# Chapter 6
# Big Data and FinTech

**Jia-Lang Seng, Yao-Min Chiang, Pang-Ru Chang, Feng-Shang Wu, Yung-Shen Yen, and Tzu-Chieh Tsai**

## 6.1 Introduction

With the prevalence of digital convergence, mobile communication, Big Data analytics, cloud computing service, and artificial intelligence, the digital financial industry is experiencing a revolutionary trend. In essence, the widespread popularity of smartphones, the intelligent computing, and the ubiquity of mobile cloud services have revolutionized the face of business worldwide, especiallyin the financial

J.-L. Seng (✉)
Department of Accounting, National Chengchi University, Taipei, Taiwan
e-mail: seng@nccu.edu.tw

Y.-M. Chiang
Department of Finance, National Taiwan University, Taipei, Taiwan
e-mail: yaominchiang@ntu.edu.tw

P.-R. Chang
Department of Risk Management and Insurance, Shih-Chien University, Taipei, Taiwan
e-mail: brchang@g2.usc.edu.tw

F.-S. Wu
Graduate Institute of Technology, Innovation and Intellectual Property Management, National Chengchi University, Taipei, Taiwan
e-mail: fswu@nccu.edu.tw

Y.-S. Yen
Department of Computer Science and Information Management, Providence University, Taichung, Taiwan
e-mail: ysyen@pu.edu.tw

T.-C. Tsai
Department of Computer Science, National Chengchi University, Taipei, Taiwan
e-mail: ttsai@cs.nccu.edu.tw



S.-H. Chen (ed.), *Big Data in Computational Social Science and Humanities*, Computational Social Sciences, https://doi.org/10.1007/978-3-319-95465-3_6

services industry, also known as financial technologies (FinTechs). The market changes to platform and infrastructure mean that data analytics related or other real-time computing requirements affect the value chain in FinTechs which in turn determines the market share in the world. The financial industry must be encountering the global changes by actively seeking for destructive innovation ideas and escalating their customer experience and service quality. With the sponsorship of Taiwan Ministry of Science and Technology's 3-year multidisciplinary research project, "Innovative and Mobile Financial Technologies," this chapter aims to document the research findings and outcome of the research project where we try to integrate the financial intelligence of structured and unstructured data, create a cloud-based mobile computing platform, and build a two-way decision support prototype system. This integrated approach adopts the academic studies and advanced technical development. The main research results documented include the news media database, rule base, asset pricing model, multi-case study series, empirical research model, data analytics algorithm development, sentiment analysis and opinion mining application, and mobile cloud computing platform (Deloitte 2016, pp. 1–5; KPMG 2015a, pp. 10–13; WEF 2015, pp. 1–3; WEF 2017).

With the emergence of service innovations and innovation integration in clusters, the focus has shifted to the development of new financial services with data science and analytics, mobile technologies, and cloud computing that diversify and customize the new functionalities available to business clients and customers. Most worldwide financial holdings corporations have faced major challenges resulting from dynamic business environments, diverse telecommunication situations, and heterogeneous data sources that must be marshaled in an integrated, real-time, and seamless manner. Today, businesses interact with heterogeneous business models, business processes, and workflows. The results have been upheaval, chaos, and disruption. Therefore, it is crucial to develop comprehensive, cross-disciplinary, and in-depth resolution of services, models, and technologies to address the above issues. Facing the massive amount of dynamically changing and heterogeneous data sources, financial advisors, managers, investors, and other stakeholders need real-time, accurate, comprehensive, high-frequency, and interactive decision support systems to make investment decisions as services. The goal is to increase the wealth of clients and to improve the corporations' competitive edge. In this chapter, we present four main research results in the following sections. They are (1) intelligent investment models in the capital market, (2) text analytics and sentiment analyses of financial news, (3) innovative financial service strategies, and (4) mobile cloud technologies to implement these proposed solutions.

In Sect. 6.2, we incorporate news sentiment into an asset pricing model to examine various views of stock price reactions to investment, macroeconomic, and political news in 2013–2014. The results show that there is positive relationship between investment news sentiment and TW50 return that is there is the statistical significance. However, regarding the effect of macroeconomic news on stock prices, these results are very different from the results of previous articles in the literature. Political news sentiment has statistically significant negative effects on certain financial stocks listed on the TW50 index.

In Sect. 6.3, we focus on the financial news sentiment analysis conducted by using software. We analyze the content of the news, develop various customized word lists, build a dictionary, and refine scoring rules that evaluate news items in the contexts of various research topics. The result is statistically significant but the result may be affected by the subjective judgment of readers. Our dictionary captures positive words and negative words, but it has limited grasp of what the words in news articles convey.

In Sect. 6.4, we investigate how securities firms and information and communication technology (ICT) development firms respond to the threats and opportunities of coming era of FinTechs. From the results, we thought that securities firms may need to rethink the way they divine customers' needs and interact with customers and building their own in-house innovative capabilities in order to be capable of quickly responding to customers' needs and to social and economic changes, and the ICT development firms may need to reposition their products and services for the financial services industry and may need to update their ICT technologies. In this section, we also find that the relationship between securities firms and ICT development firms may be changing as the securities firms appear to intend to develop mobile systems and service innovations by themselves.

In the final section, we propose a framework that can handle complex data structures and large amounts of data in a short time using cloud computing. The result shows the response times of different task assignment algorithms for different intensities of burst traffic.

## 6.2 Text Mining News for Stock Price Predictions

### *6.2.1 Introduction*

The efficient market and rational investor hypotheses in the literature imply that the price of a security reflects the information available to investors concerning the value of that security. Security pricing is a topic that is still debated among academics and practitioners, and the efficient market hypothesis has been re-examined by numerous researchers (Tetlock 2007, pp. 1139–1168; Fama and French 2015, pp. 1–22). In the traditional asset pricing model, researchers use a single factor or multiple factors to predict actual asset price, such as market premium, economic factors, firm size, and book-to-market ratio (Fama and French 1992, pp. 427–465; Fama and French 2015, pp. 1–22). However, Baker and Wurgler (2006, pp. 1645–1680), Brown and Cliff (2005, p.405–440), and Kumar and Lee (2006, pp. 2451–2486) show that other factors such as investor sentiment can predict stock returns in the cross-section. Therefore, in this study, we add news sentiment as another factor to construct a new security pricing model to examine the effect of news sentiment on stocks.

Some recent studies suggest that media coverage may have statistically significant relationships to asset prices even when it does not involve hard, breaking news (Tetlock 2007, pp. 1139–1168; Kearney and Liu 2014, pp. 171–185; Fang and Peress 2009, pp. 2023–2052) and report that firms with less media coverage have higher required rates of return. Firms that experience an exogenous reduction in analyst coverage have higher required rates of return as well as less efficient pricing and lower liquidity. Another study uses the textual sentiment to price actual security assets to show that textual sentiment can produce statistically significant abnormal returns (Loughran and McDonald 2011, pp. 35–65). Therefore, we include news sentiment into an asset pricing model. We argue that using media coverage as a proxy for investor sentiment in an asset pricing model may improve the performance of that asset pricing model.

## *6.2.2 Literature Review*

### 6.2.2.1 Asset Pricing

Asset pricing was first explored by Sharpe (1964, pp. 425–442), who uses a Markowitz model to estimate average returns and risks of market portfolios; this is called a capital asset pricing model (CAPM). This single-factor model suggests that the security risk expected return depends on beta and the risk premium of the market portfolio. Apart from the market premium, multifactor models add other factors such as business-cycle risk, monetary policy, and the ratio of economic growth to the price security, as arbitrage pricing models (APMs) do. In 1993, Fama and French design a three-factor model that considered market premium, firm size, and book-to-market equity ratio to price securities. In 2015, to improve their model's performance regarding price-risky assets, Fama and French add another two factors to their three-factor model, namely operating profitability and investment.

