

Market Research Analytic System using the Public Data and Social Network Services

Yong-Jang Cho, Yujin Song, Mucheel Kim, Ka Lok Man, and Seungmin Rho

Abstract—As the Korean government has provided a lot of support to SMEs and start-up companies, the number of people who support them has increased, but the failure rate is still high. The main factor of failure is that the market offers products that are not desired. In this paper, we propose a regional market research analytics system using public data and social network services. The proposed system is divided into recommendation, region, and job service, and the analyzed data is visualized as a table or a map.

Index Terms—Recommendation System, Market Research Analytics, public data, Social Network Services

I. INTRODUCTION

MANY people are trying to start their businesses in various forms, from small shops that can start with small capital to realizing goals of "self-employment", to agencies, chain stores, and start-ups of large-sized SMEs (Small and medium-sized enterprises) [1].

The government is making efforts to succeed in founding such as education for start-up, policy funds, and statistical data [2]. However, many people are failing due to various reasons such as lack of funds, cost issues, and so on (Figure 1). It is not easy to start a business according to the statistics released by the Korean Government (Small and Medium Business Administration) [3]. The survival rate for start-up companies is on average 50%. In particular, the survival rate was the lowest in 19.5% of patients under 30's of age.

In this study, we intend to build a market research system that can extract the needs of users and provide a supply system for them. Then we might analyze the various public data provided by the government and public institutions, collect and analyze the blog information that users have posted opinions about the products. Furthermore we have constructed an analysis system that enables the founders to easily conduct market research on the customized entrepreneurial items.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1A09919551)

Yong-Jang Cho is with Divisions of Mediasoftware, Sungkyul University, Anyang, Korea (e-mail: dydwkd48670@gmail.com).

Yujin Song is with Divisions of Mediasoftware, Sungkyul University, Anyang, Korea (e-mail: uzini_@naver.com).

Ka Lok Man is with Department of Computer Science and Software Engineering Xi'an Jiaotong-Liverpool University, Suzhou, China (e-mail: ka.man@xjtlu.edu.cn).

Mucheel Kim is with Department of Computer & Software Engineering, Wonkwang University, Korea (email : mucheel.kim@gmail.com).

Seungmin Rho is with Divisions of Mediasoftware, Sungkyul University, Anyang, Korea (corresponding author to provide phone: 031-467-8348; e-mail: smrho@sungkyul.edu)..

II. THE PROPOSED MARKET RESEARCH ANALYTIC SYSTEM

In this chapter, we have developed a recommendation service for entrepreneurial information by using public data in Gyeonggi province, Korea. It is divided into regional and occupational services.

A. Recommendation Service

The recommendation service for entrepreneurial information crawls the latest information of social network services including blogs, and analyzes the occurrence frequencies of the words in the information. This can be used to visualize and display entrepreneurial information tailored to the current trend [4-5].

In order to gather data necessary for this research, we crawled various social network service information and blog data. The collection tool used to parse and crawl web data using the jsoup library provided by Java. It is a Java HTML parser that parses the necessary parts of HTML on the web through conditional statements.

Data collected through blogs is hard to understand and analyze by computer as informal data. To support this, morphological analysis is performed. The execution of the morpheme analysis was performed using Komoran which is Korean morpheme analyzer API. In the results of the morphological analysis, the relationship between terms was analyzed based on the occurrence frequency of terms.

For this purpose, the morphologically analyzed data was processed by writing a counter algorithm to analyze the occurrence frequency between words, and analyzed by dividing them by region, Gyeonggi Province, Korea, and rate of change. As shown in (Figure 3), by using the proposed method, not only the frequency of occurrence of the keyword but also the rate of change can be analyzed to understand how the keyword changes every week.

B. Occupation, local service

In this section, public data related to small businesses supported by Gyeonggi province were used to provide information on entrepreneurship for job and regional services. The characteristics of the public data were analyzed using the foodservice business, school and teaching school, professional and university status, bus stop information by region, population by age and sex. In addition, for visualization of information with map APIs which is provided google.

The collected public data is stored in a database, and the loaded data is analyzed according to the conditions of the region and industry to provide analysis information on the map. In addition, it provides a function that can recommend

the locations for start-up recommendation. The recommendation algorithm uses the distance relation with the institute, school, apartment, etc. in the surrounding facilities to grasp the floating population of the surrounding commercial area. In order to solve the disadvantage that the simple distance relation can be located in mountains and rivers which are not suitable for the business start position, it is analyzed by including the position information of the bus stop together, and the position.

