Classification of Research Citations (CRC)

Bilal Hayat Butt
(bilal.hayat@nu.edu.pk)

Next 15 minutes..

1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
Computer Science Department at FAST-NUCES, Karachi, Pakistan.
2. Citation Analysis

3. Classification of Citations

4. Classification Schemes

5. Dataset Collection

6. Overview of CRC

7. Results

8. Conclusion and Future Work
Citation Analysis

Diagram:

- Root Paper
  - Citing Papers
    - Target Data
1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
Classification of Citations

• Nanba (2000)
  – Classification into 3 categories, using cue words.
  – Discusses a prototype system called PRESRI.
  – **Limitation**: Targeted cue words. Domain specific
Classification of Citations

• Nanba (2000)
  – Classification into 3 categories, using cue words.
  – Discusses a prototype system called PRESRI.
  – Limitation: Targeted cue words. Domain specific

• Tanguy (2009)
  – An automated technique to identify the citation with 80% precision.
  – Uses linguistic cues and natural language processing techniques for analysis.
  – Limitation: APA style.
Classification of Citations (cont..)

• Cohen (2006)
  – Looks for the cited areas which contains the reviews of the drug.
  – **Limitation**: Domain specific
Classification of Citations (cont..)

• Cohen (2006)
  – Looks for the cited areas which contains the reviews of the drug.
  – **Limitation**: Domain specific

• CRC
  – Generalized approach for classification of scientific research papers in different disciplines.
  – We are using generalized sentiment lexica used in sentiKLUE Evert (2014).
  – **Limitation**: Annotation from domain experts
Next

1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
Research Disciplines

• Humanities and Social Science (HSS) Research
  – Negative citations are less than 10%.
  – Multi category classification scheme is needed.
  – Qualitative results in HSS research.

• Scientific Research
  – Provides quantitative results.
  – Critique and rebuttal is viable not rare.
## Classification Schemes

2. Continuing a Research from point where Cited paper finished.
3. Citing paper to use its ideas, definitions, terms in a Research
4. Citing a paper to refer to data also used in Current Research.
5. Citing a paper to refer to data it used and to draw similarities from the Data used.
6. Citing Paper contains Data and Material used throughout different phases
7. Citing Paper to adopt part/full methodology it adopted for a certain task.
8. Citing paper verified/proved a statement or enlightens with its details.
## Classification Schemes

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cited Paper provides historical facts regarding undergoing Research Question.</td>
</tr>
<tr>
<td>2</td>
<td>Continuing a Research from point where Cited paper finished.</td>
</tr>
<tr>
<td>3</td>
<td>Citing paper to use its ideas, definitions, terms in a Research</td>
</tr>
<tr>
<td>4</td>
<td>Citing a paper to refer to data also used in Current Research.</td>
</tr>
<tr>
<td>5</td>
<td>Citing a paper to refer to data it used and to draw similarities from the Data used.</td>
</tr>
<tr>
<td>6</td>
<td>Citing Paper contains Data and Material used throughout different phases</td>
</tr>
<tr>
<td>7</td>
<td>Citing Paper to adopt part/full methodology it adopted for a certain task.</td>
</tr>
<tr>
<td>8</td>
<td>Citing paper verified/proved a statement or enlightens with its details.</td>
</tr>
<tr>
<td>9</td>
<td>Citing Paper evaluated positively.</td>
</tr>
<tr>
<td>10</td>
<td>Citing Paper evaluated negatively.</td>
</tr>
<tr>
<td>11</td>
<td>Ongoing Research giving proof of statement in Cited Paper.</td>
</tr>
<tr>
<td>12</td>
<td>Ongoing Research giving rebuttal of statement in Cited Paper.</td>
</tr>
<tr>
<td>13</td>
<td>Giving a new interpretation to the findings/statements in Cited Paper.</td>
</tr>
</tbody>
</table>

Inter-annotator disagreement for overlapping categories.
Sentiment Types

• In our approach we have grouped these citation types into three generalized sentiment types:

  – **TYPE-I**: Positive Sentiment
  – **TYPE-II**: Negative Sentiment
  – **TYPE-III**: Neutral Sentiment
Next

1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
## Dataset collection

<table>
<thead>
<tr>
<th></th>
<th>Type-I</th>
<th>Type-II</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers Dataset</td>
<td>109</td>
<td>41</td>
<td>150</td>
</tr>
</tbody>
</table>

Istanbul, Turkey. June 29 2015
## Dataset collection

<table>
<thead>
<tr>
<th>Papers Dataset</th>
<th>Type-I</th>
<th>Type-II</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>109</td>
<td>41</td>
<td>150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference Style</th>
<th>IEEE</th>
<th>APA</th>
<th>AMA</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>In-Text Citation</th>
<th>IEEE</th>
<th>APA</th>
<th>AMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1], [1,2]</td>
<td>Name (year)</td>
<td>[1], [1,2] or 1-2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentage of papers</th>
<th>IEEE</th>
<th>APA</th>
<th>AMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>67%</td>
<td>15%</td>
<td>18%</td>
<td></td>
</tr>
</tbody>
</table>
Data Extraction

Citing Papers

Cited Paper

References

Cited Area
Data Extraction Process

Start → Root Paper → Scrapper → Google Scholar → Retrieve

Target Data → Cited Area → Reference (number) of Root Paper → Title of Root Paper

Extraction Process:
- For every Root Paper, search title using Scrapper from Google Scholar.
- Retrieve the reference (number) of the root paper.
- Use the title to find the title of the root paper.
- Extract data from the target area.
Next

1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
Overview of CRC
Overview of CRC

Citing Papers

Grace (2012)
Citing Paper Number: CP1

Kelley (2013)
Citing Paper Number: CP2

Grace (2012)
Citing Paper Number: CP3

Xu (2012)
Citing Paper Number: CP4

Data Extraction

Root Paper

Title: Android Permissions Demystified. Felt (2011)

Abstract
Android provides third-party applications with an extensive API that includes access.
Overview of CRC

<table>
<thead>
<tr>
<th>Citing Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grace (2012)</td>
</tr>
<tr>
<td>Citing Paper Number: CP1</td>
</tr>
<tr>
<td>Kelley (2013)</td>
</tr>
<tr>
<td>Citing Paper Number: CP2</td>
</tr>
<tr>
<td>Grace (2012)</td>
</tr>
<tr>
<td>Citing Paper Number: CP3</td>
</tr>
<tr>
<td>Xu (2012)</td>
</tr>
<tr>
<td>Citing Paper Number: CP4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Citing Area</th>
</tr>
</thead>
</table>
| To address these challenges, we apply and extend Felt et al. [16] to derive a list of API calls that are of interest for our analysis. In particular, we take a similar approach by analyzing the Android documentation, source code, and disassembled bytecode to understand the standard.
| Felt et al. in their Android Permissions Demystified paper attempt to further explain permissions to developers [6]. However, neither of these papers in plore end-users understanding of permissions.
| Another interesting system in this category is Stowaway [41], which aims to identify instances of app over-privilege, where an app requests more permissions than it uses. Such over-privilege itself could be used as an input to RizRanker in the future.
| Stowaway [23] is a tool that detects over-privilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting over-privilege, and static analysis is used to determine which calls an application invokes.

Data Extraction

Root Paper

Title: Android Permissions Demystified. Felt (2011)

Abstract

Android provides third-party applications with an extensive API that includes access...........
Overview of CRC

Citing Papers

Grace (2012)
Citing Paper Number: CP1

Kelley (2013)
Citing Paper Number: CP2

Grace (2012)
Citing Paper Number: CP3

Xu (2012)
Citing Paper Number: CP4

Citing Area

To address these challenges, we apply and extend Felt et al. [15] to derive a list of API calls that are of interest for our analysis. In particular, we take a similar approach by analyzing the Android documentation, source code and disassembled bytecode to monotonistically compute the standard.

Felt et al., in their Android Permissions Demystified paper, attempt to further explain permissions to developers [6]. However, neither of these papers explore end-users understanding of permissions.

