Design for Disaster-Tolerant and Dependable Network Architecture

Yasunori Owada, Masugi Inoue, Ryu Miura, Hiroaki Harai, and Hiroyuki Tsuji

National Institute of Information and Communications Technology
4-2-1, Nukui-Kitamachi, Koganei-shi, Tokyo, 184-8795, Japan,
{yowada, inoue, ryu, harai, tsuji}@nict.go.jp

ABSTRACT

This paper lays out requirements of disaster-tolerant and dependable networks and introduces the design of a network consisting of NerveNet[2,3] access points which create mesh-topological regional access networks with wireless and wired links, an unmanned aircraft wireless communication system, a satellite communication system, and sensors. We implemented a model of the network in a discrete event simulator with a multi-agent human behavior model. Human mobility behaviors and behaviors of the network under disaster conditions can be simulated at the same time, allowing us to investigate how they affect each other. We conducted simulation scenarios which consist of 9,600 human agents or 1,000 human agents and communication systems in a suburban area (the extent is 4km vertically and 6km horizontally). Simulations we conducted show that the proposed combined network system can provide a safety confirmation service globally and locally during a disaster.

Keywords: Discrete event simulation, BAN, UAV, Wi-Fi, DTN, mesh network, disaster-tolerant, dependable network

1 INTRODUCTION

A non-human impact of the massive Great East Japan Earthquake and ensuing tsunami that struck on March 11, 2011 was that they obliterated or damaged not only houses and buildings but also lifelines across a wide-ranging area. Communication infrastructure such as fixed telephone lines, fiber lines for the Internet, and mobile communications were also critically damaged at the center of the affected area. In some areas, such communication infrastructure was recovered in a few days, yet others were unavailable for several weeks. The disruption to communications created delays in tsunami evacuation, critically hindered efforts to grasp the status of damage, disrupted transportation, and confused aid delivery.

The major reasons for these communication disruptions can be categorized as follows:

1. Physical damage to core and branch wired infrastructure or access network hardware infrastructure
2. Physical damage to application server or network control server hardware
3. Call loss by limiting acceptable call ratio from congestion control to avoid system overload

Items 1 and 2 originate from the direct damage by the disaster, i.e., network and server infrastructure damage due to the earthquake or tsunami. Item 3 stems from traffic concentrated in the network system architecture. A congested network management server renders the system unavailable even if the network is not damaged. In such a case, the entire network system is affected so that even users outside the disaster area will not be able to use the network. In fact, in the case of this earthquake many in Japan experienced cell phone outage at that time due to the mobile carrier operator's access restrictions. Mobile mail service and packet data communication was available at that time but the communication speed was extremely slow and mail delivery was delayed for a few minutes to several hours. There are two main reasons for these problems: the shape of access network topology and the dependency on a certain server to control the network system. In a country with high seismic activity like Japan, redundancy and improvement of availability are necessary for the wired and mobile network infrastructure that acts as lifelines, like water or electricity supply.

Considering the above, we have studied a disaster-tolerant and dependable network system that can be used even in disaster situations when part of the system is damaged and only a few network systems are available.

This paper is organized as follows. Section 2 defines the system requirements for the disaster-tolerant and dependable network and describes the target deployment scenarios. Section 3 shows our approach and architecture design. It is difficult to evaluate effectiveness of communication system with human behavior and variation of human behavior by communications, because they interact with each other. To evaluate the designed network works during a disaster considering human behavior, we implemented designed network protocol models and human behavior model in the discrete event simulator. Section 4 introduces the simulation model and scenarios, and then describes the simulation results. Section 5 concludes the paper.

2 SYSTEM REQUIREMENTS

Our target system needs dependability. According to [1], dependability consists of five attributes: availability, reliability, safety, confidentiality, and maintainability. We defined the system requirements along with these attributes as follows:

Availability

- System continues working even if there is partial damage to the network system

- Network connectivity is maintained even if links are partially cut
- Paths switch autonomously if link disconnection is detected, and connectivity is stably maintained
- System works locally even if global connectivity is unavailable
- Not only real-time communication but also delay/disruption tolerant communication is supported
- Users can use the system via their everyday terminals

Reliability

- Long mean time between failure (MTBF), short mean time to repair (MTTR)
- Hardware is readily available on the market

Safety

- Network can typically be available for public safety applications
- Network can be used for confirming someone's safety, checking vital signs, or identifying someone's location during an emergency or disaster

Confidentiality

- Network system has the ability to transmit confidential data securely
- Network system has the ability to control data accessibility

Maintainability

- Network can be deployed, removed, and recovered easily

Two scenarios are considered in deploying the system. One is for on-demand deployment during a disaster. A local government or rescue team may deploy the network system rapidly and widely in a disaster area, and it provides both local and global communication service to residents while the carrier network is recovering infrastructure. In such a case, the network system should be simple, easy deployable, and easily operable so that even a layman can perform the necessary tasks, since on-demand and temporary systems are not frequently used. It is also important to prepare scenarios for providing network services to residents during a disaster in the absence of prior experience and instruction.

