

Cloud Based Agriculture Safety Inspection with Multiple Standard Sources

Ming-Shen Jian, Hao-Yi Xu, Jenn-Feng Sheen, Yi-Ling Ye

Department of Computer Science and Information Engineering, National Formosa University,
No.64, Wunhua Rd., Huwei Township, Yunlin County 632, Taiwan
jianms@nfu.edu.tw, eco30246@gmail.com

Abstract— In this paper, the Cloud Based Agriculture Safety Inspection with Multiple Standard Sources is proposed which including the web crawler, pesticide residues inspection standard and algorithm. Each pesticide residues inspection result of individual agriculture based on the standard, and the mass spectrometer inspection results will be judged as qualified and unqualified. The pesticide residues inspection algorithm consists of quality management algorithms and the inspection filter. Each sample which is checked and inspected will be followed a set of quality management data. Based on the results of web crawler, the quality management data of the sample will be decided whether the sample conforms to government regulations or not. Due to the computing load, all the procedure of the proposed Pesticide Residues Inspection System will be established on virtual machine for individual sample. The cost and manual operation of system maintenance can be reduced.

Keywords— Pesticide Residues, Cloud Computing, Agriculture Safety, Virtualization, Web Crawler.

I. INTRODUCTION

Recently, a lot of food safety storm caused and happened. The government is also actively working to solve this problem, for fruit and vegetable products on the further strict checks.

Pesticides used in various crops are required to apply to the Bureau of Animal and Plant Health Inspection and Quarantine for registration. The related efficacy of pesticide and residue testing and other information should be also registered after the review of expert. Then, the relevant web information can be found in the Bureau of Animal and Plant Health Inspection and Quarantine Pesticide Information Service Network [1-2].

In addition to registration, the Food and Drug Administration in Taiwan has also established the “permissible standard” for pesticide residues online [3]. The pesticide registered should be lower than a certain concentration then cannot harm the people. To this end, the county government not only regular

sampling of pesticide residues in crops, but also check the pesticide allowable (permitted) states of the qualified agricultural products.

In the Taiwan norms, its pesticide use and norms are given respectively, 1. pesticide tolerance by the Food and Drug Administration. 2. Pesticide recommendation or not, for the Bureau of Animal and Plant Health Inspection and Quarantine. However, the farmers may use the pesticide which may not be recommended by the government. Therefore, for all types of pesticides should be checked and inspected.

However, considering the variety of pesticides and the mixed using of these pesticides, to inspect all the pesticides and compare with the standard or recommendation of government becomes a difficult issue. Today, in order to deal with huge information processing, the cloud computing with the large computing power is proposed. All the computing resource are virtualized and integrated into the resource pool. When the user needs to the computing resources, The corresponding virtual machine based on the cloud computing platform can be rapidly established [4].

However, due to the possibility that the source of information may be scattered across different websites or databases, to obtain all the information distributed at different sources is an important problem. For example, pesticide testing, information sources, norms and other information will be scattered in the Food and Drug Administration and the Council of Agriculture and other government units. Furthermore, the information may be changed dynamically. In order to keep the information in time when single event

triggers and other data updates, the network crawler[5] is a viable approach.

This paper will provide a customized service to achieve a set of lightweight systems for pesticide verification. The proposed system can not only automatically obtain the information via network, but also reduce the time of inspection and the mistake of manual operation. Based on the using of online crawler instant accessing information on the Internet, the inspectors do not have to manually check the government to regulate the pesticide maximal residue level (MRL) and recommended pesticide.

In addition, according to the inspection list, the possible used pesticide checking list is huge. Based on the on demand configured cloud virtual machine, the computing performance can be better and the processing time can be reduced.

The rest of this paper is organized as follows. Section 2 presents the related technologies. Section 3 shows the proposed Cloud Based Agriculture Safety Inspection with Multiple Standard Sources. In the section 4, the conclusion is presented.

II. RELATED TECHNOLOGIES

In this section, the related technologies such as cloud computing, web crawler, and the Chromatography-Tandem Mass Spectrometry will be introduced.

