ĐẠI HỌC QUỐC GIA TP.HCM **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**TRƯỜNG ĐẠI HỌC Độc lập – Tự do – Hạnh phúc  
 CÔNG NGHỆ THÔNG TIN**

**ĐỀ CƯƠNG ĐỀ TÀI LUẬN VĂN KỸ SƯ**

1. **Tên đề tài hoặc hướng nghiên cứu**

Tên tiếng Việt: Nghiên cứu các phương pháp biểu diễn dữ liệu lớn dưới dạng đồ thị chuyển động.

Tên tiếng Anh: Research on some methods for big data visulisation using motion chart.

1. **Ngành và mã ngành đào tạo**

Ngành: Truyền thông và Mạng máy tính.

Mã ngành: 52480102.

1. **Họ tên học viên, khóa học và đợt bảo vệ khóa luận**

SV1: Đinh Thành An, khóa 2011 – 2016, đợt 1.

SV2: Lê Văn Đức Đạo, khóa 2011 – 2016, đợt 1.

**Người hướng dẫn**:ThS. Nguyễn Quang Minh.

**Địa chỉ email, điện thoại liên lạc của người hướng dẫn:**

Email: minhnq@uit.edu.vn, tuanna@uit.edu.vn.

Điện thoại: 0937428080,09322150.

1. **Tổng quan tình hình nghiên cứu**

Đối với trong nước, trong quá trình nhóm tác giả tìm hiểu thì hướng nghiên cứu có liên quan tới đề tài này chưa được thực hiện. Các hướng nghiên cứu mới chỉ dừng lại ở việc sử dụng sử dụng các dữ liệu lớn có sẵn trong các hệ thống lưu trữ chứ chưa thực hiện việc thể hiện dữ liệu theo thời gian thực hay tự tạo và lưu trữ các dữ liệu vào một hệ thống riêng.

Đối với ngoài nước, đã có một số hướng nghiên cứu có liên quan tới đề tài này thông qua các ứng dụng như GoogleFinance của nhà phát hành Google, YahooFinance của Yahoo. Tuy nhiên, những ứng dụng này đều không có những tính năng mà mục tiêu của đề tài hướng tới. Như việc ứng dụng CloudMesh, ZeroPc hay Cloud Save đều không hỗ trợ việc phân tách, ghép nối tập tin cũng như mã hóa, giải mã và hash tập tin. Bên cạnh đó, các ứng dụng này chỉ hỗ trợ việc tải lên tập tin cho đúng một tài khoản ODS cũng như không quan tâm tới dung lượng lưu trữ còn lại của từng tài khoản đó.

1. **Tính khoa học và tính mới của đề tài**

Ngày nay, dữ liệu ngày càng gia tăng cách nhanh chóng do đó dẫn đến việc xử lý thông các dữ liệu này ngày càng khó khăn hơn. Đặc biệt trong ngành chứng khoán với những số liệu khổng lồ trong từng giờ đòi hỏi cần phải có các công cụ phân tích theo thời gian thực một cách sinh động các dữ liệu này. Dựa vào các công cụ này người dùng có thể có cái nhìn sinh động hơn về những gì đang diễn ra đối với thị trường.

1. **Mục tiêu, đối tượng và phạm vi nghiên cứu của đề tài**

* Mục tiêu nghiên cứu: Xây dựng một Hệ thống nhằm:
* Thu thập các dữ liệu theo thời gian để có thể biểu diễn dữ liệu phục vụ cho việc phân tích thị trường.
* Hỗ trợ người dùng theo dõi diễn biến của các mã chứng khoán có liên quan.
* Hỗ trợ người dùng đưa ra quyết định về việc mua bán cổ phiếu.
* Hỗ trợ việc phân tách tập tin theo dung lượng lưu trữ còn lại của các tài khoản ODS.
* Ứng dụng vào thực tiễn.
* Đối tượng nghiên cứu:
* Các công cụ lấy dữ liệu từ web.
* Tìm hiểu các thuật ngữ và dữ liệu của thị trường chứng khoán.
* Nghiên cứu hệ thống xử lý dữ liệu lớn.
* Xây dựng hệ thống webserver theo thời gian thực.
* Phương thức tải xuống và tải lên.
* Ngôn ngữ lập trình Java và nền tảng JavaFX trong xây dựng GUI.
* Phạm vi nghiên cứu:
* Các công cụ lấy dữ liệu từ web như là scrapy.
* Hệ thống xử lý dữ liệu lớn bigdata Hadoop
* Nghiên cứu các trang web về thông tin thị trường chứng khoán.

1. **Nội dung, phương pháp dự định nghiên cứu**

* Giai đoạn 1
* Kết quả dự kiến:

Xây dựng trang web thể hiện các dữ liệu của thị trường chứng khoán một cách sinh động, có phân tích theo từng loại mã chứng khoán và có các option hỗ trợ người dùng ra quyết định mua bán.

* Phương pháp thực hiện:

Đề tài được thực hiện bằng việc tiến hành khảo sát, nghiên cứu các công cụ như là scrapy, nghiên cứu các trang web về thị trường chứng khoán ở Việt Nam và nghiên cứu cách triển khai hệ thống Hadoop.

* Giai đoạn 2
* Kết quả dự kiến:

Xây dựng giao diện người dùng thân thiện, thuận tiện cho người sử dụng đồng thời đảm bảo tính thẩm mỹ, bắt mắt người sử dụng.

Hoàn thiện chức năng, cũng như hiệu năng của chương trình.

