Health Security: Towards a Hybrid Information Network for Coordinating Foodborne Disease Outbreak

Liaquat HOSSAIN a, Muhammad Rabiul HASAN b, Shahadat UDDIN c

a Professor, Information Management,
Division of Information and Technology Studies,
The University of Hong Kong

b Honorary Professor, Complex Systems, School of Civil Engineering,
Faculty of Engineering and IT,
The University of Sydney, Australia

PhD Researcher, Complex Systems Research Centre & Project Management Program
Faculty of Engineering & IT,
University of Sydney, Australia

c Lecturer, Complex Systems Research Centre & Project Management Program
Faculty of Engineering & IT,
University of Sydney, Australia

Abstract. Foodborne disease outbreaks are increasingly seen greater concern to the public health authorities. It has also become a global research agenda to identify improved pathways to coordinating outbreak detection. Furthermore, there is a significant need for timely coordination of potential foodborne disease outbreak in order to reduce the number of infected individuals and the overall impact on public health security. Therefore, this study aims to offer a new approach for coordinating foodborne disease outbreak. Firstly, it identifies the current coordination processes and challenges together with brief analysis of the problems. Secondly, it explores the social media surveillance strategies, usage, and their esteemed power of information dissemination. Finally, based on social media and hierarchical coordination approach, a hybrid information network model is proposed for improving coordination of outbreak detection.

Keywords: Public Health Security, Hybrid Information Networks, Foodborne Disease Outbreak, Coordination, Coordination Process, Social Media Analytics.

Introduction

In recent years, we are observing the increased frequency of foodborne disease outbreak, which continues to be a major public health issue [1]. The annual number of foodborne outbreaks reported to many national health authorities has increased sharply in recent years. The growing infection rate of foodborne illness (caused by consumption of contaminated food, pathogenic bacteria, viruses, or parasites) resulted millions of people being infected resulting in large number of deaths. Each year an estimate of 76 million illnesses and 5000 deaths are caused by foodborne outbreaks in
the United States alone [2] [3], however, this rate has been gradually decreased over the years 2009-2010, [4] [5]. Australia records approximately 5.4 million cases of foodborne disease per year, which costs A$1.2 billion annually [6]; and in England and Wales, foodborne disease causes 2,366,000 cases, 21,138 hospitalizations, and 718 deaths each year [3]. Based on the reported number of cases in the U.S, England, and Australia, it is easy to assume that the burden of foodborne disease outbreak is probably the same or even worse in many other countries in the world.

Although foodborne diseases are commonly affecting people, but only a small portion of these infections are formally reported to health authorities mainly because of ineffective coordination processes and poor participation by the public. As the nature of foodborne disease outbreak changes over time due to the change in food production, consumption, and distribution process, the surveillance process of foodborne outbreak detection needs to be improved [7]. This paper proposes a hybrid information network model suggesting an improved way of coordinating foodborne disease outbreak detection at earlier stage. The objectives of the paper are to identify the hierarchical coordination processes, challenges and their impact towards public health security, to explore emerging social media based surveillance for early detection to foodborne outbreak and finally, to understand how both hierarchical coordination network and social media based approach work together as a hybrid network for foodborne disease outbreak detection.

1. Definition of Coordination

For the broader definition of coordination, according to Malone & Crowston [8], it includes almost everything that happens during all session when actors work together, setting goals, performing all related activities in order to achieve the goals. A unique feature of effective coordination is to be able to focus explicitly on the elements of coordination. Based on the definition above, coordination can be seen as act of managing interdependencies among activities accomplished in order to achieve a goal [8].

There are some primary components of coordination such as goals, activities, actors/tools, resources, and interdependencies which needs to be integrated within the context of preparedness and response to coordinating foodborne outbreak. The generic processes of these components such as identifying goals, setting control plan, sequencing activities, assigning activities to actors, allocating resources, and synchronizing activities, make coordination easily understandable and applicable to diverse context [8].