Another deployment scenario is an owned network for a local government. The network infrastructure is developed and deployed by the local government and the network service is publically provided to residents by using its own infrastructure. Local governments have long installed an announcement broadcasting network system called the Municipal Disaster Management Radio Communication Network. This network system is only used for public and emergency announcements to residents such as tsunami alerts. Announcement information is transmitted via a radio signal and delivered to a receiver equipped with a loudspeaker, which then transmits the announcement to residents. But this wireless network system is not suited for modern lifestyles because (1) only for broadcasting, (2) noisy for the residents near the loudspeaker, and (3) announcements cannot always reach indoors since modern homes have high levels of soundproofing.

Municipal Disaster Management Radio Communication Network hardware is aging and many local governments are under pressure to replace their systems. It is also possible to rent or lease a network for such alert broadcasting system from a carrier, but the burdensome monthly cost for the network bear the local governments' shoulder for a long time. So it is a realistic solution to construct and install a government owned network, but legacy owned network was low utilization and low cost effectiveness because it was designed only for a specific system. A disaster-tolerant and dependable network system is therefore a good replacement candidate if the maintenance cost is almost the same as the current one. Since the network system supports the virtual network operator (VNO) service, the network infrastructure can also be shared not only for government's alert broadcasting use but also for various users such as residents, application service providers, and network service providers. It is important for the network to be a platform and that everyone can customarily use the network infrastructure.

From these requirements and this deployment scenario, we are developing a disaster-tolerant, easily installable, easily operable, and dependable local network service platform. The next section introduces specific network system design and implementation for achieving this.

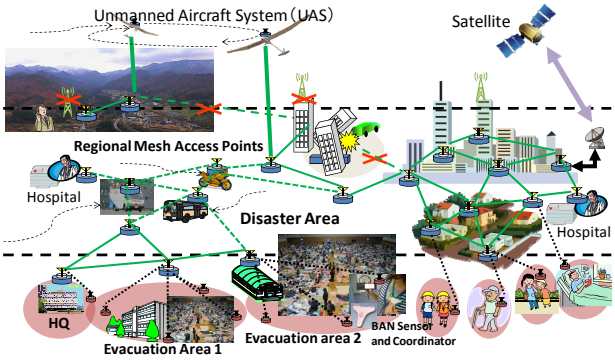

Figure 1. Disaster-tolerant and dependable network

3 DESIGN OF DISASTER TORELANT, DEPENDABLE NETWORK SYSTEM

The network system we are designing and implementing consists of components as a satellite system, wireless and wired hybrid distributed mesh regional access network system, wireless communication system for unmanned aircraft, and sensor networks. Figure 1 provides a network overview. NerveNet [2,3], which we have been proposed, is used for wireless and wired hybrid distributed mesh regional access network system. A NerveNet segment consists of up to 100 NerveNet access points (NNAPs), with one access point covering a 100 m to 500 m radius. It is possible to scale service coverage area by connecting

each segment. The following sections detail NerveNet's major features.

3.1 Mesh topology management and path switching

The NNAP consists of a CPU board with flash memory (with a Linux operating system), a L2 VLAN switch that supports IEEE802.3u and IEEE802.1Q, a wireless-LAN-access-point interface, and multiple wireless LAN/Ethernet bridge interfaces for communicating with other NNAPs.

The Ethernet cable, wireless LAN/Ethernet bridge interface, optical fiber with media converter, and other various link media can be attached to the NNAP, and form mesh topology through various links. A wireless LAN is applicable for the on-demand deployment scenario since it can operate with no license band and is commercially available at a reasonable price. IEEE802.11j and Ethernet service over WiMAX are also candidates for inter-NNAP links.