### 6.2.2.2 News Sentiment

Recent studies find that media coverage and investors sentiment are substantially related to asset prices. Models of investor sentiment predict that low sentiment will generate downward price pressure, and unusually high or low values of sentiment will generate high volume. Antweiler and Frank (2004, pp. 1259–1294) study messages in Internet chat rooms focused on stocks, characterizing the content of the messages as "buy," "sell," or "hold" recommendations to investigate the relationship between the message and return. Tetlock (2007, pp. 1139–1168) suggests that measures of media content serve as a proxy for investor sentiment or noninformational trading; his empirical results show that media volume (media coverage) has positive relationships with idiosyncratic volatility, past-year return, and past-year absolute return but a negative relationship with the book-to-market ratio. Kurov (2010, pp. 139–149) also finds that monetary policy shocks strongly influence investor sentiment in bear market periods. Garcia (2013, pp. 1267–1300) uses the proportions of positive and negative words to examine the effect of sentiment on asset prices and finds that daily news content helps predict stock returns, particularly during recessions. Kearney and Liu (2014, pp. 171–185) indicate that researchers use positive and negative verbal information based on textual sentiment, which may lead to notable differences in risk-adjusted stock returns for complementary econometric modeling of financial market effects. They further claim that textual sentiment or the tone of qualitative information has noteworthy effects on stock prices and returns and that negative sentiment has the strongest influence.

### *6.2.3 Methodology*

We construct a multifactor model to examine the effects of news on stock prices. News events are gauged by collecting 59,223 investment articles, 20,090 macroeconomic articles, and 49,848 politics articles from January 1, 2013, to December 31, 2014, and stock market returns from January 2, 2013, to December 31, 2014, as independent variables. We adopt the stock returns of TSEC-FTSE Taiwan 50 Index (TW50) listed companies as dependent variables. The model can be expressed as follows:

$$r_i = \beta_0 + \beta_1 r_m + \beta_2 \Delta s_{\mathrm{inv}} + \beta_3 \Delta s_{\mathrm{mac}} + \beta_4 \Delta s_{\mathrm{poli}} + \varepsilon_i$$

where $r_i$ is the return on the price of stock $i$, $r_m$ is the stock market return, $\Delta s_{\mathrm{inv}}$ is the difference of investment news score, $\Delta s_{\mathrm{mac}}$ the difference of macroeconomic news score, and $\Delta s_{\mathrm{poli}}$ is the difference of political news score.

### *6.2.4 Empirical Results*

Table 6.1 provides an overview of three group's news sentiment scores and the stock market index. The market index changed from 7616 to 9569 in 494 business days. In Table 6.2, the TW50-listed stock returns have strong, statistically significant positive correlations with stock market returns. The positive relation of investment news sentiment to TW50 return is statistically significant. However, regarding the effect of macroeconomic news on stock prices, these results are very different from the results of previous articles in the literature. Political news sentiment has statistically significant negative effects on certain financial stocks listed on the TW50 index, namely stocks 2883 (China Development Financial), 2884 (E.SUN Financial Holding Company), and 2891 (CTBC Financial Holding Co., Ltd.).

**Table 6.1** Summary statistics of news observations

| | Market index | Investment news | Politics news | Macroeconomic news |
|---|---|---|---|---|
| *N* | 494 | 59,223 | 20,090 | 49,848 |
| Mean | 8544.21 | 0.38 | 0.34 | 0.04 |
| Median | 8466.38 | 0.00 | 0.00 | 0.00 |
| SD | 523.13 | 2.04 | 4.92 | 1.08 |
| Min | 7616.64 | −38.00 | −56.00 | −28.00 |
| Max | 9569.17 | 36.00 | 54.00 | 36.00 |

We select news from three categories—investment, politics, and macroeconomic news, and stock market index—from January 2, 2013, to December 31, 2014.

### *6.2.5 Conclusion and Discussion*

This research examines various views of stock price reactions to investment, macroeconomic, and political news in 2013–2014. The TW50-listed stock prices are highly related to the stock market index and to the investment news, but only the stock prices in the banking industry relate to political news. In future work, we can classify various stocks according to the industries used by the Taiwan Stock Exchange Corporation (TWSE) and calculate the cross-sectional price return for each day.

This table shows the TW50 stock alphas, coefficients, *t*-values, and adj-$R^2$ values from January 2, 2013, to December 31, 2014, in a time series regression. The model is

$$r_i = \beta_0 + \beta_1 r_m + \beta_2 \Delta s_{\text{inv}} + \beta_3 \Delta s_{\text{mac}} + \beta_4 \Delta s_{\text{poli}} + \varepsilon_i$$

where the right-hand variable is stock price return and the left-hand variables are four-factor differences of investment news, macroeconomic news, political news, and stock market returns.

## 6.3 Financial News Sentiment Analysis and Application

### *6.3.1 Introduction*

This section explores and consolidates valuable information from analysis of financial news based on sentiment analysis and opinion mining. We analyze sentiments in news content and apply this sentiment analysis. We extract subjective vocabulary that reflects emotion or attitude. Moreover, we develop a dictionary and propose a method for calculating a score from specific emotionally subjective vocabulary used in the news.

We focus on using our content analysis techniques to extract information from a news database and construct an automatic news-scoring application based on our