2. Significance of Coordination in Foodborne Outbreaks

For a system to efficiently detect or prevent foodborne disease outbreak, there should be established coordination. Therefore, the process of data gathering efforts and information sharing networks related to food contamination and foodborne disease need to be effectively coordinated. For instance, all processes of early detection and surveillance of disease outbreaks should be easily completed in order to achieve common goals.
Based on currently available literature, an effective coordination is not found in foodborne disease surveillance where a hazard is unknown [9] [1]. Therefore, this gap leads to a very weak evidence, potentially may lead to uninformed decisions for food safety and public health [1]. As the number of reported cases of foodborne outbreak is gradually increasing and number of people affected is also rapidly growing, it is now a great concern for the health authorities globally to find a potentially very effective coordination strategy in order to reduce the number of infections and protect public health.

![Figure 1. Foodborne (Salmonella) outbreak trend in Australia](image)

Looking at foodborne outbreak trend in Australia particularly for salmonella outbreak (figure 1), it suggests a gradual increase since 1991. The raw data is collected from National Notifiable Diseases Surveillance System, Australia and by performing time series analysis of the data, we are able to show a linear trend that is not only increasing in recent years but over decades. Therefore, it is indeed significant to have an effective coordination network for foodborne disease outbreak management.

2.1 Current Coordination Processes and Approaches

Looking at U.S. coordination systems for outbreak detection, primary data is held when physicians and hospitals as part of surveillance networks collect specific data and/or submit samples for laboratory tests. Local public health agencies also collect primary data whereas public health workers examine patents and collect additional information from infected individuals after receiving reports from clinical health care system. If there is a single case which sufficiently requires raising a notification or if multiple cases of disease infections found, the front-line clinicians may perform the functions of the outbreak detection. There are three other components which also involve in detecting outbreak: (i) local or state health agency, (ii) academic or professional organizations, and (iii) federal government [10].
The state with legal power and authority is the primary entity for collecting cases and finding pattern. The Local public health agency derives its police power from the state depending on the state’s constitution, regulation, and custom [10]. This public health agency also can be administered by municipal or state government. The primary data are submitted individually or in batches to the local public health agency from the clinical health care system through phone calls, email, or faxes. Additional data may also be received from individuals. Formally recorded reports are sent to Centers for Disease Control and Prevention (CDC) through the National Electronic Telecommunications System for Surveillance (NETSS). Academic or professional organizations who involve with research or public health surveillance activities, also participate in outbreak detection (e.g. real-time outbreak and disease surveillance system, University of Pittsburg [10]). The federal government (includes CDC, and highly specialized laboratories and networks) contributes to outbreak detection by providing specialized research facilities, experts, and advanced laboratories. Overall, public health workers from different level of authorities frequently communicate about outbreaks occurring and the potential outbreaks. This both formal and informal communication and data sharing necessarily exist among different authorities because infectious disease outbreaks spread beyond the governmental boundaries [10].

![Image](image.png)

**Figure 2.** The follow of outbreak information coordination at NORS, CDC [11]

Outbreaks are reported to CDC using the National Outbreak Reporting System (NORS), which gathers information of outbreak. This reporting includes the date and location of the outbreak, the number of people infected and the related pathogen that caused the outbreak. The data gathered from outbreak investigations are checked, standardized, and analyzed by CDC experts to provide information for necessary decision and learning. Figure 2 shows the coordination processes at NORS: when people are exposed to a pathogen, get sick and seek treatment, then the health department is notified of a possible outbreak. The health department conducts an outbreak investigation. Afterwards, health department enters outbreak information into NORS. Then CDC checks data for accuracy and analyzes. Finally, data are summarized and published [11].

Looking at multistate disease outbreak coordination, a guideline for improving coordination and communication for foodborne outbreak investigation has been developed by the outbreak coordination workgroup for the National Food Safety System [7]. This is a usual and recognized reference on interagency collaborative activities in multistate disease outbreak inspections. Representatives from federal, state,
and local environmental, agricultural, health, and regulatory agencies; and experts from the aspects of epidemiology, laboratory, and environmental investigations, all are in the coordination workgroup to contribute together towards common goals. This workgroup holds frequent conference calls, at least twice in a week between CDC and all states involved in outbreak investigation [7]. If any issue or event occurs in one or few states, an email or small call can be used depending the agenda and regulatory agency. The interaction may take place between CDC and federal food safety regulatory agency at the stage of identification of contamination event. However, notification of the outbreak should take place as early as possible once foodborne transmission is suspected. For press inquiries about the foodborne outbreaks, they go through the Office of Communication, Division of Media Relations. Majority of inquiries are handled by professionals at the Office of Communication and some are referred to a designated CDC epidemiologist or to the specific regulatory agency [7].

