Categorization of Mergers and Acquisitions in Japan Using Corporate Databases: A Fundamental Research for Prediction

B. Shao[1], K. Asatani[1], I. Sakata[1]
[1]Department of Technology Management for Innovation, School of Engineering, The University of Tokyo, Japan
(shaobohua6@gmail.com, info@ipr-ctr.t.u-tokyo.ac.jp)

Abstract – **Mergers and Acquisitions (M&A) are recognized important strategy for corporate growth. In practice, M&A business consumes much energy and M&A success rate is not high. Hence, scientific M&A recommendation research is needed under such condition. This paper, focusing on M&A categorization, is a fundamental research for M&A prediction and recommendation. In this paper, we used M&A data, financial data and corporate data for M&A analysis. Based on them, we designed 13 features and used K-means clustering to separate M&A cases. The 13 features are of acquirer features, target features and their relationship features. We grouped M&A cases into 5 clusters and found different characteristics in these 5 clusters. Results in this paper show that these features will be effective for future M&A prediction and recommendation.**

Keywords – **Artificial intelligence (AI), clustering, M&A, prediction, recommendation.**

I. INTRODUCTION

M&A are the transaction of the ownership of organizations and have coincided with the existence of companies for a long time. M&A gained popularity in the United States since late 19th century due to dramatic effects for companies. Historically, M&A in Japan showed countercyclical trend [1] due to unique situations in Japan. However, several reforms (known as "Financial Big Bang in Japan") from 1997 to 2006 have largely changed these situations. Recently, according to a report of Ministry of Economy, Trade and Industry (METI) on 2018's tax reform [2], the shareholders of purchased companies can defer taxes on revenue during M&A, which helps promote M&A in Japan.

Currently, M&A business is still conducted by human labor and scientific approaches are not prevailing for selecting suitable M&A partners. In order to promote economy, industrial transformation and social welfare, M&A recommendation system is on demand. This research is part of the research project "Research on the feasibility of the M&A target recommendation for practical use". The ultimate goal of this research project is to develop a win-win M&A recommendation system. This work focuses on M&A categorization, which serves as fundamental research work for the project.

Previous literature about M&A preferred social science approaches, such as case studies, empirical research and theory research. This research, by novel AI methods, deals with the rare M&A prediction and recommendation research topic. As far as we know, this is the first attempt in M&A research field. In addition, we provide a new M&A clustering method, which is especially suitable for M&A in Japan after the Financial Big Bang in Japan.

This paper is arranged as follows: Section II is Previous Literature, Section III is Data, Section IV is Methodology, Section V is Results and Discussion, and Section VI is Conclusion.

II. PREVIOUS LITERATURE

In order to comprehend the current M&A research trend, we conducted the M&A on previous literature by using an Academic Landscape system. The Academic Landscape system ([3], [4]) is a system for analyzing academic papers and their citation networks, developed by the Innovation Research Policy Center, the University of Tokyo. We inputted the query " "Merger* and Acquisition*" or "merger* and acquisition*" " as the topic for the Web of Science search engine, 3466 papers were fetched on May 3, 2018. Then we downloaded the bibliometric information of these papers for analysis and we summarized the results shown in Fig. 1. From Fig. 1, we found that finance, management and technology are the most popular topics in M&A research.

Several papers mentioned the degree of integration of M&A ([5], [6]). Cloodt *et al*. [7] investigated performance after M&A in high-tech industries and concluded that the integration of companies with moderately related knowledge bases will increase innovative performance during the first couple of post-M&A years. Bower [8] argued 5 possible M&A scenarios and they are the overcapacity M&A, the geographic roll-up M&A, the product and market extension M&A, the M&A as R&D and the industry convergence M&A. It is suggested that companies should learn from M&A through constant M&A activities and M&A experience contributes to successful M&A a lot [9]. In addition, in finance research, Dickson [10] argued that different cash flow patterns correspond to different stages in the firm life cycle.

This research is sponsored by UZABASE, Inc. and The University of Tokyo Edge Capital, Co., Ltd.

