

Detection of Interviewer Falsification in Statistics Indonesia's Mobile Survey

Yusep Rosmansyah¹, Ibnu Santoso², Ariq Bani Hardi³, Atina Putri¹, and Sarwono Sutikno¹

¹School of Electrical Engineering and Informatics, Bandung Institute of Technology
Bandung, Indonesia ²Polytechnic of Statistics, ³BSSN, Jakarta, Indonesia
yusep@stei.itb.ac.id, ibnu@stis.ac.id, ariq.bani@bssn.go.id, 23517042@std.stei.itb.ac.id,
ssarwono@stei.itb.ac.id

Abstract: Interviewer falsification is an important issue faced by institutions conducting censuses and surveys around the world, including Statistics Indonesia. This study discusses several methods to systematically detect interviewer falsification and validation using data mining techniques so that human supervisors can take further actions. After analyzing relevant features and conducting experiments, the results showed that unsupervised classification algorithm using simple 2-means clustering achieved 70.5% accuracy, while the supervised classification using logistic regression improved the accuracy to 88.5%. A greater level of accuracy is still needed to be pursued in further research, but the current results are certainly better than the traditional method which has almost no falsification detection method at all.

Keywords: Data Mining; Interviewer Falsification; Mobile Survey; Statistics

1. Introduction

Interviewer falsification occurs when a field enumerator (surveyor) violates enumeration Standard Operating Procedures (SOP) by filling part of or entire questionnaire by herself without any interaction with a respondent [1]. Interviewer falsification is an important problem faced by Statistics Indonesia (Badan Pusat Statistik, BPS) and other national statistical agencies. In addressing this issue, BPS runs tiered supervision and extensive monitoring program in all surveys and census activities. This tiered supervision aims to ensure that field enumerators work in accordance with the SOP so that the resulting data are genuine and reliable.

An example of falsification was recorded in the official report on quality monitoring of SP2010 Population Census. It informed that there were indications of interviewers falsifying data in the surveillance area [2]. National newspaper sites [3-6] also reported the falsification incidents carried out by enumerators when surveying field data. This falsification issue is also a problem faced by other statistical agencies, even in developed countries such as the U.S. and Germany [7-13]. Simmons, et al [26] also showed that 1 in 5 available international surveys contains data that indicates falsification.

Research on data falsification becomes very important considering the magnitude of impacts that can be caused. The effect of data falsification can lead to serious consequences for the statistics generated from survey data. Bredl, et al [14] showed that a very small proportion (2,5%) of falsified data could reduce the effects of training on the gross wages log multivariate regression by about 80%. The bigger difference between the actual variable value with the value of the falsified variable value, the bigger the survey bias [12,15,16].

2. Related Work

The study of interviewer falsification by survey methodologists was formally begun in 1945, marked by a paper published by Crespi [17]. Crespi explained the reasons why an interviewer falsified data, such as long questionnaires, complex questions, hard-to-reach respondents, and external factors, such as weather and condition of society. Crespi also gave suggestion to survey organizers to design the survey carefully with the goal to make data falsification unattractive to enumerators.

Received: July 16th, 2019. Accepted: September 25th, 2019

DOI: 10.15676/ijeei.2019.11.3.2

The survey methodologists then began to investigate the possibility of using survey data itself for detecting data falsification. Statistical methods were able to provide an overview of the characteristics of falsified data, such as following the Benford's laws [14,15], the high rate of non-response [10], the absence of the contact information of the respondents [10], and low data variance [18].

In addition to using the survey data, other studies used metadata to detect data falsification, such as date stamp, time stamp, and behavioral data. The number of studies of detecting falsification using metadata was still very limited [19]. The latest research by Birnbaum, et al. [20], which introduced a supervised learning technique in the interviewer's behavior log data, showed acceptable results. An important advantage of approaches relying on metadata analysis is that they can be applied to a vast range of surveys [19].

This study combines several methods to detect interviewer falsification.