**Table 6.2** TW50 stocks 4-factor regression

| Stock id | Intercept | $\Delta s_{\text{inv}}$ | $t$-value | $\Delta s_{\text{mac}}$ | $t$-value | $\Delta s_{\text{poli}}$ | $t$-value | $r_m$ | $t$-value | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1101 | 0.053 | 0.632*** | (3.99) | −0.003 | (−0.07) | 0.209 | (0.80) | 0.618*** | (10.73) | 0.19 |
| 1102 | 0.039 | 0.276** | (2.51) | 0.013 | (0.50) | 0.091 | (0.51) | 0.387*** | (9.68) | 0.15 |
| 1216 | 0.030 | 0.450** | (2.65) | −0.025 | (−0.61) | −0.100 | (−0.36) | 0.541*** | (8.77) | 0.13 |
| 1301 | 0.005 | 0.234* | (1.80) | 0.036 | (1.14) | −0.193 | (−0.90) | 0.541*** | (11.46) | 0.21 |
| 1303 | 0.050 | 0.283* | (1.65) | 0.017 | (0.41) | 0.585** | (2.08) | 0.632*** | (10.17) | 0.17 |
| 1326 | 0.000 | 0.519** | (3.63) | 0.040 | (1.17) | 0.307 | (1.31) | 0.623*** | (11.99) | 0.22 |
| 1402 | 0.018 | 0.262** | (2.09) | 0.018 | (0.61) | −0.315 | (−1.53) | 0.435*** | (9.56) | 0.16 |
| 1476 | 0.290 | 0.581 | (1.53) | −0.094 | (−1.04) | −0.788 | (−1.27) | 0.494*** | (3.59) | 0.03 |
| 2002 | 0.010 | 0.374** | (3.79) | 0.038 | (1.60) | 0.162 | (1.00) | 0.436*** | (12.16) | 0.23 |
| 2105 | 0.045 | 0.512** | (3.19) | 0.055 | (1.42) | 0.065 | (0.25) | 0.434*** | (7.45) | 0.10 |
| 2207 | 0.181 | 0.773** | (3.05) | −0.003 | (−0.05) | 0.121 | (0.29) | 0.983*** | (10.66) | 0.18 |
| 2227 | 0.112 | 0.256 | (1.08) | −0.009 | (−0.16) | −0.092 | (−0.24) | 0.466*** | (5.40) | 0.05 |
| 2301 | 0.018 | 0.598** | (3.15) | 0.028 | (0.61) | 0.545* | (1.75) | 0.482*** | (6.99) | 0.09 |
| 2303 | 0.074 | 0.508** | (2.52) | 0.110* | (2.27) | −0.294 | (−0.89) | 0.768*** | (10.48) | 0.18 |
| 2308 | 0.144 | 0.477** | (2.15) | −0.033 | (−0.62) | −0.194 | (−0.53) | 0.585*** | (7.26) | 0.09 |
| 2311 | 0.107 | 0.327 | (1.61) | 0.042 | (0.86) | −0.349 | (−1.05) | 0.618*** | (8.38) | 0.12 |
| 2317 | 0.056 | 0.452** | (3.27) | 0.017 | (0.52) | 0.141 | (0.62) | 0.588*** | (11.69) | 0.21 |
| 2325 | 0.120 | 0.409** | (2.06) | 0.051 | (1.08) | −0.022 | (−0.07) | 0.614*** | (8.50) | 0.12 |
| 2330 | 0.093 | 0.731*** | (4.70) | 0.072* | (1.92) | −0.323 | (−1.27) | 0.845*** | (14.98) | 0.32 |
| 2354 | 0.021 | 0.418** | (2.69) | 0.033 | (0.88) | −0.349 | (−1.37) | 0.506*** | (8.98) | 0.14 |
| 2357 | 0.057 | 0.246 | (1.14) | 0.043 | (0.83) | −0.347 | (−0.98) | 0.525*** | (6.70) | 0.08 |
| 2382 | 0.069 | 0.760** | (3.82) | 0.075 | (1.57) | 0.349 | (1.07) | 0.601*** | (8.30) | 0.12 |
| 2395 | 0.185 | 0.332 | (1.29) | 0.121* | (1.95) | 0.116 | (0.28) | 0.694*** | (7.45) | 0.10 |
| 2408 | 0.301 | 0.218 | (0.46) | −0.019 | (−0.17) | −0.527 | (−0.68) | 0.650*** | (3.80) | 0.02 |
| 2409 | 0.061 | 0.392 | (1.47) | 0.041 | (0.63) | −0.074 | (−0.17) | 0.759*** | (7.84) | 0.11 |
| 2412 | 0.023 | 0.105 | (1.49) | −0.006 | (−0.36) | 0.039 | (0.33) | 0.186*** | (7.23) | 0.09 |
| 2454 | 0.098 | 0.463** | (2.20) | 0.066 | (1.30) | −0.220 | (−0.64) | 0.607*** | (7.96) | 0.11 |

(continued)

**Table 6.2** (continued)

| Stock id | Intercept | $\Delta s_{\mathrm{inv}}$ | $t$-value | $\Delta s_{\mathrm{mac}}$ | $t$-value | $\Delta s_{\mathrm{poli}}$ | $t$-value | $r_m$ | $t$-value | Adj-$R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2474 | 0.141 | 0.404 | (1.55) | −0.044 | (−0.71) | −0.655 | (−1.54) | 0.769*** | (8.15) | 0.12 |
| 2801 | 0.056 | 0.420** | (3.84) | 0.033 | (1.25) | 0.110 | (0.61) | 0.473*** | (11.89) | 0.22 |
| 2880 | 0.043 | 0.217** | (2.23) | 0.008 | (0.33) | −0.186 | (−1.17) | 0.454*** | (12.90) | 0.25 |
| 2881 | 0.097 | 0.548** | (3.63) | 0.003 | (0.07) | −0.274 | (−1.10) | 0.676*** | (12.33) | 0.24 |
| 2882 | 0.124 | 0.349** | (2.29) | 0.021 | (0.58) | −0.186 | (−0.74) | 0.698*** | (12.61) | 0.24 |
| 2883 | 0.079 | 0.481** | (3.44) | 0.033 | (0.99) | −0.524** | (−2.28) | 0.533*** | (10.49) | 0.19 |
| 2884 | 0.091 | 0.250** | (1.90) | 0.003 | (0.11) | −0.430** | (−1.99) | 0.433*** | (9.07) | 0.15 |
| 2885 | 0.036 | 0.602*** | (4.08) | 0.050 | (1.42) | 0.230 | (0.95) | 0.715*** | (13.34) | 0.26 |
| 2886 | 0.040 | 0.281** | (2.32) | 0.009 | (0.32) | −0.248 | (−1.25) | 0.492*** | (11.20) | 0.20 |
| 2887 | 0.077 | 0.374** | (3.06) | 0.028 | (0.94) | −0.059 | (−0.29) | 0.486*** | (10.94) | 0.19 |
| 2890 | 0.056 | 0.376** | (3.14) | 0.034 | (1.19) | −0.437** | (−2.22) | 0.499*** | (11.47) | 0.22 |
| 2891 | 0.078 | 0.579*** | (4.61) | 0.022 | (0.73) | −0.065 | (−0.31) | 0.578*** | (12.67) | 0.24 |
| 2892 | 0.048 | 0.203** | (2.32) | −0.003 | (−0.15) | −0.161 | (−1.13) | 0.445*** | (14.06) | 0.29 |
| 2912 | 0.114 | 0.343* | (1.91) | 0.070 | (1.62) | −0.129 | (−0.44) | 0.440*** | (6.74) | 0.08 |
| 3008 | 0.254 | 0.723** | (2.30) | −0.054 | (−0.72) | −0.050 | (−0.10) | 0.835*** | (7.33) | 0.10 |
| 3045 | 0.030 | 0.527** | (3.36) | 0.035 | (0.92) | 0.330 | (1.28) | 0.397*** | (6.96) | 0.09 |
| 3474 | 0.581 | 0.674 | (1.52) | −0.049 | (−0.46) | −0.743 | (−1.02) | 0.496*** | (3.08) | 0.02 |
| 3481 | 0.018 | 0.707** | (2.46) | 0.103 | (1.49) | −0.627 | (−1.33) | 0.810*** | (7.77) | 0.11 |
| 4904 | 0.027 | 0.156 | (0.98) | 0.014 | (0.35) | 0.510* | (1.95) | 0.295*** | (5.11) | 0.05 |
| 4938 | 0.172 | 0.238 | (0.94) | 0.064 | (1.06) | 0.113 | (0.27) | 0.486*** | (5.31) | 0.04 |
| 5880 | 0.033 | 0.169** | (2.35) | 0.003 | (0.19) | −0.062 | (−0.53) | 0.305*** | (11.68) | 0.21 |
| 6505 | −0.034 | 0.300** | (2.04) | −0.070* | (−1.98) | −0.001 | (0.00) | 0.589*** | (11.01) | 0.20 |
| 9904 | 0.077 | 0.748** | (3.41) | 0.040 | (0.77) | 0.543 | (1.51) | 0.624*** | (7.82) | 0.11 |

***, **, and * are respectively the statistical significance levels at 1%, 5%, and 10%

scoring rules. In addition, on the basis of an analysis of commonly used words in the news, our study develops various customized word lists of terms oriented toward macroeconomics, politics, and investment.

### *6.3.2 Literature Review*

Data analytics is a process in which computer programs apply mathematical and statistical methods to extract new and previously unknown information from textual data. Unlike numerical and financial data, textual data contains not only the effect, but also the possible causes of the event. The ability to exploit textual information successfully could increase the quality of the input data and improve the understanding of issues.