Normally, frame looping occurs when VLAN switches from the mesh topology. To avoid frame looping, most VLAN switches support the Spanning Tree Protocol (STP). STP blocks some VLAN switches' ports autonomously when looping is detected, and then logical tree topology is created over a mesh physical link topology. STP is disabled in NNAP and multiple VLAN tree topologies are configured over a mesh network topology. First, the NerveNet daemon process on NNAP listens and generates a specific L2 VLAN frame, which can be reachable only for the next hop (neighboring) NNAP for link detection. Then, each NNAP knows what is in the neighborhood and which VLAN switch port has a link to the other NNAP's VLAN switch port. Each NNAP broadcasts its own link state information to all NNAPs by using its link-local IPv6 address and TCP-based hop-by-hop bucket-bridge manners. After that, every NNAP knows the entire network topology.

This bucket-bridging data flooding is used for data synchronization among NNAPs. One NNAP (network manager: NM) calculates the VLAN tree topology from the mesh network topology on behalf of the NNAPs in the network. A VLAN ID is assigned to every VLAN tree topology. NM then floods the calculation result to all NNAPs. When the NNAP receives the result, it changes its own VLAN switch configuration based on the calculation result. Finally, multiple VLAN paths are established between each NNAP.

The source NNAP selects the minimum cost of the VLAN tree to communicate to another NNAP. The VLAN switch of the intermediate NNAP only switches frames based on the IEEE802.1Q VLAN tag header and its own MAC-address-forwarding table. At this point, all NNAPs can communicate with each other using a MAC address (or upper-layer address like an IPv4/v6 address). In the case of link disconnection, the NNAP that detects the link disconnection floods a link error message to all NNAPs. When the NNAP receives a link error message and it is using a VLAN path including a disconnected link as a source node, the minimum cost of the VLAN tree is calculated from the backup tree candidates, and there is a switch to the new VLAN path.

3.2 Direct communications between mobile terminals

The NNAP provides Wi-Fi service to mobile terminals, acting as a default gateway for them. Smartphones, such as iPhone and Android-based, are the assumed mobile user terminal. Internet access is available via NNAP for the terminals. Inter-group information sharing and distribution features based on its distributed name and address resolution for direct communication are also available in NerveNet. These information-sharing and distribution details are given in 3.3. We are developing means of providing application programming interfaces (APIs) for these NerveNet-specific network features since anyone can easily create applications using these features. The features of distributed name and address resolution, and session continuity on NNAP handover, are implemented based on the SIP protocol, but no SIP server is required and every NNAP acts as an SIP proxy. The SIP URI for mobile terminals (used for terminal ID), IP address, and associating NNAP address are shared among NNAPs. Therefore the NNAP can forward the terminal's SIP request to the end terminal directly, and can resolve the IP address from the SIP-URI.

3.3 Inter-group information sharing and distribution among NerveNet

The NNAP shares and synchronizes network link state information among NNAPs. Inter-group information sharing and distribution among user terminals use NNAP APIs and memory space as cache storage. This provides a delay/disruption tolerant network (DTN) to the mobile terminals. Database APIs are provided for a smartphone's application software development kit (SDK) and the application can create and maintain its own database. Information sharing and distribution is achieved by synchronizing the application database among mobile terminals that belong to certain groups. We provide database API extensions to synchronize the databases through NerveNet. Terminals can insert, delete, or update their own database regardless of whether it is connected to the network. When the terminal is connected to NerveNet, the application database begins synchronization to the latest information. To ensure data record consistency, data record deleting and updating is prohibited for anyone but the owner. Every data record has a common record header consisting of Record ID, Owner ID, Group ID, Timestamp,

Life Time, and Discard Flag. The system assumes that time synchronization for every mobile terminal has been performed by GPS or NTP or other time synchronization mechanisms.

The latest differential information for other terminals is synchronized when the mobile terminal is connected to NerveNet. At the same time, its own update differential information is sent to the NNAP, and is flooded in NerveNet. The NNAP temporarily caches the update information and automatically volatizes over time. This achieves DTN-like data synchronization for assisting the NNAP infrastructure.

