

火星救援机器人·产品需求文档

Neruthes

0.1.0 (2021-08-20)

1 背景

在一些 Telegram 群里，经常有不同群友先后将相同的频道消息、新闻链接分享在群里。这造成了一些问题。一方面，已经看过的人会觉得刷屏打扰。另一方面，后分享的人不知道群友们已经围绕该新闻产生了哪些讨论。

因此，我们需要开发一个火星救援机器人（以下简称奥尔加），来解决这些困扰。

2 术语定义

| 术语 | 定义 |
|----------|------------------------|
| 频道 | Telegram 平台上的 Channel |
| 频道消息 | 源头由频道发布的消息 |
| 网页链接 | 使用 HTTP/HTTPS 协议头的 URL |
| 群内直接消息 | 在群内由群成员发布的群内消息 |
| 群内转发频道消息 | 群成员将频道消息转发到群内产生的群内消息 |
| 群内转发用户消息 | 群成员将用户消息转发到群内产生的群内消息 |
| 群内消息 | 以上三类消息的统称 |
| 救援 | 提示一个群内消息的发送者「你火星了」 |
| 待救对象 | 一条需要被救援的群内消息 |
| 珠玉 | 导致待救对象需要被救援的已有群内消息 |

3 工作环境

奥尔加将会以普通用户的身份加入群内，持续监听群内消息，并记录消息历史，以便在恰当的时机救援。

4 救援目标

1. 频道消息重复救援

如果一个群内转发频道消息指向的频道消息在过去一段时间内曾出现在其他群内转发频道消息中，那么救援本次的发送者。

2. 网页链接重复救援

如果一个群内直接消息包含的某个网页链接在过去一段时间内曾出现在其他群内直接消息中，那么救援本次的发送者。

5 基本原则

1. 时效性原则

检测范围仅为过去 48 小时内；即使火星，只要之前最近一次不在过去 48 小时内，就不救援。超出 48 小时的数据没有保留的必要。

2. 存在性原则

尝试救援时，必须找到一个珠玉，即尚未删除；如果最新的珠玉已经删除，那么找较早的珠玉。如果找不到珠玉，那么认为这个消息其实没火星。

3. 特殊性原则

根据网页链接重复检测尝试救援时，如果该网页链接的 pathname 部分（包含开头的斜线）和 search 部分（包含开头的问号）的总和长度小于或等于 6，那么认为这个消息其实没火星。

6 救援方式

6.1 频道消息重复救援

回复待救对象如下文本：

你火星了！这条消息是第 [RepeatedTimes] 次来到本群了，快去爬楼。第一次出现是在：[ChatMsgUrl]

6.2 网页链接重复救援

回复待救对象如下文本：

你火星了！这个链接是第 [RepeatedTimes] 次来到本群了，快去爬楼。第一次出现是在：[ChatMsgUrl]

7 常见问题

1. 复读行为是否会触发？
不会。
2. 哪些消息内的网页链接会被救援？
只救援群内直接消息。