Content analysis[1] is a wide and heterogeneous set of manual or computer-assisted techniques for contextualized interpretations of documents produced by communication processes (in the strict sense of that phrase) or by signification processes, of which the ultimate goal is the production of valid and trustworthy inferences. Tetlock (2007, pp. 1139–1168) examines investor sentiment by measuring the pessimism index from the GI dictionary. Tetlock et al. (2008, pp. 1437–1467) find that negative words in firm-specific news stories predict low firm earnings. Past studies are also commonly known as document-level sentiment classification because the whole document is considered as a basic information unit.

Hu et al. (2012, pp. 674–684) indicate that research on sentiment analysis uses an automatically generated sentiment lexicon, in which a list of seed words is used to determine whether a sentence contains positive or negative connotations. Then, the polarity (positive or negative) of an opinion is determined on the basis of the words in the document. Xu et al. (2011, pp. 743–754) classify technologies of sentiment analysis into two categories: unsupervised approaches and supervised approaches. The unsupervised approaches usually create a sentiment lexicon and determine a document's polarity by counting that document's positive and negative phrases. The supervised approaches use labeled data to train certain classifiers to predict unlabeled data. The semantic orientation approach (Zhang et al. 2013, pp. 851–860) performs classification on the basis of positive or negative sentiment words (or phrases) contained in each evaluated item (this operates on several levels: document, sentence, or attribute). The lexicon is crucial to the semantic orientation approach. However, the speed at which vocabulary items are collected is less than the speed at which vocabulary is generated daily on the Internet. In addition, constructing different vocabularies for different domains is indeed challenging because of the polysemous nature of the terms. Therefore, the task of vocabulary construction is difficult because of the different areas of domain knowledge. Four challenges are encountered when creating domain-specific texts (Ittoo and Bouma 2013, pp. 2530–2540): silence, the absence of knowledge resources, complex terms, and noise (informal or ungrammatical language). Domain-specific texts are sparse and do not provide sufficient statistical evidence to facilitate the detection of terms. This phenomenon, whereby infrequent but vital terms are rejected or missed, known

[1] https://en.wikipedia.org/wiki/Content_analysis.

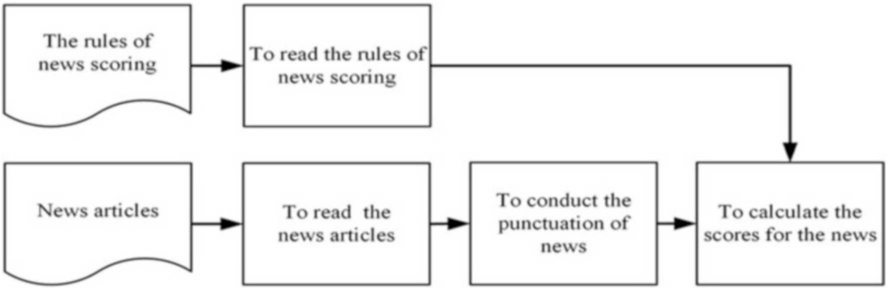


**Fig. 6.1** News scoring process

as silence, affects the recall of term extraction. Most extraction techniques fail to identify long phrases or expressions. Furthermore, ambiguity arises when a domain-specific corpus is expressed in a terse language. Another common type of linguistic incoherence is the omission of punctuation symbols, such as periods (".") to indicate the end of sentences. As mentioned previously, the lexicon is pivotal and must include comprehensive, domain-specific, and multiword forms, enabling expanded research on financial news (Fig. 6.1).

## 6.3.3 *Methodology*

### 6.3.3.1 Data Collection

Our news data source contains a total of 129,161 news items from a news database.[2] For classifying each news item on the basis of its content, we choose some news seed words that relate to the macroeconomics, politics, and investment. For each news item in the corpus of news from the news database,[3] we calculate the optimal group for the news item using category seed words and a vector space model (Manning et al. 2008, pp. 2–4). Each news article is placed into the category with the least distance between the seed words and the news text. Concurrently, we calculate the relevant scores for each news article based on similarity with news category seed words. The higher an article's scores, the more relevant that article is to the category (Fig. 6.2).

### 6.3.3.2 Research Design

We use content analysis techniques to extract information from the news data. Initially, we select 1535 news articles from the news database.[4] We separate those news articles into three subsets as follows: 1092 news articles related to

[2]Knowledge Management Winner, http://kmw.chinatimes.com/.
[3]Knowledge Management Winner, http://kmw.chinatimes.com/.
[4]Knowledge Management Winner, http://kmw.chinatimes.com/.

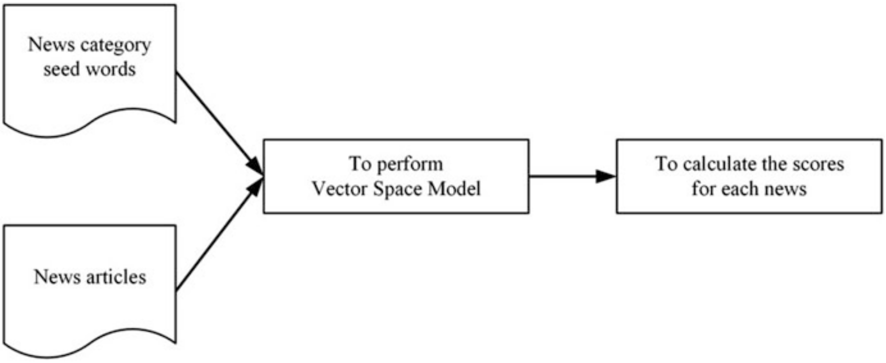


**Fig. 6.2** Calculating relevant scores for each news item

macroeconomics, 129 news articles related to politics, and 314 news articles related to investment. Furthermore, we select the most relevant 130 news articles from these three types of news; we apply a computer program to read the news artificially, and use our dictionary to count the positive words and negative words objectively. The system develops various customized word lists and creates scoring rules by dividing the scale into ten deciles: five negative and five positives. The score interval for a news article is configured between the most negative decile and the most positive decile. We filter news through the three basic news categories at first and then analyze the features of each news category to build the sentiment lexicon.

Finally, the system applies the rules to compute news scores automatically for each news article. The application automatically reads both news and scoring rules from the data source. Thus, through content analysis, we analyze the content of sentences according to the scoring rules and score each news article.

### *6.3.4 Research Results*

By analyzing the news data through sentiment analysis, we build the scoring rules and construct an automatic news-scoring application. Following (Liu 2012, pp. 1–167), we use a numeric score expressing the strength (i.e., intensity). The maximum score is $+5$ and the minimum score is $-5$; news content indicating a preferable situation receives a positive score, whereas news content indicating an undesirable situation receives a negative score.

For example, if the news refers to positive GDP growth, the macroeconomic news score is $+2$; if the news mentions a plunge of stock prices, the macroeconomic news score is $-5$. If a political news article refers to the Sunflower Student Movement of Taiwan, the political score for that article is $-4$; by contrast, an article that mentions economic integration is assigned a score of $+1$. A news item that mentions the reappearance of the Golden Cross is scored at $+5$ on the investment scale, whereas an article about a plunge of stocks at the close of a trading session receives a −5 investment news score. The tone of a news item is the affect or emotional feeling that the news item communicates regarding macroeconomic news, political news, and investment news.

The score interval is divided into ten segments, five of which are below zero and five above; the scoring of each news article depends on the amounts of positive and negative words.

### *6.3.5 Conclusion and Discussion*

This section focuses on the financial news sentiment analysis conducted by using software. We analyze the content of the news, develop various customized word lists, build a dictionary, and refine scoring rules that evaluate news items in the contexts of various research topics. These rules are applied by an automatic news-scoring application. Social media influences human behavior and causes fluctuations in the market. The price changes in financial instruments (such as stocks, bonds, or mutual funds) are consequences of actions taken by investors, reflecting their perceptions of events commented on in social media. By use of our lexicon dataset and our scoring rules, we can investigate whether the emotional messages in the social media context correlate with the price changes of financial instruments, and make it possible to forecast stock price fluctuations in advance. The empirical result is statistically significant but the result may be affected by the subjective judgment of readers. Our dictionary captures positive words and negative words, but it has limited grasp of what the words in news articles convey. Consequently, future research can expand the quantity of data that is considered and the range of years from which news is sampled to gain additional insights.