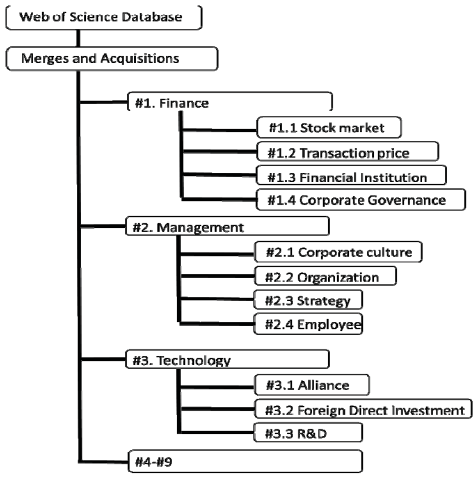

Fig. 1. Academic Landscape (M&A literature)

Mori *et al.* [11] used support vector machine (SVM) to predict reciprocal business partners. They designed supplier's features, customer's features and their relationship's features. Finally, they also developed a web system for predicting the business partners. Results from this method showed success with 77% *F*-value for finding reciprocal business partners.

III. DATA

In this paper, we used data from UZABASE, Inc. for analyzing M&A. It has collected 1.5 million M&A deals and information of 5 million companies. We used M&A, financial and company information in this research.

In M&A information, companies are divided into 4 roles. They are "acquirer", "ultimate acquirer", "seller" and "target". The difference between ultimate acquirers and acquirers is that ultimate acquirers are mainly Private Equity funds and venture capital firms whereas acquirers are ordinary companies. The difference between targets and sellers is that targets are organizations for sale whereas sellers are the owners or shareholders of targets. In this research, we focused only on "acquirer" and "target" because they are typical M&A roles.

M&A are grouped into 7 types in this data, they are "Merger of Equals", "Fund Buy-out", "Acquisition", "Minority Stake", "Joint Venture", "MBO" and "Demerger". Fig. 2 shows quantities of the 7 types of M&A in Japan in a time-series way. From it, we observe "Minority Stake", "Joint Venture" and "Acquisition" are the most M&A cases. In this research, we only focused on "Acquisition" because it is a typical M&A pattern.

In the financial database, we mainly used the Balance Sheet, the Profit and Loss Statement and the Statement of Cash Flow of every company across all industries. The yearly financial data can be fetched from 1989 to now and we analyzed the finance information within this timespan.

For M&A analysis, we also have access to a chart of 3692 Tokyo Stock Listed companies' information. In this research, we concentrated on M&A among these companies because they are typical Japanese companies.

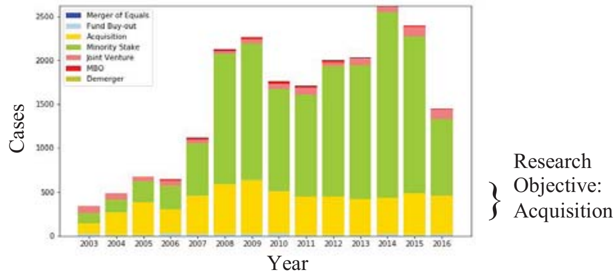

Fig. 2. Deal type of M&A in Japan (2003-2016)

In this chart, we have financial items such as net sales, net sales growth rate and cash and cash equivalent for analysis. We also have industrial classification data for companies, the SPEEDA codes. The labelling of the SPEEDA code is done by business analysts and a machine learning algorithm with accepted criteria. The SPEEDA code system consists of 9 digits (3 parts) in total. The first 3 digits show the general classification. The second 3 digits show the medium classification. The last 3 digits show the detailed classification.

IV. METHODOLOGY

A. AI methods

In order to classify M&A well in this research, we used the K-means clustering [12], the Principal Components Analysis (PCA) [12] and the t-SNE visualization [13]. Since M&A are very complex, this research only serves the first step of the whole project. The K-means clustering is widely used in the machine learning field and we used it for the first trial for analyzing and understanding M&A phenomenona. PCA is a typical method in statistics and we used PCA to find effective features. The t-SNE method is a well-known method for projecting high-dimension space onto 2-dimension surface. We used the t-SNE method to show the distribution of M&A cases.