3. Methodology and Exploration of Features to Detect False Data

This research adopted the Benford's Law to characterize the falsified data in surveys. For mobile application prototype development, extreme programming method was adopted as an agile process development [21]. Firstly, features to detect false data were explored and defined architectural spikes of the planned mobile application. Next, the application concept was depicted as release planning. Thereafter, the iterations of development and implementation processes were performed until a stable version was achieved. The stable version of the application was evaluated using user acceptance test. Moreover, the results were also used as a survey result to analyze.

Features that can be used to detect false data are summarized in the following sections.

A. Metadata Features

Metadata can be defined as a set of data that provide information about other data. In this case, our data are the collected data by the enumerator, and the metadata can be in the forms of the following sections.

1. Issues, Controversies, Problems

Theoretically, honest interviews were conducted in proper locations. An indication of a falsified interview is if it is conducted in the wrong location. Thus, the location plays an important role in detecting falsified interview. Tablet and smartphone devices nowadays are equipped with a Global Positioning System (GPS) to detect the location of the device. Coordinate obtained from the GPS device is expressed in latitude and longitude coordinate.

Location detection is very useful if the coordinate of a respondent is available. The proximity of the respondent's coordinate and enumerator's device's coordinate can be computed. BPS does not have respondent's coordinate data, but BPS has a working area map with geolocation. The smallest unit of working area map is a census block map. One census block consists of 80-100 households. Census block map as shown in Figure 1 is standard equipment that enumerators must carry while collecting field data. The usage of maps is aimed at making sure that the data is obtained from the correct location. Census block map is polygon-shaped and also available in digital form. Digital census block maps were saved in the form of shapefiles.

A common approach in examining location can be conducted by checking whether or not a coordinate of the interview was inside the census block polygon map. If the coordinate is not in the polygon or far outside, the honesty of the interview is suspected. In addition to seeing the point in polygon, the movement of the tablet phone device can also be examined. The honesty of interviews was suspected if many interviews occurred in the same location. Technically, the detection of coordinate points inside the shapefile polygon can be conducted using the iPython and OSGeo library. Using latitude and longitude coordinates as input, the output of the program generates which area in the polygon contains the coordinate.

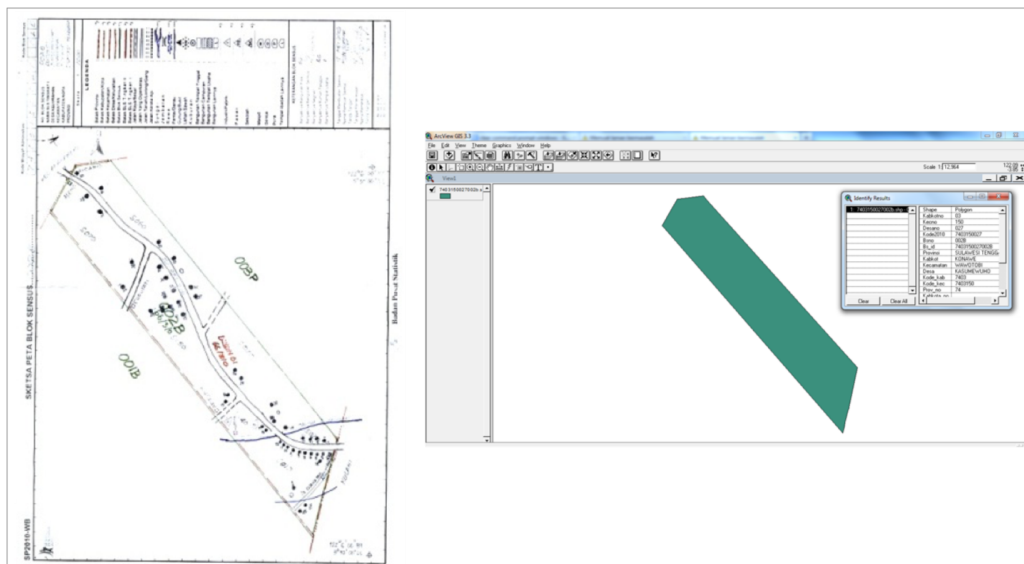


Figure 1. Analog map (left) and digital map (right)