The recommendation algorithm is provided along with statistical information on the map, which is obtained by analyzing the occupation and the area. Recommendation services provide analyzed statistical information including personnel, age, competitors, and public facilities.

III. CONCLUSIONS

As interest in entrepreneurship has increased recently, many people consider entrepreneurship based on various types of items. In this study, we propose a business information recommendation service that can grasp current market situation for each business item and recommend

business item by job and region.

As a result of this research, it is necessary to secure diversity of information by analyzing user information of various social network services by expanding the result of this study, which is limited to blog data. In addition, based on the collected information, it is necessary to construct a multi-layered entrepreneur information service by analyzing trend information on entrepreneurial items.

REFERENCES

- [1] H.O. Jeong, Employment. The guidebook of entrepreneurship has diversified., The Korean Publishing Journal, Vol. 71, 1990, pp. 16-17.
- [2] Ministry of SMEs and Startups, Korea, "Survival rate of start-up companies", <http://www.mss.go.kr>
- [3] CB insights, "Startup Failure", <https://www.cbinsights.com>
- [4] Y. Kang, A Study on the Technique for Public Open Data Analysis, Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol. 5, No. 6, 2015, pp. 29-36.
- [5] Y. Kang, H. Kim, A Study on Geographic Visualization of Public Data Using Google API in Cloud Computing Environment, Journal of the Korean Cartographic Association, Vol 14, No. 1, 2014, pp.1-15

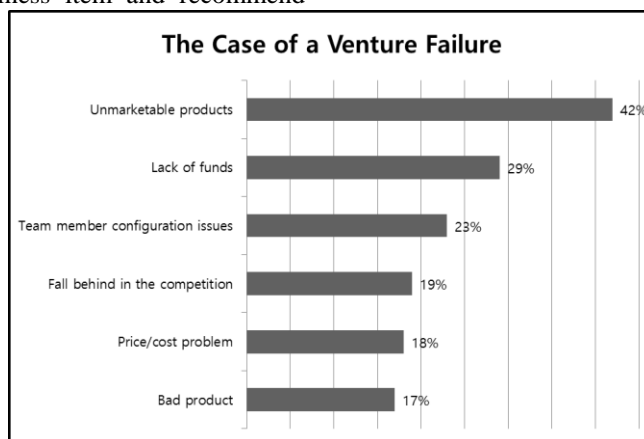


Fig. 1. Statistics of the case of a venture failure

※ 안양치과추천 20년 경력만큼 능숙한 진료 ※	안양
안양 박달동중국집 <칭니> 안양 중식은 바로 여기 추천	치과
안양마사지 : 석수동... 스와디차크라 해독관리 추천	추천
[안양 미용실 추천] 안양일본가 라브리지헤어살롱 매직의 신세계!!	년
본사대표 인기강사 직강 안양오피학원 추천	경력
안양가발 효도선물로 적극 추천~	진료
안양 평촌 맛집/안양 맛집 추천 "인기명 평촌점" 회식장소 추천 !!	안양
안양 카페 추천 카페우와	박달동
[안양돼지갈비맛집] '백년식당서서갈비' 안양맛집추천	중국집
남자파마 / 남자머리 / 남자머리추천 / 안양미용실 / 김가혜어.	안양
안양 냉면 맛집 여름엔 여기지!	중식
곰스603 안양IT밸리점 :: 안양수제버거, 안양맛집추천 !!	추천
안양교정치과추천 ▶알니부터 전체교정까지	안양
[안양데이트추천] 고양이카페 "고양이숲"	마사지
안양 아로마 마사지 추천 ~ ~	석수동
키웬 안양점 : 피부재생에 좋은 카복시스켈링 추천~!	해동

