**SOCIAL MEDIA DRIVEN BIG DATA ANALYSIS FOR DISASTER SITUATION AWARENESS: A TUTORIAL**

**ABSTRACT:**

Electronic communications play a key role in the assessment of situation in the event of disasters and accordingly dispatching of aid and rescue resources. These communications are shifting more and more towards social media postings, particularly using the twitter platform. Extracting intelligence from the available data involves several challenges that we discuss in this paper. This includes (a) filtering out irrelevant data, (b) fusion of heterogeneous data generated by the social media and other sources, and (c) working with partially geo-tagged social media data in order to deduce the needs of the affected people. Bigdata techniques are essential to accomplish this because of large volume of data, much of which is not very relevant. Spatial analytics of the data plays a key role in understanding the situation but available only sparsely because many users do not want to be tracked. We also discuss the role of edge computing handling this analytics in a scalable manner.

**INTRODUCTION:**

Situational awareness is crucial in a disaster scenario and is often difficult to come by due to difficulty in obtaining the necessary information in coherent manner and organizing it. Part of the difficulty arises due to potential damage to and overloading of communications networks, but to a large extent it is not clear a priori what information is most relevant and how it should be gathered. Since disasters evolve over hours and days, tracking situational awareness becomes even more challenging. Lately, social media has emerged as a primary means for informing the ground realities and expressing the needs by people caught in the disasters. Since much of the social media access in disasters occurs from smartphones, it is possible, in theory, to find the spatial location of the data origin, but in reality location information is rather spotty due to privacy concerns. Twitter has established itself as the disaster communication vehicle of choice due to its modest networking requirements, ease of use, and brevity.

For example, after 2011 Japanese earthquake there were more than 5,500 tweets per second after the disaster. Twitter has been used for a wide variety of disaster scenarios, including the three major Hurricanes in 2017, namely Harvey, Maria and Irma that

affected Carribian and US East coast [1], 2019 Pan-European Floods [2], and 2019 US midwestern floods [3]. Fig. 1(a)-(b) shows the distribution of earthquake related tweets (with keywords ‘earthquake’, and ‘jishin’ which means disaster in Japanese) in the Kumamoto Earthquake that struck at Kumamoto City of Kumamoto Prefecture in Kyushu Region, Japan in 2016. The density of these keywords shows close correlation with the shake map observed to the east of Kumamoto City obtained by the Geospatial Information Authority of Japan and National Disaster Institute for Earth Science and Disaster Resilience [5]. Fig. 1(c) shows the power outage related geo-tagged tweets from New York city during the occurrence of Hurricane Sandy in 2012. The storm hit New York city hard on Oct. 29th night, leaving hundreds of thousands without power [6].

Fig. 1(d) shows the intensity map of the affected areas in the Eastern USA, demonstrating that New York and New Jersey areas were worst affected by the storm. In fact the regions of Lower Manhattan from Madison Square to the tip of the island was hit the hardest, with more than 0.24 million people without power as of noon on Nov. 1st. The distribution of the such disaster related tweets was well correlated with the actual areas of damage, which shows the usefulness of the tweet analysis. In addition, various network performance related data, such as network usage, call drops, bandwidth utilization, signal strength measurements etc. can also be obtained from the radio access network (RAN), the core network (CN), and Internet service providers (ISPs).

A significant amount of such data can be accumulated and used to gain valuable insights into where and how the network repair or capacity addition should be scheduled. In this paper our main objective is to explore situational awareness in a disaster area through various types of big data analysis. The paper is organized as follows. Section II describes the background of spatial big data analysis. Section V-A–IV summarize the possibilities and challenges of big data analytics in emergency management networks. We have demonstrated an application of big data analysis in section VI through a case study. The paper is concluded in section VII.

**EXISTING SYSTEM**

social media has emerged as a primary means for informing the ground realities and expressing the needs by people caught in the disasters. Since much of the social media access in disasters occurs from smartphones, it is possible, in theory, to find the spatial location of the data origin, but in reality location information is rather spotty due to privacy concerns. Twitter has established itself as the disaster communication vehicle of choice due to its modest networking requirements, ease of use, and brevity. The storm hit New York city hard on Oct. 29th night, leaving hundreds of thousands without power [6]. Fig. 1(d) shows the intensity map of the affected areas in the Eastern USA, demonstrating that New York and New Jersey areas were worst affected by the storm. In fact the regions of Lower Manhattan from Madison Square to the tip of the island was hit the hardest, with more than 0.24 million people without power as of noon on Nov. 1st. The distribution of the such disaster related tweets was well correlated with the actual areas of damage, which shows the usefulness of the tweet analysis

**PROPOSED SYSTEM**

In addition, various network performance related data, such as network usage, call drops, bandwidth utilization, signal strength measurements etc. can also be obtained from the radio access network (RAN), the core network (CN), and Internet service providers (ISPs). A significant amount of such data can be accumulated and used to gain valuable insights into where and how the network repair or capacity addition should be scheduled. In this paper our main objective is to explore situational awareness in a disaster area through various types of big data analysis.

**SYSTEM STUDIES**

An important outcome of preliminary investigation is the determination that the system request is feasible. This is possible only if it is feasible within limited resource and time. The different feasibilities that have to be analyzed are

• OPERATIONAL FEASIBILITY

• ECONOMIC FEASIBILITY

• TECHNICAL FEASIBILITY

**Operational Feasibility**

Operational Feasibility deals with the study of prospects of the system to be developed. This system operationally eliminates all the tensions of the Admin and helps him in effectively tracking the project progress. This kind of automation will surely reduce the time and energy, which previously consumed in manual work. Based on the study, the system is proved to be operationally feasible.

**Economic Feasibility**

Economic Feasibility or Cost-benefit is an assessment of the economic justification for a computer based project. As hardware was installed from the beginning & for lots of purposes thus the cost on project of hardware is low. Since the system is a network based, any number of employees connected to the LAN within that organization can use this tool from at any time. The Virtual Private Network is to be developed using the existing resources of the organization. So the project is economically feasible.

**Technical Feasibility**

According to Roger S. Pressman, Technical Feasibility is the assessment of the technical resources of the organization. The organization needs IBM compatible machines with a graphical web browser connected to the Internet and Intranet. The system is developed for platform Independet environment. Java Server Pages, JavaScript, HTML, SQL server and Web Logic Server are used to develop the system. The technical feasibility has been carried out. The system is technically feasible for development and can be developed with the existing.

**INPUT DESIGN**

Input Design plays a vital role in the life cycle of software development, it requires very careful attention of developers. The input design is to feed data to the application as accurate as possible. So inputs are supposed to be designed effectively so that the errors occurring while feeding are minimized. According to Software Engineering Concepts, the input forms or screens are designed to provide to have a validation control over the input limit, range and other related validations.

This system has input screens in almost all the modules. Error messages are developed to alert the user whenever he commits some mistakes and guides him in the right way so that invalid entries are not made. Let us see deeply about this under module design.

Input design is the process of converting the user created input into a computer-based format. The goal of the input design is to make the data entry logical and free from errors. The error is in the input are controlled by the input design. The application has been developed in user-friendly manner. The forms have been designed in such a way during the processing the cursor is placed in the position where must be entered. The user is also provided within an option to select an appropriate input from various alternatives related to the field in certain cases.

Validations are required for each data entered. Whenever a user enters an erroneous data, error message is displayed and the user can move on to the subsequent pages after completing all the entries in the current page.

**OUTPUT DESIGN**

The Output from the computer is required to mainly create an efficient method of communication within the company primarily among the project leader and his team members, in other words, the administrator and the clients. The output of VPN is the system which allows the project leader to manage his clients in terms of creating new clients and assigning new projects to them, maintaining a record of the project validity and providing folder level access to each client on the user side depending on the projects allotted to him. After completion of a project, a new project may be assigned to the client. User authentication procedures are maintained at the initial stages itself. A new user may be created by the administrator himself or a user can himself register as a new user but the task of assigning projects and validating a new use rests with the administrator only.The application starts running when it is executed for the first time. The server has to be started and then the internet explorer in used as the browser. The project will run on the local area network so the server machine will serve as the administrator while the other connected systems can act as the clients. The developed system is highly user friendly and can be easily understood by anyone using it even for the first time.