2. Date Stamps

Date stamp is the date when the interview is conducted. In a survey using paper as a medium, usually there is a form to fill designating the date of the interview and date of supervision, as shown in Figure 2. Date stamps are very useful and can provide information on how many days an enumerator could complete all questionnaires. If too many interviews were completed in one single day, then the enumerator is suspected to data falsification. Indications of data falsification can also be assessed when a large number of interviews are completed close to the deadline.

III. KETERANGAN PETUGAS		
Uraian	Pencacah	Pengawas/Pemeriksa
1. Nama	RAHMIN	NIP
2. Jabatan	1. Staf BPS Provinsi ③ KSK 3 2. Staf BPS Kab/Kota 4. Mitra	1. Staf BPS Provinsi 3. KSK 2 2. Staf BPS Kab/Kota 4. Mitra
3. Tanggal Pencacahan/Pemeriksaan	Tanggal 04 Bulan 07	Tanggal 17 Bulan 07
4. Tanda Tangan		

Figure 2. A copy of paper-based date stamp

Using a paper questionnaire, the date stamp data can be faked very easily. Enumerators conducting interviews can fill the date field at will. Supervisory control mechanism is required to supervise the enumerators. Fortunately, using tablet phones, date stamps can be recorded using the Internet time server and delivered concurrently with the delivery of data. Although it can also be faked by changing local time, date stamp can be accurate when taken from the Internet time server.

3. Time Stamps

Time stamps inform at what time the interviews were conducted. Indications of data falsification arise if some interviews have too tight time stamp intervals, or interviews occur at unusual time (e.g. between 12 pm to 6 am).

Time stamp metadata can be obtained more accurately using a tablet phone as a data collection instrument. Just like date stamp, time stamp information can also be done on a paper questionnaire, but can easily be falsified by enumerators.

4. Duration

Duration is associated with the length of the interview describing how long the interview was conducted. In some literatures, duration is considered as a strong predictor for detecting falsified interview. The honest interview usually takes place in the normal duration. Indication of data falsification arises if the interview is completed too quickly. Even if the enumerator does not falsify data but duration is too short, the quality of data is questioned because there a possibility that the enumerator does not probe the respondent deep enough. When the interview is completed far too long from normal time, the concerned enumerator may require additional training on how to ask questions effectively and how to use the interviewing time efficiently.

Duration data can be recorded for other interviewing details, such as duration of the entire interview, duration of the questioning, and duration of the respondents answering. Difficult questions can also help predict falsification because it normally takes a longer time for the respondent to answer properly.

5. Behavioral Data

Behavioral data explain the behavioral patterns of interaction between an enumerator and a respondent. For example, an enumerator who swiped too many times in a short period might only interact with the device, instead of the respondent.

Behavioral data can be recorded from number of swipes, number of clicks, next button pressed, previous button pressed, editing an answer, press the help button, and many others. Recent notable research by Birnbaum [24] used behavioral data to detect interviewer falsification. He proved that using supervised learning on the behavioral data could provide high accuracy even if the enumerators know how the system works.

B. Data Features

Starting in the 1980s, researchers began to investigate the possibility of the property from the survey data alone to detect data falsification. Some of the well-known techniques are the use of Benford's law [14,15], variability computation [18], unusual combinations of answers, and the number of skips.

Benford's Law states that the probability of the numbers into the First Significant Digit (FSD) to follow a particular distribution, namely:

$$P(\text{first digit is } x) = \log_{10} \left(1 + \frac{1}{x} \right) \quad (1)$$

where $x = 1, 2, \dots, 9$. This formula makes the chance of ascending FSD distributed in a logarithmic scale. Research conducted by George and Schechter [22] showed that Benford's law can be used as a tool to detect falsified data.