* Phương pháp thực hiện: Đề tài được thực hiện bằng việc tiến hành khảo sát, nghiên cứu xu hướng của người sử dụng về giao diện, các thao tác người dùng khi tương tác với ứng dụng, phân tích kết quả thu được cũng như phân tích các ứng dụng hiện có để hoàn thiện ứng dụng.

1. **Kế hoạch bố trí thời gian nghiên cứu**

|  |  |  |
| --- | --- | --- |
| Nội dung | Thời gian | Công việc |
| Đăng kí đề tài. | 17/09/2015 | - Gặp giảng viên hướng dẫn.  - Đăng kí đề tài. |
| Khảo sát, nghiên cứu và phân tích thực trạng. | 22/09/2015 –28/09/2015 | -Khảo sát tìm hiểu các công cụ liên quan (Scrapy, d3js, Hadoop)  -Tìm hướng đi cho đề tài |
| Nghiên cứu các kiến thức về thị trường chứng khoán Việt Nam. | 29/09/2015 –05/12/2015 | Nghiên cứu các thuật ngữ, thuật toán được dùng trong chứng khoán |
| Xây dựng web server bằng cách sử dụng ngôn ngữ python ( tích hợp scrapy ) | 29/09/2015 –12/10/2015 | - Xây dựng web server  - Tích hợp scrapy để thực hiện các yêu cầu của đề tài |
| Nghiên cứu các dữ liệu lấy được từ website | 13/10/2015 –26/10/2015 | - Nghiên cứu các dữ liệu lấy được từ các website chứng khoán  - Phân tích và lựa chọn các dữ liệu phù hợp với yêu cầu của đề tài |
| Triển khai hệ thống Hadoop | 12/10/2015 –25/10/2015 | - Nghiên cứu lớp tập tin (File) của java có trong Hadoop.  - Triển khai việc lưu và lấy thông tin từ Hadoop. |
| Triển khai giao diện người dùng | 26/10/2015 –08/11/2015 | - Nghiên cứu phát triển giao diện bằng JavaSwing.  - Thiết kế và xây dựng giao diện với câu lệnh trong JavaSwing. |
| Báo cáo 50% tiến độ. | 09/11/2015 –15/11/2015 | - Chuẩn bị báo cáo 50%.  - Gặp Giảng viên hướng dẫn (GVHD), báo cáo 50% và nhận được nhiều hướng dẫn quan trọng về giao diện vào bảo mật từ GVHD. |
| Nghiên cứu phương thức mã hóa tập tin. | 24/11/2015 –30/11/2015 | Sử dụng phương thức mã hóa AES để mã hóa tập tin. |
| Nghiên cứu phương thức hash tập tin. | 01/12/2015 –07/12/2015 | - Nghiên cứu các giải thuật hash  - Ứng dụng giải thuật SHA-256 để hash tập tin. |
| Thiết kết giao diện người dùng với JavaFX. | 08/12/2015 –22/12/2015 | - Nghiên cứu về JavaFX.  - Thiết kế giao diện trong Photoshop.  - Xây dựng các tập tin CSS. |
| Gặp GVHD báo cáo tiến độ. | 12/12/2015 | - Báo cáo về giao diện mới.  - Nhận được hướng dẫn về giao diện và chức năng hệ thống từ GVHD.  - Nhận được hướng dẫn về trình bày báo cáo từ GVHD. |
| Chuyển code vào giao diện người dùng với JavaFX. | 15/12/2015 –29/12/2015 | - Nghiên cứu cách hoạt động, tương tác với giao diện với JavaFX.  - Đưa các phần xử lý vào trong giao diện với JavaFX.  - Xử lý các tương tác người dùng trên giao diện với JavaFX. |
| Gặp GVHD báo cáo tiến độ. | 09/01/2016 | - Báo cáo về ứng dụng và bản thảo báo cáo.  - Nhận được hướng dẫn thêm về giao diện và chức năng chương trình từ GVHD.  - Nhận được hướng dẫn về trình bày báo cáo từ GVHD. |
| Hoàn thiện chức năng cũng như giao diện người dùng. | 29/12/2015 – 12/01/2016 | - Tiếp tục hoàn thiện và xây dựng giao diện người dùng.  - Thiết kết lại các biểu tượng, các mục hiển thị cho phù hợp.  - Xử lý lỗi phát sinh. |
| Hoàn thiện báo cáo. | 05/01/2016 –18/01/2016 | Nhận thông báo từ Khoa và chuẩn bị đầy đủ các thủ tục để nộp báo cáo và các vấn đề liên quan. |

1. **Tài liệu tham khảo**

**[1].** Wales, Jimmy; Sanger, Larry, "Drobox (Service)," Wikimedia Foundation, Inc., 15 January 2001. [Online]. Available: http://en.wikipedia.org/wiki/Dropbox\_(service). [Accessed 3 December 2014].

**[2].** R. Boyd, Getting Started with OAuth 2.0, Sebatopol: O'Reilly Media, Inc., 2012.

[d3js.org](http://d3js.org/)

<http://scrapy.org/>

<http://bost.ocks.org/mike/nations/>

*TP. HCM, ngày 13 tháng 09 năm 2015*

**NGƯỜI HƯỚNG DẪN SINH VIÊN KÝ TÊN**

Sinh viên 1 Sinh viên 2

**NGUYỄN QUANG MINH ĐINH THÀNH AN LÊ VĂN ĐỨC ĐẠO**