3.4 Emergency repair for connectivity to outer network

Satellite communication might be the only candidate for a global access line when outer network connectivity has been completely severed due to a disaster. In such a case, NerveNet can enlarge satellite global access network to the regional area. It is easy to inter-connect with NNAP and other network-system-like satellite systems because NNAP provides an Ethernet interface for connecting them. For when the NerveNet segment has been damaged and divided into partitions, we are examining the possibility of using an unmanned aircraft system (UAS) to bridge these network partitions wirelessly. These temporary, expedient measures are at most for the period while the global access line recovers, but can provide global and local network service quickly even as a best-effort service. We have developed a high-speed wireless communication system for airplane and ground communications [4,5], but the size of aircraft we are targeting for use has a nearly 3-4 m wing length and a payload of almost half a kilogram. Therefore the wireless system hardware, including data processing, routing unit, and battery, needs to be downsized. We also need to develop dynamic routing or a bridging protocol to maintain a route through the path between the UAS and NerveNet. This protocol design is currently under development.

3.5 Cooperation with sensors

Information is required for people in disaster areas to find out about their family's safety, lifeline damage, tsunami alerts, where to procure water and relief supplies, etc. In addition, not only during a disaster but also in ordinary times, it is becoming increasingly important in Japan's aging society to gather information on and monitor human biomedical signals. The smart utility network (SUN), body area network (BAN), and other wireless sensor network protocols have recently been under development and standardization, but thorough discussion has not taken place on how to send these sensor data to the specific server passing through a global network, where to cache them when the intermediate network is disconnected, and how to transmit them securely. The backhaul network requirement for the sensors also varies for the applications. We focus on the BAN sensor and healthcare applications to use safety-confirmation applications in combination as a first step.

4 HUMAN BEHAVIOR AND COMMUNICATION SYSTEM MODELING FOR DISASTERS

Network system simulation studies have long been conducted on evaluating the network system, but these are insufficient in evaluating how the system responds to human behavior, or how social activity is changed by these communications. We try to find how our network contributes to better human behavior via simulation by analyzing performance of network depending on variation of human behavior and subsequent human behavior. This section presents the initial step toward this procedure.

We considered the relationship between information and human behavior, and are trying to implement simulation models and scenarios that can simulate both human behavior and communication systems concurrently and cooperatively. The implementation is based on the Scenargie simulator [6], a commercially available discrete event simulation framework. We implement human behavior models, NerveNet network system models, UAS and satellite communication system models, BAN sensor models, and safety confirmation applications such as a VoIP telephone system, e-mail, and intra-group information sharing application using NerveNet-specific features. IEEE802.11 and IEEE802.15.6 BAN wireless system are available in the Scenargie simulator, so we perform detailed implementation of the IEEE802.1Q VLAN feature and NerveNet protocol model and application model. The satellite link is defined as an abstracted link that includes delay, packet loss, and bandwidth. IEEE802.11 MAC and PHY are also assumed as links between UAS and NerveNet. Human behavior is modeled based on a multi-agent simulation system (MAS). We defined several human behavior models and attributes of each human (agent). Each agent can make decisions on how to behave based on its own environmental situation, behavior model attribute, and conditions.

4.1 Simulation assumption and parameters

We assume a large earthquake occurring in Koganei city in suburban Tokyo (near our office) in the morning (around 8 a.m.). Figures 2 to 4 shows the simulation area, and Fig. 4 also shows NNAP position as blue boxes, network topology as doted lines. The communication system assumptions and parameters, human agent behavior assumptions, and safety confirmation application assumptions and parameters are as follows:

Network system assumptions

Table 1. Simulation parameters

	Parameter	Value
Satellite	Wireless system	Packet loss: 0% Bandwidth: 6 Mbps RTT: 800 milliseconds (fixed)
UAS	Wireless system	IEEE802.11a (W56) 11 channels available 6 Mbps fixed rate Tx power 20 dBm Rx sensitivity -85 dBm Packet capture threshold 10 dB Omni-directional antenna (0 dBi)
	Database synchronization	Enabled
NerveNet Access Point	Wireless interface for AP-AP communication	IEEE802.11a (W56), 11 channels available 6 Mbps fixed rate Tx power 20 dBm Rx sensitivity -85 dBm Packet capture threshold 10 dB Omni-directional antenna (0 dBi)
	Wireless interface for AP-STA communication	IEEE802.11g 3 channels available 6 Mbps fixed rate Tx power 20 dBm Rx sensitivity -85 dBm Packet capture threshold 10 dB Omni-directional antenna (0 dBi)
	L2 VLAN Switch	IEEE802.1Q TAG VLAN IEEE802.3u

We assume that the cell phone system and wired infrastructure are completely broken and only a part of our proposed system is working. Our office (shown in Fig. 4) has a satellite communication system, so this system is the only global access line. Some NerveNet links are also damaged with two divided segments (one consisting of two NNAPs in Koganei Park, and the other consisting of the others). The unmanned aircraft communication system appears soon after the disaster and serves as a support to bridge these two NerveNet segments in this scenario.