In addition to using Benford's law, researchers also measured levels of variance. Enumerator who falsified data with simple thought tends to choose the same answers to certain questions. The combination of answers that are not prevalent or rare can also be used as indicators, for example the data of people who drink alcohol but did not smoke.

Number of skips or null variables can also be used as a predictor when the enumerator who fakes interviews chose the short route in questions. Shorter route means that there are many skips thus many null variables. However, the absence of all short routes does not always mean that enumerators falsified data. It could be the actual field conditions is a homogenous area. A short route of interview can be used as an indicator when an enumerator tends to complete the task in mostly short-route interview so that the interview can be quickly conducted.

4. Characteristics of Falsified Data

Birnbaum [24] derived 15 characteristics of falsified data. These characteristics can then be mapped with mobile survey application features to detect them. The results of this mapping can be seen in Table 1.

Table 1. Mapping of falsified data characteristics and application features to detect them

No	Characteristic	Metadata Features					Data Features			
		1	2	3	4	5	6	7	8	9
1	Fast interview	-	-	√	√	-	-	-	-	-
2	Many missing units*)	-	-	-	-	-	-	-	-	-
3	Interview surge	-	√	-	-	-	-	-	-	-
4	Close to deadline	-	√	-	-	-	-	-	-	-
5	Short paths through survey	-	-	-	-	-	-	-	-	√
6	Unusual time of day	-	-	√	-	-	-	-	-	-
7	Many incomplete interviews*)	-	-	-	-	-	-	-	-	-
8	Missing telephone number	-	-	-	-	√	-	-	-	-
9	Low data variance	-	-	-	-	-	-	√	-	-
10	Few missing units*)	-	-	-	-	-	-	-	-	-
11	Bad fit to Benford's law	-	-	-	-	-	√	-	-	-
12	Unusual data	-	-	-	-	-	-	-	√	-
13	Rare response combination	-	-	-	-	-	-	-	√	-
14	Low time variance	-	-	-	√	-	-	-	-	-
15	Long interview	-	-	-	√	-	-	-	-	√

Note: 1=Location, 2=Date Stamp, 3=Time Stamp, 4=Duration, 5=Behavioral Data, 6=Benford's Law, 7=Variability, 8=Data rarity, 9= Skips

*) Unlike paper-based questionnaire, the mobile survey does not allow incomplete interviews and missing values

It can be seen for the mapping that: 3 falsified data properties such as "Many missing units", "Many incomplete interviews", and "Few missing units can be solved by the application of a mobile survey", 7 other characteristics can be addressed using the metadata, and the rest can be addressed using the response data. Overall, the characteristic of falsified data can be detected simply by using features especially date stamp, time stamp, timing data, and simple data response analysis. Other features such as the location, and behavioral data as mentioned in the previous section can be used to make falsification identification more accurate.

5. Implementation

The prototype of the interviewer falsification detection system is evaluated for the Environmental Concern Behavioral Survey or "*Surat Pernyataan Pengelolaan Lingkungan Hidup*" (SPPLH). SPPLH was chosen because of two reasons. First, it is a real survey conducted by BPS and second, survey completion time is relatively short, around a month with duration time for a survey is about 10-20 minutes.

An SPPLH survey questionnaire consists of 13 blocks and 36 questions. The questionnaire has many conditional questions that only asked if certain answers were given to the previous question.

A. Data Features

Because this implementation is applied to the experimental environment, some features that do not make sense to use in experimental situations such as date stamp and time stamp were not used. The features used in this study are described in Table 2.