Fig. 2. Keyword Extraction for recommendations

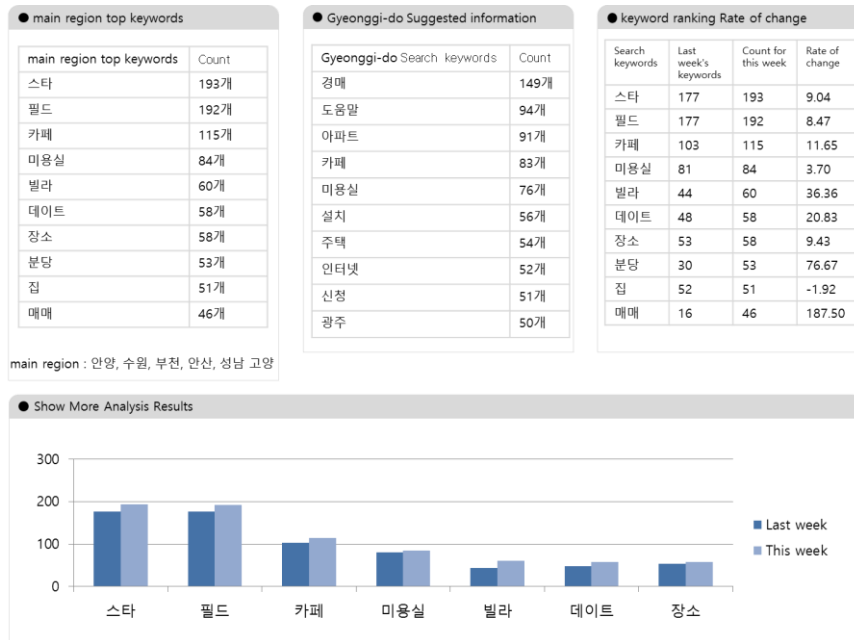


Fig. 3. The Statistics of Market Research Analysis

Result screen Professional information

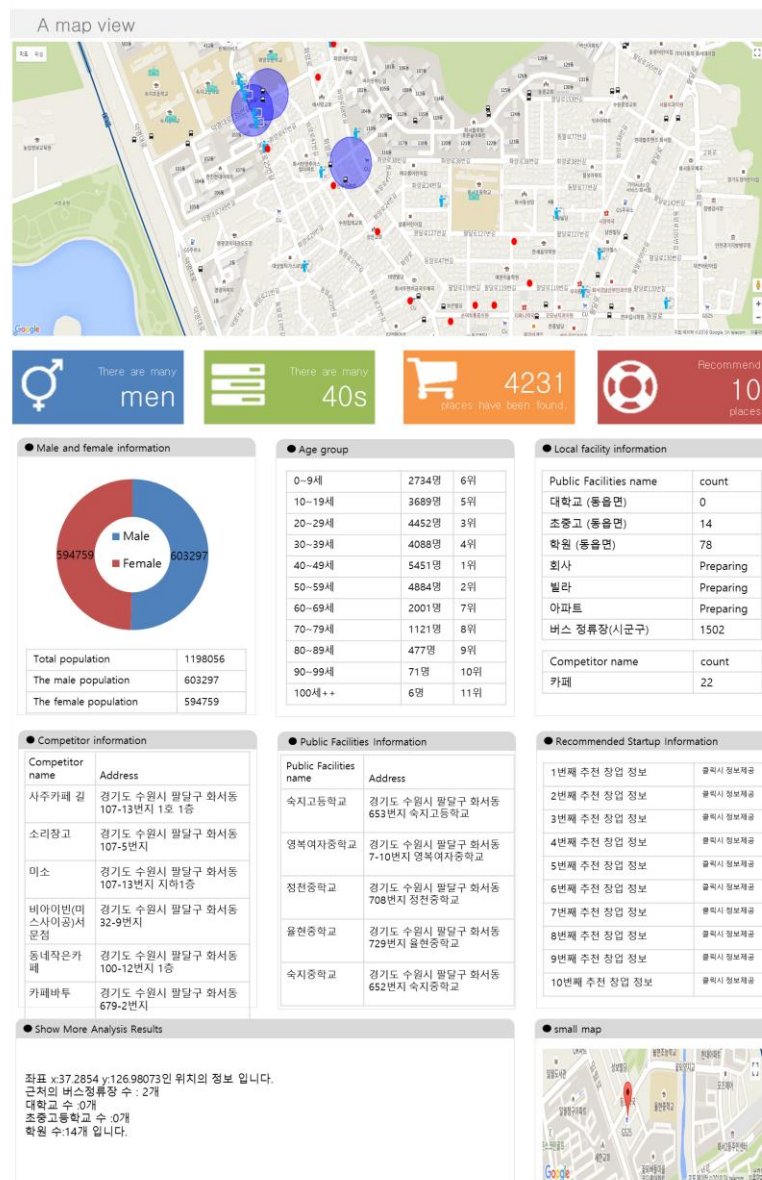


Fig. 4. The proposed Market Research Analytic System