ADVANCES IN BUSINESS AND MANAGEMENT FORECASTING

Edited by Kenneth D. Lawrence
and Ronald K. Klimberg

ADVANCES IN BUSINESS
AND MANAGEMENT
FORECASTING

VOLUME 12

ADVANCES IN BUSINESS AND MANAGEMENT FORECASTING

ADVANCES IN BUSINESS AND MANAGEMENT FORECASTING

Series Editors: Kenneth D. Lawrence and Ronald K. Klimberg

Recent Volumes:

Volume 1: Advances in Business and Management Forecasting: Forecasting Sales

Volume 2: Advances in Business and Management Forecasting

Volume 3: Advances in Business and Management Forecasting

Volume 4: Advances in Business and Management Forecasting

Volume 5: Advances in Business and Management Forecasting

Volume 6: Advances in Business and Management Forecasting

Volume 7: Advances in Business and Management Forecasting

Volume 8: Advances in Business and Management Forecasting

Volume 9: Advances in Business and Management Forecasting

Volume 10: Advances in Business and Management Forecasting

Volume 11: Advances in Business and Management Forecasting

ADVANCES IN BUSINESS AND MANAGEMENT
FORECASTING VOLUME 12

ADVANCES IN BUSINESS AND MANAGEMENT FORECASTING

EDITED BY

KENNETH D. LAWRENCE
New Jersey Institute of Technology, Newark, NJ, USA

RONALD K. KLIMBERG
Saint Joseph's University, Philadelphia, PA, USA

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2018

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78743-070-9 (Print)
ISBN: 978-1-78743-069-3 (Online)
ISBN: 978-1-78743-256-7 (Epub)

ISSN: 1477-4070 (Series)

CONTENTS

LIST OF CONTRIBUTORS

Matthew Butler	University of York
Catherine Cardamone	Saint Joseph's University
Okan Duru	Texas A&M University
Kathleen Campbell Garwood	Saint Joseph's University
Zhimin Huang	Adelphi University
Sudhanshu Joshi	Doon University
Ronald K. Klimberg	Saint Joseph's University
Kenneth D. Ko	Pepperdine University
Stephan Kudyba	School of Management, New Jersey Institute of Technology
Kenneth D. Lawrence	School of Management, New Jersey Institute of Technology
Sheila M. Lawrence	School of Management and Labor Relations, Rutgers, The State University of New Jersey
Matthew Lindsey	Stephen F. Austin State University
Virginia M. Miori	Saint Joseph's University
Son Nguyen	Bryant University
Yasen Nuermaiti	University of Science and Technology of China
Alan Olinsky	Bryant University

Robert Pavur	University of North Texas
John Quinn	Bryant University
Shalu Rathi	Doon University
Samuel Ratick	Clark University
Manu Sharma	University of Petroleum and Energy Studies
Harvey Smith	ARI
Feng Yang	University of Science and Technology of China

EDITORIAL BOARD

EDITORS-IN-CHIEF

SENIOR EDITORS

SECTION A
FORECASTING APPLICATIONS

THE EFFECT OF RELEASED INFORMATION ON SEARCHING FOR MISSING CHILDREN: THE CASE OF THE BABY BACK HOME NETWORK

Feng Yang, Yasen Nuermaiti and Zhimin Huang

ABSTRACT

The problem of missing children draws much attention of both governmental and nongovernmental organizations in China due to huge numbers of the missing children. According to the records of Baby Back Home network (BBHNet), a professional website to search missing children in China, 1,666 missing children have been found via releasing information on BBHNet; however, there are still 30,561 families searching for their children and 24,603 missing children are searching for biological parents through this website and have not succeeded yet. What is the difference between successful and unsuccessful cases in the aspect of released information? Motivated by this question, our research proposes to determine the crucial information in the process of searching missing children. A logistic regression model was developed on the data summarized from 500 succeed cases and 500 cases which have not succeed yet from BBHNet for forecasting success rate of searching missing children. The model identifies that the differences in terms of released information, number of children, address, and natural geographical features are the three most crucial factors that cause differences in the result of searching. This research can be used as a guide for improving the success

Advances in Business and Management Forecasting, Volume 12, 3–17

ISSN: 1477-4070/doi:10.1108/S1477-407020170000012002

rate of searching missing children and provide reference for developing a missing children information-sharing platform.