B. Cash flow features

We designed the following 8 features, shown in TABLE I, for analyzing financial items in the financial database. We used the following easy and practical method, Equation (1), to define and calculate the free cash flow.

$$\textit{Free cash flow} = \textit{Cash flow from operating activities} + \textit{Cash flow from investing activities} \quad (1)$$

We made about 70,000 vectors from Japanese companies. Each vector represents yearly financial information of a company. A vector consists of the 8 features. By using PCA for all vectors, we concluded that net sales and cash reserves are important in financial items. The detailed results of this analysis are shown in section V A.

TABLE I
8 FRATURES

Name	Financial items
cfoa	Cash flow from operating activities
cfia	Cash flow from investing activities
cffa	Cash flow from financing activities
freecf	Free cash flow
plnetsales	Net sales
oprofit	Ordinary profit
cfb	Cash & Cash Equivalent - Beginning
cftd	Changes in cash flow

TABLE II
FIRM FEATURES

Name of features	Definition
$growth_{3y}^{*}$	Whether the firm *'s sales were growing or not for the period of three years in succession before the M&A announcement Values: 1: growing more than 2 years ; 0: others
$cashathand_{4y}^{*}$	Whether the firm *'s "cash at hand" was abundant (cash/sales >0.25) or not for the period of four years in succession before the M&A announcement Values: 1: 3 or 4 years' cash is abundant; 0: others
$industrygrowth_{3y}^{*}$	Whether sales of the whole industry (SPEEDA general classification of the firm *) were growing or not for the period of three years in succession before the M&A announcement Values: 1: growing more than 2 years ; 0: others
$scale^{*}$	Whether the firm *'s sales of 4 years before M&A announcement were above the industry sales median (SPEEDA general classification) Value: 1: acquirer's sales of more than 2 years were above the median; 0: others
$hist_{3y}^{*}$	Value: 1: the firm * had M&A experience in the period of three years in succession before the M&A announcement; 0: others
$hist_{bool}^{*}$	Value: 1: the firm * had M&A experience before the M&A announcement; 0: others

C. M&A features

We used K-means clustering for analyzing M&A and we designed features for acquirers, targets and their relationships. Based on the findings from cash flow, we used only net sales and cash reserves for designing financial M&A features.

In order to have a macro-indicator, the feature "industry growth" is introduced. The "industry growth" is the percentage change of "industry total sales" during two successive years. An "industry total sales" is the sum of all yearly net sales of Tokyo Stock Exchange listed companies labelled with the same general industrial classification code.

The values of M&A features are binary, i.e., 0 or 1. We adopted this practice for two reasons. One is that we did not know deeply about the current situation of M&A in Japan and this indirect data processing practice helped us have new findings in M&A. Another one is that we can reduce the bias from different data scales.

TABLE II and III shows the definition of features for firms and for the firm relationships respectively. The left column records names of features and the right column records definitions of features. The superscript of these features are determined by firm roles, such as "acquirer" or "target". TABLE IV shows all features for analysis according to their affiliations. The left column is the acquirer's features. The middle column is the target's features. The right column is the relationship features. Based on these features, we extracted information from the database and finally we made 109 intact M&A case vectors.

TABLE III
FIRM RELATIONSHIP FEATURES

Name of features	Definition
dis_3	If the acquirer and the target share the same general industrial classification (the first 3 digits in SPEEDA codes), the value is 1; otherwise: 0.
dis_6	If the acquirer and the target share the same medium industrial classification (the first 6 digits in SPEEDA codes), the value is 1; otherwise: 0.
$salesmean_{bool}$	If the mean of the acquirer's sales of 4 years in succession before M&A announcement was larger than that of the target, then the value is 1; otherwise 0.