## 6.4 Development of Mobile Banking Service Innovation and Its Effects on Securities Firms

### *6.4.1 Introduction*

Banks increasingly struggle to enhance their mobile banking capabilities as they confront growing pressure from customers, emerging technological innovations, and growing competition from new market entrants (KPMG 2015a, pp. 1). Recently, new types of financial technologies (abbreviated as FinTechs) have been receiving avid attention from all over the word (Chuen and Teo 2015, pp. 24–37). The concept of FinTechs emphasizes firms' usage of new technologies to provide new financial services in a more effective and efficient way. Banks and other financial services companies attach great importance to momentous innovations from FinTechs and attempt to comprehend threats and find responses to some fundamental questions (Gulamhuseinwala et al. 2015, pp. 16–23).

The study mainly focuses on two problems. First, we investigate the effects of FinTechs on both securities firms and information and communication technology (ICT) development firms in Taiwan. Second, we address how these securities firms and ICT development firms endeavor to innovate and change their business models to adapt to the new age of FinTechs.

## *6.4.2 Literature Review*

### 6.4.2.1 Global Development Trend of Mobile Banking

In the East, giants in the internet industry, such as Alibaba and Tencent, are rising to become providers of banking services with branchless banks such as Ant Financial and WeBank. These communication technologies not only enhance the financial services sector but also provide wider access to banking and financial services (Chuen and Teo 2015, pp. 24–37).

KPMG (2015b, pp. 1–10) indicates that in the early twenty-first century, smartphones with WAP support enable the use of the mobile web. Various mobile devices and tablets can access bank websites and services. After 2010, mobile banking apps capitalize on the burgeoning success of the iPhone and the speedy growth of Android smartphones. Bank customers are directed to mobile-based websites or apps. After initially offering basic portfolio of banking through mobile, mobile banking has evolved from basic service to include a broad, rich set of capabilities.

### 6.4.2.2 FinTech Development and Strategy from a Global Perspective

FinTech firms combine innovative business models and technologies to enable, enhance, and disrupt financial services (Gulamhuseinwala et al. 2015, pp. 16–23). FinTechs also involve innovative financial services or products delivered through new technology. Consumer expectations alter with advances in technology (particularly advances in mobile and Internet technologies). Customers expect benefits from widespread technologies that have been globally adopted (Chuen and Teo 2015, pp. 24–37). FinTechs must meet customers' expectations efficiently; customers must find FinTechs easy to use.

Even though FinTech development is in an early stage, successful FinTech development is obviously difficult in the current situation of extreme competition. Financial service firms face challenges from competitors from both inside and outside of the securities industry (Chuen and Teo 2015, pp. 24–37). Regarding new entrants to the FinTech market, these start-ups appear to have begun to "unbundle" banking services and carve out business in some of the established banks' most profitable business lines (KPMG 2015c, pp. 1–33). The Internet Finance Guidelines indicate that China is creating both a financial market infrastructure and a regulatory framework that is specific to FinTechs. In fact, FinTechs-related services are booming in China with numerous peer-to-peer lending providers (Arner and Barberis 2015, pp. 78–91). From the global perspective, most young banking customers appear to be willing to use new mobile banking services; the average age of mobile banking users is in the mid-1930s (KPMG 2015a, pp. 1–8; KPMG 2015b, pp. 1–10).

Some traditional companies are beginning either to cooperate with partners or to implement outright acquisitions to respond to FinTechs. The money transfers and payment services provided by FinTechs are also an essential part of the customer journey for numerous popular e-commerce sites, which are designed to eliminate conflicts and enhance conversion rates at the purchase stage (Gulamhuseinwala et al. 2015, pp. 16–23).

Taiwan's 2016 Financial Supervisory Commission (FSC) "FinTechs Development Strategy White Paper" (FSC 2016, pp. 109) recommends the Taiwanese securities industry to pursue goals such as raising online order rates to 70%, improving automated trading mechanisms (such as robo-advisors and online fund sales platforms), strengthening cloud services for securities and futures, and deepening Big Data applications.

## *6.4.3 Methodology*

### 6.4.3.1 Research Framework

The study investigates how FinTechs affect both securities firms and ICT development firms and how these firms change in response to the maturation of FinTechs. Thus, a conceptual framework for this study is proposed as follows (Fig. 6.3).

### 6.4.3.2 Research Approach and Research Subjects

The development of FinTechs is still in its infancy. Additionally, very few studies in this area have been completed. Thus, the use of a case study approach for obtaining a clear understanding is appropriate (Yin 1994).

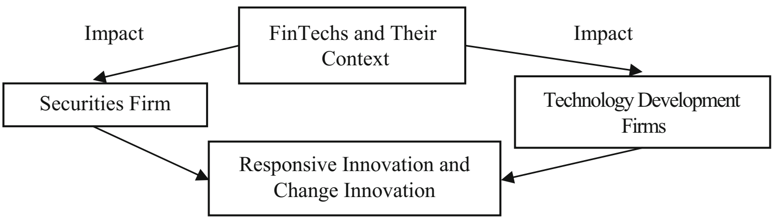


**Fig. 6.3** Conceptual framework for the study

We eventually chose one top securities firm, JihSun Securities, and two ICT development firms, Mitake Corporation and SysJust Corporation, as the major companies to be investigated. The latter two firms occupy more than 95% of the market for special ICT and financial service software in Taiwan. We select nine managers from these three firms for our in-depth interviews.

### 6.4.4 Research Results

#### 6.4.4.1 Effects of FinTech

The interviewee from JihSun mentions that FinTechs have numerous effects on the JihSun company. For example, they may face competition from new entrants that might be considered outsiders, such as Alibaba. Additionally, a securities firm faces more pressure from customers who are equipped with advanced handheld devices and who request prompt feedback regarding securities. Furthermore, as the consumers and marketing actions become more individualized, the firm needs to find an acceptable way to identify particular needs. Finally, the interviewee pointed out that government policies are also crucial influencing factors; financial services firms ask for help from government because FinTechs intensify global competition. The interviewees from the ICT development firms indicated their originally dominated markets appear to be affected because financial services companies plan to build internal innovative capabilities to deal with the threats and opportunities from FinTechs. Additionally, they must update their relevant technological capabilities, such as online transaction platforms and Big Data, which are changing much faster than before. The major effects of FinTechs on both securities firms and ICT development firms are listed in Table 6.3.

**Table 6.3** Effects of FinTech

| | Impacts |
|---|---|
| FinTechs and its context | *Securities firm*<br>• More competitions from new entrants and global competitors<br>• Deregulation in areas of online account, online payment, online order, personal information protection, prediction risk notice, forbiddance in investment suggestions, etc.<br>• Provision of new business development fund and consulting services from government to support firms to develop new services<br>• Increased customers' bargaining power<br>*ICT development firms*<br>• Gradually losing customers from and markets of financial service sector<br>• Facing new and broader technologies, such as online transaction platform, APP technology, robo-advisor, Big Data, and cloud service |

#### 6.4.4.2 Responsive Innovations and Changes

Facing the challenges from FinTech, both financial holdings and ICT companies are innovating technologies, services, and products supporting their digital strategy. The interviewee from the securities firm mentioned that his company is putting more resources into the new areas of robo-advisor Big Data, apps, and third-part payments. He further indicates that they must analyze customers' behaviors more quickly and accurately by using these new technologies to provide superior services. To reach the aforementioned goals, the company has tried hard to establish internal innovative capabilities. For instance, they hire some new employees with particular expertise in new areas. The securities firm also puts great effort into training its employees. They attempt to cooperate with several universities either to conceive new service ideas or to develop new technologies in an effective and efficient way. However, the ICT development companies seem to realize the changes experienced by their customers. They must upgrade their technological capabilities as FinTechs generate some new technologies, which are quite different from the traditional "information" technologies. Furthermore, they must reposition their products and markets as both technologies and customers undergo tremendous changes. The detailed results from the case studies are listed in Table 6.4.

