						jinankjain	claudijd					
31	278	Port option doe	Closed	Yes	99b423c127cb80	reorder argument	bin/ssh_scan					12/05/2016	12/12/2016	7	reorder	arguments				gerhard-tinned	agaurav77	claudijd				Pull Request
32	207	Add a telemetry	Open	No								09/20/2016		#NUM!				feature	test	claudijd						
33	31	Tool recommen	Closed	No								03/01/2016	08/03/2016	155				question		claudijd	claudijd					Workaround
34	66	nit:scan_engine	Closed	Yes	4f15744075d6f8	Fix bug in host k	lib/ssh_scan/scan_engine.rb					08/11/2016	08/11/2016	0	fix	bug	class			claudijd	claudijd					
35	354	ssh_scan fails t	Closed	No								03/25	03/26	1						ihirlmann	claudijd					Newer version
36	9	Convert policie	Closed	Yes	ae6ea30d76d8c1	Move to JSON p	bin/ssh_scan	lib/ssh_scan/policy.rb				02/18/2016	02/19/2016	1		move	json			claudijd	claudijd		24, 31			Pull request
37	136	Add output whe	Closed	No								08/30/2016	10/27/2016	58				feature	help wanted	claudijd	claudijd		197			Fixed in another i
38	29	Add IPv6 suppo	Closed	No								02/24/2016	08/23/2016	181				enhancement	help wanted	claudijd	claudijd					
39	353	ssh_scan for ac	Open	No								03/24		#NUM!						peterns						
40	202	Lessons learne	Closed	No								09/20/2016	09/28/2016	8				bug	lessons learned	claudijd	claudijd			203		
41	279	The verbosity -	Closed	Yes	a0f3c3f53f65334	Fix #279 - verbos	bin/ssh_scan	spec/ssh_scan/intergration.sh				12/05/2016	12/05/2016	0	fix	bug				gerhard-tinned	claudijd					
42	168	Break fingerprin	Closed	Yes	1fde190870a7d4	break fingerprnt	spec/ssh_scan/b	spec/ssh_scan/b	spec_ssh_scan/b	spec/ssh_scan/banner/mixed_spe		09/06/2016	09/07/2016	1	break			fingeprint	test	claudijd	agaurav77					Pull request
43	211	make gem versi	Closed	Yes	969ac211a100dc	make update che	/bin/ssh_scan	lib/ssh_scan/update.rb				09/20/2016	09/25/2016	5	make	update	feature	bug		agaurav77	agaurav77					Pull request
44	223	Blog Post Ideas	Open	No								09/26/2016		#NUM!				help wanted	lessons learned	claudijd						
45	303	Add GitHub bot	Closed	No								12/22/2016	01/08	17						rishabhs95	rishabhs95					No plan to work o
46	191	ssh_scan/lib/ssl	Closed	Yes	be20b5a5f90c5d	Fix exception har	lib/ssh_scan/client.rb					09/19/16	09/19/2016	0	fix	exception		bug	help wanted	claudijd	claudijd					
47	38	Write up a Jenki	Closed	Yes	1874338b5e5042	Merge pull reque	bin/ssh_scan	lib/ssh_scan/version.rb				03/02/2016	03/04/2016	2						claudijd	claudijd					Pull request
48	122	Lessons learne	Closed	No								08/23/2016	09/20/2016	28				bug	help wanted	claudijd	claudijd		202			
49	209	automate updat	Open	No								09/20/2016		#NUM!				automation	fingerprint	agaurav77						Pull request not a
50	127	Add unit-test co	Closed	Yes	ed1eab1e2b51a	Fixed #127	spec/ssh_scan/policy_manager_spec.rb					08/29/2016	09/01/2016	3	fixed			help wanted	test	claudijd	flash1452	jinankjain				
51	32	Error when poli	Closed	Yes	6b136dedfdcdec	Fix compression	lib/ssh_scan/policy_manager.rb					03/02/2016	03/02/2016	0	fixed	compression				gerhard-tinned	claudijd					Pull request
52	74	Add support for	Closed	Yes	08f2f5ecc2798fd	Fixes #74	lib/ssh_scan/clie	lib/ssh_scan/sca	lib/string_ext.rb			08/15/2016	08/30/2016	15	fixed			enhancement	help wanted	claudijd	jinankjain					
53	351	There seems to	Closed	Yes	8a17985beebf0a	Fix duplicate fing	lib/ssh_scan/scan_engine.rb					03/24	03/24	0	fix	duplicate	bug	bug		claudijd	claudijd					Pull request
54	358	License	Closed	Yes	7cf50eba1c2636	Add License	LICENSE					04/17/2017	04/20	3	add					ImNtReal	claudijd					Pull request