Human agent behavior assumptions

We defined three types of humans: worker, student, and others. In normal situations, each agent starts moving to its destination, which is preconfigured: office for workers, school for students. Others remain at home. The destination office or school is selected randomly. Five transportation methods are defined: walking, bicycle, car, bus, and train. Each agent selects one transportation method based on its situation and metrics (minimizing travel time, minimizing cost, etc.). Road and traffic lights and building information are imported from a commercially or openly available GIS database. Every agent's movement is restricted on the road, in parks, and in buildings. Traffics lights are also simulated. An agent driving a car stops when the traffic light in the direction of travel turns red. Every agent driving a car moves based on the car-flow model.

The behavior pattern changes after the disaster occurs. We defined two patterns: "go home" and "evacuate to nearest evacuation area." The local government designates evacuation areas, so we apply these in the simulation scenario. If the nearest evacuation area is occupied, the agent moves to the larger evacuation area.

Safety confirmation applications

Each agent belongs to a group consisting of three agents (assumed to be a family) and starts trying to contact the others after the disaster. The agents first try to call via a VoIP phone with the help of the NerveNet direct communication feature, and then try to send e-mail. Finally, they send a message to their own group via the NerveNet inter-group information-sharing application. In the first step of the VoIP call, the agent tries to establish a TCP session to the destination. If the session is established, the peer nodes exchange RTP over UDP packets for a while. After ending the RTP exchange, both nodes calculate the RTP packet reception ratio. Communication is assumed success when the reception ratio is above the threshold.

E-mail transmission and reception are only available for users who have global connectivity. In this scenario, global connectivity is available for users in the NerveNet segment that has a satellite link.

Inter-group information sharing application is a DTN-like inter-group message exchange as described in 3.3. The NNAP and UAS cache all messages and temporarily synchronize. The UAS also delivers the messages in both NerveNet segments by synchronizing the databases.

4.2 Simulation results

Figure 2 shows a simulation result on 9,600 people's behavior without the communication system. The human behavior model is based on the joint metropolitan disaster drill by Tokyo and four cities held on October 29, 2011 at Koganei Park (the same area as in these simulations). More than 8,000 people participated.

The behavior pattern changes when the earthquake occurs in the simulation. Some agents try to go home while some others try to evacuate to the nearest school and then move to a much larger evacuation area if the school's capacity is full. Many people try to move at the same time when the disaster occurs, which causes a heavy traffic jam (Fig. 3). Sizeable Koganei Park exists at the center of the simulation area in this scenario, so a large amount of people ultimately evacuated there.

The human behavior details are not implemented in this simulation, so the model's adequacy, the level of reality of the simulation behavior, and the extent the model should be implemented need to be verified in detail. Various types of statistical information on the Great East Japan Earthquake are becoming available, so we need to simulate based on this real disaster environment and real statistical data, and verify the adequacy of the simulation in relation to the real data.

Figure 4 shows the simulation result on human behavior and communication systems, NerveNet, UAS, satellite, and safety confirmation applications. A total of 1,000 people

with Wi-Fi terminals were simulated. The colors of the people illustrate the safety confirmation achievement ratio: red indicating 0%-49%, yellow is 50%-99%, and blue is 100%. The group member is not always in the simulation area. We assume that half of the group members are not in the simulation area, so they must communicate via the satellite system in this scenario. Even the satellite link is extremely narrowband, and not every NerveNet segment is connected to the satellite in this scenario, but more than half the people achieved at least one safety confirmation with the help of NerveNet, satellite, UAS, and safety confirmation applications, including inter-group information sharing.

5 CONCLUSION

This paper lays out disaster-tolerant and dependable network requirements and introduces a network architecture design consisting of NNAPs which create mesh topology regional access networks combined with wireless and wired links, an unmanned aircraft wireless communication system, satellite communication system, and sensors. We interconnected these network systems and implemented these models into a discrete event simulator with a multi-agent human behavior model. Human behavior in a disaster situation and communication systems can be concurrently simulated in this simulation, so it is possible to see how the former affects the latter, and vice versa. We also confirmed that the proposed combined network system could provide a safety confirmation service globally and locally even if the fixed and mobile infrastructure is not available during a disaster. This simulator enables simulation of a disaster situation under various assumptions. The final goal of this simulation is to simulate a metropolitan disaster situation and evaluate the utility of the communication systems, what kind of communication system is required, and the level of performance. We will continue working to improve interactions between human behavior and communication systems seeking greater reality and efficacy, together with verifications based on comparison with real disaster statistics.

