

Simulate the Diffusion Process of Adopting AI Chatbots for Increasing Physical Activity

For PHS 7045

Lingyi Fu

Introduction

Artificial intelligence (AI)-driven Chatbots (AI chatbots), developed using natural language processing techniques, are capable of conversing with humans in a natural language format. Recently, AI chatbots have been used to increase physical activity. On the basis of traditional mobile technologies (mHealth), such as wide-reaching and accessibility, portability, and privacy, this new technique benefits the mHealth intervention to be more interactive, autonomous, personalized, supportive, and more natural and intuitive.

The adoption of an AI chatbot is a critical factor that influences the efficacy of using an AI chatbot to change health behaviors. A high adoption can promote health behavior change. Diffusion of Innovation (DOI) theory is a social science theory that explains how new ideas, products, or behaviors spread through a population over time. This theory emphasizes the effect of the social system on the adoption of an innovation, and it also describes five stages of the population adopting an innovation, including innovators, early adopters, early majority, late majority, and laggards. Innovators are people who want to be the first to try the innovation. Early adopters are people who represent opinion leaders. The early majority are rarely leaders, but they do adopt new ideas before the average person. Late majority are skeptical of change, and will only adopt an innovation after it has been tried by the majority. Laggards are bound by tradition and very conservative. People in different stages

Underpinned by the diffusion of innovation theory, this project aims to simulate the diffusion process of adopting an AI chatbot within a social network.

Solution Plan

The innovators play an important role in activating a social system to adopt an innovation. Therefore, in order to simulate the diffusion process, the critical step is to know who is better to be the innovators to achieve a high adoption rate. There are many approaches to finding

out the innovators. In Valente & George's (2020) study, they proposed centrality measures, such as in-degree centrality, betweenness centrality, closeness centrality, bridging approach, inv constraint approach, marginal approach, and random approach. However, these approaches did not consider the strength of ties between nodes. The strength of ties can tell us how easily an AI chatbot can be diffused between two nodes. This project will simulate a diffusion process based on the strength of ties between nodes.

Step 1. Simulate correlation matrix

The first step is to simulate a correlation matrix among 25 nodes with the r^2 value ranging from -1 to 1. The positive correlation between two nodes represents one node that may promote the use of the AI chatbot of another. While the negative correlation represents one node may hinder the use of the AI chatbot of another.

Step 2. Initialize network structure based on the adjacency matrix

The second step is to create a network graph based on an adjacency matrix derived from the correlation matrix at the first step. The nodes of the network represent each individual. The line links two nodes represents the connection of two individuals. The thickness of the line represents the strength between two nodes, with a thick line representing large correlation weight and a thin line representing small correlation weight. The color of the line represents the positive or negative correlation with blue representing positive correlation while red representing negative correlation.

Step 3. Calculate the centrality of each individual

Then, based on the simulated network, I calculated the centrality of each individual in the network. The centrality of each node in the simulated network was calculated using the strength metric, which represents the sum of the weights that an individual associates with others in the network. Individuals with higher strength are more likely to influence others to adopt AI chatbots.

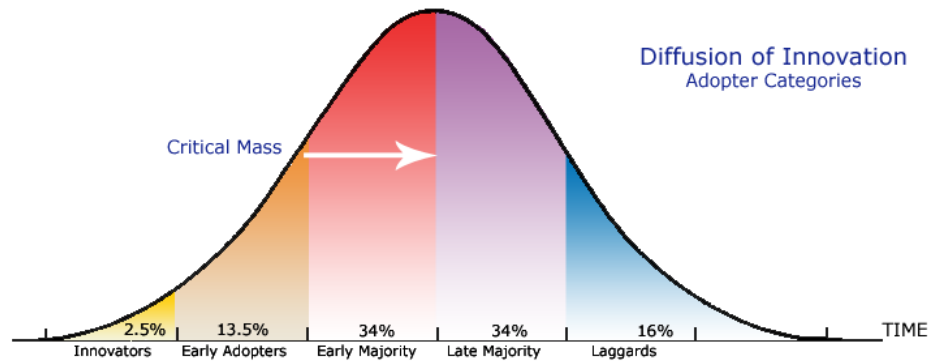
Step 4. Simulate diffusion process

Lastly, I simulated the diffusion process. I first assigned the number of individuals among each category of adopters based on the diffusion of innovation theory (see Figure 1 below).