55	288	Add a more det	Closed	No								12/09/2016	12/12/2016	3				help wanted		claudijd	claudijd					Different approac
56	277	Missing compli	Closed	Yes	f1dbb10401a60b	Fix compliance p	lib/ssh_scan/scan_engine.rb					12/03/2016	12/12/2016	9	fix	bug		bug	help wanted	claudijd	claudijd					Pull request
57	296	Provide a cachi	Open	No								12/15/2016		#NUM!				help wanted		claudijd						Pull request not a
58	256	Investigate a ru	Closed	Yes	d9fb3e3f366fe44	Add rubocop as	.rubocop.yml	.rubocop_todo.ym	.travis.yml	ssh_scan.gemspec		10/27/2016	12/02/2016	36	add					claudijd	rishabhs95					
59	104	Define default c	Closed	Yes	58513a46cd000f	Make SSHScan:	lib/ssh_scan/clie	lib/ssh_scan/con	lib/ssh_scan/protocol.rb			08/17/2016	08/29/2016	12	make	scratch		enhancement	help wanted	claudijd	claudijd		74			Pull request
60	154	Add SSH versio	Closed	Yes	602d33fd833138	Added SSH versi	lib/ssh_scan/polic	lib/ssh_scan/polic	spec/ssh_scan/policy_manager_spec.rb			09/02/2016	09/10/2016	8	added	control	fixes	fingerprint	help wanted	claudijd	yashmehrotra					
61	84	Abstract host k	Closed	Yes	3f24b4f22a7acd	take out host key	lib/ssh_scan/cryp	scan_engine.rb	crypto_spec.rb			08/15/2016	09/16/2016	32	take	fingerprinting	host	enhancement	help wanted	claudijd	agaurav77		77			Pull request
62	193	SQLite prototyp	Closed	No								09/19/2016	09/20/2016	1						claudijd	claudijd		197			Fixed in #197
63	49	Create a queue	Open	No								08/03/2016		#NUM!				enhacement	help wated	claudijd			258, 259	228	0.1.0	
64	10	Determine serv	Closed	No								02/18/2016	02/19/2016	1						claudijd	claudijd					Question
65	280	No compliance	Closed	Yes	f1dbb10401a60b	Fix compliance p	lib/ssh_scan/scan_engine.rb					12/05/2016	12/05/2016	0	fix	bug				gerhard-tinned	claudijd		277			Pull request
66	216	Extend ssh fing	Closed	Yes	66faaf2c3f813c0	Extend ssh finger	.gitignore					09/21/2016	10/25/2016	34	extend	fingerprints		enhancement	fingerprint	claudijd	royalharsh					Pull request
67	328	API authenticati	Closed	Yes	e249ef56f0ea2c5	probable repair f	config/api/config.	lib/ssh_scan/auth	lib/ssh_scan/worker.rb			02/09	03/02	21	repair			bug		agaurav77	pwnbus	claudijd	321			Pull request
68	304	Correct ruboco	Open	No								12/22/2016		#NUM!				code-styling	question	rishabhs95						
69	287	Provide support	Closed	Yes	4c70e248b4148	Add mongodb as	lib/ssh_scan/api.rb					12/09/2016	03/02	83	add			help wanted		claudijd	royalharsh					
70	64	Add a timeout c	Closed	Yes	3588c2ee1cc5f5	add timeout optio	bin/ssh_scan	lib/ssh_scan/clie	lib/ssh_scan/error.rb			08/10/2016	09/02/2016	23	add	options		enhancement	help wanted	claudijd	agaurav77					
71	320	Make JobQueue	Open	Yes	d1ab16379e5611	API caching feat	lib/ssh_scan/api.r	lib/ssh_scan/data	config/api/config.yml			01/12		#NUM!	api	caching	feature			claudijd	rishabhs95		325			Pull request not a
72	78	Add multi-threa	Closed	No								08/15/2016	08/17/2016	2				enhancement	help wanted	claudijd	claudijd					
73	116	Multi-ip scans d	Closed	Yes	c54dd9f2079408	Fixed #116	lib/ssh_scan/targ	spec/ssh_scan/target_parser_spec.rb				08/23/2016	08/23/2016	0	fixed			bug	help wanted	claudijd	claudijd					
74	274	Add valid SSL/T	Closed	No								12/03/2016	03/02	89				enhancement		claudijd	claudijd					
75	289	Need to update	Closed	No								12/09/2016	12/12/2016	3						claudijd	royalharsh					
76	273	Add urls to con	Closed	No								12/03/2016	03/02	89						claudijd	claudijd					Duplicate
77	308	How to deal wit	Closed	No								12/29/2016	03/02	63						claudijd	claudijd		310, 330			