Table 2. Features used

No	Features	Description
1	insideblock	whether or not the interview conducted in specified census block
2	totaltime	total time taken to complete the interview
3	totalquesttime	total time taken to ask questions
4	averagequesttime	average time taken to ask questions (totalquesttime divided by the number of questions that passed the screen)
5	totalanswertime	total time taken to answer questions
6	averageanswertime	average time taken to answer questions (totalanswertime divided by the number of questions that passed the screen)
7	sumcontendit	how many times enumerator change the answer in one screen
8	sumnoncontedit	how many times enumerator does not change the answer after the screen appears
9	numconditional	how many times the conditional screen appeared
10	timeconditional	total time spent in conditional questions
11	forward	how many times the forward "next" button is pressed
12	backward	how many times the "previous" button is pressed
13	help	how many of times the "calculator" is pressed

B. Mobile Application and Server-side Application Prototype Development

The mobile survey application system prototype consisted of two parts: a native mobile application and a server-side application. The mobile application was developed using extreme programming methods. Samsung Galaxy Tab 2 P3100 was a target device to run the application. It had a 7-inch screen size and run Android 4.0.3 Ice Cream Sandwich operating system. The authors used Android SDK API Level 19 and built the source code using Eclipse IDE (Integrated Development Environment) with Android ADT (Android Development Tools) plugins.

The application was equipped with mechanisms that can record data and metadata locally and remotely. The application was also fortified with data validation checking before it could be saved as clean data. Figure 3 shows the screenshots of the mobile application prototype.

The server-side application was developed using PHP language with the following functionalities:

- Authenticating user login
- Transferring the list of households that must be interviewed by enumerators
- Showing the progress in field data collection for each enumerator
- Saving all the response data and metadata uploaded by clients in a database

C. Enumerator Training and Field Data Collection

Training of the developed mobile surveyor application prototype to 30 enumerators was conducted in one day and divided into two sessions. The goal of first session was to build an understanding of concepts and definitions used in SPPLH so that the enumerators can ask questions specifically and effectively. In the second training session, the application was demonstrated and we trained the enumerators on how to use it on tablet smartphones.

Participating enumerators consisted of 30 officers from the District Coordinator of Statistics or “*Koordinator Statistik Kecamatan*” (KSK) and outsourcing enumerators who had years of experience in field data collection. The participants then divided into 15 teams, where in one team there was at least one tablet phone.