Figure 1. Diffusion of innovation theory

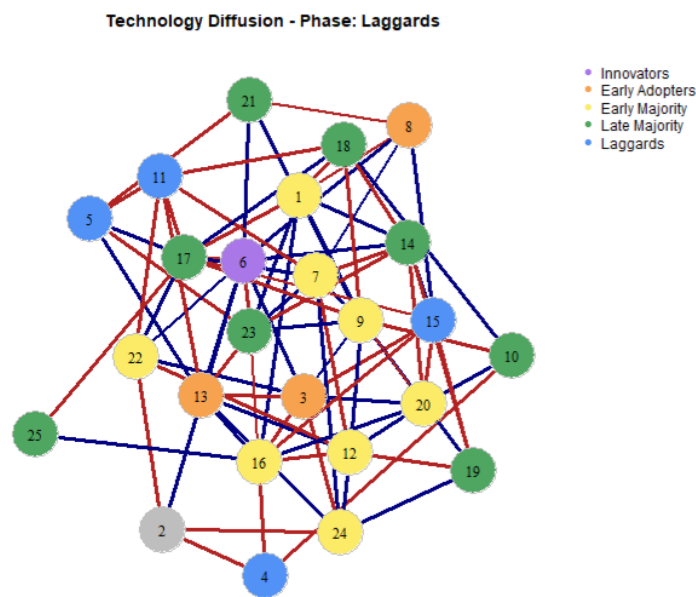
The innovators were determined 2.5% highest centrality nodes. Early adopters, early majority, late majority, and laggards were determined based on the two criteria:

- (1) These nodes were connected with the nodes at the prior stage;
- (2) These nodes ranked in the top percentage of the adopter category (e.g., 13.5% for early adopters) for weight.



Preliminary Results

The GIF below visualizes the diffusion of the process of AI Chatbots from innovators to laggards within the social network. The innovator was shown first with purple node, followed by the early adopters with orange nodes, the early majority with yellow nodes, the late majorities with green nodes, and the laggards with blue nodes. There was one node which did not assign to any group.



Next Step

- Different approaches to identifying innovators may lead to varying adoption percentages. Currently, the final adoption rate is nearly 100% because I predetermined the number of

individuals who would adopt the technology. However, this may not be entirely accurate. The next step is to determine the actual adoption rate that this approach can achieve.

- Once the adoption percentage is established, I will compare the percentage of individuals meeting the PA guidelines across different stages of adopters and non-adopters.