According to International Health Regulations 2005 Decision Instrument [12] events are detected by national surveillance system, if any case of the following diseases found: Smallpox, Poliomyelitis due to wild-type poliovirus, Human influenza caused by new subtype, Severe acute respiratory syndrome. The event should then be notified to the World Health Organization. If any event of potential international public health concern, or if a case of the following diseases found such as Cholera, Pneumonic plague, Yellow fever, Viral hemorrhagic fevers, West Nile fever, or other diseases of special national or regional concern, then the following queries should be applied in decision algorithm: (a) is the public health impact of the event serious? (b) is the event unusual or unexpected? (c) is there a significant risk for international spread? (d) is there a significant risk for international travel or trade restrictions? From the above criteria, if any two criteria are found, then the event should be notified to the World Health Organization [12]. These are the current practices of coordination processes among renowned public health organizations.

2.2 Challenges in Coordination for Foodborne Outbreaks

There are many challenges exist in current coordination systems, and operational processes of foodborne outbreaks at local, state, and national level.

Firstly, Difficulties in identification of Foodborne Pathogens: the identification of new biological pathogens indicates that there is possibility of having many agents of foodborne disease remain unknown and yet to be discovered. According to the Centers for Disease Control and Prevention (CDC), the unknown agent either to be identified or not detected in reported illness [13]. Based on the latest CDC data release (2009-2010, [14]) on foodborne disease outbreaks, total number of reported cases of illness is 29444, hospitalizations is 1184 and total death is 23. Thus, the unknown agents issue further complicates the estimation of the number of cases that are not reported to officials, due to perhaps negligence of infected individuals to seek medical treatment, physicians do not recommend for laboratory tests, or medical facilities do not report cases of illness and pathogens altogether [13].

Secondly, Problems in Population Surveillance: Effective way of information sharing is a necessary factor for planning interventions to reduce foodborne illnesses. The CDC has developed two systems: FoodNet and PulseNet that are providing required information in an efficient way. FoodNet is an active system of disease surveillance dedicated to provide information about foodborne disease outbreaks. However, it covers only 10% of American population [13]. This is a big weakness of
existing FoodNet system. The purpose of this system is to monitor population and identify sources of foodborne disease outbreak by collecting data on laboratory based confirmed occurrence.

Thirdly, Problems in PulseNet Surveillance: PulseNet communicates with a number of agencies and programs (such as National Antimicrobial Resistance Monitoring System (NARMS), FoodNet an active surveillance system for foodborne disease, and OutbreakNet which is used to enhance communication during foodborne outbreaks) involved in activities related to foodborne disease surveillance. However, it faces great challenge of insufficient funds, staffing shortage, isolate report submission by clinical laboratories, lack of standard communication with partners, and lack of quality exposure information [15].

2.3 General Analysis of Coordination Problems

In the traditional coordination system, food borne disease outbreak reporting at state or national level is based on number of laboratory reported cases, related to a certain group of the population may be on a national basis. There are many cases of foodborne disease that never reach the laboratory tests or health care system. Thus, the number of the cases reporting at the health care system actually represents a fraction of the total amount of infected cases [1]. There are also many cases that individuals do not feel the necessary to visit physicians and report any symptom of disease.

The existing coordination processes cannot effectively provide surveillance of majority of population in a timely manner. In most of the cases, the foodborne outbreak is reported after getting laboratory confirmation with few similar cases. PulseNet, FoodNet, HealthMap surveillance are facing many internal problems such as insufficient funds, staffing shortage, isolate report submission by clinical laboratories, lack of standard communication with partners, limited capacity of coordination, even sometimes less than one quarter population is under surveillance networks. The surveillance structure in many countries is weak and the dimension of disease outbreak is still unknown [1]. Looking at the developing countries, laboratory resources and skills to identify foodborne pathogens are not commonly found, even for the developed countries, laboratory based surveillance is not well developed and having lack of simple and reliable laboratory confirmed report [12] [1]. Therefore, a timely successful coordination structure of foodborne disease outbreak still remains unknown.