A. Cloud Computing

Cloud Computing was proposed to provide the huge computing resource via network for rapidly establishing and using [4]. All the physical computing resource will be integrated into the single resource pool. The virtual machines can be created corresponding to the client users' requirements based on network. Therefore, all the information or data will be delivered, processed, and exchanged by network. In other words, based on the on demand requirements and the configured virtual machines, the repeatedly computing works can be distributed and parallel processed. Therefore, based on the on demand defined configuration, the same configuration of the virtual machine can be rapidly and repeatedly established when users need. Furthermore, when the virtual machine is not needed anymore, the virtual machine can be closed

and all the computing resource will be returned to the resource pool.

B. Web Crawler

A web crawler [5] is a web application service that uses a program to search for data on the Internet and get relevant data to the local end for subsequent applications. Web crawlers are primarily targeted at the content of web pages on the Internet, including tagged HTML file content. As a result of the Council of Agriculture and the Food and Drug Administration is pesticide testing related information, are in the form of web pages to provide content. So the appropriate design of web crawlers, cyclical comparison and update the relevant web content has become a feasible way to collect scattered source data.

In addition, due to the provisions of pesticide testing, including the allowable amount of detection and recommended medication which defined by the inspection are Bureau of Animal and Plant Health Inspection and Quarantine and Food and Drug Administration, any replacement or changing of the list should be checked. In other words, the agriculture safety inspection system need to manually or eventually update and compare the contents of the list of checklist. If the sample is under the inspection process when the checklist list changes, to check the current date of the corresponding list of the inspection for verifying the pesticide test pass or not, the standard should be subject to the specification and the date of inspection. In other words, web crawlers for the repeatedly content gaining based on network is needed.

C. Chromatography-Tandem Mass Spectrometry

For the agriculture inspection, today, the Chromatography-Tandem Mass Spectrometry technology is used. A mass spectrum measures the masses within a sample. A mass spectrum is a plot of the ion signal as a function of the mass-to-charge ratio. These spectra are used to determine the elemental or isotopic signature of a sample, the masses of particles and of molecules, and to elucidate the chemical structures of molecules, such as peptides and other chemical compounds. [7]

III. SYSTEM ARCHITECTURE AND IMPLEMENTATION

The system architecture of the system is shown in Figure 5. Whenever the information of a new sample is uploaded through Web Services, a corresponding virtual machine with on demand configuration and embedded program will be established. In other words, each inspection sample will be assigned one virtual machine for further processing.

In the virtual machine, there will be the verification system including the agriculture safety inspection algorithm and the web crawler. The results and the information from the web crawler will be stored in the cloud database.