TABLE IV
M&A FEATURES

Acquirer's features	Target's features	Relationship features
$growth_{3y}^{a}$	$growth_{3y}^{t}$	dis_3
$cashathand_{4y}^{a}$	$cashathand_{4y}^{t}$	dis_6
$industrygrowth_{3y}^{a}$	$industrygrowth_{3y}^{t}$	$salesmean_{bool}$
$scale^{a}$	$scale^{t}$	
$hist_{3y}^{a}$		
$hist_{bool}^{a}$		

[a] indicates an acquirer and [t] indicates a target

TABLE V
THE FIRST AND SECOND COMPONENTS BY PCA

Name	n1, the first principal component	n2, the second principal component
cfoa	0.069	0.124
cfia	-0.030	0.267
cffa	-0.005	-0.024
freecf	0.039	0.391
plnetsales	**0.980**	-0.164
oprofit	0.040	0.026
cfb	0.169	**0.770**
cftd	0.036	0.374
Explained variance ratio	69.04%	17.49%

D. Random Samples

In order to have further understanding of M&A, we used the 3792 listed companies again and generated random samples for comparison. By random combination, we generated about 7,000,000 pairs. Next, we randomly selected 1,000 pairs and made features according to TABLE II, III and IV and deleted any data flaw. Finally, we made 778 vectors.

V. RESULTS AND DISCUSSION

A. Cash flow analysis

The results of PCA to cash flow features are shown in TABLE V. The columns show coefficients (weights) of the 8 features in the first and second principal components respectively. The last row is about explained variance ratios of the two principal components. The two principle components totally explain 86.5% of the variance and we found that "plnetsales" and "cfb" have the highest coefficients in the two components respectively. This indicates that net sales and cash reserves effectively explain financial situations of companies.

B. M&A clustering

Firstly, we constructed a vector set of both M&A cases and random samples. Next, we used K-means to separate the vector set into 5 clusters. We adopted 5 clusters because this is an empirical and intuitive practice. We visualized the 5 clusters by the t-SNE method and the cluster numbers range from 0 to 4, as shown in Fig. 3. Since the t-SNE method projects dots of high-dimension space onto 2-dimension surface, locations of dots in 2-dimension surface are of no importance. Instead, relative distances between dots in t-SNE visualization map matter much [13]. Thus, values of horizontal and vertical axes in Fig. 3 are unimportant t-SNE parameters. From Fig. 3, we also found most dots are clearly separated. The feature values of the centroid of each cluster, shown in TABLE VI, are calculated by the average feature values of all vectors (dots) which belong to that cluster. The columns represent feature values of each cluster centroid. The last 5 rows are about M&A and random sample quantities.

From TABLE VI, we found several M&A trends in each cluster. In general, creditor-based M&A and countercyclical phenomenon become fewer in Japan. In the cluster 0, there are only 2 M&A cases and up to 157 random samples and this cluster has the fewest M&A cases. We further examined the 2 M&A cases and found that these two acquirers had abundant cash but no M&A experience, according to the values of "cashathand$_{4y}^{a}$" and "hist$_{bool}^{a}$" whereas targets were growing, according to "growth$_{3y}^{t}$". This phenomenon indicates managerialism [14] and investment principles in corporate finance. Hence, we assume and name this cluster as "abundant cash, having M&A for effectively using cash reserves". In the cluster 1, there are 18 M&A and 245 random samples. In this cluster, according to the value of "salesmean$_{bool}$", we found that scales of acquirers are relatively large and scales of targets are relatively small. We also scrutinized these M&A cases and assume that these acquirers were willing to purchase small companies around their business boundaries (according to values of "dis$_3$" and "dis$_6$") so as to extend their business boundaries or to integrate resources on supply chains and value chains [15]. In the cluster 2, there are 29 M&A and 134 random samples. The features "hist$_{bool}$" and "hist$_{3y}$" are all 1, meaning that acquirers in this cluster have abundant M&A experience. The features "dis$_3$" and "dis$_6$" are nearly 0, meaning acquirers were buying companies from different industries. Hence, we assume that these acquirers were making full use of their experience [9] and having multiple aims, such as growth, extension and R&D substitute. In the cluster 3, there are only 4 M&A cases and 210 random samples. We noticed that the value of "salesmean$_{bool}$" of this cluster is extremely low. This phenomenon indicates that small companies purchased large companies. In this cluster, M&A have special situations, for example, the acquirer holds over but nearly half of all shares of the target. In the cluster 4, we observed the most M&A cases. There are up to 56 M&A cases and only 32 random samples. Hence, we name this cluster "normal M&A cluster". We also indicate that M&A in this cluster are for industrial reorganization, according to values of "dis$_3$" and "dis$_6$". Industrial reorganization has two perspectives, one is to reduce overcapacity and ill-managed companies [8] and another is to achieve economies of scale and scope [16]. In the cluster 4, we manually found over half cases are consistent with the latter perspective. Intuitively, the determinants of either two reorganization scenarios are beyond our current 13 features and this point needs further study in the future research.