**SOFTWARE ENVIRONMENT**

**Java Technology**

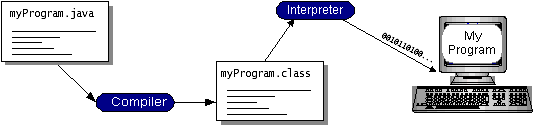
Java technology is both a programming language and a platform.

**The Java Programming Language**

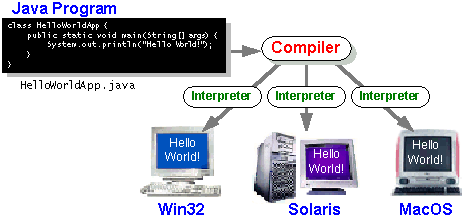
The Java programming language is a high-level language that can be characterized by all of the following buzzwords:

* + - Simple
    - Architecture neutral
    - Object oriented
    - Portable
    - Distributed
    - High performance
    - Interpreted
    - Multithreaded
    - Robust
    - Dynamic
    - Secure

With most programming languages, you either compile or interpret a program so that you can run it on your computer. The Java programming language is unusual in that a program is both compiled and interpreted. With the compiler, first you translate a program into an intermediate language called *Java byte codes* —the platform-independent codes interpreted by the interpreter on the Java platform. The interpreter parses and runs each Java byte code instruction on the computer. Compilation happens just once; interpretation occurs each time the program is executed. The following figure illustrates how this works.



You can think of Java byte codes as the machine code instructions for the *Java Virtual Machine* (Java VM). Every Java interpreter, whether it’s a development tool or a Web browser that can run applets, is an implementation of the Java VM. Java byte codes help make “write once, run anywhere” possible. You can compile your program into byte codes on any platform that has a Java compiler. The byte codes can then be run on any implementation of the Java VM. That means that as long as a computer has a Java VM, the same program written in the Java programming language can run on Windows 2000, a Solaris workstation, or on an iMac.



**6.2 Java Platform**

A *platform* is the hardware or software environment in which a program runs. We’ve already mentioned some of the most popular platforms like Windows 2000, Linux, Solaris, and MacOS. Most platforms can be described as a combination of the operating system and hardware. The Java platform differs from most other platforms in that it’s a software-only platform that runs on top of other hardware-based platforms.

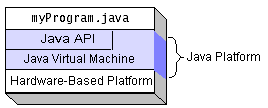
The Java platform has two components:

* The *Java Virtual Machine* (Java VM)
* The *Java Application Programming Interface* (Java API)

You’ve already been introduced to the Java VM. It’s the base for the Java platform and is ported onto various hardware-based platforms.

The Java API is a large collection of ready-made software components that provide many useful capabilities, such as graphical user interface (GUI) widgets. The Java API is grouped into libraries of related classes and interfaces; these libraries are known as *packages*. The next section what Can Java Technology Do? Highlights what functionality some of the packages in the Java API provide.

The following figure depicts a program that’s running on the Java platform. As the figure shows, the Java API and the virtual machine insulate the program from the hardware.



Native code is code that after you compile it, the compiled code runs on a specific hardware platform. As a platform-independent environment, the Java platform can be a bit slower than native code. However, smart compilers, well-tuned interpreters, and just-in-time byte code compilers can bring performance close to that of native code without threatening portability.

**What Can Java Technology Do?**

The most common types of programs written in the Java programming language are *applets* and *applications*. If you’ve surfed the Web, you’re probably already familiar with applets. An apple is a program that adheres to certain conventions that allow it to run within a Java-enabled browser.

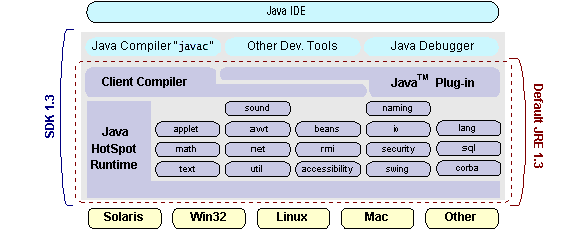
However, the Java programming language is not just for writing cute, entertaining applets for the Web. The general-purpose, high-level Java programming language is also a powerful software platform. Using the generous API, you can write many types of programs.

An application is a standalone program that runs directly on the Java platform. A special kind of application known as a *server* serves and supports clients on a network. Examples of servers are Web Servlets proxy servers, mail servers, and print servers. Another specialized program is a *servlet* A servlets can almost be thought of as an applet that runs on the server side. Java Servlets are a popular choice for building interactive web applications, replacing the use of CGI scripts. Servlets are similar to applets in that they are runtime extensions of applications. Instead of working in browsers, though, serviettes run within Java Web servers, configuring or tailoring the server.

How does the API support all these kinds of programs? It does so with packages of software components that provides a wide range of functionality. Every full implementation of the Java platform gives you the following features:

* **The essentials**: Objects, strings, threads, numbers, input and output, data structures, system properties, date and time, and so on.
* **Applets**: The set of conventions used by applets.
* **Networking**: URLs, TCP (Transmission Control Protocol), UDP (User Data gram Protocol) sockets, and IP (Internet Protocol) addresses.
* **Internationalization**: Help for writing programs that can be localized for users worldwide. Programs can automatically adapt to specific locales and be displayed in the appropriate language.
* **Security**: Both low level and high level, including electronic signatures, public and private key management, access control, and certificates.
* **Software components**: Known as JavaBeans can plug into existing component architectures.
* **Object serialization**: Allows lightweight persistence and communication via Remote Method Invocation (RMI).
* **Java Database Connectivity (JDBCTM)**: Provides uniform access to a wide range of relational databases.

The Java platform also has APIs for 2D and 3D graphics, accessibility, servers, collaboration, telephony, speech, animation, and more. The following figure depicts what is included in the Java 2 SDK.



**How Will Java Technology Change My Life*?***

We can’t promise you fame, fortune, or even a job if you learn the Java programming language. Still, it is likely to make your programs better and requires less effort than other languages. We believe that Java technology will help you do the following:

* **Get started quickly**: Although the Java programming language is a powerful object-oriented language, it’s easy to learn, especially for programmers already familiar with C or C++.
* **Write less code**: Comparisons of program metrics (class counts, method counts, and so on) suggest that a program written in the Java programming language can be four times smaller than the same program in C++.
* **Write better code**: The Java programming language encourages good coding practices, and its garbage collection helps you avoid memory leaks. Its object orientation, its JavaBeans component architecture, and its wide-ranging, easily extendible API let you reuse other people’s tested code and introduce fewer bugs.
* **Develop programs more quickly**: Your development time may be as much as twice as fast versus writing the same program in C++. Why? You write fewer lines of code and it is a simpler programming language than C++.
* **Avoid platform dependencies with 100% Pure Java**: You can keep your program portable by avoiding the use of libraries written in other languages. The 100% Pure Java Product Certification Program has a repository of historical process manuals, white papers, brochures, and similar materials online.
* **Write once, run anywhere**: Because 100% Pure Java programs are compiled into machine-independent byte codes, they run consistently on any Java platform.
* **Distribute software more easily**: You can upgrade applets easily from a central server. Applets take advantage of the feature of allowing new classes to be loaded “on the fly,” without recompiling the entire program.

**ODBC**

Microsoft Open Database Connectivity (ODBC) is a standard programming interface for application developers and database systems providers. Before ODBC became a *de facto* standard for Windows programs to interface with database systems, programmers had to use proprietary languages for each database they wanted to connect to. Now, ODBC has made the choice of the database system almost irrelevant from a coding perspective, which is as it should be. Application developers have much more important things to worry about than the syntax that is needed to port their program from one database to another when business needs suddenly change.