**Table 6.4** Responsive innovations and changes

| Securities firm | ICT development firms |
|---|---|
| *New product and service innovation*<br>**New products**<br>• Currency trading, funds<br>**New services**<br>• Online orders, TSM service (for credit card), VIP service fee discount, or investment services seminar, openness of information about investment targets; e.g., managed funds-push notifications, personalized advice recommend, third-part payment<br>*Technology and equipment resources*<br>• Investment on software, such as APP, Big Data analysis, and robo-advisor system<br>• Investment on ICT infrastructure<br>*Human resource*<br>• Hiring and training the human talents with skills in APP technology<br>*External cooperation for innovation*<br>• Cooperating with universities | *New product and service innovation*<br>**New products**<br>• Security mobile order platform, database, cloud testing centers, third-part payment system<br>**Service-mobile security**<br>• Message order, e.g., LINE, voice order, mobile banking order, web banking order; APP order, e.g., Mitake Estock, cloud service, analysis of big data; e.g., customer behavior, investment, customized service<br>*Technology and equipment resources*<br>• Investment on ICT infrastructure, IT resource for APP<br>*Human resource*<br>• Hiring and training human talents in APP technology and Big Data.<br>*External cooperation for innovation*<br>• Cooperate with securities firms to design the online platform |

### *6.4.5 Conclusion and Discussion*

This study investigates how securities firms and ICT development firms respond to the threats and opportunities of coming era of FinTechs. From the results of preliminary qualitative case studies, several conclusions can be drawn: (1) Securities firms may need to rethink the way they divine customers' needs and interact with customers because those consumers are equipped with much more advanced and fast-changing handheld devices than earlier consumers had. (2) Securities firms may need to consider building their own in-house innovative capabilities in order to be capable of quickly responding to customers' needs and to social and economic changes. (3) The relationship between securities firms and ICT development firms may be changing as the securities firms appear to intend to develop mobile systems and service innovations by themselves. (4) The ICT development firms may need to reposition their products and services for the financial services industry and may need to update their ICT technologies.

## 6.5 Real-Time Financial Service Framework

### *6.5.1 Introduction: Creative Financial Services*

The advances of cloud computing and mobile devices have changed human lifestyles in numerous respects. We propose a framework that can handle complex data structures and large amounts of data in a short time using cloud computing. To provide ubiquitous mobile financial services, we develop an adaptive task assignment algorithm on our framework to overcome the instability of wireless mobile networks so that the financial system can deliver essential messages to a large number of customers instantly.

### *6.5.2 Methodology and Research Results*

#### 6.5.2.1 Cloud Computing and Big Data

Cloud Computing and Big Data Platform Comparison

There are numerous Big Data platforms based on the MapReduce model (Dean and Ghemawat 2008, pp. 107–113), such as Hadoop, Disco, and Spark. Although Spark is a heavyweight framework, Spark is a practical choice to manipulate Big Data because it is known for in-memory computing. In this chapter, we propose a real-time framework for solving stock market prediction algorithms, as shown in Fig. 6.4.

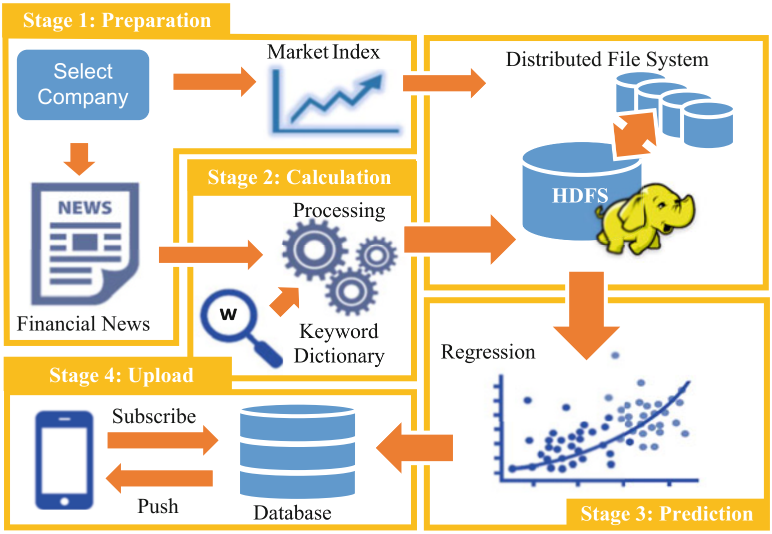


**Fig. 6.4** Four stages of stock market prediction algorithms

**Stage 1: Preparation** To make a precise prediction, we collect data from the Internet through the web crawler. The data includes news, volumes, stock quotes, price–book ratios (PBRs), price–earnings ratios (PERs), and market indexes.

**Stage 2: Calculation** The key point of the model is news-processing algorithms. First, we filter the news by topics appropriate to finance. Second, we segment compound Chinese sentences into Chinese words. Third, we search through the keywords in a news file. Keywords are divided into two categories: positive words and negative words. According to the keywords in the content, we score each news item.

**Stage 3: Prediction** A regression analysis is used for our stock market prediction model. In our case, we try to discover the relationship between news and stock trends.

**Stage 4: Upload** The prediction results for stock trends are saved in our database. We build a back-end web service to send this information to users instantly.

### Simulation Result of Platforms Comparison

In the simulation, we provide results using a Python language model for normal Python, Disco, and Spark. We focus on the calculation step, because the other steps only take few seconds to complete processing. In our MapReduce strategy, each news item is processed line by line and in one job.

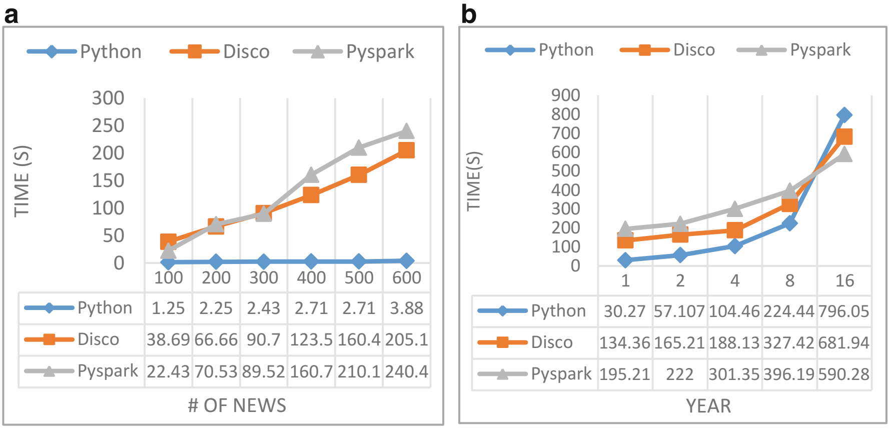


**Fig. 6.5** (**a**) Calculation time for real cases (600 news/MB). (**b**) Calculation time for Big Data sets (60,000 news/year)

Figure 6.5a shows the real calculation times of news-processing steps on different platforms. In our experiment, we found that we collected only 600 news items per day performed by the financial news web crawler. On the Python platform, it takes less than 5 s to finish this step; however, on the Disco and Spark platforms, it takes more than 200 s to finish the job. First, the heavy-duty Big Data platforms (Spark and Disco) require long calculation times because our news datasets are not big enough to fit the MapReduce model. A typical new item has only 30 lines and produces a small dataset that performs badly in the MapReduce model.