3. Social Media Analytics and Outbreak Surveillance

Social media analytics has become a popular term to researchers [16] [17]. The media analytics is to develop and evaluate tools and frameworks in order to collect, monitor, and analyze social media data, based on specific requirements from a selected application. Social media analytics research serves many objectives including facilitating conversations and interaction between online communities and extracting useful patterns and intelligence to serve active contributors in ongoing communications [18].

There are some challenges associated with social media analytics. First, it contains an enriched dataset or metadata, which have been haphazardly treated in data- and text-mining literature. Second, it is an example of human-centered computing with their own unique emphasis and strategy on social interactions among users. Therefore,
context-dependent user profiling and user-needs presentation as well as many types of human computer interaction considerations must be reconsidered. Finally, social media has issues of semantic inconsistency, conflicting evidence, lack of structure, inaccuracies in data and dynamic data volume with rapid increment [18].

Traditionally, CDC relies on outpatient reporting and laboratory test results nationwide. The confirmation of outbreaks by CDC takes about 2 weeks after they occur, where social media can create immediate concerns and it can detect earlier than traditional methods [19]. To identify Flu Trends, CDC collaborated to Google and Google launched Google Flu Trends, a website1 that allows people to find flu-related search activity against reported illness displayed graphically on a map. The CDC monitors Google Flu Trends and considers as a potential source for early detection of flu outbreaks. It is important to understand that online search behavior might have no assurance on whether an outbreak is really occurring. In addition, Google Flu Trends system tries to have media bias by modeling set of search terms over time to see which ones remain frequent and stable [19].

Comparing to Google, Social media is more open where users do not use search terms. For example, in Twitter (a rapid growing microblogging platform), each tweet is 140 characters long and has enough contextual information rather than what is offered by search terms in “the Google Flu Trends” site. Corley et al [20] have conducted a study based on social medial surveillance where blog data are used for the identification of Influenza trend. Blogs are classified into a ranked tier structure based on specific focus at how successful the blog is in terms of creating and spreading an element (e.g. video, image, text) over the internet. The results extracted from blog posts are found significantly correlated with the CDC ILINet data.

3.1 Use of Social Media Surveillance for Disease Outbreak

Although, it is not yet common that social media cover public health threats, some researchers already have used social media as platforms for retrieving health data. In a study, Schmidt [19] found that blog and Facebook postings cover two environmental health issues. The results from social media data well matched with what appeared in official reports on the same topics. The researcher concluded that what most people were saying was fairly accurate compared to official reports. Many researchers [21] [20] [22] [23] have validated the results of social media data with the traditional health reports. All of those researchers found strong correlation between social media data and traditionally reported data in terms of disease outbreaks.

Quincey and Kostkova [24] have conducted a preliminary study on the use of Twitter for early warning of disease outbreak detection. It was for 7 days observation and the program was started at 14:00 on Thursday 7th May 2009 and continued until 14:00 on Thursday 14th May 2009. During this seven days period, there were a total of 135,438 tweets posted by 70,756 unique users that contained the word “flu”. The researchers used a set of common keywords to find flu related tweets and concluded that twitter or social media can be used to detect potential disease outbreak. This is not the end yet; the coordination problems are still there and could not be solved just by social media surveillance alone. The traditional (hierarchical) coordination approach would remain and researchers need to find a better solution.

1 http://www.google.org/flutrends/
3.2 Hybrid Information Networks: A Way Forward

A review of surveillance systems for outbreak detection has incorporated analysis of a large group of surveillance systems in U.S. and gathered information on the system type, data origin, data type, disease scope, geographic scope, and main partners of each system [10]. The outcome of the analysis showed how they relate and depend on each other for the outbreak detection. The results identified the components of the formal coordination network which include clinical healthcare system, state, the federal government, professional organizations, and collaborating governmental organizations. The state here includes all the local public health agencies which receive data from clinical healthcare system via phone calls, faxes, or electronic data exchanges. After that the reports are sent to CDC at the federal level via the National Electronic Telecommunications System for Surveillance (NETSS) for the purpose of data aggregation. The standardization of the process of data collection is done through the National Electronic Disease Surveillance System (NEDSS). The federal government has role in outbreak detection by providing specialized research facilities, experts, and advanced laboratories [10]. Similar surveillance process is also found in developing countries. Indonesia employs the Early Warning Outbreak Recognition System (EWORS) that includes patients to be presented at hospitals, clinics or emergency departments. Data are collected through standardized questionnaire and filed out in EWORS system for daily analysis. The counted signs and symptoms are combined and compared to baseline counts by automated algorithm to see the infectious disease of national importance. If an outbreak is suspected Ministry of health take initiative for investigation [12]. Germany has employed electronic surveillance system for infectious disease outbreak. All locally identified notifiable diseases are verified by local health departments based on national case definitions and the case reports are sent electronically through the state health departments to the national surveillance unit [25].