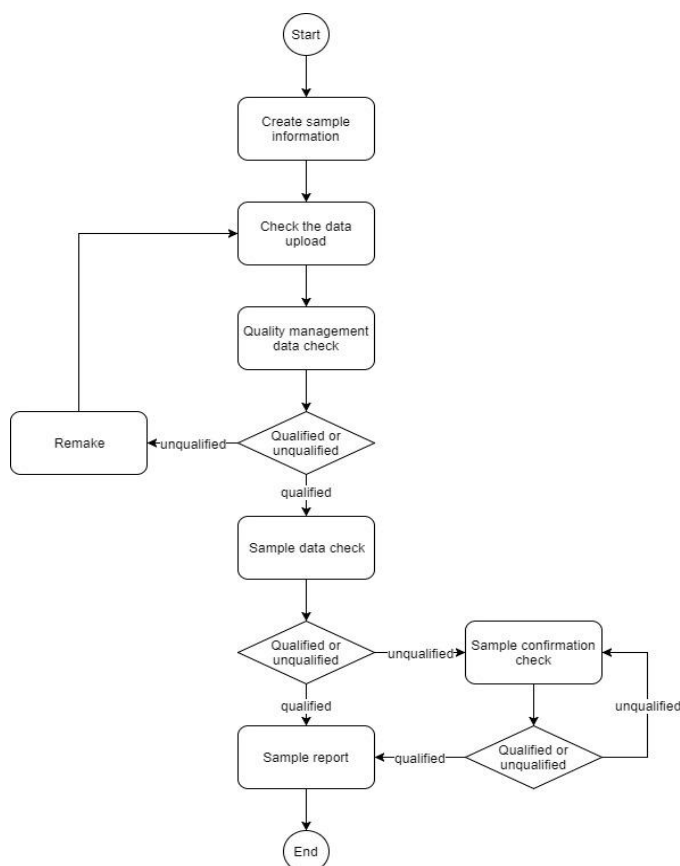


Figure 1. Flowchart of the agriculture safety inspection

The total amount of detected pesticides defined by the government is 310. Considering the recommended drug list of pesticides made by government, the mixed pesticide should be also considered. Therefore, the possible used pesticide in the mixed pesticide can be C_n^m , where m is the total amount of registered pesticide. In this paper, the value of m is 310. The n is the possible number of mixed pesticide which can be $1 \leq n \leq m$. Hence,

the worst case of the total possible amount or types of the mixed pesticide can be shown as:

$$\sum_{n=1}^{n=m} C_n^m$$

Especially, when the system inspects one agriculture sample, this computing workload will be needed to be done corresponding to the found the pesticide. In other words, the possible computing load will be increased and heavier if more types of pesticide are used. Considering the worst case, the computing workload will be huge. Therefore, in this paper, the cloud computing is used. Each virtual machine with the on demand configured proposed system will be established for each agriculture inspection sample.

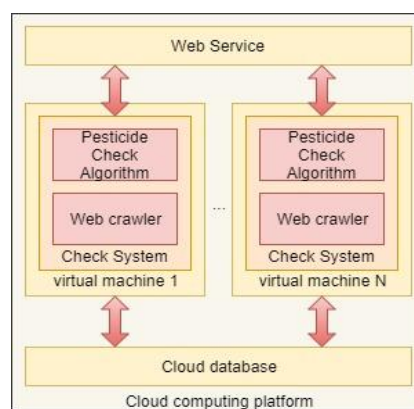


Figure 2. System Architecture

For example, under the prerequisites for the crop as a wax apple (*Syzygium samarangense*), if the Cyprodinil pesticide and Fludioxonil pesticide are used individually, the checking results from the Bureau of Animal and Plant Health Inspection and Quarantine will be "non-recommended". However, when the Cyprodinil pesticide is mixed with the Fludioxonil pesticide, the combination pesticide is listed as a "recommended medication" in the committee list by the Bureau of Animal and Plant Health Inspection and Quarantine, as shown in Figure 6.

檢體(2)代碼	檢出項目英文名稱	檢出值(3)	檢出濃度(ng/mL)	檢出倍率(5)	檢出濃度(ppm)	容許量	備註	許可證	推薦用藥
L41	Cyprodinil	0.994	100	1	0.1	1.0	合格	有許可證	不推薦
L69	Fludioxonil	0.984	100	1	0.1	1.0	合格	有許可證	不推薦
CYPRODINIL + FLUDIOXONIL									有許可證/推薦

Figure 3. Combination of Pesticide Check Results

In other words, not only a single pesticide needs to be checked about the maximal residue level and recommend by the government, but also a variety of

pesticides recommended for mixed pesticides should be also checked the recommend by the government.

Therefore, when the total number of examined samples is larger, the demand for computing resources is quite large. In this paper, "a sample corresponds to a virtual machine" is proposed. Utilize pre-set and pre-defined cloud virtual machine ISO image file as standard for establishing virtual machine with the proposed system.