Keywords: Missing children searching; success rate; logistic regression; forecast

INTRODUCTION

With the increasing number of missing children and the rapid growth of no-child families, it is essential to focus on the problem of missing children. Trafficking and adoption are the main reasons of missing children. Many children were given to adoption for the reasons such as the state birth planning policy and economic pressures. Sexism against females and disabilities also can be reasons for parents to give their children for adoption. Due to the birth policy strictly limiting the number of children for each family, Chinese laws banned fetal gender authentication techniques without medical need. Some families abandoned their daughters so they could have a son, and families with no child were willing to adopt those abandoned daughters.

On the other hand, some children become missing children as victims of child-trafficking or being lost. The shortage of brides caused by gender imbalance and the lack of care from parents are the main causes of child trafficking. Trafficking in China has many purposes: purchasing women as brides, purchasing males as heirs, trading for unwanted female children, and the use of people for slave labor, commercial sex, or prostitution (Tiefenbrun & Edwards, 2009). According to estimates, about 70,000 children are trafficked each year in China, and 30% of whom have been sold by their own parents because of financial hardship (Shen, Antonopoulos, & Papanicolaou, 2013). Regions where the enforcement of the one-child policy is rather lax are mainly source area of child-trafficking and where this enforcement is tough are marketing areas, and the motivating factors for this black industry are lack of the heirs and brides (*Financial Times*, December 3, 2015. Available at: https://www.ft.com/content/9704cbdc-8eaa-11e5-8be4-3506bf20cc2b).

Nowadays, China has become much fairer with 30 years' development of economy after reform and opening up. Nowadays, for the families whose children are missing, the biological parents have sufficient ability to search for their children. At the same time, the development of information technology and the internet makes information transmission more convenient, and makes it easier for missing children to search for their biological family. Then China has formed a tide of tracing missing children or parents on this background.

The Chinese government and nongovernmental organizations made valiant efforts to rescue the missing children. At the government and policy level,

China enacted laws and established Anti-trafficking Office in the Ministry of Public Security to coordinate and guide the national anti-trafficking. China also established National Anti-trafficking DNA Database and rescued more than 4,000 trafficked children by 2015 according to the data from the State Council Information Office of the People Republic of China. Additionally, China Central Television (CCTV) produced a TV program named "Wait for Me" for searching for lost relatives from 2014 and shows this program on CCTV Channel 1, which is the most popular TV channel in China. Nongovernmental organizations established information-sharing platforms such as Network for Searching Notices of China (http://www.zgxrqsw.cn) and Baby Back Home network (BBHNet; http://www.baobeihuijia.com) to promote the searching of missing children.

Missing children's information in this chapter is characterized by gender, age, photo, house address, natural geographical features around house, and so on (shown in Table 1). Although releasing information on website and searching by matching information is a common way in searching missing children currently, information formats and details vary from website to website and notice to notice. Due to the lack of uniform standards or official information database, different information formats caused inconvenience for information sharing between different websites. Present researches concerned on searching for missing children mainly focused on forensic age progression (Lampinen, Arnal, & Adams, 2012) and DNA comparison (Qian, 2012). However, the number of missing children has been found by DNA comparison and image recognition merely reached 82 and 1, respectively, making up 16.4% and 0.2% of a total of 500 found children. Additionally, the number of families and children who have placed notices on the websites without clear photograph and available DNA profiles are in majority. Although information matching is a widely used method for searching missing children currently, there are few studies available on released information format and details and their effects on success rate. Our research, based on BBHNet, aims to distinguish the information which plays a crucial role in searching missing children and this research will provide a reference for developing a Chinese missing children alert system similar to the AMBER Alert in America.

