1. When extracting links contained in a website, setting these links as URL type may cause that two different websites with the same host IP are treated as the same URL. This may be the reason that 33 or 34 are both ok for testing case. I prefer to declare these links as String type when comparing difference of them.

2. Sometimes when trying the last testing case, it will cause timeout error. This situation happens when my laptop is not charged. The reason may be that URLs are got one by one. I think concurrency using Actor is a better choice.