REFERENCES

[1] A. Avizienis, J.-C. Laprie, B. Randell, "Fundamental Concepts of Dependability," Research Report No 1145, LAAS-CNRS, April 2001.

[2] M. Inoue, M. Ohnishi, C. Peng, R. Li, Y. Owada, "NerveNet: A Regional Platform Network for Context-Aware Services with Sensors and Actuators," IEICE Trans. Commun., Vol. E94-B, No.3, pp.618-629, Mar. 2011.

[3] M. Inoue, M. Ohnishi, H. Morino, T. Sanefuji, Y. Owada "Fast Recovery from Link Failures and Blackout of Managed Wireless Mesh for NerveNet," IEEE GLOBECOMM, Dec. 2010.

[4] M. Suzuki, et. al., "High Altitude Wireless Network for Disasters," Space Japan Review, No. 51, Aug/Sep. 2007. (http://satcom.jp/English/e-51/index.html)

[5] R. Miura, et. al., "Challenges to the Lifeline Wireless Communications –Latest Trials Using Airship and UAV-," IEEE WRECOM, Rome, Italy, Sep. 2007.

[6] Scenargie simulator, http://www.spacetime-eng.com/

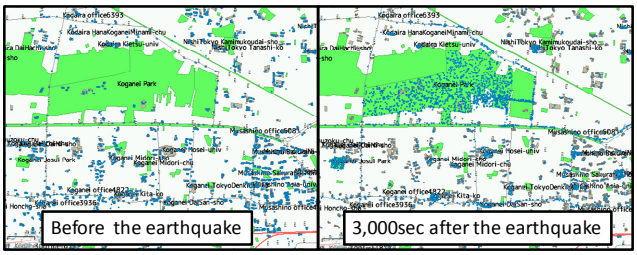

Figure 2. 9,600 human agent behavior simulation results before (left) and after (right) the disaster without communication system.

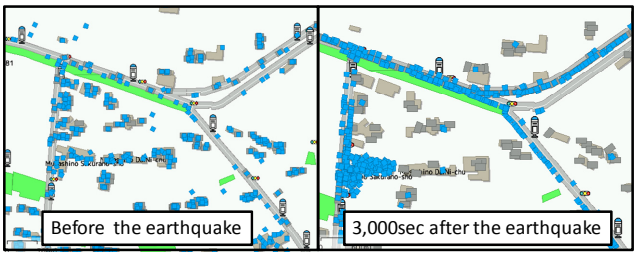

Figure 3. Traffic before (left) and after (right) the disaster.

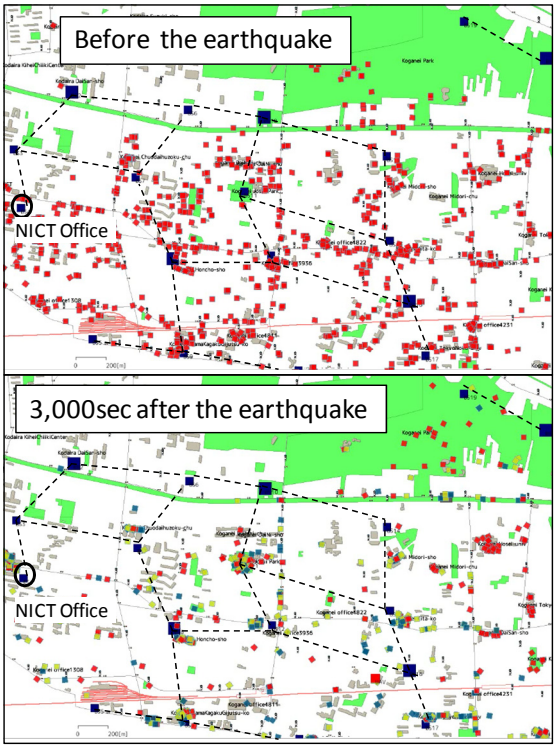

Figure 4. Human behavior and disaster-tolerant communication system simulation results with 1,000 agents (agent color means safety confirmation achievement); top figure shows before the disaster and bottom figure shows 3,000 seconds after the disaster.