Another interesting system in this category is Stowaway [41], which aims to identify instances of app over-privilege, whereas an app requests more permissions than it uses. Such over-privilege itself could be used as an input to RiskRanker in the future.

Stowaway [23] is a tool that detects overprivilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting overprivilege, and static analysis is used to determine whether each API can lead to the demystification of an application invokes.

Tokenized Sentences

CP1 Label: Negative
Tokenize into Sentence
To address these challenges, we apply and extend Felt et al. [15] to derive a list of API calls that are of interest for our analysis. In particular, we take a similar approach by analyzing the Android documentation, source code and disassembled bytecode to monotonistically compute the standard.

Label: Negative
Tokenize into Sentence
Felt et al., in their Android Permissions Demystified paper, attempt to further explain permissions to developers [6]. However, neither of these papers explore end-users understanding of permissions.

Label: Positive
Tokenize into Sentence
Another interesting system in this category is Stowaway [41], which aims to identify instances of app over-privilege, whereas an app requests more permissions than it uses.

Label: Positive
Tokenize into Sentence
Stowaway [23] is a tool that detects overprivilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting overprivilege, and static analysis is used to determine whether each API can lead to the demystification of an application invokes.
Overview of CRC

Citing Papers

- Grace (2012)
  Citing Paper Number: CP1

- Kelley (2013)
  Citing Paper Number: CP2

- Grace (2012)
  Citing Paper Number: CP3

- Xu (2012)
  Citing Paper Number: CP4

Citing Area

To address these challenges, we apply and extend Felt et al. [16] to derive a list of API calls that are of interest for our analysis. In particular, we take a similar approach by analyzing the Android documentation, source code and disassembled bytecode to non-perturbative un infuriating the domain.

Kelley et al. in their Android Permissions Demystified paper, attempt to further explain permissions to developers [5]. However, neither of these papers explore end-users understanding of permissions.

Another interesting system in this category is Stowaway [41], which aims to identify instances of app over-privilege. The paper suggests that app requests more permissions than it uses. Such over-privilege itself could be used as an input to RiskRanker in the future.

Xu et al. [23] is a tool that detects over-privilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting over-privilege, and static analysis is used to determine which calls an application invokes.

Data Extraction

Root Paper

Title: Android Permissions Demystified. Felt (2011)
Abstract
Android provides third-party applications with an extensive API that includes access

Tokenized Sentences

CP1 Label: Negative
Tokenize into Sentence
To address these challenges, we apply and extend Felt et al. [16] to derive a list of API calls that are of interest for our analysis. In particular, we take a similar approach by analyzing the Android documentation, source code and disassembled bytecode to non-perturbative un infuriating the domain.

Label: Negative
Tokenize into Sentence
Kelley et al. in their Android Permissions Demystified paper, attempt to further explain permissions to developers [5]. However, neither of these papers explore end-users understanding of permissions.

Label: Positive
Tokenize into Sentence
Another interesting system in this category is Stowaway [41], which aims to identify instances of app over-privilege, where an app requests more permissions than it uses.

Label: Positive
Tokenize into Sentence
Stowaway [23] is a tool that detects over-privilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting over-privilege, and static analysis is used to determine which calls an application invokes.

Diagram:

Sentence + Label

Feature Extractor

Lexicon

Naïve-Bayes Classifier

+ve

Positive
Cited Papers

-ve

Negative
Cited Papers

Istanbul, Turkey. June 29 2015
CRC - CLBib, ISSI 2015
26
Example Root Paper

Android permissions demystified

AP Felt, E Chin, S Hanna, D Song... - Proceedings of the 18th ..., 2011 - dl.acm.org
Abstract Android provides third-party applications with an extensive API that includes access to phone hardware, settings, and user data. Access to privacy-and security-relevant parts of the API is controlled with an install-time application permission system. We study Android ...