BBHNet is an online public service organization established in this context in which a tide of tracing missing children or parents has formed. BBHNet has gained wide recognition and praise for 1,666 cases in which people found their family with the help of volunteers from this network. At the same time, according to the notices released on this website, there are still 30,561 families searching for their children, and 24,603 missing children are searching for biological parents through this website. How did the 1,666 cases succeed? And how can users release searching information more effectively? The aim of this chapter is to discover what kind of information released is helpful to more quickly and efficiently find a missing child (Data from BBHNet up to September 28, 2016).

Table 1. Variables Selected from Cases.

Variables	Type	Implication
Status	Binary variable	1 implies the missing children has been found, 0 implies has not been found yet.
Gender	Binary variable	1 implies the missing children is male, 0 implies female.
Age	Continuous variable	Age of children at the time of loss.
Photo	Binary variable	1 implies the released information of the missing children contains a photo, 0 implies that no photo has been released.
Characteristic	Binary variable	1 implies the released information of the missing children from initiative one (from missing children or their birth families) contains description of another (from missing children or their birth families).
Economic	Continuous variable	Economic level for missing children's birth family characterized by provincial regional per capita GDP of 2010. National per capita GDP is used to characterize the economic levels of families for whom information about the ancestral province is not available.
Number of children	Continuous variable	Number of children for the missing children birth family mentioned in the released information.
Address	Binary variable	1 implies the released information of missing children contains home address, 0 implies no information about address.
Natural geographical features	Binary variable	1 implies the released information of missing children contains natural geographical features, 0 implies no description.
Released time	Continuous variable	Time span from releasing information on the website to been found for successful case or to now for cases in which the missing children has not been found.
Missing time	Continuous variable	Time span from missing to been found for successful cases or the time from missing to now for cases in which the missing children has not been found.

Through summarizing 1,666 successful cases, we found that the average time of 1,620 people who have accurate records as missing children is 23.561 years, and the longest one reached 88 years. We found that 837 cases of trafficked children, accounted for 50.24%, 399 cases of adoption, and 71 cases of abandonment accounted for 23.95% and 14.7%, respectively. We examined the details and released channel of information from cases which had been found and still searching, to explore the effect of the missing children information such as gender, age, economic level of family, and accuracy of this information to the success rate. The results as the response of the regression model follow binary distribution. Logistic regression is a special case of generalized linear modeling where estimation in frequentist settings is based on maximum likelihood estimation (Brimacombe, 2016). A comparison method of DNA profiles

for missing children based on cloud computing also has been presented (Qian, 2012). According to statistical data of these successful cases, DNA was used in the last step of searching of missing children 229 times to verify the relationship of missing children and their family members by blood, and this using frequency is more than 82 times in which DNA used in the first step to find suspected children from a total of 500 cases. We developed a logistic regression model with information such as gender, age, photo, house address, natural geographical features around house as independent variables, and results of searching as the dependent variable to determine which kind of information will improve the success rate in searching missing children.

By establishing this regression model, identifying factors that play a crucial role to success rate will help us to release targeted information. This model also can be used to forecast the probability of success of searching and to shorten the expected searching time for a specific case, and the forecast searching time can provide references for developing assistance policies for these families and missing children. This research will give priority list to release the key information to increase the probability of finding missing children, shorten the searching time, and ease the pain caused by such losses.

The remainder of this paper is organized as follows. The following section presents an overview of the relevant methods that underlie the study. Data, variables, and model are introduced in the next section. Results and discussion are presented in the subsequent section, and conclusion in the final section.

LITERATURE REVIEW

Child alert system is an important channel broadcasting and disseminating information to the public while searching for missing children (Pashley, Enhus, & Leys, 2010). Child alert system originated in a case happened in America in which a girl named Amber Hagerman was abducted and murdered in 1998. Amber alert was signed into federal law by President George W. Bush in 2003 (Zgoba, 2004). The United Kingdom, Greece, France, the Netherlands, and Germany have also launched alert systems (Pashley et al., 2010).

Although China does not have a national child alert system, researchers exerted great efforts in the field of researching missing children. Atkinson (2010) developed a simulation-based approach for searching missing children. Forensic age progression is another approach based on outdated photographs to estimate the current appearance of missing children (Lampinen et al., 2012).