A. Pesticide Check Algorithm

Pesticide check algorithm, mainly including quality management algorithms and the inspection filter, is proposed and followed in this paper. A quality management algorithm is used for quality management, data checking, and the inspection filter. In addition, the data checking is not only used in the quality management but also used in the sample data check.

```

if($BQCData[$Bser][$i][4]>5)
{
  $BQCans[$Bser]=$BQCData[$Bser][$i];
  //判断空白样品
}

if($SQData[$Sser][$i][4]>5)
{
  $SQCans[$Sser]=$SQData[$Sser][$i];
  //判断溶剂
}

if($FQCData[$Fser][$i][14]<85||$FQCData[$Fser][$i][14]>115)
{
  $FQCans[$Fser]=$FQCData[$Fser][$i];
  //判断含量初始值
}

if($TQCData[$Tser][$i][11]<1088||$TQCData[$Tser][$i][12]<3)
{
  $TQCans[$Tser]=$TQCData[$Tser][$i];
  //判断灵敏度
}

if($Qans[$Q123][16]>20||$Qans[$Q123][15]<60||$Qans[$Q123][15]>140)
{
  $Q12ans[$Q12I]=$Qans[$Q123];
  //判断浓度是否合格
}

```

Figure 4. Quality Management Check

Figure 7, the solvent and blank steps are used to check 310 types or kinds of pesticides. If the value of ppb field is greater than 5 ppb, the checking field will be marked as unqualified. Calibration curve is used to check whether the R% field value of 310 pesticides will be less than or equal to 85, or greater than 115. Then, according to the checking results, the corresponding inspection will be marked as unqualified or qualified. Spike blank is used to check whether the S / N field value of 310 pesticides is less than 10 or not. According to the checking results, the corresponding inspection will be marked as unqualified or qualified. Furthermore,

if S / N field value is less than 3, the inspection result will be marked as unqualified.

After the QR and QF conversion, the proposed system will check the R% and RPD% value. As shown in Figure 8, QR and QF will be checked: if the R% field value is less than 60 or greater than 140, the corresponding inspection will be marked as unqualified. QR and QF will be checked that: if the RPD% field value is greater than 20, in line with the corresponding inspection will be marked as unqualified. QR and QF need to convert the %Rec field of each pesticide into R% and RPD%, shown as follows:

$$R\%_{QR} = \frac{\%Rec_{Q1} + \%Rec_{Q2}}{2}$$

$$RPD\%_{QR} = \frac{|\%Rec_{Q1} - \%Rec_{Q2}|}{R\%_{QR}}$$

$$R\%_{QF} = \frac{\%Rec_{Q1} + \%Rec_{Q3}}{2}$$

$$RPD\%_{QF} = \frac{|\%Rec_{Q1} - \%Rec_{Q3}|}{R\%_{QF}}$$

When the quality management checking reports that there is one unqualified pesticide, inspectors will all have to start from scratch. Quality management checks which are qualified will enter the concentration check. All the 310 pesticides for each sample should be filtered. As shown in Figure 8, the concentration check filter for each sample is greater than or equal to 7 ppb for each pesticide.

```

if($jsonData[$data_i][4]>7)
{
  $arr[$count][4]=$jsonData[$data_i][4];
  //浓度合格筛选
}

$result[$i][7]= CatchTolerance($skindname,$scrotype,$conversion[$j][1]);
//网络爬虫获取MRL资讯
if($result[$i][6]>$result[$i][7])
{
  $result[$i][8]='不合格';
}
else
{
  $result[$i][8]='合格';
}
//输出浓度与MRL对比

```

Figure 5. Sample check

B. Web crawler

The web crawler is mainly used to obtain the information from the two websites: the Food and Drug Administration, and the Bureau of Animal and Plant Health Inspection and Quarantine. Respectively, recommended medication and the maximal residue level of pesticide. Respectively, recommended medication and the maximal residue level of pesticide. The crawled data will be

compared with the pesticide concentration of the sample.

Recommended medication is searched to obtain two parts of the information. The first is the license checking and the second is the recommended crop. License checking is to check whether the founded pesticides in the inspection are still validly registered by any manufacturer. Pesticide recommended crop is to check whether the sample of pesticide is found in the list of recommended pesticide used to crops. In this step, the checking is a pure string matching. For example: corn alias with white corn, corn is miscellaneous grains. So the proposed system will use the network crawler to obtain the list of pesticide recommended crops which let inspectors choose the correct alias or category collectively. Finally, the corrected results by the inspectors will be stored in the cloud database. Next time the same sample uploaded into the system can be managed well.