Through the ODBC Administrator in Control Panel, you can specify the particular database that is associated with a data source that an ODBC application program is written to use. Think of an ODBC data source as a door with a name on it. Each door will lead you to a particular database. For example, the data source named Sales Figures might be a SQL Server database, whereas the Accounts Payable data source could refer to an Access database. The physical database referred to by a data source can reside anywhere on the LAN.

The ODBC system files are not installed on your system by Windows 95. Rather, they are installed when you setup a separate database application, such as SQL Server Client or Visual Basic 4.0. When the ODBC icon is installed in Control Panel, it uses a file called ODBCINST.DLL. It is also possible to administer your ODBC data sources through a stand-alone program called ODBCADM.EXE. There is a 16-bit and a 32-bit version of this program and each maintains a separate list of ODBC data sources.

From a programming perspective, the beauty of ODBC is that the application can be written to use the same set of function calls to interface with any data source, regardless of the database vendor. The source code of the application doesn’t change whether it talks to Oracle or SQL Server. We only mention these two as an example. There are ODBC drivers available for several dozen popular database systems. Even Excel spreadsheets and plain text files can be turned into data sources. The operating system uses the Registry information written by ODBC Administrator to determine which low-level ODBC drivers are needed to talk to the data source (such as the interface to Oracle or SQL Server). The loading of the ODBC drivers is transparent to the ODBC application program. In a client/server environment, the ODBC API even handles many of the network issues for the application programmer.

The advantages of this scheme are so numerous that you are probably thinking there must be some catch. The only disadvantage of ODBC is that it isn’t as efficient as talking directly to the native database interface. ODBC has had many detractors make the charge that it is too slow. Microsoft has always claimed that the critical factor in performance is the quality of the driver software that is used. In our humble opinion, this is true. The availability of good ODBC drivers has improved a great deal recently. And anyway, the criticism about performance is somewhat analogous to those who said that compilers would never match the speed of pure assembly language. Maybe not, but the compiler (or ODBC) gives you the opportunity to write cleaner programs, which means you finish sooner. Meanwhile, computers get faster every year.

**JDBC**

In an effort to set an independent database standard API for Java; Sun Microsystems developed Java Database Connectivity, or JDBC. JDBC offers a generic SQL database access mechanism that provides a consistent interface to a variety of RDBMSs. This consistent interface is achieved through the use of “plug-in” database connectivity modules, or *drivers*. If a database vendor wishes to have JDBC support, he or she must provide the driver for each platform that the database and Java run on.

To gain a wider acceptance of JDBC, Sun based JDBC’s framework on ODBC. As you discovered earlier in this chapter, ODBC has widespread support on a variety of platforms. Basing JDBC on ODBC will allow vendors to bring JDBC drivers to market much faster than developing a completely new connectivity solution.

JDBC was announced in March of 1996. It was released for a 90-day public review that ended on June 8, 1996. Because of user input, the final JDBC v1.0 specification was released soon after.

The remainder of this section will cover enough information about JDBC for you to know what it is about and how to use it effectively. This is by no means a complete overview of JDBC. That would fill an entire book.

**JDBC Goals**

Few software packages are designed without goals in mind. JDBC is one that, because of its many goals, drove the development of the API. These goals, in conjunction with early reviewer feedback, have finalized the JDBC class library into a solid framework for building database applications in Java.

The goals that were set for JDBC are important. They will give you some insight as to why certain classes and functionalities behave the way they do. The eight design goals for JDBC are as follows:

1. **SQL Level API**

The designers felt that their main goal was to define a SQL interface for Java. Although not the lowest database interface level possible, it is at a low enough level for higher-level tools and APIs to be created. Conversely, it is at a high enough level for application programmers to use it confidently. Attaining this goal allows for future tool vendors to “generate” JDBC code and to hide many of JDBC’s complexities from the end user.

1. **SQL Conformance**

SQL syntax varies as you move from database vendor to database vendor. In an effort to support a wide variety of vendors, JDBC will allow any query statement to be passed through it to the underlying database driver. This allows the connectivity module to handle non-standard functionality in a manner that is suitable for its users.

1. **JDBC must be implemented on top of common database interfaces**  
    The JDBC SQL API must “sit” on top of other common SQL-level APIs. This goal allows JDBC to use existing ODBC-level drivers through the use of a software interface. This interface would translate JDBC calls to ODBC and vice versa.
2. **Provide a Java interface that is consistent with the rest of the Java system**

Because of Java’s acceptance in the user community thus far, the designers feel that they should not stray from the current design of the core Java system.

1. **Keep it simple**

This goal probably appears in all software design goal listings. JDBC is no exception. Sun felt that the design of JDBC should be very simple, allowing for only one method of completing a task per mechanism. Allowing duplicate functionality only serves to confuse the users of the API.

1. **Use strong, static typing wherever possible**

Strong typing allows for more error checking to be done at compile time; it also, appears at runtime.

1. **Keep the common cases simple**

Because more often than not, the usual SQL calls used by the programmer are simple SELECTs, INSERTs, DELETE and UPDATE, these queries should be simple to perform with JDBC. However, more complex SQL statements should also be possible. Finally, we decided to proceed with the implementation using Java Networking. And for dynamically updating the cache table we go for the MS Access database.

Java has two things: a programming language and a platform.

Java is a high-level programming language that is all of the following

Simple Architecture-neutral

Object-oriented Portable

Distributed High-performance

Interpreted multithreaded

Robust Dynamic

Secure

Java is also unusual in that each Java program is both compiled and interpreted. With a compile you translate a Java program into an intermediate language called Java byte codes the platform-independent code instruction is passed and run on the computer. Compilation happens just once; interpretation occurs each time the program is executed. The figure illustrates how this works.

**JavaProgram**

**Compilers**

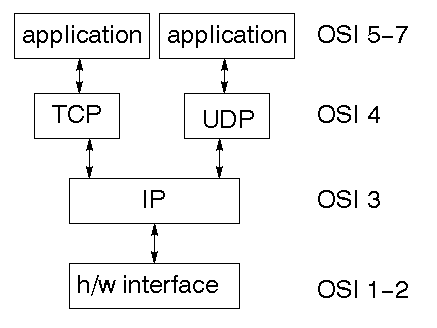
**Interpreter**

**My Program**

You can think of Java byte codes as the machine code instructions for the Java Virtual Machine (Java VM). Every Java interpreter, whether it’s a Java development tool or a Web browser that can run Java applets, is an implementation of the Java VM. The Java VM can also be implemented in hardware. Java byte codes help make “write once, run anywhere” possible. You can compile your Java program into byte codes on my platform which has a Java compiler. The byte codes can then be run in any implementation of the Java VM. For example, the same Java program can run Windows NT, Solaris, and Macintosh.

**Networking**

**TCP/IP stack :**The TCP/IP stack is shorter than the OSI one:



TCP is a connection-oriented protocol; UDP (User Datagram Protocol) is a connectionless protocol.

**IP datagram’s**

The IP layer provides a connectionless and unreliable delivery system. It considers each datagram independently of the others. Any association between datagram’s must be supplied by the higher layers. The IP layer supplies a checksum that includes its own header. The header includes the source and destination addresses. The IP layer handles routing through the Internet. It is also responsible for breaking up large datagram’s into smaller ones for transmission and reassembling them at the other end.

**UDP**

UDP is also connectionless and unreliable. What it adds to IP is a checksum for the contents of the datagram and port numbers. These are used to give a client/server model - see later.

**TCP**

TCP supplies logic to give a reliable connection-oriented protocol above IP. It provides a virtual circuit that two processes can use to communicate.

**Internet addresses**

In order to use a service, you must be able to find it. The Internet uses an address scheme for machines so that they can be located. The address is a 32-bit integer which gives the IP address. This encodes a network ID and more addressing. The network ID falls into various classes according to the size of the network address.

**Network address**

Class A uses 8 bits for the network address with 24 bits left over for other addressing. Class B uses a 16-bit network address. Class C uses 24-bit network addressing and class D uses all 32.