Qian (2012) presented a comparison method of DNA profiles for missing children based on cloud computing. DNA fingerprint comparison is the most efficient and accurate method for searching missing children, on the basis of the DNA profiles of children and parents available. However, some children have been missing for many years and their parents' DNA profiles may not be

available, because one or both of the parents have died or their parents divorced (Lampinen et al., 2012).

The above advanced DNA comparison methods are helpful in this context that the DNA profiles of parents and children are available. Forensic age progression technology can be exercised only under such conditions that the photos of one person could be provided. However, out of 1,000 cases we studied, the number of cases in which the photo is available is 547, and only two children were recognized by photo comparison from a total of 500 children. The number of children found by DNA comparison and image recognition added up to 82, and had not reached 1/5 of the successful cases.

Based on the above analysis, we assume that DNA comparison and forensic age progression technology are helpful but not applicable to all cases. Releasing information on website and searching by matching information is a common way of searching for a missing person currently. The purpose of our research is to identify factors that play a crucial role to the success rate and develop a logistic regression for the effect of released information on searching missing children to improve the success rate of searching missing children.

DATA, VARIABLE, AND MODEL

As already mentioned, the main aim of this chapter is identifying the differences of cases in which the missing children had not been found compared to successful cases with regard to the released information. In order to identify differences between successful cases and unsuccessful cases on the aspect of released information, we selected 500 successful cases randomly in which missing children have been found after releasing notices on the BBHNet. We found that these 500 missing children have been found corresponding to 635 notices, of which 310 notices were from missing children searching for their biological families, 55 notices were from families searching for missing children, and in 135 cases where, both missing children and their family released notices searching for each other.

Then we classified these 500 missing children cases to family searching for baby and baby searching for family by released time of notices, for example, one case will be classified as the type of "baby searching for family" if the released time of searching information from the children is earlier than the time from family, otherwise, will be classified as the type of "family searching for baby".

To compare with 500 successful cases, we randomly selected 500 cases in which the missing children have not been rescued, these 500 cases consist of 395 cases in which missing children are searching for family and 105 cases in which families are searching for children.

We selected missing children's information as shown in Table 1 via summarizing these cases such as gender, age, photo, and characteristics, which are helpful to characterize the missing children. We also selected economic indicators such as economic level and number of children, since financial pressure is one of the reasons for missing (Shen et al., 2013); additionally, searching for missing children requires sufficient financial resources. We also suppose that the more children will reinforce financial pressures, and extra children may lead to more pressures to birth family from birth control policy. Birth family address will be used as a guide for searching birth family of missing children.

By analyzing these successful cases, we recognized that natural geographical features are helpful information in searching for missing children. Although China covers a large territory, complicated in topography, having various climate and custom, regional differences are outstanding in these aspects. So, it can give a clue to find the missing children. For example, if one child says he saw a banana grove around home before, this child is more likely to be from tropical and subtropical regions of South China where banana cultivation in China is mainly based. We defined natural geographical features around a house.

Time span of searching characterized by missing time and releasing time is regarded as an influence since information about missing children may be invalidated with time.

Since some variables such as characteristics and natural geographical features are in the type of text, additionally, great gap in continuous variables, the data need to be preprocessed in the purpose of being useful for further analysis. First, we valued the photo, characteristics, address, and natural geographical features as 1 or 0 according to the presence or absence of the related information (Table 2).

Table 2. Descriptive Statistics.