Each batch of sample inspection, from the beginning of the sample inspection to the output report, the working day is up to 7 working days. The government may change the maximal residue level of pesticide or recommended information. In order to ensure the quality of the audit, the virtual machine must re-acquire the maximal residue level of pesticide and the information of the recommended medication from time to time and before the report.

After the Internet crawler accessing information, the qualified pesticide comparison can be directly formatted as the report. Otherwise, inspectors must start to confirm the inspection process. The results must be checked by a professional inspector. Finally, the report can be produced.

C. Implementation

Figure 9 is the operator page which uploads the batches to the cloud database by uploading the file. The Farmer Information field represents the information added to each sample. The dilution ratio field represents the sample weighing information. The Quality Management Check field represents the quality management check. Sample check includes web crawler, unit conversion, comparison of web crawler results, and sample checking.

Figure 6. Operator Page

Figure 10 shows the results of the network crawler's acquisition of MRL. The resulting page format is ashx. Therefore, the result is presented in the form of an array, where tolerance is the pesticide maximal residue level.

Figure 7. MRL Results

Figure 11 is the result of the web crawler getting the recommended medication, and the web page format is aspx. In the web crawler access time need to pass the hidden field in order to obtain the information required by the system. Because the web crawler program language is used PHP, so the echo display, the html tag will be presented in the form of web pages.

Figure 8. Recommended Pesticide Results

Figure 12 shows the results of the system, MRL, notes, permits, and recommended pesticides which comparing the results of the web crawler information searching. The form below is an alias correction. Currently, in the verification of the

system, the web crawlers obtain the corresponding government information and the list about less than half a minute. The system web crawler results will be stored in the cloud database for subsequent work. In addition, fortunately, the current types of pesticides used by farmers are not much. In other words, by using the proposed Cloud Based Agriculture Safety Inspection with Multiple Standard Sources, if there is no manual mistake, the proposed system can almost immediately get the final report.

儀器代碼	檢出項目英文名稱	濃度(%)	濃度(μg/mL)	檢出濃度(μg/mL)	檢出濃度(ppm)	容許量	備註	許可證	推薦用途
L3	Athamycin	0.974	100	1	0.1	0.01	不合格	有許可證	不推薦
L4	Acephate	0.999	100	1	0.1	0.5	合格	有許可證	不推薦
L10	*Ametyn	0.998	100	1	0.1	0.01	不合格	有許可證	不推薦

☐ 蔬菜	☐ 十字花科蔬菜類	☐ 小葉菜類	☐ 豌豆	☐ 十字花科小葉菜類	☐ 芥藍
☐ 甜料小葉菜類	☐ 甘藷	☐ 豆薯	☐ 樹薯	☐ 山藥	☐ 十字花科根菜類
☐ 牛蒡	☐ 甜料根菜類	☐ 芋	☐ 甜料根菜類	☐ 番薯	☐ 甜子
☐ 枸杞	☐ 瓜菜類	☐ 菜蓴	☐ 葡萄	☐ 桑樹	☐ 柑桔類
☐ 花木	☐ 菊	☐ 茉莉	☐ 茶	☐ 硬布子	☐ 藥料作物
☐ 作物園	☐ 雜糧	☐ 甘蔗			

Figure 9. System Check Results

IV. CONCLUSIONS

Based on the proposed Cloud Based Agriculture Safety Inspection with Multiple Standard Sources, the information from the government can be updated by web crawler in time. According to the requirements of government, the huge computing workload can be managed by the proposed Cloud Based Agriculture Safety Inspection. The inspection time and the mistake of manual operation can be reduced. Furthermore, the worst case of computing can be managed well.