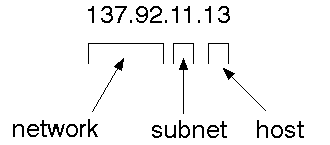
**Subnet address**

Internally, the UNIX network is divided into sub-networks. Building 11 is currently on one sub network and uses 10-bit addressing, allowing 1024 different hosts.

**Host address**

8 bits are finally used for host addresses within our subnet. This places a limit of 256 machines that can be on the subnet.

**Total address**



The 32-bit address is usually written as 4 integers separated by dots.

**Port addresses**

A service exists on a host and is identified by its port. This is a 16-bit number. To send a message to a server, you send it to the port for that service of the host that it is running on. This is not location transparency! Certain of these ports are "well known".

**Sockets**

A socket is a data structure maintained by the system to handle network connections. A socket is created using the calling socket. It returns an integer that is like a file descriptor. In fact, under Windows, this handle can be used with Read File and Write File functions.

#include <sys/types>

#include <sys/socket>

Int socket (int family, int type, int protocol);

Here "family" will be AF\_INET for IP communications, the protocol will be zero, and the type will depend on whether TCP or UDP is used. Two processes wishing to communicate over a network create a socket each. These are similar to two ends of a pipe - but the actual pipe does not yet exist.

**6.3 JFree Chart**

is a free 100% Java chart library that makes it easy for developers to display professional quality charts in their applications. Extensive feature set includes:

A consistent and well-documented API, supporting a wide range of chart types;

A flexible design that is easy to extend, and targets both server-side and client-side applications;

Support for many output types, including Swing components, image files (including PNG and JPEG), and vector graphics file formats (including PDF, EPS, and SVG);

JFree Chart is "open source" or, more specifically, [free software](http://www.gnu.org/philosophy/free-sw.html). It is distributed under the terms of the [GNU Lesser General Public License](http://www.gnu.org/licenses/lgpl.html) (LGPL), which permits use in proprietary applications.

***1.* Map Visualizations**

Charts showing values that relate to geographical areas. Some examples include (a) population density in each state of the United States, (b) income per capita for each country in Europe, and (c) life expectancy in each country of the world. The tasks in this project include:

Sourcing freely redistributable vector outlines for the countries of the world, states/provinces in particular countries (USA in particular, but also other areas);

Creating an appropriate dataset interface (plus default implementation), a rendered, and integrating this with the existing XY Plot class in JFree Chart;

Testing, documenting, testing some more, documenting some more.

***2.* Time Series Chart Interactivity**

Implement a new (to JFree Chart) feature for interactive time series charts --- to display a separate control that shows a small version of ALL the time series data, with a sliding "view" rectangle that allows you to select the subset of the time series data to display in the main chart.

**3. Dashboards**

There is currently a lot of interest in dashboard displays. Create a flexible dashboard mechanism that supports a subset of JFree Chart types (dials, pies, thermometers, bars, and lines/time series) that can be delivered easily via both Java Web Start and an applet.

**4. Property Editors**

The property editor mechanism in JFree Chart only handles a small subset of the properties that can be set for charts. Extend (or reemployment) this mechanism to provide greater end-user control over the appearance of the charts.

**J2ME (Java 2 Micro edition):-**

Sun Microsystems defines J2ME as "a highly optimized Java run-time environment targeting a wide range of consumer products, including pagers, cellular phones, screen-phones, digital set-top boxes, and car navigation systems." Announced in June 1999 at the JavaOne Developer Conference, J2ME brings the cross-platform functionality of the Java language to smaller devices, allowing mobile wireless devices to share applications. With J2ME, Sun has adapted the Java platform for consumer products that incorporate or are based on small computing devices.

**1. General J2ME architecture**



J2ME uses configurations and profiles to customize the Java Runtime Environment (JRE). As a complete JRE, J2ME is comprised of a configuration, which determines the JVM used, and a profile, which defines the application by adding domain-specific classes. The configuration defines the basic run-time environment as a set of core classes and a specific JVM that run on specific types of devices. We'll discuss configurations in detail in the profile that defines the application; specifically, it adds domain-specific classes to the J2ME configuration to define certain uses for devices. We'll cover profiles in depth in the following graphic depicts the relationship between the different virtual machines, configurations, and profiles. It also draws a parallel with the J2SE API and its Java virtual machine. While the J2SE virtual machine is generally referred to as a JVM, the J2ME virtual machines, KVM and CVM, are subsets of JVM. Both KVM and CVM can be thought of as a kind of Java virtual machine -- it's just that they are shrunken versions of the J2SE JVM and are specific to J2ME.

1. **Developing J2ME applications**

Introduction In this section, we will go over some considerations you need to keep in mind when developing applications for smaller devices. We'll take a look at the way the compiler is invoked when using J2SE to compile J2ME applications. Finally, we'll explore packaging and deployment and the role pre-verification plays in this process.

1. **Design considerations for small devices**

Developing applications for small devices requires you to keep certain strategies in mind during the design phase. It is best to strategically design an application for a small device before you begin coding. Correcting the code because you failed to consider all of the "gotchas" before developing the application can be a painful process. Here are some design strategies to consider:

\* Keep it simple. Remove unnecessary features, possibly making those features a separate, secondary application.

\* Smaller is better. This consideration should be a "no-brainer" for all developers. Smaller applications use less memory on the device and require shorter installation times. Consider packaging your Java applications as compressed Java Archive (jar) files.

\* Minimize run-time memory use. To minimize the amount of memory used at run time, use scalar types in place of object types. Also, do not depend on the garbage collector. You should manage the memory efficiently yourself by setting object references to null when you are finished with them. Another way to reduce run-time memory is to use lazy instantiation, only allocating objects on an as-needed basis. Other ways of reducing overall and peak memory use on small devices are to release resources quickly, reuse objects, and avoid exceptions.

**4. Configurations overview**

The configuration defines the basic run-time environment as a set of core classes and a specific JVM that run on specific types of devices. Currently, two configurations exist for J2ME, though others may be defined in the future:

\* **Connected Limited Device Configuration (CLDC)**is used specifically with the KVM for 16-bit or 32-bit devices with limited amounts of memory. This is the configuration (and the virtual machine) used for developing small J2ME applications. Its size limitations make CLDC more interesting and challenging (from a development point of view) than CDC. CLDC is also the configuration that we will use for developing our drawing tool application. An example of a small wireless device running small applications is a Palm hand-held computer.

\* **Connected Device Configuration (CDC)**is used with the C virtual machine (CVM) and is used for 32-bit architectures requiring more than 2 MB of memory. An example of such a device is a Net TV box.

**5. J2ME profiles**

**What is a J2ME profile?**

As we mentioned earlier in this tutorial, a profile defines the type of device supported. The Mobile Information Device Profile (MIDP), for example, defines classes for cellular phones. It adds domain-specific classes to the J2ME configuration to define uses for similar devices. Two profiles have been defined for J2ME and are built upon CLDC: KJava and MIDP. Both KJava and MIDP are associated with CLDC and smaller devices. Profiles are built on top of configurations. Because profiles are specific to the size of the device (amount of memory) on which an application runs, certain profiles are associated with certain configurations. A skeleton profile upon which you can create your own profile, the Foundation Profile, is available for CDC.

**Profile 1: KJava**

KJava is Sun's proprietary profile and contains the KJava API. The KJava profile is built on top of the CLDC configuration. The KJava virtual machine, KVM, accepts the same byte codes and class file format as the classic J2SE virtual machine. KJava contains a Sun-specific API that runs on the Palm OS. The KJava API has a great deal in common with the J2SE Abstract Windowing Toolkit (AWT). However, because it is not a standard J2ME package, its main package is com.sun. Java We'll learn more about the KJava API later in this tutorial when we develop some sample applications.

**Profile 2: MIDP**

MIDP is geared toward mobile devices such as cellular phones and pagers. The MIDP, like KJava, is built upon CLDC and provides a standard run-time environment that allows new applications and services to be deployed dynamically on end-user devices. MIDP is a common, industry-standard profile for mobile devices that is not dependent on a specific vendor. It is a complete and supported foundation for mobile application

development. MIDP contains the following packages, the first three of which are core CLDC packages, plus three MIDP-specific packages.