Code

```
1 #install.packages("igraph")
2 library(igraph)
3 library(animation)
4
5 # Step 1: initialize network
6 set.seed(42) #reproduction
7 num_nodes <- 25 # Number of nodes in the network
8
9 # Generate a random corr matrix
10 cormatrix <- matrix(runif(num_nodes2, min = -1, max = 1), num_
   nodes)
11 cormatrix[lower.tri(cormatrix)] <- t(cormatrix)[lower.tri(cormatrix
   )] #Make symmetric
12 diag(cormatrix) <- 0 #set diag as 0
13
14 # transfer corr matrix to adjacency matrix
15 threshold <- 0.7 # Only keep edges with a correlation > 0.7 based
   on the correlation cut-off
16 adj_matrix <- abs(cormatrix) > threshold
17 adj_matrix[upper.tri(adj_matrix)] <- FALSE
18
19 #initial weights
20 initial_graph <- graph_from_adjacency_matrix(adj_matrix, mode = "
   undirected", weighted = TRUE)
21 edge_weights <- cormatrix[adj_matrix]
22
23 #define the parameters of the network layout
24 fixed_layout <- layout_with_graphopt(initial_graph)
25 E(initial_graph)$weight <- edge_weights
26 E(initial_graph)$width <- abs(E(initial_graph)$weight) * 4
27 E(initial_graph)$color <- ifelse(E(initial_graph)$weight > 0, "
   #000080", "#B22222")
28 V(initial_graph)$size <- 20
29 V(initial_graph)$color <- "gray"
30 V(initial_graph)$label.color <- "black"
```

```

31 #plot(initial_graph, layout = fixed_layout)
32
33 # step 2: define the number of nodes in each phase based on
   diffusion of innovation theory
34 n_innovators <- round(num_nodes*0.025)
35 n_early_adopters <- round(num_nodes*0.135)
36 n_early_majority <- round(num_nodes*0.34)
37 n_late_majority <- round(num_nodes*0.34)
38 n_laggards <- round(num_nodes*0.16)
39 numer_phases <- c(n_innovators, n_early_adopters, n_early_majority,
   n_late_majority, n_laggards)
40
41 #Step3: simulate the diffusion process
42
43 #calculate weights of each node
44 node_strength <- strength(initial_graph, vids = V(initial_graph),
   mode = "all", weights = E(initial_graph)$weight)
45
46 simulate_diffusion <- function(initial_graph, total_phase, node_
   strength, numer_phases, node_colors, phase_colors){
47   total_phase <- total_phase
48   top_indices <- NaN
49   used_nodes <- list()
50   adoption_history <- list()
51
52   for (i in 1:total_phase){
53     num_phase <- numer_phases[i]
54
55     if (i == 1){
56       top_indices <- order(node_strength, decreasing = TRUE)[1:num_
         phase]
57
58     }else{
59       connected_nodes <- unique(unlist(lapply(top_indices, function
         (x) {
60         c(neighbors(initial_graph, x, mode = "all"))
61       })))
62
63       #print(connected_nodes)
64       #remove the used nodes
65       if (any((used_nodes) %in% connected_nodes)){
66         connected_nodes <- setdiff(connected_nodes, used_nodes)
67       }else{

```

```

68     connected_nodes <- connected_nodes
69   }
70   weights_connected_nodes <- node_strength[connected_nodes]
71   medium <- order(weights_connected_nodes, decreasing = TRUE)
72     [1:num_phase]
73   top_indices <- connected_nodes[medium]
74   #print(top_indices)
75 }
76 # Save the adoption picture
77 top_indices <- top_indices
78 used_nodes <- unlist(append(used_nodes, top_indices))
79 adoption_history[i] <- list(top_indices)
80 }
81 return(adoption_history)
82 }
83
84 total_phase <- 5
85 adoption_history <- simulate_diffusion(initial_graph, total_phase,
86   node_strength, numer_phases, node_colors, phase_colors)
87 adoption_history
88 # Step 4: create GIF changing the color of the nodes
89 #initial and phase color
90 node_colors <- rep("gray", vcount(initial_graph))
91 phase_colors <- c("#AA77E9", "#F7A24F", "#FBEB66", "#4EA660", "#5292
92   F7")
93
94 phases <- c("Innovators", "Early_Adopters", "Early_Majority", "Late
95   _Majority", "Laggards")
96
97 saveGIF({
98   for (i in 1:total_phase){
99     color_phase <- phase_colors[i]
100     top_indices <- unlist(adoption_history[i])
101     phase <- phases[i]
102     node_colors[top_indices] <- color_phase # Change the color of
103       other nodes with different color
104
105     plot(initial_graph,
106       vertex.color = node_colors,
107       layout = fixed_layout,
108       vertex.frame.color = "gray",

```

```

106         main = paste("Technology_Diffusion_Phase:", phase))
107     legend("topright", legend = phases, col = phase_colors, pch =
        19, bty = "n")
108 }
109 }, movie.name = "technology_diffusion_revised_20_nodes.gif",
        interval = 2, ani.width = 800, ani.height = 600)

```