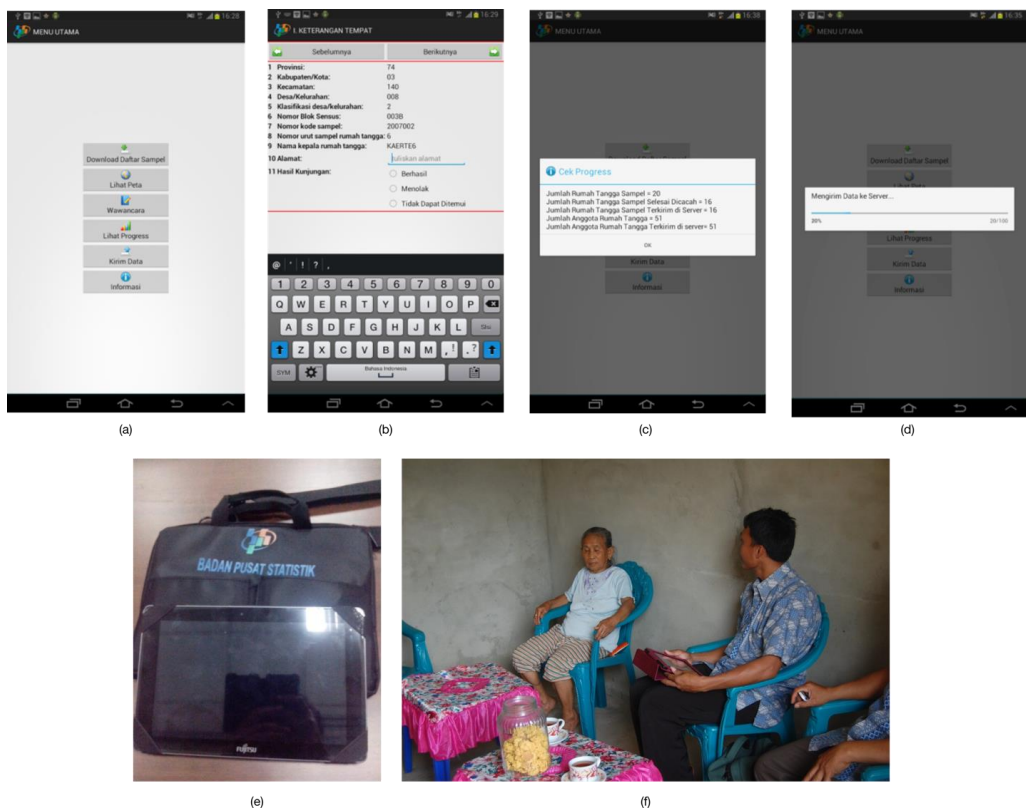


Figure 3. Screenshots of mobile survey SPPLH application prototype (a-d) and photos of the device in use (e-f)

Each enumerator was assigned to collect the data of 20 households with 10 data obtained from the actual enumeration and the other 10 data obtained from falsification. Of 10 falsified data, 5 were made during the training when the enumerators did not have knowledge and experience in an actual interview. The remaining 5 falsified data was conducted after the interview with 10 actual households in order to gain more knowledge and experience. The final result is the data for a total of 300 actual households and 300 falsified households. This kind of method was also applied by Birnbaum [24] in his study.

6. Testing and Evaluation

A. User Acceptance Test

A user acceptance survey was conducted to see whether the application easy to use. Out of 30 enumerators, 25 people (83%) agreed that the application was easy to use without significant constraints, and the remaining 5 (17%) stated that the application is quite easy to use. No user commented that the application was difficult to use.

B. Evaluation of Features

The evaluation of the 13 features used (except the inside block feature, as it will only work on a real site) was conducted by calculating the correlation using point-biserial correlation. The point-biserial correlation coefficient is used to estimate the degree of relationship between nominal scale and an interval (or ratio) scale (Brown, 1988). In this case, the point-biserial correlation coefficient used to measure the correlation between numeric variables (from features) and categoric variables (honest interview and falsified interview). To calculate the correlation for each item, the following formula can be used:

$$r_{pbi} = \frac{M_p - M_q}{S_t} \sqrt{pq} \quad (2)$$

where r_{pbi} is the point-biserial correlation coefficient; M_p is mean of whole-test for answering item correctly; M_q is mean of whole-test for answering item incorrectly; S_t is standard deviation for the whole test; p is proportion of correct answers; and q is proportion of incorrect answers. Table 3 shows the correlation values for each feature.

Table 3. Feature correlation to data falsification

No	Feature	Correlation	Description
1.	totaltime	0.539	Strong
2.	averageanswertime	0.506	Strong
3.	totalanswertime	0.499	Strong
4.	totalquesttime	0.489	Moderate
5.	averagequesttime	0.470	Moderate
6.	help	0.467	Moderate
7.	timeconditional	0.382	Moderate
8.	sumcontendit	-0.126	Weak
9.	forward	0.079	None or very weak
10.	backward	0.061	None or very weak
11.	sumnoncontedit	0.057	None or very weak
12.	numconditional	0.055	None or very weak

Table 3 data showed that 7 features had moderate to strong correlation with interviewer falsification. The rest had none or very weak correlation. Table 3 also shows that the features that have strong correlation with interviewer falsification were mostly associated with the duration such as *totaltime*, *averageanswertime*, *totalanswertime*, *averagequesttime*, *totalquesttime*, and *timeconditional*. Other feature shows that most likely, enumerators who falsified data rarely used help button. Then, each feature was evaluated using both unsupervised and supervised classification.

B. Unsupervised Classification

Unsupervised classification can be used when there is no information about labeling data. k -means clustering algorithm parts data into k clusters based on their features. k is the number of clusters. In this domain, the interviews can be divided into two clusters, namely honest interview and falsified interview. The tradeoff of unsupervised classification compared to supervised classification is its accuracy.