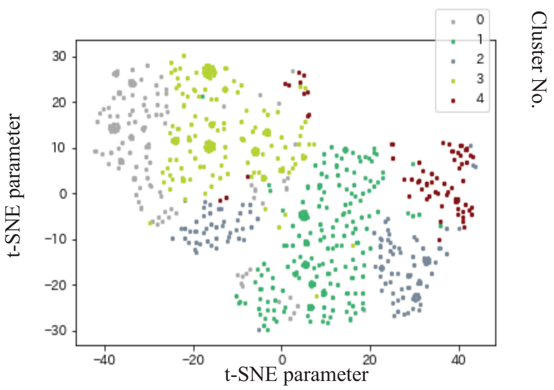

Fig. 3. t-SNE visualization of the M&A cases and random samples in 5 clusters

VI. CONCLUSION

This paper records a fundamental research part for the future win-win M&A recommendation system. We applied PCA to analyzing financial items and found that net sales and cash reserves are important features in financial items. We employed the K-means clustering to separate M&A cases into 5 clusters and provided our insights on these clusters.

The success of M&A categorization contributes to academia because there are many categorization standards already and our results support several previous theories. Our standard well fits M&A in Japan in the past decade.

As for M&A research methods, data mining and case studies are often used. In this research, we only employed data mining and we will request comments from M&A experts and conduct M&A case studies so as to improve the accuracy and precision in the future. We used indirect,

TABLE VI
CENTRIOD INFORMATION OF EACH CLUSTER

cluster_id	0	1	2	3	4
$growth_{3y}^{a}$	0.692^{***}	0.627^{*}	0.669	0.673	0.705^{**}
$cashathand_{4y}^{a}$	1.000^{**}	0.186	0.282^{***}	0.000^{*}	0.159
$industrygrowth_{3y}^{a}$	0.629^{*}	0.631	0.650^{***}	0.640	0.727^{**}
$scale^{a}$	0.164	0.833^{**}	0.706	0.332	0.830^{***}
$hist_{3y}^{a}$	0.000^{*}	0.000^{*}	1.000^{**}	0.000^{*}	0.352^{***}
$hist_{bool}^{a}$	0.157^{*}	0.395	1.000^{**}	0.215	0.659^{***}
dis_{3}	0.119^{***}	0.08^{*}	0.098	0.107	1.000^{**}
dis_{6}	0.025^{***}	0.000^{*}	0.000^{*}	0.023	0.943^{**}
$growth_{3y}^{t}$	0.704^{**}	0.650	0.650	0.696^{***}	0.545^{*}
$cashathand_{4y}^{t}$	0.277^{***}	0.369^{**}	0.252	0.210^{*}	0.273
$industrygrowth_{3y}^{t}$	0.610^{*}	0.677	0.693^{**}	0.645	0.682^{***}
$scale^{t}$	0.654^{***}	0.331^{*}	0.558	0.836^{**}	0.614
$salesmean_{bool}$	0.094	0.989^{**}	0.589	0.023^{*}	0.807^{***}
Pattern	Abundant cash	Big fish eats small fish	Abundant experiences	Small fish eats big fish	Normal M&A
M&A cases	2^{*}	18	29^{**}	4	56^{***}
Random samples	157	245^{**}	134	210^{***}	32^{*}
Ratio of M&A to Random Samples	$1.27\%^{*}$	7.35%	$21.64\%^{***}$	1.90%	$175.00\%^{**}$
Total M&A cases	109				
Total Random samples	778				

*, **, *** indicates the lowest value, the highest value and the second highest value in the row questioned, respectively.

binary values for designing features and we will try other features and values as well in the future. We hope this research will improve the efficiency of M&A business in the nearing future.