Second, launching Spark and Disco applications incurs several seconds of start-overhead. As mentioned earlier, each new job is considered as a single task, and when the datasets are small, numerous instances of start-overhead add up into a major performance issue.

To determine the effects of expanding the news dataset in our model, we combine an entire year's worth of news into one news file. In Fig. 6.5b, the Big Data platform (Spark and Disco) curves show slow growth, and the Python curve shows a linear relationship. However, Python shows positive exponential growth when the resource limit is reached.

According to the simulation results, it is crucial for researchers to choose an appropriate data processing platform. For example, if the dataset is too small to be computed, using regular high-end programming languages is the optimal choice. There are various practical applications that fit the MapReduce models. MapReduce is a flexible, highly fault-tolerant, and distributed processing framework, which can process massive data efficiently. We compare the performance between Pyspark and Disco. The module of Pyspark uses a Java Virtual Machine (JVM), and the module of Disco uses Erlang. If you want to implement your project on Big Data platform, you must consider the sizes of datasets and their effects upon start-overhead. In our case, the start-overhead of Pyspark is higher than Disco.

#### 6.5.2.2 Mobile Messaging System to Overcome Bursty Traffic and Network Instability

To deliver numerous instant messages to customers, we develop an adaptive task assignment algorithm on our service framework to overcome bursty traffic and mobile network instability.

System Architecture

The delivery service is built on the top of a proven and widely used open source XMPP project, Openfire, with a Connection Manager (CM) Module that can be horizontally scaled to meet increasing demands, as shown in Fig. 6.6.

System Behavior and Bottleneck

After the Load Balancer selects a CM to build the keep-alive connection with an incoming user, the messages sent to that user go through the same connection path until that user goes offline. Because each CM queue is a single FIFO queue and the mobile network is unstable, when bursty traffic happens, performance is greatly affected by how the Load Balancer selects the CM to build connections with users. Thus, our goal is to predict the network delay for an adaptive task assignment algorithm to reduce the average response time when the system handles bursty traffic in an unstable network environment.

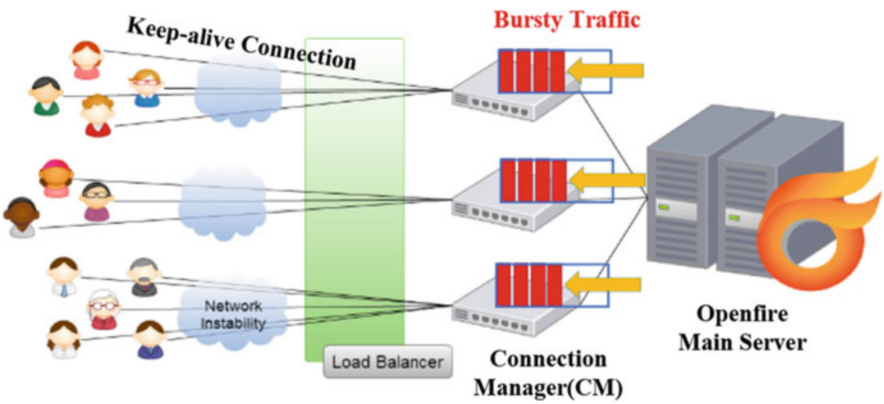


**Fig. 6.6** System architecture

Adaptive Task Assignment Algorithm

We predict future network delays through short-term historical data by using Queueing Theory (Harchol-Balter 2013) to model our system and to predict the system performance. The task assignment algorithm that we propose can predict future network delays from historical data; it is named SeeFuND. Consider that there is an interval of time between the algorithm processing time and the actual arrival time. In the interval of the same $\boldsymbol{CM_k}$ queue, there is an average probability of 0.5 that messages for other existing users might cut-in in front of the incoming message. We denote that probability as $\boldsymbol{q_{cut\_}in^k}$. Given this phenomenon, a correction term is added in the equation. Therefore, the waiting time $\boldsymbol{W^k_{new}}$ is predicted as follows: $\boldsymbol{W^k_{new}} = \left(\boldsymbol{q^k_{now}} + \boldsymbol{q_{cut\_}in^k}\right) \times \overline{\boldsymbol{S}}_{\boldsymbol{k}} + \boldsymbol{W^k_0}$ where $\boldsymbol{q_{cut\_}in^k} = \frac{\boldsymbol{user^k_{now}}}{\boldsymbol{2}}$ and $\boldsymbol{W^k_0}$ is the remaining service time.

The SeeFuND task assignment algorithm calculates the expected waiting time $\boldsymbol{W^k_{new}}$ of each CM queue, then the queue with the shortest expected waiting time is selected to connect to incoming user.

Emulation Results

We compare the performance of different task assignment policies like Round-Robin (RR) (Xu and Huang 2009), Random (Buot 2006, pp. 395–396), Shortest Queue First (SQF) (Teo and Ayani 2001, pp. 185–195), and Least User (Connection) First (LUF) (Choi et al. 2010, pp. 127–134). Virtual servers are provided by Digital Ocean's IaaS; user robots are deployed evenly in different datacenter regions in several countries. Each host is run on an Ubuntu14.04 Linux server with 1 GB RAM, a single-core processor, and 30 GB of disk storage. Other factors in the emulating environment are the same as before.

In Fig. 6.7, the result shows the response times of different task assignment algorithms for different intensities of bursty traffic. The nonbursty traffic state is such that the arrival of messages to each queue follows a Poisson process. The arrival rate of low-intensity bursty traffic is 40 messages/s per queue, and the arrival rate of high-intensity bursty traffic is 150 messages/s per queue.

We found that when the intensity of bursty traffic increases from 40 messages/s per queue to 150 messages/s per queue, the SeeFuND algorithm shows a minimally increasing response time (Seng et al. 2017).

### 6.5.3 Implementation: Mobile Application

For combining mobile devices and cloud computing, we develop an Android application to demonstrate our research. Users can connect to our cloud server and