REFERENCES

- [1] https://pesticide.baphiq.gov.tw/web/Insecticides_MenuItem5_3.aspx
- [2] https://pesticide.baphiq.gov.tw/web/LawsDetailView.aspx?law_code=1&sn=103
- [3] <https://consumer.fda.gov.tw/Law/PesticideList.aspx?nodeID=520>
- [4] Ming-Shen Jian, Jun-Yi Wu, Jing-Yan Chen, Yue-Jyun Li, Yi-Cheng Wang, Hao-Yi Xu "IOT base smart home appliances by using Cloud Intelligent Tetris Switch," Proc. of IEEE ICACT, pp. 589–592, Feb. 2017.
- [5] Chandni Saini, Vinay Arora, "Information retrieval in web crawling: A survey," Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on.
- [6] www.fda.gov.tw/tc/includes/GetFile.ashx?mID=19&id=64555
- [7] Downard, Kevin "Mass Spectrometry - A Foundation Course". Royal Society of Chemistry, 2004



Ming-Shen Jian is an assistant professor of Dept. Computer Science and Information Engineering at National Formosa University. Ming-Shen Jian's current research interests are in the area related to IOT application, Big Data, optimal solution, Intelligent System, and cloud computing. He received B.S. degree at Electrical and Control Engineering in National Chiao Tung University and Ph.D. at the Department of Computer Science and Engineering of

the National Sun Yet-Sen University in Taiwan, 2007, investigating resource management in 3G mobile communication systems.



Hao-Yi Xu is a master degree student of Dept. Computer Science and Information Engineering at National Formosa University. His current research interests are in the area related to Big Data and cloud computing. He received B.S. degree at Computer Science and Information Engineering at National Formosa University, 2016.



Jenn-Feng Sheen BIOGRAPHY:

Dr. Jenn-Feng Sheen obtained PhD degree in analytical chemistry from National Taiwan University (2005). In 2007, he joined department of biotechnology of National Formosa University (Yulin County, Taiwan) as an Assistant Professor. Before academic, he has been worked in a contracted laboratory company, Mithra Bioindustry Co. (Taipei County, Taiwan) for 10 years. He was served as laboratory director for both analytical and drug abuse laboratories (2002-2007), and R&D manager (2005-2007). He currently serves as the director of the analytical service center of agriculture and biotechnology products in National Formosa University (2012- at present).

RESEARCH INTERESTS:

Dr. Sheen research is mainly focused on the analytical method development for drug substance, metabolites and agricultural compounds by using the chromatography tandem mass spectrometers. The applications were mainly aimed in the pharmaceutical and food safety analyses, such as the preclinical studies, antibiotics or pesticides residues. The improvements in combined preparation methods, difficult compounds and LC-MS interfaces are especially interested.

SELECTED PUBLICATIONS:

A discharge adaptor interface for use in liquid chromatography/mass spectrometry, Sheen JF, Ho YH. Rapid Commun Mass Spectrom. 2011 Dec 30; 25(24):3675-82

Application of pentafluorophenyl hydrazine derivatives to the analysis of nabumetone and testosterone in human plasma by liquid chromatography-atmospheric pressure chemical ionization-tandem mass spectrometry., Sheen JF, Her GR., Anal Bioanal Chem. 2004 Dec;380(7-8):891-7.

Analysis of neutral drugs in human plasma by fluoride attachment in liquid chromatography/negative ion electrospray tandem mass spectrometry., Sheen JF, Her GR. Rapid Commun Mass Spectrom. 2004; 18(17):1911-8



Yi-Ling Ye

National Formosa University, Taiwan

Yi-Ling Ye was born in Tainan, Taiwan, in 1971. She received the master and the Ph.D. degree of Immunology from the National Taiwan University (NTU), Taipei, Taiwan, in 1995 and 2002, respectively. She worked with Dr. Bor-Luen Chiang studying the immune mechanism and therapeutic strategies for autoimmune and asthma. Since 1999, Yi-Ling Ye worked as assistant professor in the Department of Medical Technology, Chung Hwa University of Medical Technology in Tainan, Taiwan. Her teaching interests include medical microbiology, immunology and cell biology. In 2008, she worked as associated professor in Department of Biotechnology, National Formosa University. She also was the chief of Education and Extension Division, Analytical Service of Agriculture and Biotechnology Product, National Formosa University Recently. The major research topics include the microfluidic microfluidic disk for cell separation and manipulation of the immune-modulator effect by herbal medicine.