RE: FWD: RE: FWD: RE: RE: Never Break the Chain

Background

At DER8.9, an email was recently sent out announcing a new office space, which has taken employees by surprise. The announcement especially took the email infrastructure team by surprise.

Company: DER8.9
Mailing List: onsite-staff@der8-9.com
Email Subject: IMPORTANT: Office Relocation Announcement

Dear DER8.9 Team,
We hope this message finds you well. We're writing to inform you about an important change that will affect all of us at the DER8.9 office in Windport.

Starting next month, our main office will be relocating to the new Voltview Park in the downtown district. This decision comes after months of deliberation and is in line with our goals to be more accessible and to foster an environment of growth and collaboration.

New Address: Voltview Towers, 33 Perlman Street, Windport.
Moving Date: 15th of next month. Please ensure your personal workspace is packed up by the 13th.
Transportation: We're aware that this move might change your daily commute. We're currently in discussions with transport services for potential company shuttle options. More details soon.

We understand that this might raise a lot of questions and concerns. Please feel free to reply to this email or reach out to your respective department heads for further clarification.
Thank you for your understanding and cooperation. Together, we will make this transition as smooth as possible.

Warm regards,
DER8.9 Management Team

Who knew that switching over to a new system for mailing and distribution lists could cause such a big problem for DER8.9? These lists play a central role in communications, and it just so happens that the new system appears to have had some configuration issues...

Everyone had the capability to reply-all to the onsite-staff mailing list.

Users have been replying-all with anything from genuine questions or “stop sending me this” to just reveling in the chaos of replying to “who’s ready to carpool” requests. Some are even forwarding the message to remote colleagues to share the news! Needless to say, this has triggered a massive reply-all storm that has overtaken inboxes and overloaded mail servers.

RE: FWD: RE: FWD: RE: RE: Never Break the Chain

Objective

This scenario will provide a very simplified mathematical model that has been outlined to track and predict email activity resulting from the announcement. Please note the following caveats:

Model Simplification: This model assumes everyone reads and reacts to emails instantly.

Email System Overload: The scenario doesn't overtly account for potential system overloads due to high email traffic. It's assumed the email system can handle the volume without issues, but this concept is explored in the questions section.

Recursive Replies: We're simplifying the behavior by assuming each reply-all triggers another immediate round of reply-all responses. In reality, not everyone would engage like this.

Your mission, if you choose to accept it, is to build out this model given the provided information and initial conditions to answer the outlined questions at the end of this document which can be submitted as flags to the scoreboard. Try using Octave, MATLAB, or whatever you prefer.

Mathematical Model

Parameters and Variables:

M: Total members on the original mailing list.

- M = 800

$E(t)$: The cumulative number of reply-all emails sent up to time t.

$F(t)$: The cumulative number of emails forwarded outside the mailing list up to time t.

$N(t)$: Cumulative total emails generated (both replies and forwards) up to time t.

p: Initial probability that a member of the mailing list will reply-all.

- p = 0.04 (4%)

q: Initial probability that a member of the mailing list will forward the email.

- q = 0.38 (38%)

r: Probability that a recipient of a forwarded email going out of their way to change the “To” field from the address of the forwarder to the onsite-mailing list causing a reply-all. This is very low, as this may only happen when an employee forwards the email to another remote employee who was not on the onsite-staff list and the employee would have to go out of their way to reply all to the on-site list.

- r = 0.01 (0.1%)

ρ: Re-engagement probability, the likelihood that someone who has already replied will reply again.

- ρ = 0.08 (8%)

α: Fraction of the initial email population that might consider forwarding the email.

- α=0.3 (30%)

δ: Decay factor representing a 5% reduction in engagement every time step.