Cited by 593 Related articles All 30 versions Cite Save
Does all 593 citations carry equal weight?
Qualitatively? No
Quantitatively? Yes
Citing Papers

Android permissions demystified

Dissecting android malware: Characterization and evolution
Abstract—The popularity and adoption of smartphones has greatly stimulated the spread of mobile malware, especially on the popular platforms such as Android. In light of their rapid growth, there is a pressing need to develop effective solutions. However, our defense ... Cited by 678 Related articles All 27 versions Cite Save

Abstract In this paper, we present a systematic study for the detection of malicious applications (or apps) on popular Android Markets. To this end, we first propose a permission-based behavioral fingerprinting scheme to detect new samples of known Android ... Cited by 350 Related articles All 13 versions Cite Save More

Android permissions: User attention, comprehension, and behavior
AP Felt, EH Ha, S Eisenman, A Haney, E Chin - Proceedings of the ... - dl.acm.org
Abstract Android's permission system is intended to inform users about the risks of installing applications. When a user installs an application, he or she has the opportunity to review the application’s permission requests and cancel the installation if the permissions are... Cited by 270 Related articles All 17 versions Cite Save

Detecting repackaged smartphone applications in third-party android marketplaces
W Zhou, Y Zhou, X Jiang, Z Wang - Proceedings of the second ACM conference on ... - dl.acm.org
Abstract Recent years have witnessed an incredible popularity and adoption of smartphones and mobile devices, which is accompanied by large amount and wide variety of feature-rich smartphone applications. These smartphone applications (or apps), typically organized in ... Cited by 212 Related articles All 8 versions Cite Save More

Systematic Detection of Capability Leaks in Stock Android Smartphones.
MC Grae, Y Zhou, Z Wang, X Jiang - NDSS, 2012 - ncsu.edu
Abstract Recent years have witnessed a meteoric increase in the adoption of smartphones. To manage information and features on such phones, Android provides a permission-based security model that requires each application to explicitly request permissions before it... Cited by 212 Related articles All 10 versions Cite Save More

Towards Taming Privilege-Escalation Attacks on Android.
S Bujigal & Davi, A Dimitrakos, T Fischer, AR Sadeghi... - NDSS, 2012 - core.ac.uk
Abstract Android's security framework has been an appealing subject of research in the last few years. Android has been shown to be vulnerable to application-level privilege escalation attacks, such as confused deputy attacks, and more recently, attacks by colluding ... Cited by 184 Related articles All 15 versions Cite Save More

Riskranker: scalable and accurate zero-day android malware detection
M Grae, Y Zhou, D Zhang, S Zhou, X Jiang - Proceedings of the 10th... - dl.acm.org
Abstract Smartphone sales have recently experienced explosive growth. Their popularity also encourages malware authors to penetrate various mobile marketplaces with malicious applications (or apps). These malicious apps hide in the sheer number of other normal ... Cited by 175 Related articles All 8 versions Cite Save

Aurasim: Practical Policy Enforcement for Android Applications.
B Xu, H Sandel, E Anderson - USENIX Security Symposium, 2012 - usenix.org
Abstract The increasing popularity of Google's mobile platform Android makes it the prime target of the latest surge in mobile malware. Most research on enhancing the platform's security and privacy controls requires extensive modification to the operating system... Cited by 14 Related articles All 6 versions Cite Save More
Example Citing paper

Aurasium: Practical Policy Enforcement for Android Applications

Rubin Xu
Computer Laboratory
University of Cambridge
Cambridge, UK
Rubin.Xu@cl.cam.ac.uk

Hassen Saidi
Computer Science Laboratory
SRI International
Menlo Park, USA
hassen.saidi@sri.com

Ross Anderson
Computer Laboratory
University of Cambridge
Cambridge, UK
Ross.Anderson@cl.cam.ac.uk