Using WEKA software [23], when associated with the label on the clustering tab using simple k -means clustering where $k = 2$, the result showed that the incorrect cluster instance of the cluster was 29.5%. This means that the accuracy obtained reached 70.5%.

C. Supervised Classification

Supervised classification can be used when there is information about labeling data for each instance. Compared to unsupervised classification, this supervised classification offers better accuracy.

The common method to evaluate the classifier is the k -fold cross-validation, where the data is divided into k parts where $k-1$ part is used for training data and 1 part for test data. The process then iterated as many as the number of k and the final accuracy is determined by computing average accuracy.

The classifier used in this study was logistic regression. Logistic regression is suitable as binary classifier and offers simplicity, speed, and scalability. The resulting accuracy in our case reached 88.5%.

7. Conclusion

It can be concluded that the proposed features were able to detect interviewer falsification of data. The most correlated feature to detect interviewer falsification was duration-related features such as total time of interview, respondent's answer time, and enumerator's question time. Another feature correlated to detect interviewer falsification was "help", where enumerators who falsify data tend not to use the help button and fill the data directly. For future work, it is suggested that a more advanced monitoring dashboard in the server side is developed.

8. References

- [1] Harrison, P. (1947). Public Opinion Quarterly. *A British view on cheating*, 11 (1), 172–173.
- [2] Badan Pusat Statistik Republik Indonesia (in press). Laporan Monitoring BPS RI. *Laporan Monitoring Kualitas Sensus Penduduk 2010*.
- [3] Okezone News. (2013). Kerja tak Beres 5 Petugas Sensus Jaktim Dipecat. Retrieved March 28, 2013, from <http://news.okezone.com/read/2010/05/19/337/334389/kerja-tak-beres-5-petugas-sensus-jaktim-dipecat>
- [4] Detik Finance. (2013). Kepala BPS Marah Petugas Sensus Penduduk Teledor. Retrieved March 28, 2013, from <http://finance.detik.com/read/2010/05/30/095823/1366067/4/kepala-bps-marah-petugas-teledor-sensus-penduduk>
- [5] Joglosemar. (2013). Membongkar Manipulasi Sensus. Retrieved March 28, 2013, from <http://revisi.joglosemar.co/berita/membongkar-manipulasi-sensus-14977.html>
- [6] Detik Finance. (2013). Kepala BPS: Ada saja Petugas yang Mengarang Data. Retrieved March 28, 2013, from <http://finance.detik.com/read/2010/05/30/175607/1366205/4/kepala-bps-ada-saja-petugas-yang-mengarang-data>
- [7] Schreiner, I., Pennie, K., & Newbrough, J. (1998). Interviewer falsification in census bureau surveys. *American Statistical Association (Survey Research Methods Section)* (pp. 491–496).
- [8] Koch, A. (1995). Fake Interview: Results of the Interviewer Control in ALLBUS 1994. *ZUMA-Nachrichten: Vol.36* (pp.89–105).
- [9] Kiecker, P. & Nelson, J. E. (1996). Do interviewers follow telephone survey instructions?. *Journal of the Market Research Society: Vol. 38* (pp. 161-176).
- [10] Hood, C., & Bushery, M. (1997). Getting more bang from the reinterviewer buck: Identifying 'at risk' interviewers". *American Statistical Association (Survey Research Methods Section)* (pp. 820–824).
- [11] Turner, C., Gribbe, J., Al-Tayyip, A., & J. Chromy. (2002). Falsification in epidemiologic surveys: Detection and remediation (prepublication draft). *Technical Papers on Health and Behavior Measurement, No. 53*. Washington DC: Research Triangle Institute.
- [12] Murphy, J., Baxter, R., Eyerman J., Cunningham D., & Kennet, J. (2004). A system for detecting interviewer falsification. *American Association for Public Opinion Research 59th Annual Conference*.
- [13] Lawrence, C., & Love, E. (2010). Characteristic of Falsified Interviews. *Section on Survey Research Methods – JSM*.
- [14] Bredl, S., Winker P., & Kotschau, K. (2008). *A statistical approach to detect cheating interviewers*. (Technical Rep No.39). University Giessen, Center for International Development and Environmental Research.
- [15] Schrapler, J. & Wagner, G. (2003). Identification, characteristics and impact of faked interviews in surveys - an analysis by means of genuine fakes in the raw data of SOEP. *IZA Discussion Paper Series* (pp. 969).
- [16] Schnell, R. (1991). The impact of counterfeit interviews on survey results. *Journal of Sociology 20 (1)* (pp. 25–35).