\* java.lang

\* java.io

\* java.util

\* javax.microedition.io

\* javax.microedition.lcdui

\* javax.microedition.midlet

\* javax.microedition.rms

**SYSTEM DESIGN**

**UML DIAGRAMS**

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for it to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: aMeta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

GOALS:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.

2. Provide extendibility and specialization mechanisms to extend the core concepts.

3. Be independent of particular programming languages and development process.

4. Provide a formal basis for understanding the modeling language.

5. Encourage the growth of OO tools market.

6. Support higher level development concepts such as collaborations, frameworks, patterns and components.

7. Integrate best practices.

###### USE CASE DIAGRAM:

In the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Role of the actors in the system can be depicted.

**CLASS DIAGRAM:**

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

**SEQUENCE DIAGRAM:**

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

**ACTIVITY DIAGRAM:**

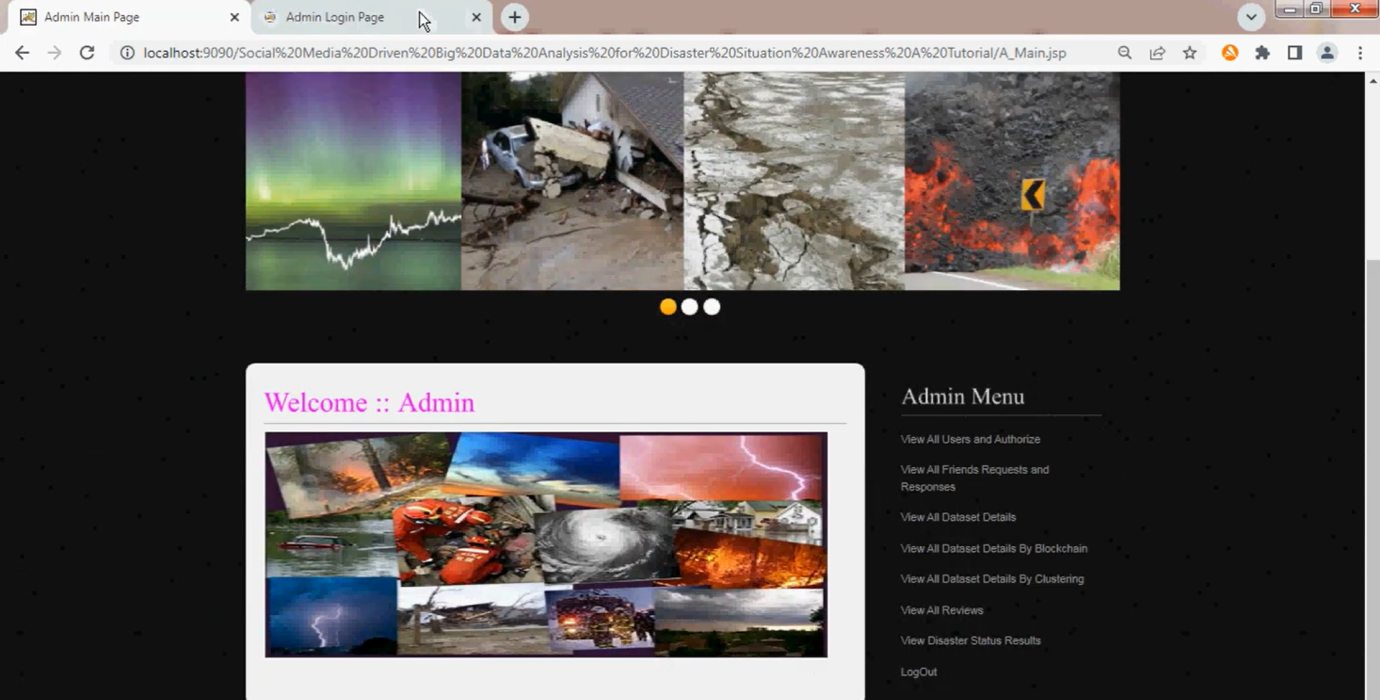
Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