Abstract
The increasing popularity of Google’s mobile platform Android makes it the prime target of the latest surge in mobile malware. Most research on enhancing the platform’s security and privacy controls requires extensive modification to the operating system, which has significant usability issues and hinders efforts for widespread adoption. We develop a novel solution called Aurasium that bypasses the need to modify the Android OS while providing much of the security and privacy that users desire. We automatically repackage arbitrary applications to attach user-level sandboxing and policy enforcement code, which closely watches the application’s behavior for security and privacy violations such as attempts to retrieve a user’s sensitive information, send SMS covertly to premium numbers, or access malicious IP addresses. Aurasium can also detect and prevent cases of privilege MacAfee [29], making it the most assaulted mobile operating system during that period. While much of the initial wave of Android malware consisted of trojans that masquerade as legitimate applications and leak a user’s personal information or send SMS messages to premium numbers, recent malware samples indicate an escalation in the capability and stealth of Android malware. In particular, attempts are made to gain root access on the device through escalation of privilege [37] to establish a stealthy permanent presence on the device or to bypass Android permission checks.

Fighting malware and securing Android-powered devices has focused on three major directions. The first one consists of statically [20] and dynamically [12, 36] analyzing application code to detect malicious activities before the application is loaded onto the user’s device. The second consists of modifying the Android OS in.
Reference


Regular Expression

• After finding the Title of the root paper, our regular expression extracts the data till the preceding full stop of the previous reference. As shown in the diagram.

• Then starting from the full stop, we check what is coming after it i.e. whether is it a bracket ([1],IEEE) or a number followed by a full stop (1.,AMA) or author’s name/year Felt, A. P., (2010) (APA).

• Thereby assigning the reference to be either IEEE,AMA or APA via the regular expression.
**Dynamic Analysis** Despite its limitations, dynamic analysis remains the preferred approach among researchers and antivirus companies to profile malware and extract its distinctive features. The lack of automated ways to explore all the state space is often a hindering factor. Techniques such as multipath exploration [31] can be useful. However, the ability of mobile malware to load arbitrary libraries might limit the effectiveness of such techniques. The honeynet project offers a virtual machine for profiling Android Applications [36] similar to profiling desktop malware. **Stowaway [23]** is a tool that detects overprivilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting overprivilege, and static analysis is used to determine which calls an application invokes.

6 Conclusion and Future Work

We have presented Aurasium, a technology that protects users of the web from malicious and untrusted applications. Some of the security solutions proposed in this paper do not require rooting and device root access.

Aurasium allows us to take full advantage of an application. This allows for the modification of policies at runtime. By using...
Dynamic Analysis  Despite its limitations, dynamic analysis remains the preferred approach among researchers and antivirus companies to profile malware and extract its distinctive features. The lack of automated ways to explore all the state space is often a hindering factor. Techniques such as multipath exploration [31] can be useful. However, the ability of mobile malware to load arbitrary libraries might limit the effectiveness of such techniques. The honeynet project offers a virtual machine for profiling Android Applications [36] similar to profiling desktop malware. Stowaway [23] is a tool that detects overprivilege in compiled Android applications. Testing is used on the Android API in order to build the permission map that is necessary for detecting overprivilege, and static analysis is used to determine which calls an application invokes.

6 Conclusion and Future Work

We have presented Aurasium, a technology that protects users of the technology from malicious and untrusted software. The security solutions proposed do not require rooting and device-specific policies at runtime. By using
Feature Extraction

Cited Area

Tokenization

Lexical Processing

Feature Extractor

[Diagram showing the process of feature extraction with steps involving cited area, tokenization, and lexical processing]
1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
## Results

<table>
<thead>
<tr>
<th></th>
<th>Successfully extracted</th>
<th>Issue with pdf-to-text conversion</th>
<th>Referencing Style not followed</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers Dataset</td>
<td>130</td>
<td>5</td>
<td>15</td>
<td>150</td>
</tr>
</tbody>
</table>

Istanbul, Turkey. June 29 2015

CRC - CLBib, ISSI 2015
## Results

<table>
<thead>
<tr>
<th></th>
<th>Successfully extracted</th>
<th>Issue with pdf-to-text conversion</th>
<th>Referencing Style not followed</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers Dataset</td>
<td>130</td>
<td>5</td>
<td>15</td>
<td>150</td>
</tr>
</tbody>
</table>