	N	Minimum	Maximum	Mean	Std. deviation
Status	1,000	0	1	0.500	0.500
Gender	1,000	0	1	0.490	0.500
Age	1,000	0	1	0.124	0.150
Photo	1,000	0	1	0.550	0.498
Characterization	1,000	0	1	0.880	0.330
Economic	1,000	0	1	0.282	0.206
Number of children	1,000	0	1	0.134	0.203
Address	1,000	0	1	0.410	0.492
Natural geographical features	1,000	0	1	0.280	0.447
Missing time	1,000	0	1	0.196	0.213
Releasing time	1,000	0	1	0.323	0.161
Valid *N* (listwise)	1,000				

To cope with the gap between the continuous variables, we normalized the continuous variables by the following equation:

$$x_{ij} = \frac{y_{ij} - y_{\min}}{y_{\max} - y_{\min}}, y_{\max} = \max_{j=1} y_{ij}, y_{\min} = \min_{j=1} y_{ij}, i = 1, 2, \ldots, n \quad (1)$$

Binary logistic regression is a frequently applied procedure used to predict the probability of occurrence for some binary outcome using one or more continuous or categorical variables as predictors (Ramos, Ollero, & Suárez-Llorens, 2017). The advantage of logistic regression over usual regression models is that this model allows variables to be either continuous or discrete, or any combination of both types (Lee, 2005). In our research, the dependent variable (status) is a binary variable representing the presence or absence of finding. The independent variables are the combination of both binary (gender, photo, characteristics, address, and natural geographical features) and continuous (age, economic level, number of children, and time span) types. Where the dependent variable is binary, the logistic link function is applicable (Atkinson et al., 1998). To cope with these variables, we performed logistic regression analysis. Relation between outcome (status) counted by Y and its probability of occurrence is presented in the following form:

$$p(Y = 1) = \frac{1}{1 + e^{-z}} \quad (2)$$

where p is the probability that the dependent variable Y is 1, implying that the missing child has been found. Z is the linear combination of a set of predictor variables which related to outcome (status). Z is expressed in the following form:

$$z = b_0 + b_1 x_1 + b_2 x_0 + \ldots + b_n x_n \quad (3)$$

where b_0 is a constant in this model, and b_i ($i = 1, 2, \ldots, n$) are regression parameters estimated by maximum likelihood estimation.

RESULTS AND DISCUSSION

With the aim of identifying the differences of cases in which the missing children had not been found compared to successful cases in regard with released information, we performed the binary logistic regression method, estimated logistic model as shown in Fig. 1, model summary and estimated parameters are tabulated in Table 3 and Table 4, respectively. As shown in Table 3, the Cox and Snell R^2 value is 0.642, Nagelkerke R^2 is 0.857, which indicates that this logistic regression identifies key factors which affect the success probability of searching for the missing children successfully. Since $\chi^2(8) = 1028.215$, $p < 0.05$, full model

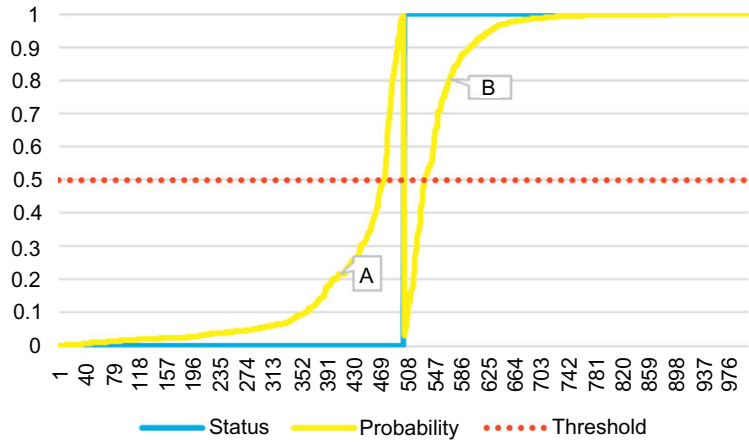

Fig. 1. Comparison of Real Status and Regressed Model.

Table 3. Model Summary.

−2 Log-likelihood	Cox & Snell R^2	Nagelkerke R^2
353.978[a]	0.644	0.858

Table 4. Variables in the Equation.