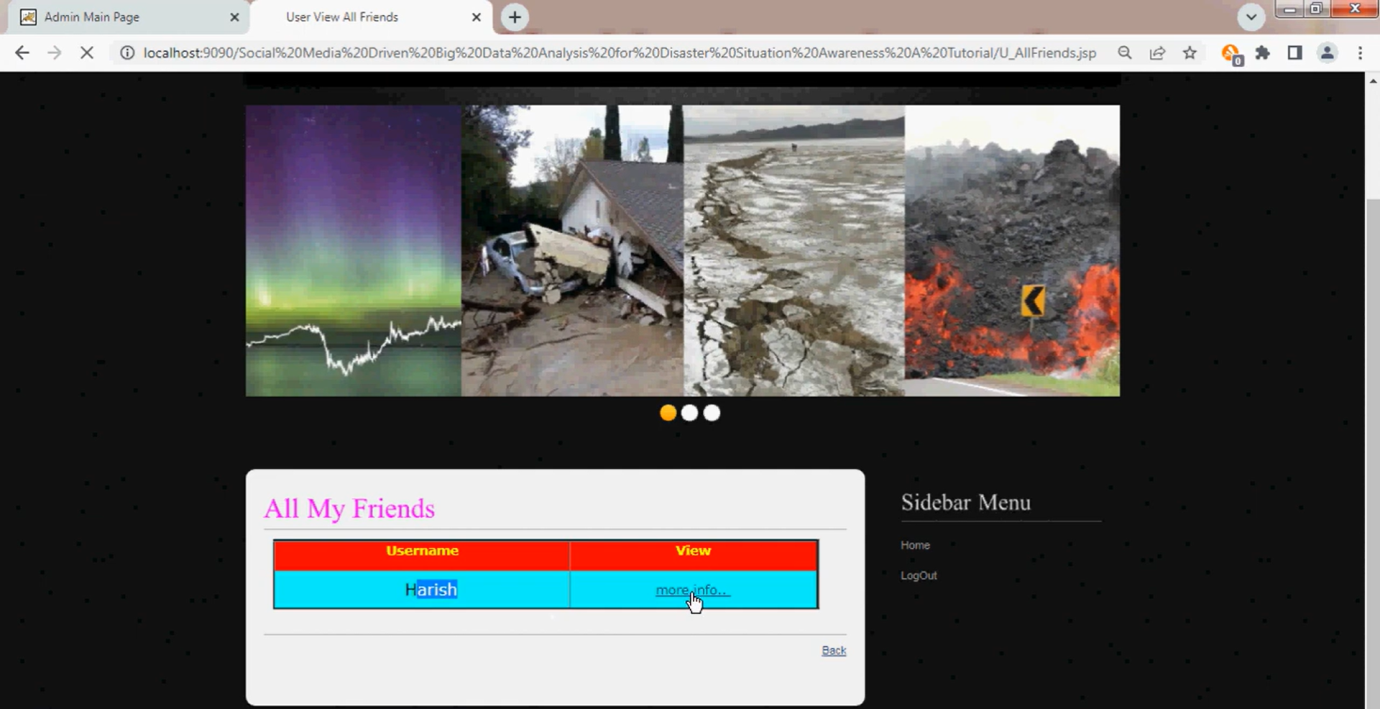
**SYSTEM TESTING**

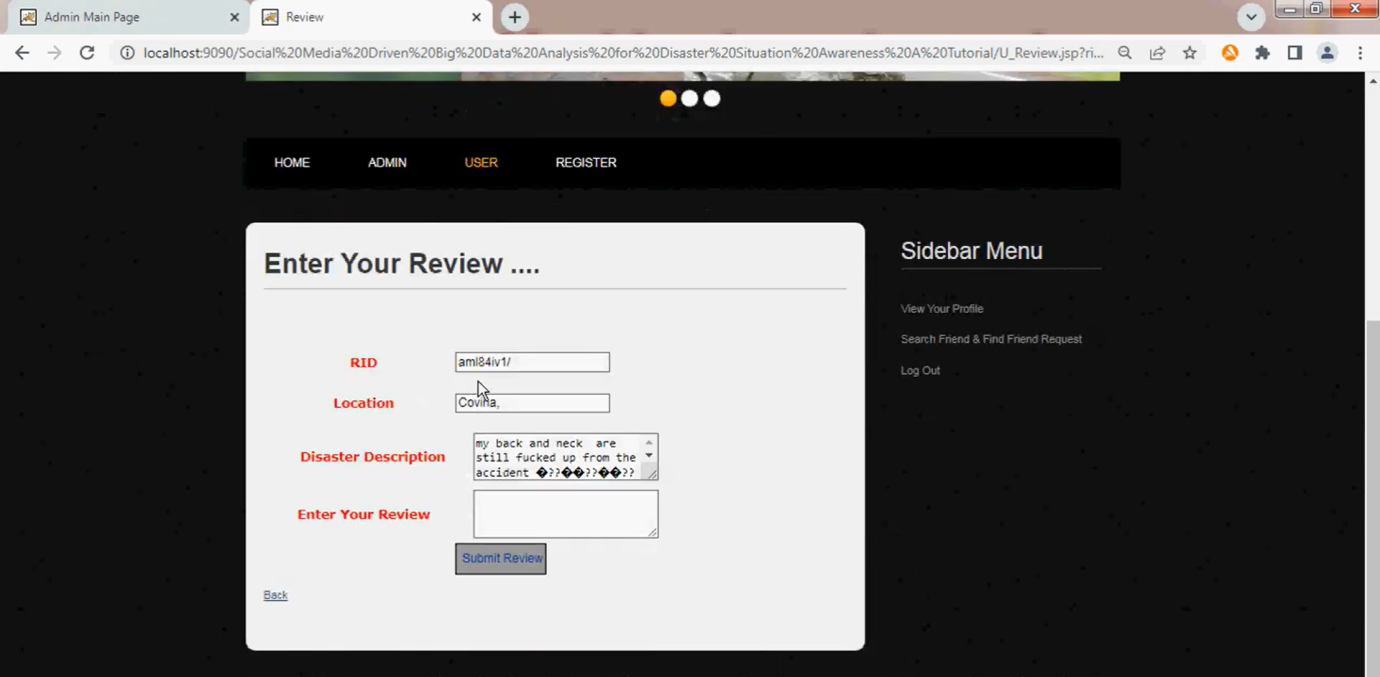
Software once validated must be combined with other system elements (e.g. Hardware, people, database). System testing verifies that all the elements are proper and that overall system function performance is

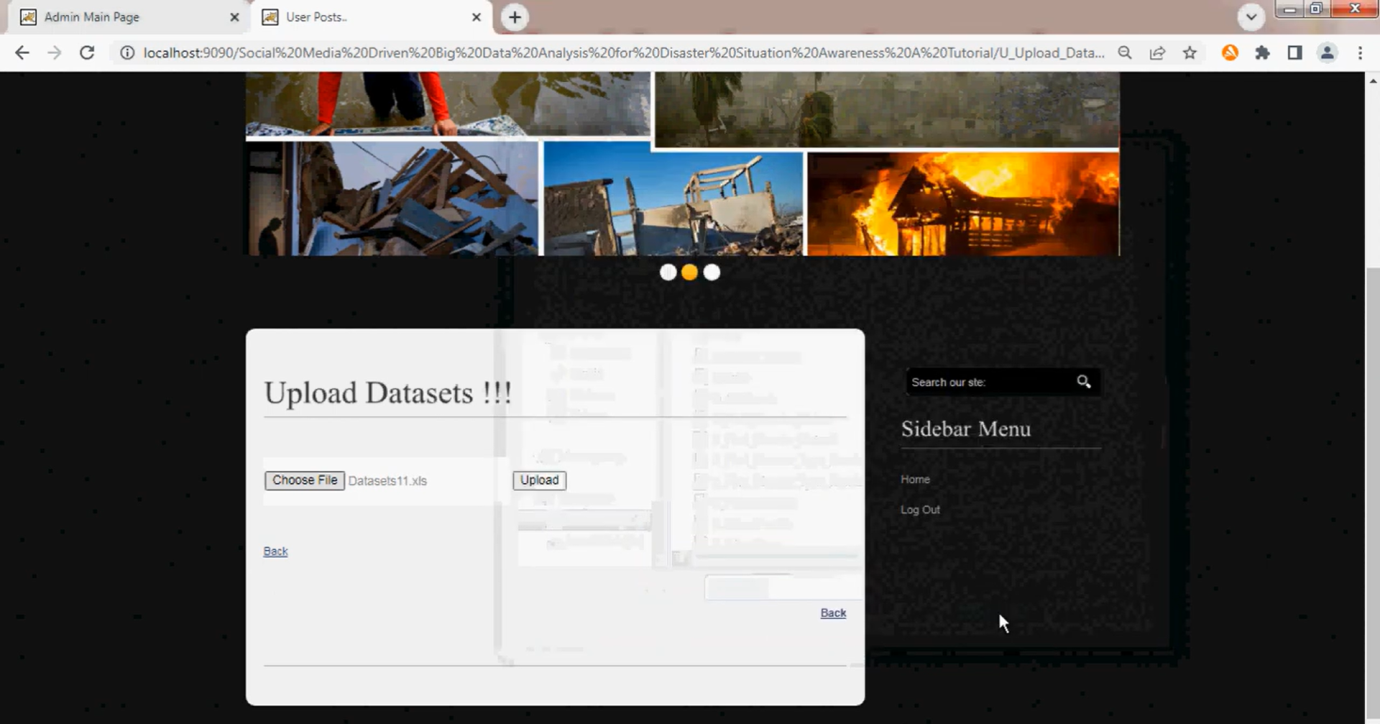
achieved. It also tests to find discrepancies between the system and its original objective, current specifications, and system documentation.

**RESULT:**

****

****

****

****

**CONCLUSIONS:**

During disasters, the data relevant for situational assessment is generated from many different sources including social media use by the affected people (usually Twitter), direct communications with other, possibly unaffected users who put the information on the social media, and observations by the deployed monitoring infrastructure, etc. Different types of disaster related data may come from various sources: such as from the user’s end, from the social medias or from the network operators. The data collected from these sources contains a lot of irrelevant or weakly relevant information, and it becomes necessary to use big data techniques to extract intelligence from them. Spatial information and context is crucial for this, and the paper focuses on several opportunities and challenges of spatial big-data analytics with partial spatial information to extract situational awareness of the disaster. We hope this article will spur further research and results in solutions to many of these issues.

**REFERENCES:**

[1] F. Alam, F. Ofli, M. Imran, and M. Aupetit, “A twitter tale of three hurricanes: Harvey, irma, and maria,” in Proc. 15th Intl. Conf. on Information Systems for Crisis Response and Management, ISCRAM 2018 (B. Tomaszewski and K. Boersma, eds.), vol. 2018-May, pp. 553– 572, 1 2018.

[2] V. Lorini, C. Castillo, F. Dottori, M. Kalas, D. Nappo, and P. Salamon, “Integrating social media into a pan-european flood awareness system: A multilingual approach,” ArXiv, vol. abs/1904.10876, 2019.

[3] “Midwest farmers take to twitter to document flood disaster,” May 2019.

[4] J. Yee, “Hurricane sandy: A chance to identify vulnerabilities, learn from the past, and increase future resilience.” https://research.library.fordham.edu/environ theses/5, 2013.

[5] Asian Disaster Reduction Center, “2016 Kumamoto Earthquake Survey Report (Preliminary).” http://www.adrc.asia/publications/201604 KumamotoEQ/ADRC 2016KumamotoEQ Report 1.pdf. Accessed: March, 2017.

[6] “Hurricane Sandy 2012.” https://www.huffingtonpost.com/2012/10/31/ hurricane-sandy-new-york-city-power-outage-map n 2050380.html. Accessed: January, 2018.

[7] S. Shekhar, P. Zhang, Y. Huang, and R. R. Vatsavai, “Trends in Spatial Data Mining,” Data mining: Next generation challenges and future directions, pp. 357–380, 2003.

