|  |
| --- |
| **Requirements Specification**  **Citation-Based Patent Tree and Valuation** |

###### **2XB3 - Final Project**

**Development Team:**

Daniel Cefaratti

Jim Wu

Maxim Vasiliev

Susie Yu

Saadi Salahuddin

This application regards the domain of patents, and assessment of their value network structure. Patents are documents which are written by an individual or group to describe an invention, and if approved, give exclusive rights to the assignee of the patent to commercial use of the invention. A typical patent consists of a header and body. Of concern to us is information in the header, specifically the Patent Number, and “Related Patent Documents” list. A patent which build upon the ideas mentioned in a previously approved, or currently pending patent may list that previous patent as a “Related Patent”. As such, patents which inspire more ideas described in future patents could be considered more valuable, and would have a higher occurrence rate as a “Related Patent”.

As fundamental components of corporations’ or individuals’ intellectual property (IP) portfolio, patents are important to consider whenever a capital assessment of their owner is made, or when calculating financial metrics related to the use of a patent as an IP tool. End users may include financial analysts wishing to quantify the intellectual capital of a corporation, such as prior to a merger or acquisition. Lawyers wishing to determine the market value of an invention contained within a patent as part of fair royalty calculations made to assess feasibility of using the particular patent invention may find this application useful. Or entrepreneurs wishing to gain insight into the patent network structure of a particular industry. The application could also be part of a larger financial or intellectual property software suite.

Information contained within all patents is regularly published by the United States Patent and Trademark Office (USPTO). Individual patents are accessible electronically by searching for their patent number. The resultant information includes the Related Patents list. Searching for a patent and extracting out its Related Patents can be done programmatically to build a patent citation network. This could be represented graphically, and/or the number of times a specific patent is mentioned by patents up its citation network can then be deduced to produce a score.

Functionally, upon execution the application prompts the user to input a patent number, and either the “tree” or “score” setting, after which the main calculation is performed. If the tree setting was selected, the application produces a graphical network showing consisting of nodes representing patents, and directed edges representing which patents are related. If the score setting is selected, the application simply outputs a count of the number of patents which are in the input patent’s citation tree. One possible change or additional feature might be to consider the patents’ authors, and the creation of a network in which individuals are nodes and edges represent the existence of a patent where two individuals were co-authors. This would involve extracting author data from patent queries as well.

Non-functional considerations include limits on USPTO database accession. Theoretically, a network of all patent citations is quite feasible since in 2014 only 326,033 patents were granted. This places the number of nodes for the last 2 decades at under 10 million. However, making that many requests each time the application is executed is not feasible or efficient. The application should therefore be divided into an active and passive and passive component. The passive component queries and build the patent network over several days, then upon completion adds new patents as they are granted. Meanwhile the active component displays the network or calculates the score upon execution by the user.

Since 1976 the USPTO standardized their requirements for patent applications. This ensures the data collected from USPTO since 1976 is reliable.