	B	S.E.	Wald	*df*	Sig.
Gender	0.619	0.292	4.488	1	0.034
Age	−0.164	1.165	0.020	1	0.888
Photo	2.448	0.320	58.389	1	0.000
Characterization	−1.036	0.363	8.150	1	0.004
Economic	−2.482	0.820	9.152	1	0.002
Number of children	10.463	1.217	73.954	1	0.000
Address	2.659	0.314	71.643	1	0.000
Natural geographical features	3.937	0.548	51.686	1	0.000
Missing time	−5.466	0.778	49.405	1	0.000
Releasing time	−0.446	1.105	.163	1	0.687
Constant	−1.419	0.582	5.944	1	0.015

is statistically significant, thus, the model is successful in identifying the differences of cases in which the missing children had not been found compared to successful cases in regard to the released information. Additionally, this model classified 94% of cases correctly as shown in Fig. 1.

Table 5. Data for Forecasting.

Gender	Photo	Characterization	Economic	Number of children	Address	Natural geographical features	Missing time
0	1	1	0.502216	0	0	0	0

As shown above, from 472 to 531, the estimated probabilities of succeed cases are less than 0.5 or the estimated probabilities of cases which have not succeeded are greater than 0.5. Actually, the cases will be judged as absence of the incident while the logistics forecasted value is less than 0.5. These 60 cases are judged as classified incorrectly from a total of 1,000 cases.

Results of this regression are shown in Table 5. First, the estimates on age and released time are nonsignificant upon 5% significance level, implying that the age of children at the time of missing is not a significant variable on the success rate of searching missing children. Actually, though the age of children at the time they go missing is an important characteristic of a child, effects of missing age on the success rate of searching missing children are not statistically significant, this implies that the information obtained from the missing age of children has no significant difference among all cases. Releasing time is another nonsignificant estimator in this regression, showing that the time the information has been released on website is not a more crucial factor than the information itself to find the missing children.

Estimates on gender, photo, and number of children, address, and natural geographical features are significant and positive. This positive and significant estimate on gender implies that males can be found more easily than females, indicating that the willingness of parents to search for missing male children is more than that shown for missing female children. Notices released for searching for missing male children (16,608) accounted for 52.5% of total notices (31,661), more than notices released for searching for the missing female children (15,053); at the same time, the notices released for searching for families from females (15,700) accounted for 60.8% of total notices (25,796), which are significantly more than notices from males (10,092). Positive estimates on photo, address, and natural geographical features imply that the release of these relevant information will be helpful in finding the missing children. Photos can be used in image recognition. Due to the place name repetition, searching for the missing children by address can be confusing, such as place names Dongcheng district: Beijing, Sichuan, Guangdong, and Inner Mongolia, the four provincial areas which have large differences in natural geographical features all have places named as Dongcheng district. Natural geographical features can be used in conjunction with address to handle the place name repetition; additionally, natural geographical features can be used as an address in the cases in which information about address is absent.

As for the number of children of the missing children's family, this positive estimate seems to imply that the more children there are in a family will increase the success rate of searching for missing children. This is contrary to the initial assumption that the more children in a family will reinforce financial and political pressure which will prevent families from searching for the missing children. This abnormal result may imply that the missing children who remember more other children in a family could provide more other accurate information such as address and natural geographical features and this will be helpful for searching.

Estimates on missing time are negative and significant, implying that the more recent a child goes missing, the higher the probability that they can be found, and this is due to the likelihood that information about the missing children may be invalidated with time. Characteristics and economic with negative estimates, the negative estimates may be on account of characteristics are ambiguous and inaccurate. The economic level is also an important influencing factor for searching for missing children. In this chapter, the negative estimates due to the economic developed area optimized and combined various resources and natural geographical features have been changed in a large range such as urban reformation and so on.

According to the significance of estimates, we performed these significant estimates in to Eq. (2) as the following form:

$$\begin{aligned} Z = & -1.419 + \text{Gender} \times 0.619 + \text{Photo} \times 2.448 - \text{Characteristic} \times 1.036 \\ & - \text{Economic} \times 2.482 + \text{Number of children} \times 10.463 + \text{Address} \times 2.659 \\ & + \text{Natural geographical features} \times 3.937 - \text{Missing time} \times 5.466 \end{aligned} \quad (4)$$