[8] J. Wang, Y. Wu, N. Yen, S. Guo, and Z. Cheng, “Big data analytics for emergency communication networks: A survey,” IEEE Communications Surveys & Tutorials, vol. 18, no. 3, pp. 1758–1778, 2016.

[9] Z. Jiang, “A survey on spatial prediction methods,” IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 9, pp. 1645–1664, 2018.

[10] A. McGovern, N. Troutman, R. A. Brown, J. K. Williams, and J. Abernethy, “Enhanced spatiotemporal relational probability trees and forests,” Data Mining and Knowledge Discovery, vol. 26, no. 2, pp. 398–433, 2013.

[11] F. Wu, Z. Li, W.-C. Lee, H. Wang, and Z. Huang, “Semantic annotation of mobility data using social media,” in Proceedings of the 24th International Conference on World Wide Web, pp. 1253–1263, 2015.

[12] F. Wu and Z. Li, “Where did you go: Personalized annotation of mobility records,” in Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 589–598, 2016

. [13] S. Chawla, S. Shekhar, W. Wu, and U. Ozesmi, “Modeling spatial dependencies for mining geospatial data,” in Proceedings of the 2001 SIAM International Conference on Data Mining, pp. 1–17, SIAM, 2001.

[14] M. L. Stein, Interpolation of spatial data: some theory for kriging. Springer Science & Business Media, 2012.

[15] Z. Jiang, “A survey on spatial prediction methods,” IEEE Transactions on Knowledge and Data Engineering, vol. PP, pp. 1–1, 08 2018.

[16] S. Shekhar, C. Lu, and P. Zhang, “Detecting graph-based spatial outliers.,” Intelligent Data Analysis, vol. 6, no. 5, pp. 451–468, 2002.

[17] S. Shekhar and S. Chawla, Spatial databases: a tour, vol. 2003. prentice hall Upper Saddle River, NJ, 2003.

[18] C. Yujun, P. Juhua, D. Jiahong, W. Yue, and X. Zhang, “Spatial– temporal traffic outlier detection by coupling road level of service,” IET Intelligent Transport Systems, vol. 13, no. 6, pp. 1016–1022, 2019.

[19] J. Haslett, R. Bradley, P. Craig, A. Unwin, and G. Wills, “Dynamic graphics for exploring spatial data with application to locating global and local anomalies,” The American Statistician, vol. 45, no. 3, pp. 234–242, 1991.

[20] N. Cressie, “Statistics for spatial data: Wiley series in probability and statistics,” 1993.

[21] R. Webster and M. Oliver, Software for spatial data analysis in 2D., vol. 48. European Journal of Soil Science, 1997.

[22] R. Haining, Spatial data analysis in the social and environmental sciences. Cambridge University Press, 1993.

[23] L. Anselin, “Exploratory spatial data analysis and geographic information systems,” New tools for spatial analysis, vol. 54, 1994.

[24] L. Anselin, “Local indicators of spatial association-lisa,” Geographical analysis, vol. 27, no. 2, pp. 93–115, 1995.

[25] S. Shekhar, M. R. Evans, J. M. Kang, and P. Mohan, “Identifying patterns in spatial information: A survey of methods,” WIREs Data Mining and Knowledge Discovery, vol. 1, no. 3, pp. 193–214, 2011.

[26] G. Zheng, S. L. Brantley, T. Lauvaux, and Z. Li, “Contextual spatial outlier detection with metric learning,” in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 2161–2170, 2017.

[27] M. Zala, L. Rushirajsinh, M. Mehta, B. Brijesh, M. Zala, and R. Mahipalsinh, “A Survey on Spatial Co-location Patterns Discovery from Spatial Datasets,” International Journal of Computer Trends and Technology, vol. 7, no. 3, pp. 137–142, 2014.

[28] Y. Huang, S. Shekhar, and H. Xiong, “Discovering colocation patterns from spatial data sets: a general approach,” IEEE Transactions on Knowledge and data engineering, vol. 16, no. 12, pp. 1472–1485, 2004.

[29] V. Estivill-Castro and I. Lee, “Data mining techniques for autonomous exploration of large volumes of geo-referenced crime data,” in Proc. of the 6th International Conference on Geocomputation, pp. 24–26, 2001.

[30] A. S. Fotheringham, C. Brunsdon, and M. Charlton, Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons, 2003.

[31] R. Caruana, “Multitask learning,” Machine learning, vol. 28, no. 1, pp. 41–75, 1997.

[32] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise.,” in Knowledge Discovery in Databases, vol. 96, pp. 226–231, 1996.

[33] Z. Cai, J. Wang, and K. He, “Adaptive density-based spatial clustering for massive data analysis,” IEEE Access, vol. 8, pp. 23346–23358, 2020.

[34] X. Zhou, H. Zhang, G. Ji, and G. Tang, “A multi-density clustering algorithm based on similarity for dataset with density variation,” IEEE Access, vol. 7, pp. 186004–186016, 2019.

[35] B. Wu and B. M. Wilamowski, “A fast density and grid based clustering method for data with arbitrary shapes and noise,” IEEE Transactions on Industrial Informatics, vol. 13, no. 4, pp. 1620–1628, 2016.

[36] Y.-m. Cheung and Y. Zhang, “Fast and accurate hierarchical clustering based on growing multilayer topology training,” IEEE transactions on neural networks and learning systems, vol. 30, no. 3, pp. 876–890, 2018.

[37] F. Ferstl, M. Kanzler, M. Rautenhaus, and R. Westermann, “Timehierarchical clustering and visualization of weather forecast ensembles,” IEEE transactions on visualization and computer graphics, vol. 23, no. 1, pp. 831–840, 2016. [

38] M. Bendechache, N.-A. Le-Khac, and M.-T. Kechadi, “Hierarchical aggregation approach for distributed clustering of spatial datasets,” in 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), pp. 1098–1103, IEEE, 2016.

[39] A. Woodley, L.-X. Tang, S. Geva, R. Nayak, and T. Chappell, “Using parallel hierarchical clustering to address spatial big data challenges,” in 2016 IEEE International Conference on Big Data (Big Data), pp. 2692–2698, IEEE, 2016.

[40] J. Wang, C. Zhu, Y. Zhou, X. Zhu, Y. Wang, and W. Zhang, “From partition-based clustering to density-based clustering: Fast find clusters with diverse shapes and densities in spatial databases,” IEEE Access, vol. 6, pp. 1718–1729, 2017.

[41] J. Yang and S. Puri, “Efficient parallel and adaptive partitioning for load-balancing in spatial join,” in 2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 810–820, 2020.

[42] C. Lu, D. Chen, and Y. Kou, “Detecting spatial outliers with multiple attributes,” in proceedings of 15th IEEE International Conference on Tools with Artificial Intelligence, pp. 122–128, 2003.

[43] N. Hubballi, B. K. Patra, and S. Nandi, “Ndot: Nearest neighbor distance based outlier detection technique,” in Pattern Recognition and Machine Intelligence, pp. 36–42, Springer, 2011.

[44] Y. Chen, D. Miao, and H. Zhang, “Neighborhood outlier detection,” Expert Systems with Applications, vol. 37, no. 12, pp. 8745–8749, 2010.

[45] M. Zhou, T. Ai, G. Zhou, and W. Hu, “A visualization method for mining colocation patterns constrained by a road network,” IEEE Access, vol. 8, pp. 51933–51944, 2020.

[46] A. M. Sainju, D. Aghajarian, Z. Jiang, and S. K. Prasad, “Parallel gridbased colocation mining algorithms on gpus for big spatial event data,” IEEE Transactions on Big Data, 2018.

[47] X. Hu, G. Wang, and J. Duan, “Mining maximal dynamic spatial colocation patterns,” IEEE Transactions on Neural Networks and Learning Systems, 2020.

[48] B. D. Ripley, “The second-order analysis of stationary point processes,” Journal of applied probability, vol. 13, no. 2, pp. 255–266, 1976.