Accuracy: 86%
## Results

### Papers Dataset

<table>
<thead>
<tr>
<th></th>
<th>Successfully extracted</th>
<th>Issue with pdf-to-text conversion</th>
<th>Referencing Style not followed</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers Dataset</td>
<td>130</td>
<td>5</td>
<td>15</td>
<td>150</td>
</tr>
</tbody>
</table>

**Accuracy: 86%**

### Category Performance

<table>
<thead>
<tr>
<th>Category</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>0.84</td>
<td>0.94</td>
<td>0.89</td>
<td>109</td>
</tr>
<tr>
<td>Type II</td>
<td>0.25</td>
<td>0.10</td>
<td>0.14</td>
<td>21</td>
</tr>
<tr>
<td>Average/Total</td>
<td>0.75</td>
<td>0.81</td>
<td>0.77</td>
<td>130</td>
</tr>
</tbody>
</table>

Istanbul, Turkey. June 29 2015
## Results

<table>
<thead>
<tr>
<th></th>
<th>Successfully extracted</th>
<th>Issue with pdf-to-text conversion</th>
<th>Referencing Style not followed</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers Dataset</td>
<td>130</td>
<td>5</td>
<td>15</td>
<td>150</td>
</tr>
</tbody>
</table>

### Accuracy: 86%

<table>
<thead>
<tr>
<th>Category</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>0.84</td>
<td>0.94</td>
<td>0.89</td>
<td>109</td>
</tr>
<tr>
<td>Type II</td>
<td>0.25</td>
<td>0.10</td>
<td>0.14</td>
<td>21</td>
</tr>
<tr>
<td>Average/Total</td>
<td>0.75</td>
<td>0.81</td>
<td>0.77</td>
<td>130</td>
</tr>
</tbody>
</table>

### Accuracy: 80%
Next

1. Introduction
2. Citation Analysis
3. Classification of Citations
4. Classification Schemes
5. Dataset Collection
6. Overview of CRC
7. Results
8. Conclusion and Future Work
Conclusion

• Help the research scholars by minimizing the time required to find the relevant research, for a topic of interest.

• Classify the citations into Positive, Negative and Neutral.

• Our preliminary results show an accuracy of 80%.

• We assert that the technique can be generalized to classification of scientific research papers.
Future Work

- Currently, support for Type-III papers in CRC is in progress, and we are working on a lexicon dictionary with neutral words.

- In future, we plan to provide a web portal to assist the research scholars in automatically searching and downloading citing papers, for a root paper, and classification of citing papers into sentiment categories.
Proposed UI

Welcome to CRC home

Upload new File

Welcome to CRC home

["signature include And more designed function needs designed" => positive]

Positive Words
Negative Words
Total Words count

Overall Words Distribution

Istanbul, Turkey. June 29 2015

CRC - CLBib, ISSI 2015
Thank you

Questions
Concerns
Suggestions
## Distribution in Feature-Set.

<table>
<thead>
<tr>
<th>No.</th>
<th>Test Type</th>
<th>Window (distribution of Type I papers)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Train=21, Test=21</td>
<td>0-21</td>
<td>81%</td>
</tr>
<tr>
<td>2</td>
<td>Train=21, Test=21</td>
<td>22-43</td>
<td>77%</td>
</tr>
<tr>
<td>3</td>
<td>Train=21, Test=21</td>
<td>44-65</td>
<td>71%</td>
</tr>
<tr>
<td>4</td>
<td>Train=21, Test=21</td>
<td>66-87</td>
<td>85%</td>
</tr>
<tr>
<td>5</td>
<td>Train=21, Test=21</td>
<td>88-109</td>
<td>67%</td>
</tr>
<tr>
<td>6</td>
<td>Train=21, Test=21</td>
<td>110-125</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>(78%)</td>
</tr>
</tbody>
</table>
## Result with random sampling

<table>
<thead>
<tr>
<th>No.</th>
<th>Test Type</th>
<th>Window (distribution of Type I papers)</th>
<th>Accuracy</th>
</tr>
</thead>
</table>
References