- [17] Crespi, L. (1945). The cheater problem in polling. *Public Opinion Quarterly* 9 (4) (pp. 431–445).
- [18] Shcafer, C., Schrapler, J.P., Muller, K.R., & Wagner, G.G. (2004). Automatic identification of faked and fraudulent interviews in surveys by two different methods. *ASA Section on Survey Research Methods* (pp. 4318-4325).
- [19] Bredl, S., Storfinger, N., & Menold, N. (2011). *A literature review of methods to detect fabricated survey data*. (Technical Rep). ZEU, Universitat Gieben.
- [20] Birnbaum, B., Gaetano, B., Abraham, D. F., Brian D., & Anna, R. K. (2013). Using Behavioral Data to Identify Interviewer Fabrication in Surveys. *The 13th Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2911-2920).
- [21] Wells, Don. (2013). Extreme Programming. Retrieved April 20, 2013, from <http://www.extremeprogramming.org/>
- [22] Judge, G., & Schechter, L. (2013). Detecting Problems In Survey Data Using Benford's Law. Retrieved March 28, 2013, from <http://www.aae.wisc.edu/ljschechter/benford.pdf>
- [23] Mark, H., Eibe, F., Geoffrey, H., Bernhard, P., Peter, R., & Ian, H. W. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations, Volume 11:1*.
- [24] Birnbaum, B. (2012). *Algorithmic Approaches to Detecting Interviewer Fabrication in Surveys*, doctoral dissertation, University of Washington.
- [25] Brown, J. D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*, Cambridge, Cambridge University Press.
- [26] Simmons, K., Mercer, A., Schwarzer, S., & Kennedy, C. (2016). Evaluating a new proposal for detecting data falsification in surveys. *Statistical Journal of The IAOS* (32) (pp. 327-338).



Yusep Rosmasnyah received a B.S. degree from Bandung Institute of Technology, Indonesia, and both the M.S. and Ph.D. degrees from the University of Surrey, U.K. He has been a researcher and faculty member at the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia. His current research interest includes mobile learning technologies and cybersecurity.
email: yusep@stei.itb.ac.id



Ibnu Santoso has completed Diploma studies in Computational Statistics Study Program, College of Statistics in 2007 and Master of Informatics at Bandung Institute of Technology in 2014. Worked in the Integration of Statistical Data Processing, Statistics Indonesia year 2008-2014. He began teaching at the Polytechnic of Statistics STIS in 2014 and since then has studied and taught courses in Algorithm and Programming and Data Mining and Knowledge Management.
email: ibnu@stis.ac.id



Ariq Bani Hardi is a Master's Student at the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia. He received a scholarship from the National Cyber and Crypto Agency (BSSN Republik Indonesia). His main research interests are related to the design and development of security of the mobile application, cybersecurity, and applied cryptography.
email: ariq.bani@bssn.go.id



Atina Putri is a Master's Student at the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia. Her main research interests are related to the design and development of web and mobile application and learning technologies.

email: 23517042@std.stei.itb.ac.id



Sarwono Sutikno is an Advisor at Corruption Eradication Commission (KPK Republik Indonesia). He is also an associate professor at the School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia. He received a Doctor Engineering degree from the Tokyo Institute of Technology.

email: ssarwono@stei.itb.ac.id