For the improvement of surveillance method, social media based approach is a part of new generation of public health surveillance [26] [19]. It is a supplementary method to the existing public health surveillance networks, focuses on monitoring of infectious diseases by leveraging data from internet news, blog posts, and other web and social media [27] [28] [26]. Researchers from Harvard University and the University of California have used social network analysis for flu outbreak prediction and found that social media can detect outbreak earlier than traditional tracking methods [19]. There are significant correlations found between social media data and traditional reporting data by the researchers [28] when they have tested for influenza-like illness using social media surveillance. The methodology was to gather blog posts, perform content analysis and data mining to compare with traditional reporting data. Through link extraction identify the web and social media communities and release information on outbreaks. All these findings from the both hierarchical and social media surveillance approaches have derived the following Hybrid Information Network model for disease outbreak.
In this model, there would be combination of hierarchical networks and social media networks in order to work together towards foodborne disease outbreak detection. This model has demand to be developed and implemented for the greater success in public health security. Today’s world is talking about the use of social networks for many scientific innovations, and using this media for detecting potential disaster outbreak is one of them. For an effective surveillance approach, social media is getting popularity and being tested by quite number of public health researchers. There are some recent applications of social media use could be found in the field of natural disaster and emergency management research [29] [30] [31]. However, they are not fully applying hybrid approach in emergency situation, even researchers [29] argued to bypass or eliminate formal liaison structures used previously for knowledge sharing among different agencies during a disaster. At this stage, Hybrid Information Network (HIN) for coordinating foodborne outbreak is seems to be new approach as this model employs both formal and informal networks as equally significant for decision making and it could logically provide better outcome.

In this structure (figure 3), resilient foodborne outbreak management will be tested against hierarchical and social media based network structure. It would also test whether hierarchical approach influences social media based approach and vice versa. The same network measurement would be used for testing both formal and informal network structure in this model.

For hierarchical coordination, primary data is held when people get sick and seek treatment, thus, physicians and hospitals as part of surveillance networks collect specific data from the patients or laboratories. Once physicians collect the data, health department is notified of a possible outbreak and that initiates outbreak investigation. The data gathered from the investigation are checked, standardized, and analyzed by specific experts to provide information for outbreak related decision [7] [10] [11].

On the other hand, when people get infected of foodborne disease, it quickly gets spread among friends within his circle in social networks. While many are getting affected in the same way, same time, and same area, the information is rapidly disseminated and gets attention of the public health authorities for potential foodborne outbreak. At this stage, in social media based coordination, publicly shared data on health and wellbeing issues in Facebook, Twitter or blogs are aggregated. For instance, Tweets are collected through the streaming API of Twitter [31], and similarly for Facebook and other social media data. System administrator, in this case needs to define filters in the form of keywords, set up geographic bounding boxes and a user account for which all matched tweets will be gathered as a stream. In the next step, all tweets would be categorized based on each of the keywords and find the rate of occurrence (e.g. % of total tweets) against all those keywords. If any keyword has higher frequent value than the median value, this means there is certain possibility of
occurring disease outbreak at certain level of confidence. This is one way of exploring possibility of foodborne outbreak which this study has observed. There could have few other ways to analyze tweets for better prediction and some approaches might be more appropriate, thus, need to be examined. There may need to implement a decision support system to generate outbreak alert and reduce human intervention.

Finally, after obtaining some prediction outcome from social media data, the experts would compare them with the traditionally (through hierarchical networks) captured data and reach a decision. Overall, data received from formal ( hierarchical) and informal (social media) sources could make easier for health authorities to manage foodborne outbreak detection.