- δ = 0.71

RE: FWD: RE: FWD: RE: RE: Never Break the Chain

Equations

Reply-All Emails, $E(t+1)$: The equation models the number of reply-all emails sent after each time step.

$$E(t+1) = E(t) + (p(t) \times M \times M) + (\rho \times E(t))$$

- $E(t)$ is the number of reply-all emails until time t.
- $p(t) \times M \times M$: This term represents the number of new reply-all emails sent to the members on the original mailing list. $p(t)$ is the time-dependent probability that a member of the mailing list will reply-all, and M is the total number of members on the mailing list. Since each reply-all email goes to all members, multiply by M.
- $\rho \times E(t)$: This term represents the re-engagement in the email storm. Here, ρ is the re-engagement probability, or the likelihood that someone who has already replied will reply again. Multiplying it with $E(t)$ (the cumulative reply-all emails up to time t) gives the number of additional reply-all emails due to re-engagement.

Forwarded Emails, $F(t+1)$: The number of forwarded emails after each time step.

$$F(t+1) = F(t) + (q(t) \times \alpha \times M)$$

The term $q(t) \times \alpha \times M$ represents the number of new emails being forwarded outside the original mailing list:

- $F(t)$ is the number of forwarded emails until time t
- $q(t)$: The time-dependent probability that someone on the mailing list will forward the email.
- α: Fraction of the initial email population that might consider forwarding the email.
- M: Total members on the original mailing list.

3. Total Generated Emails, $N(t+1)$**:** Total number of emails after each time step.

$$N(t+1) = N(t) + E'(t) + F'(t) + (r \times F'(t-1) \times M)$$

- N(t) is the total number of emails up to time t.
- E′(t) is the number of new reply-all emails generated at time t.
  - $E'(t) = E(t+1) - E(t)$
- F′(t): Number of new emails forwarded at time t.
  - $F'(t) = F(t+1) - F(t)$
- r×F′(t-1)×M Number of reply-all messages generated at time t due to emails forwarded at time t−1.

RE: FWD: RE: FWD: RE: RE: Never Break the Chain

Time-Dependent Probabilities Accounting for Decay in Engagement

The decay factor δ represents a reduction in engagement every minute, which can be used to model decay in the probabilities $p(t)$ and $q(t)$ as exponential functions. This means the probability will decrease as time goes on, which reflects the fact that people are less likely to engage as the storm drags on.

Time-dependent probability that a member of the mailing list will reply-all, $p(t)$:

$$p(t) = p \times \delta^t$$

Where:

- p is the initial probability that a member will reply-all (given as 4% or 0.04).

- δ is the decay factor (given as 0.71).

- t represents minutes since the start of the reply-all storm.
This equation essentially multiplies the initial reply-all probability by the decay factor raised to the power of the elapsed time, thus modeling the decay in the likelihood of reply-all as time progresses.

Time-dependent probability that a mailing list member will forward the email: $q(t)$:

$$q(t) = q \times \delta^t$$

Where:

- q is the initial probability that a member will forward the email (given as 38% or 0.38).

- δ is the decay factor (given as 0.71).

- t represents time increment step since the start of the reply-all storm.

This equation models the decay in the likelihood of forwarding the email as time progresses.

Initial Conditions:

$\boldsymbol{E(0) = 0}$ – Represents reply-all messages to the email at the start (0, there are no replies)

$\boldsymbol{F(0) = 0}$ – Represents forwards of the message at the start (0, there are no forwards)

$\boldsymbol{N(0) = 800}$ – Represents the start of the mail storm, as one announcement went out to a list of 800 employees.

RE: FWD: RE: FWD: RE: RE: Never Break the Chain

Questions:

Model these equations over a 1-hour period with each step *t* being in 1-minute increments starting at t(0).
(For example: at t(1) 1 minutes have passed, t(2) 2 minutes have passed…)

Round decimal up to the nearest integer when submitting answers as flags.

1. At what minute does the cumulative total of reply-all emails first exceed 120,000?
2. What is the total number of emails at the end of the 1-hour interval?
3. How many employees have forwarded the email outside the onsite-staff mailing list at the end of the 1-hour interval?
4. What is the number of reply-all emails that originated from forwarded emails?
5. Suppose the server would crash due to reaching a total of 1,400,000 emails generated due to this email storm. At what minute would this happen?