This model is able to forecast the probability of a given case for finding the missing children according to Eqs. (1) and (3). We performed this model on the case released on BBHNet on December 3, 2016, which case ID is 226366 (http://www.baobeihuijia.com/view.aspx?id=226366). Pei ling Zhong, a 13-year-old and 160-cm-tall girl with long hair and bangs, disappeared on November 14, 2016 at No. 53, Shapai Village, Xiapu Town, Dayawan County, Huizhou City, Guangdong Province. Her information was released on the BBHNet on December 4, 2016, 20 days after she disappeared. Case 226366 was scored as follows:

Where the forecasted probability of this case for finding this girl is 0.222, point A in Fig. 1. This is a case in which family is searching for the missing children, though natural geographical features around home and home address are available for this girl's family members, they did not release any information about natural geographical features. We suggest this family release information about natural geographical features, address, and number of children to improve the probability of finding this girl. Since the effect of releasing time to the probability of finding is statistically nonsignificant, it is better to enrich the

Table 6. Means and Average Searching Time.

Searching means	Quantity	Average searching time
DNA comparison	82	1.131073
Address matching	158	1.155026
Characteristics matching	119	1.230736
Notices matching	98	0.946581
Contacted by searched one	42	1.33105
Image recognition	1	2.293151

information for releasing complete notice than releasing fragmentary information hastily. In this case, there is an address where this girl disappeared, if this is an exact address of this girl's family, then the probability given by our model is 0.803. The crucial effect of key information such as address was shown in this case. As shown in Fig. 1, this probability can be improved from point A to point B by adding the information about the address.

We have also analyzed the main methods of successful cases that were utilized in searching and classified these means to DNA comparison, address matching, characteristics matching, notices matching, contracted by searched one and image recognition, as shown in Table 6.

DNA comparison is an advanced and principal commendatory method at the present stage, which provides the most convincing evidences. However, both parents and children DNA samples should be provided for DNA comparison. Characteristics matching limited by the accuracy of description is another useful method used frequently. Notice matching is the most time-saving method but only works if the notice is released on the same website and contains the same comparable details. There are 42 cases in which the searched one contacted the website after seeing the notice released on the website.

As address matching is the most frequently used method in these 500 successful cases, the administrative region is divided into provincial region, municipal region, county-level region, town-level region, and village in China. Different levels of address provide varying degrees of information, for instance, information that one's home address is Dongcheng district in Beijing is more accurate than the information that one's family live in Beijing. To determine the effect of accuracy of information on average search time, we analyzed the average search time of cases with various accurate levels of information about the address. As shown in Fig. 2, average search time of cases with provincial level of information about address take longer. Although village level of information about address is most specific, the average search time of cases with village level information about address is greater than the cases with county-level address and similar to cases with town-level address. A more specific address should lead to a shorter search time, however, the average search time at country-level is shorter in our results, and may be that villages and towns

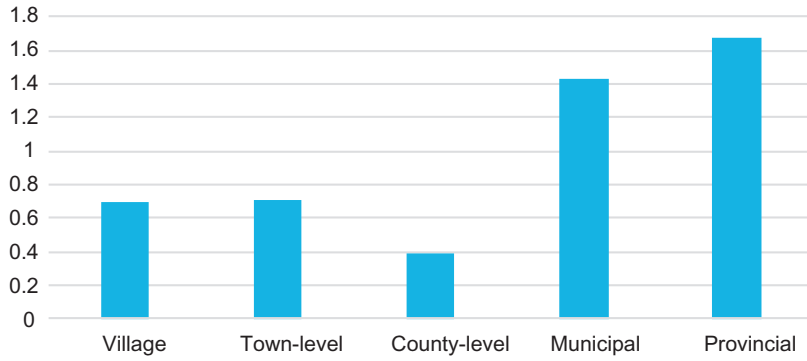

Fig. 2. Average Search Time according to Address.

experience more regionalized administration than at country level, regionalized administration in villages and towns invalidate information about villages and towns.