In order to implement this model, some principles need to be realized by health agencies at ground level. Early detection of disease outbreak expectantly be possible by this model, however providing response to the event is important too. Some might think that impact of particular disease outbreak may not be that high. Thus, convincing health agencies for alert generation is very significant and they should not wait longer to actualize the impact. For a case of outbreak caused by unknown pathogens, it may be difficult to estimate consequences. Thus, establishing principle of taking action for any potential outbreak detection would be considered as preparation of implementing this hybrid model. For technical implementation, at the preliminary stage, any of the local health agencies can set up a server to store stream tweets from specific location (as explained earlier in this section) and system administrator there can perform necessary textual data analysis in order to reach a decision. This is important to make sure that the particular health agency has traditionally reported outbreak data ready at the time of interpreting findings from social media to compare and have a sound decision on potential disease outbreak.

4. Brief Findings

This study has incorporated some sample data collection from twitter and brief analysis of the tweets in order to support the feasibility of the proposed model. There were five sets of data collected at different time slots. Total 53440 (set1 11350, set2 13776, set3 9011, set4 10444, and set5 8859) tweets were analyzed against thirteen keywords (sick, food, pain, alarm, tired, weak, poison, dead, disaster, medicine, burger, lunch, dinner). Among those keywords, six (sick, food, pain, tired, weak, poison) were given more attention at the analysis phase. As this was a sample study, thus, selecting keywords did not follow any particular method. The result shows that keywords “pain” and “poison” have occurred at a higher rate in particular period (time slot-1) compared to other different times. Similarly “tired” and “weak” also have higher rate of occurrence during the time slot-4. The remaining two keywords “sick” and “food” indicated higher occurrence rate during the period of time-slot-5 and time-slot-3 accordingly.

These results could be interpreted in the process of decision making and claiming potential outbreak occurrence. There is little challenge in choosing set of keywords and the users in social media sometimes use their own way of spelling words and expressing opinion or feelings with mix-language which may be influenced by socio-technical culture. Overall, the findings at least show the social media data (in this case tweets) could significantly offer positive supports towards decision making for potential disease outbreak alert.
5. Discussion

An understanding of foodborne outbreaks, frequencies of occurring, and their infection rates in some parts of the globe have been briefly discussed in this paper. Since last few decades, foodborne disease outbreaks remain one of the top public health issues. Thus, there are many surveillance systems and approaches exist. However, having a surveillance system does not mean that number of foodborne illness could be reduced easily, rather it depends on how effective an existing surveillance system is and how timely it can detect a potential outbreak. Having these questions in background, this paper tried to identify ways of meaningful coordination for foodborne outbreak detection. It is not a surprise, purposely social media based surveillance came into focus besides the traditional coordination approaches. While considering a crisis moment of disease outbreak, people immediately share their knowledge on the source of resources and capacity of local help in social media. This is a power of social networks today where every people acts as an actor of information dissemination process, and it works very rapidly as every single geographic location is under the coverage of a social media. Thus, hybrid information network perceived to have high significance where social media could play vital role together with the hierarchical coordination networks for effective foodborne outbreak surveillance.

With a successful implementation of Hybrid Information Networks, related information at very early stage of potential foodborne disease outbreak can be gathered and analyzed timely. The coordination process, alert, communication and feedback on any kind of disease related outbreak would be faster. Therefore, the total number of infections and epidemic could significantly be reduced due to quick time response. General public could gain more health security and national government could expectantly be able to save millions of dollar from the budget of public health supports and disease disaster management.

6. Conclusion

Complexity of coordination is widely depicted in the first half of this paper especially throughout the discussions on current coordination processes, challenges, and analysis of coordination problems in regards to foodborne disease outbreak. Later, social media analytic issue came into discussion, as not an alternative, a vital complementary tool to be used for reducing complication in coordination. Finally, the paper draws public health researchers’ attention in order to mitigate coordination dilemma through hybrid coordination approach as this study deals with a decision making scenario for disease outbreak surveillance. The overall idea, data evidences, and scenario analysis presented in this study would support the health managers or surveillance specialists to take decision on the application the hybrid method for improved disease outbreak management. The optimum outcome of this approach in coordinating foodborne outbreak would lie upon its effective implementation for receiving and disseminating information before, during, and after an outbreak occurs. The usefulness of this hybrid information network model would not be limited to foodborne outbreak coordination only; it would rather be wider and can be applied to any type of disaster outbreak management.
References


[19] Schmidt CW. Using Social Media to Predict and Track Disease Outbreaks. Environmental Health Perspectives. 2012;120.