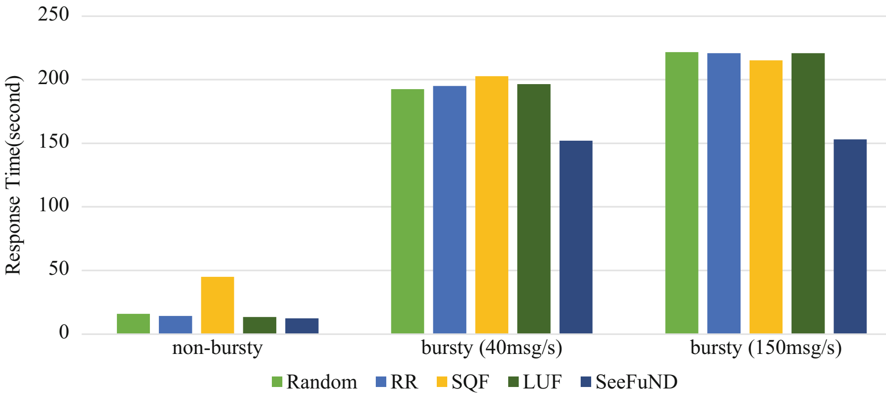


**Fig. 6.7** Response times of different task assignment algorithms for different intensities of bursty traffic

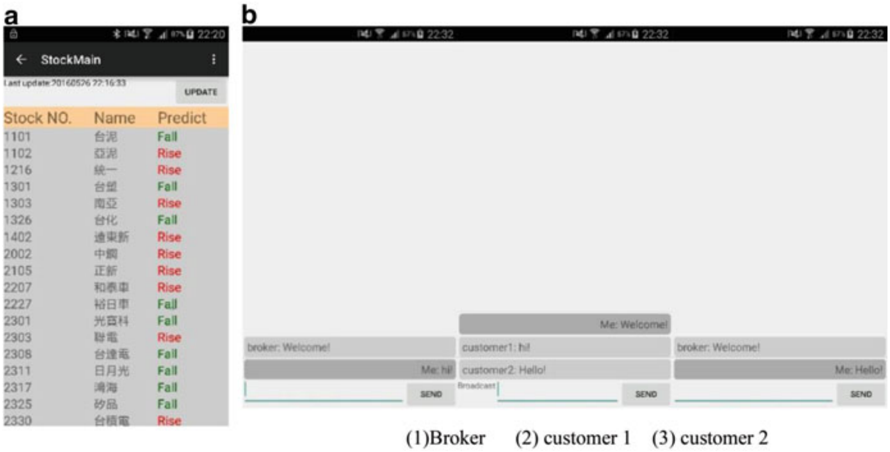


**Fig. 6.8** (**a**) Application for forecasting results. (**b**) Application for chat function which displays broadcast and unicast communications

get the service by this application. We display our forecast result, which predicts Taiwan's top 50 stocks, in Fig. 6.8a.

We provide the latest formal corporate announcements on a real-time platform named Market Observation Post System. The second part is a chat function which lets users can communicate with others directly without switching applications, as shown in Fig. 6.8b.

## 6.6 Concluding Remarks

Big Data, Artificial Intelligence (AI), Internet of Things (IoT), and Cloud Computing's widespread popularity and ubiquity have revolutionized the face of business in the world, especially for the financial service industry, such as FinTech (financial technology). With the emergence of service innovation and innovation integration, the focus has shifted to the development of novel financial services with data analytics, business intelligence, mobile technologies, and cloud computing that diversify and customize the new functionalities available to business clients and customers. Most of the worldwide financial holdings corporations have faced the biggest challenges resulting from the dynamic business environments, diverse telecommunication settings, and heterogeneous data sources that must work together in an integrated, real-time, and seamless manner. Today, businesses have accumulated large numbers of online services and data sources that run and reside on a variety of environments.

Furthermore, they have heterogeneous business models, business processes, and workflows to interact with. The results have been upheaval, chaotic, and disruptive. Therefore, it is crucial to develop a comprehensive, cross-disciplinary, and in-depth resolution of services, methods, and technologies to address the above issues. In this chapter, a set of cross-disciplinary sections are written up and devoted to describe and discuss the data analytics, service innovation, mobile technology in the financial services. We consider the next generation of innovation and integration of financial services in the financial service industry. We propose to investigate the key research issues of the innovation, intelligence, integration of new financial services over mobile technologies, and cloud computing information systems. We conduct extensive literature reviews to develop novel research models, to collect domain databases, to perform live case studies and experiments, and to create prototyping systems and evaluation. Empirical and econometric studies, data and text mining algorithms, selection and interaction models, real life case studies, mobile prototyping, and performance evaluation are performed.

## References

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance, 59*(3), 1259–1294.

Arner, D. W., & Barberis, J. (2015). FinTech in China: From the shadows? *The Journal of Financial Perspectives, 3*, 78–91.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance, 61*(4), 1645–1680.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business, 78*(2), 405–440.

Buot, M. (2006). Probability and computing: Randomized algorithms and probabilistic analysis. *Journal of the American Statistical Association, 101*(473), 395–396.

Choi, D., Chung, K. S., & Shon, J. (2010). An improvement on the weighted least-connection scheduling algorithm for load balancing in web cluster systems. In *Grid and distributed computing, control and automation* (pp. 127–134).

Chuen, D. L. K., & Teo, E. G. S. (2015). Emergence of fintech and the LASIC principles. *The Journal of Financial Perspectives, 3*, 24–37.

Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM, 51*(1), 107–113.

Deloitte. (2016). *Perspectives: Banking and securities outlook 2017*. http://www2.deloitte.com/us/en/pages/financial-services/articles/banking-industry-outlook.html

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance, 47*(2), 427–465.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1–22.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance, 64*(5), 2023–2052.

Financial Supervisory Commission. (2016). *Fintech-development strategy white paper*. Taipei: Financial supervisory commission.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance, 68*(3), 1267–1300.

Gulamhuseinwala, I., Bull, T., & Lewis, S. (2015). FinTech is gaining traction and young, high-income users are the early adopters. *The Journal of Financial Perspectives, 3*, 16–23.

Harchol-Balter, M. (2013). *Performance modeling and design of computer systems: Queueing theory in action*. Cambridge: Cambridge University Press.

Hu, N., Bose, I., Koh, N. S., & Liu, L. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems, 52*(3), 674–684.

Ittoo, A., & Bouma, G. (2013). Term extraction from sparse, ungrammatical domain-specific documents. *Expert Systems with Applications, 40*(7), 2530–2540.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis, 33*, 171–185.

KPMG. (2015a). *Mobile banking is a key selling point for a growing number of customers* (pp. 1–8). Amstelveen: KPMG International Cooperative.

KPMG. (2015b). *Digital offerings in mobile banking—The new normal* (pp. 1–10). Amstelveen: KPMG International Cooperative.

KPMG. (2015c). *Mobile banking-global trends and their impaction banks* (pp. 1–33). Amstelveen: KPMG International Cooperative.

Kumar, A., & Lee, C. (2006). Retail investor sentiment and return comovements. *The Journal of Finance, 61*(5), 2451–2486.

Kurov, A. (2010). Investor sentiment and the stock market's reaction to monetary policy. *Journal of Banking & Finance, 34*(1), 139–149.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1–167.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35–65.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Scoring, term weighting and the vector space model. In *Introduction to information retrieval* (Vol. 100, pp. 2–4).

Seng, C., Wu, T., & Chang, Y. (2017). *Innovative, integrated, mobile financial services* (MOST three-year final research project report, MOST 104-2627-E-004-001).

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442.

Teo, Y. M., & Ayani, R. (2001). Comparison of load balancing strategies on cluster-based web servers. *Simulation, 77*(5–6), 185–195.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance, 63*(3), 1437–1467.

World Economic Forum. (2015). *Global agenda council on the future of financing & capital: The future of FinTech a paradigm shift in small business finance*. Retrieved from http://www3.weforum.org/docs/IP/2015/FS/GAC15_The_Future_of_FinTech_Paradigm_Shift_Small_Business_Finance_report_2015.pdf

World Economic Forum. (2017). *Beyond Fintech: How the successes and failures of new entrants are reshaping the financial system*. Part of the Future of Financial Services series | Prepared in collaboration with Deloitte, August 2017. http://www3.weforum.org/docs/Beyond_Fintech_-_A_Pragmatic_Assessment_of_Disruptive_Potential_in_Financial_Services.pdf

Xu, K., Liao, S. S., Li, J., & Song, Y. (2011). Mining comparative opinions from customer reviews for competitive intelligence. *Decision Support Systems, 50*(4), 743–754.

Xu, Z., & Huang, R. (2009). *Performance study of load balancing algorithms in distributed web server systems* (CS213 Parallel and Distributed Processing Project Report, 1).

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Newbury Park: Sage Publications.

Zhang, Y., Dang, Y., & Chen, H. (2013). Research note: Examining gender emotional differences in web forum communication. *Decision Support Systems, 55*(3), 851–860.