78	156	Build a generic	Closed	Yes	cee17987c8d083	Simplify fingerpri	lib/ssh_scan/ban	spec/ssh_scan/banner_spec.rb				09/02/2016	09/02/2016	0						claudijd	claudijd					Pull request
79	71	--port option br	Closed	Yes	22284680e1da3	Fix net-ssh custo	lib/ssh_scan/scan_engine.rb					08/15/2016	08/15/2016	0	fix	bug				h0bi	claudijd					Pull request
80	340	Propose ssh_sc	Open	No								03/19		#NUM!						paulehoffman						
81	100	Extend port sw	Closed	No								08/17/2016	09/03/2016	17				enhancement	help wanted	claudijd	jinankjain	claudijd	111, 113			
82	8	Make sure the p	Closed	No								02/16/2016	02/18/2016	2						claudijd	claudijd					
83	183	There is a bug i	Closed	Yes	1c3456619ee610	Fix condition che	lib/ssh_scan/policy_manager.rb					09/10/2016	09/2016	0	fix	check	control	bug	help wanted	claudijd	yashmehrotra					
84	150	Add ssh_lib fing	Closed	Yes	7883449bd84de	add dropbear ss	lib/ssh_scan/ban	lib/ssh_scan/ssh_lib.rb				09/02/2016	09/07/2016	5	add	fingerprints		fingerprint	help wanted	claudijd	agaurav77	janankjain				
85	276	ssh_scan_api n	Closed	No								12/03/2016	03/02	89				enhancement		claudijd	claudijd					
86	362	Come up with a	Open	No								05/11		#NUM!						claudijd						
87	139	Add auth_meth	Closed	Yes	f33efdf2c899a39	Added auth_met	lib/ssh_scan/polic	lib/ssh_scan/polic	spec/ssh_scan/policy_manager_spec.rb			08/31/2016	09/01/2016	1	added	fix	control	enhancement	help wanted	claudijd	claudijd					
88	343	Include a list of	Closed	Yes	1c355fe5555b6d	Add ssh-badkey	examples/github.com.json					03/21	03/21	0	added	detection				claudijd	claudijd					Pull request
89	306	provide a gener	Open	No								12/22/2016		#NUM!						claudijd				307		
90	151	Add Cisco OS fi	Closed	Yes	2c3bbafcc072b1	Add Cisco finger	lib/ssh_scan/ban	lib/ssh_scan/os/	lib/ssh_scan/os/cisco.rb			09/02/2016	09/02/2016	0	add	fingerprint		fingerprint	help wanted	claudijd	claudijd					Pull request
91	205	Make sure Ubur	Closed	Yes	3f15c6e86ded85	report unknown	lib/ssh_scan/os/u	spec/ssh_scan/b	spec/ssh_scan/banner/os/ubuntu_spec.rb			09/20/2016	09/23/2016	3	report	version	identified	bug	help wanted	claudijd	agaurav77					Pull request
92	54	Add SSH finger	Closed	Yes	744c49079a43e	Add SSH fingerp	.travis.yml					08/06/2016	08/08/2016	2	add	fingerprint	detection	enhancement	help wanted	claudijd	claudijd					Pull request
93	321	Add authenticat	Closed	Yes	a8e48e801ca4e4	Allow localhost w	lib/ssh_scan/api.rb					01/13	03/02	48	allow	bypass				claudijd	claudijd					Pull request
94	290	Clean up and st	Open	No								12/12/2016		#NUM!				help wanted		claudijd						
95	161	Does not exit gr	Closed	Yes	bca0e49b396791	Fix #161	lib/ssh_scan/client.rb					09/02/2016	09/03/2016	1	fix			help wanted	enhancement	jinankjain	jinankjain					
96	309	Expose an API	Open	No								01/05		#NUM!						claudijd			322			
97	152	Add OS/sshlib f	Closed	Yes	45209d9470412	Added fingerprin	lib/ssh_scan/ban	lib/ssh_scan/os/	lib/ssh_scan/os.rb			09/02/2016	09/06/2016	4	added	fingerprint		fingerprint	help wanted	claudijd	jinankjain					
98	18	Gem installatior	Closed	No								02/19/2016	02/19/2016	0						claudijd	claudijd					Newer version fix
99	159	Add a standard	Closed	Yes	b60ad60d3358c	Added the stand	README.md	bin/ssh_scan				09/02/2016	09/06/2016	4	added	logger		feature	help wanted	claudijd	jinankjain					Pull request
100	264	ssh_scan_work	Closed	Yes	a50c47ea4d400	Move worker con	bin/ssh_scan_wo	bin/ssh_scan_wo	lib/ssh_scan/worker.rb			12/01/2016	12/01/2016	0	move	config	file	help wanted		claudijd	claudijd					

Table A3. *Data collection from Project "SSH Scan"*