[49] A. Okabe and F. Miki, “A conditional nearest-neighbor spatialassociation measure for the analysis of conditional locational interdependence,” Environment and Planning A, vol. 16, no. 2, pp. 163–171, 1984.

[50] M. Ruiz, F. Lopez, and A. P ´ aez, “Testing for spatial association of ´ qualitative data using symbolic dynamics,” Journal of Geographical Systems, vol. 12, pp. 281–309, September 2010.

[51] T. Sakaki, M. Okazaki, and Y. Matsuo, “Earthquake shakes twitter users: Real-time event detection by social sensors,” in ACM WWW, pp. 851–860, 2010.

[52] M. Mathioudakis and N. Koudas, “Twittermonitor: trend detection over the twitter stream,” in ACM SIGMOD, pp. 1155–1158, 2010.

[53] C. Li, A. Sun, and A. Datta, “Twevent: segment-based event detection from tweets.,” in CIKM, pp. 155–164, 2012.

[54] S. Petrovic, M. Osborne, and V. Lavrenko, “Streaming first story de- ´ tection with application to twitter,” in Human Language Technologies, pp. 181–189, 2010.

[55] X. Dong, D. Mavroeidis, F. Calabrese, and P. Frossard, “Multiscale event detection in social media,” Data Min. Knowl. Discov., vol. 29, no. 5, pp. 1374–1405, 2015.

[56] D. C. B. P. S. Earle and M. Guy, “Twitter earthquake detection: earthquake monitoring in a social world,” Annals of Geophysics, vol. 54, no. 6, pp. 708–715, 2012

. [57] K. Xie, C. Xia, N. Grinberg, R. Schwartz, and M. Naaman, “Robust detection of hyper-local events from geotagged social media data,” in ACM MDMKDD, 2013.

[58] A. M. MacEachren, A. R. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford, “Senseplace2: Geotwitter analytics support for situational awareness.,” pp. 181–190, 2011.

[59] A. Marcus, M. Bernstein, O. Badar, D. Karger, S. Madden, R. Miller, and O. Arenz, “Twitinfo: Aggregating and visualizing microblogs for event exploration,” in Annual conference on Human factors in computing systems, pp. 227–236, 2011.

[60] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, “Twitcident: Fighting fire with information from social web streams,” in International Conference Companion on World Wide Web, pp. 305– 308, 2012

. [61] J. Yin, S. Karimi, B. Robinson, and M. Cameron, “Esa: Emergency situation awareness via microbloggers,” in ACM CIKM, pp. 2701–2703, 2012.

[62] M. Imran, S. Elbassuoni, C. Castillo, F. Diaz, and P. Meier, “Extracting information nuggets from disaster- related messages in social media,” in ISCRAM, ISCRAM Association, 2013.

[63] N. Morrow, N. Mock, A. Papendieck, and N. Kocmich, “Independent evaluation of the ushahidi haiti project,” 2011.

[64] Y. Qu, P. F. Wu, and X. Wang, “Online community response to major disaster: A study of tianya forum in the 2008 sichuan earthquake,” in 42st Hawaii International International Conference on Systems Science (HICSS-42 2009), Proceedings (CD-ROM and online), 5-8 January 2009, Waikoloa, Big Island, HI, USA, pp. 1–11, IEEE Computer Society, 2009.

[65] L. Palen, S. Vieweg, S. B. Liu, and A. L. Hughes, “Crisis in a networked world: Features of computer-mediated communication in the april 16, 2007, virginia tech event,” Social Science Computer Review, vol. 27, no. 4, pp. 467–480, 2009.

[66] I. Shklovski, L. Palen, and J. N. Sutton, “Finding community through information and communication technology in disaster response,” in ACM CSCW (B. Begole and D. W. McDonald, eds.), pp. 127–136, ACM, 2008.

[67] L. A. S. Denis, L. Palen, and K. M. Anderson, “Mastering social media: An analysis of jefferson county’s communications during the 2013 colorado floods,” in ISCRAM (S. R. Hiltz, L. Plotnick, M. Pfaf, and P. C. Shih, eds.), ISCRAM Association, 2014

68] S. Dashti, L. Palen, M. P. Heris, K. M. Anderson, T. J. Anderson, and S. Anderson, “Supporting disaster reconnaissance with social media data: A design-oriented case study of the 2013 colorado floods,” in ISCRAM (S. R. Hiltz, L. Plotnick, M. Pfaf, and P. C. Shih, eds.), ISCRAM Association, 2014.

[69] T. Shelton, A. Poorthuis, M. Graham, and M. Zook, “Mapping the data shadows of hurricane sandy: Uncovering the sociospatial dimensions of ‘big data’,” Geoforum, vol. 52, pp. 167 – 179, 2014.

[70] A. Tiwari, C. V. D. Weth, and M. S. Kankanhalli, “Multimodal multiplatform social media event summarization,” ACM Trans. Multimedia Comput. Commun. Appl., vol. 14, pp. 38:1–38:23, Apr. 2018.

[71] S. Poria, E. Cambria, N. Howard, G.-B. Huang, and A. Hussain, “Fusing audio, visual and textual clues for sentiment analysis from multimodal content,” Neurocomputing, vol. 174, pp. 50–59, 2016.

[72] K. Liu, Y. Li, N. Xu, and P. Natarajan, “Learn to combine modalities in multimodal deep learning,” arXiv preprint arXiv:1805.11730, 2018. [73] Y. Wu, A. Pal, J. Wang, and K. Kant, “Incremental spatial clustering for spatial big crowd data in evolving disaster scenario,” in IEEE CCNC, pp. 1–8, 2019.

[74] F. Laylavi, A. Rajabifard, and M. Kalantari, “A multi-element approach to location inference of twitter: A case for emergency response,” ISPRS Int. J. Geo-Information, vol. 5, no. 5, p. 56, 2016.

[75] Z. Cheng, J. Caverlee, and K. Lee, “You are where you tweet: a contentbased approach to geo-locating twitter users,” in ACM CIKM, pp. 759– 768, 2010.

[76] J. A. de Bruijn, H. de Moel, B. Jongman, J. Wagemaker, and J. C. J. H. Aerts, “Taggs: Grouping tweets to improve global geoparsing for disaster response,” Journal of Geovisualization and Spatial Analysis, vol. 2, no. 1, p. 2, 2017.

[77] A. Schulz, A. Hadjakos, H. Paulheim, J. Nachtwey, and M. Muhlh ¨ auser, ¨ “A multi-indicator approach for geolocalization of tweets,” in ICWSM, 2013.

[78] W. Zhang and J. Gelernter, “Geocoding location expressions in twitter messages: A preference learning method,” J. Spatial Information Science, vol. 9, no. 1, pp. 37–70, 2014

. [79] M. Kulldorff, “A spatial scan statistic,” Communications in Statistics - Theory and Methods, vol. 26, no. 6, pp. 1481–1496, 1997. [80] D. B. Neill, A. W. Moore, M. Sabhnani, and K. Daniel, “Detection of emerging space-time clusters,” in ACM SIGKDD, pp. 218–227, 2005.

81] M. Kulldorff, F. Mostashari, L. Duczmal, W. Katherine Yih, K. Kleinman, and R. Platt, “Multivariate scan statistics for disease surveillance,” Statistics in Medicine, vol. 26, no. 8, pp. 1824–1833, 2007.

[82] L. Lan, V. Malbasa, and S. Vucetic, “Spatial scan for disease mapping on a mobile population,” in AAAI, pp. 431–437, 2014. [83] T. Rattenbury, N. Good, and M. Naaman, “Towards automatic extraction of event and place semantics from flickr tags,” in ACM SIGIR, pp. 103–110, 2007. [84] C. Maru, M. Enoki, A. Nakao, S. Yamamoto, S. Yamaguchi, , and M. Oguchi, “Development of failure detection system for network control using collective intelligence of social networking service in large-scale disaster,” Proc. the 27th ACM Conference on Hypertext and Social Media (HT2016), Halifax, Canada, pp. pp.267–272, July 2016.