With the problem of missing children attracting increasing attention in China, more and more channels have opened up for releasing information about missing children. In addition to traditional media such as television, radio, and newspapers, there are other new channels. Releasing information on the 404 error page is a suggested information-releasing method which has been performed by BBHNet. Some mobile applications also provide information-releasing functions such as Alipay which is the most popular third-party online payment system in China, and releases missing children information to a user who lives in the same region in which a child disappeared. Publicizing information about missing children on mineral water bottles has also been used to help find missing children.

CONCLUSION

The problem of missing children has remained severe; a number of notices have been released for families, and searches for children reached 30,561 and 24,603 to 31,661 and 25,796 within two months, increasing from 3.6% to 4.8%, respectively. At the same time, the number of successful cases only increased by 3.2% between 28 September and November 28. Children go missing for many reasons, including but not limited to economic hardship, political pressure, and sexism. Society today is different from that of the past, and poverty alleviation and economic development will relieve financial stress. The two-child policy, which was put forward by the Central Government of China, may increase labor supply (Yang, Yang, & Huang, 2016), and this policy will reduce pressure

on families with more than one children. Chinese Government and public organizations interventions may improve this situation; however, many children need love and attention from their parents and families.

Our regression model, which identifies the importance releasing information in successful cases and cases that have not yet succeded can be used to provide directions on releasing new notices for missing children, or a guide for improving and complementing notices that have been released, but have not yet had success. Indeed, further information will lead to increased chances of finding missing children, and the media is willing to publish notices on missing children; however, releasing all details via the media is unrealistic due to limitations of space and time. Therefore, our regression model can prioritize details to be released based on the estimates in Eq. (3). Calls for establishing a national official missing children information-sharing system such as AMBER Alert in the United States has recently been growing. Our research can be used as a reference to establish this system.

In terms of the release of information, our research identified some crucial differences between successful cases and those that have not yet succeeded. The basis of our research lies in guiding the release of information rather than solving the present problem. Additionally, our research is hosted on BBHNet website and it is our intention to expand this work to other platforms such as Wait for Me.

ACKNOWLEDGMENTS

The authors would like to thank the National Natural Science Foundation of China (Grant nos. 71631006, 71271195, 71322101, and 71110107024) and USTC Fund for Innovation Teams (WK2040160008) for their financial support.

REFERENCES

Atkinson, J. (2010). Simulation of Search for Missing Child [J].

Atkinson, P., Jiskoot, H., Massari, R., et al. (1998). Generalized linear modelling in geomorphology. *Earth Surface Processes and Landforms*, *23*(13), 1185–1195.

Brimacombe, M. (2016). Large sample convergence diagnostics for likelihood based inference: Logistic regression. *Statistical Methodology*, *33*, 114–130.

Lampinen, J., Arnal, J. D., Adams, J., et al. (2012). Forensic age progression and the search for missing children. *Psychology, Crime & Law*, *18*(4), 405–415.

Lee, S. (2005). Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, *26*(7), 1477–1491.

Pashley, V., Enhus, E., & Leys, M. (2010). Child Alert': public information dissemination of child disappearances.

Qian, H. J. (2012). Research and simulation of large-scale DNA fingerprint comparison method of missing children based on cloud computing [J]. *Computer Simulation*, *29*(11), 308−311.

Ramos, H. M., Ollero, J., & Suárez-Llorens, A. (2017). A new explanatory index for evaluating the binary logistic regression based on the sensitivity of the estimated model. *Statistics & Probability Letters*, *120*, 135−140.

Shen, A., Antonopoulos, G. A., & Papanicolaou, G. (2013). China's stolen children: internal child trafficking in the People's Republic of China. *Trends in Organized Crime*, *16*(1), 31−48.

Tiefenbrun, S., & Edwards, C. J. (2009). Gendercide and the cultural context of sex trafficking in China. *Fordham International law Journal*, *32*(3), 731.

Yang, F., Yang, Y., & Huang, Z. (2016). Forecasting the population growth in China based on the welfare efficiency and demographic statistics. *Advances in Business and Management Forecasting* (pp. 169−183), Bingley: Emerald Group Publishing Limited.

Zgoba, K. (2004). The Amber Alert: the appropriate solution to preventing child abduction? *The Journal of Psychiatry & Law*, *32*(1), 71−88.