ACKNOWLEDGMENT

We would like to acknowledge UZABASE, Inc. for providing data and The University of Tokyo Edge Capital, Co., Ltd. for providing advice during this research.

REFERENCES

[1] V. Mehrotra, D. van Schaik, J. Spronk and O. Steenbeek, "Creditor-Focused Corporate Governance: Evidence from Mergers and Acquisitions in Japan," *Journal of Financial and Quantitative Analysis,* vol. 46, *(4),* pp. 1051-1072, 2011. . DOI: 10.1017/S002210901100024X.

[2] Ministry of Economy, Trade and Industry, "経済産業関係税制改正について [Economy, trade and industry relations, about tax reform]," Tokyo: Ministry of Economy, Trade and Industry, 2017. [Online]. Available: http://www.meti.go.jp/main/yosan/yosan_fy2018/pdf/zeiseikaisei.pdf. [Accessed: Jan, 26, 2018]

[3] Innovation Policy Research Center, the University of Tokyo. "Academic Landscape," *Innovation Research Policy Center*, 2010-2013. [Online]. Available: http://academic-landscape.com/page/about [Accessed: May, 03, 2018]

[4] Y. Kajikawa, J. Ohno, Y. Takeda, K. Matsushima and H. Komiyama, "Creating an academic landscape of sustainability science: an analysis of the citation network," *Sustainability Science,* vol. 2, *(2),* pp. 221-231, OCT, 2007. . DOI: 10.1007/s11625-007-0027-8.

[5] J. Hagedoorn and G. Duysters, "External sources of innovative capabilities: the preference for strategic alliances or mergers and acquisitions," *Journal of Management Studies,* vol. 39, *(2),* pp. 167, 2002.

[6] S. Shibayama, K. Tanikawa, R. Fujimoto and H. Kimura, "Effect of mergers and acquisitions on drug discovery: perspective from a case study of a Japanese pharmaceutical company," *Drug Discov. Today,* vol. 13, *(1),* pp. 86-93, 2008.

[7] M. Cloodt, J. Hagedoorn and H. Van Kranenburg, "Mergers and acquisitions: Their effect on the innovative performance of companies in high-tech industries," *Research Policy,* vol. 35, *(5),* pp. 642-654, 2006.

[8] J. L. Bower, "Not all M&As are alike--and that matters," *Harvard Business Review,* vol. 79, *(3),* pp. 92, 2001.

[9] D. Harding and S. Rovit, *Mastering the Merger: Four Critical Decisions that make Or Break the Deal.* Harvard Business Press, 2004.

[10] V. Dickinson, "Cash Flow Patterns as a Proxy for Firm Life Cycle," *Account. Rev.,* vol. 86, *(6),* pp. 1969-1994, NOV, 2011. . DOI: 10.2308/accr-10130.

[11] J. Mori, Y. Kajikawa, H. Kashima and I. Sakata, "Machine learning approach for finding business partners and building reciprocal relationships," *Expert Syst. Appl.,* vol. 39, *(12),* pp. 10402, 2012.

[12] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms.* New York: Cambridge University Press, 2014.

[13] L. v. d. Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research,* vol. 9, *(Nov),* pp. 2579-2605, 2008.

[14] A. Seth, K. P. Song and R. Pettit, "Synergy, Managerialism or Hubris? An Empirical Examination of Motives for Foreign Acquisitions of U.S. Firms," *J. Int. Bus. Stud.,* vol. 31, *(3),* pp. 387-405, 2000.

[15] M. P. Ferreira, J. C. Santos, M. I. R. de Almeida and N. R. Reis, "Mergers & acquisitions research: A bibliometric study of top strategy and international business journals, 1980–2010," Journal of Business Research, vol. 67, (12), pp. 2550-2558, 2014.

[16] A. D. Chandler Jr, "The enduring logic of industrial success," Harvard Business Review, vol. 68, (2), pp. 130, 1990.