**Presentation Script:**

\*Stayko and Huy sit in front of their computers, typing\*

Jiyeong: Hey, guys. How’s it going? We haven’t seen you in ages!

Stayko: “Hi Guys, sorry, our boss at Deutsche Telekom just gave us this massive data dump on honeypots to analyze.

Jiyeong: “And what are you analyzing it for?”

Huy: Well, mainly to understand the patterns of these cyberattacks and also to generate business recommendations for our boss.

Jonathan: Ooh, I heard that WhatsApp was hacked recently and it was a pretty big deal. But is it going to affect my company or me?

Stayko: It’s hard to say. All companies and houses have computers and smart devices, like smart cameras or a smart fridge. All of these could be affected and be used as a botnet for cyberattacks, like illegal activities.

Jiyeong: “Oh wow, and did you discover anything interesting in the data?

Huy: Yes, we found some pretty fascinating stuff actually. The attack came mainly from Canada, US and China. But the most interesting thing is the frequency of these attacks. When we plot them out in a time series, there is a clear patterns. Every 2 days, the behavior of attack repeat itself.

That was an amazing discovery so we look deeper into the individual attacks and identified 2 different patterns. 1 is extremely regular. We hypothesize that these patterns mimic the behavior of the affected servers that the attackers used. The other is extremely erratic and hard to predict.

Jonathan: “That sounds pretty scary.

Huy: “Don’t worry, we’ve got you! With these findings, we’re working on a solution stacks to help solve this problem to propose to Telekom.

First, for the regular attacks, we can use Recurrent Neural Networks with Long-short-term memory units to forecast the next waves of attacks. We have implemented this model, it’s able to pretty closely predict what’s gonna happen next. With this knowledge, companies can prepare their defense for the attack ahead of time.

For the irregular attacks, based on the database and libraries of known attacks, we can use Generative Adversarial Neural Networks to simulate new and unknown attacks that which can help to improve system defense as well!

And lastly, of course, we have the traditional Machine Learning classifier to detect attacks in real time.

And the best part is, for Generative Neural Network, we don’t need to look into the traffic of our customers for data. There are no ethical concerns and the privacy of our customers is secured.”

Stayko: “Innovative and ethically fine! But is it going it scale?”

Huy:”Yes, we can deploy neural networks quite cheaply and easily with cloud services nowadays. These models don’t have big impact on network load so the clients can still enjoy high speed internet.

Stayko: “I think we’ve done good. We are in Berlin, Let’s go to a party”

**Challenge aspects**

|  |  |
| --- | --- |
| Criteria |  |
| Innovation | Pattern analysis  Crowd intelligence for intrusion detection |
| Effective/Scalable | Detect attacks fast  Virtual network of honeypots  High certainty for predictions |
| Technological Approach | System architecture graphic  RNN w. LSTM  GAN  ML classification |
| Ethical aspect | Criminals can use ML for improving their attacks  No analysing user data without consent  → offer the service as paid solution  Anonymity |
| Presentation | Present as a group in a role-play |

**Solution Architecture**

1. Find patterns in incoming honeypot traffic and use them to detect the presence of malware in existing networks business and home
2. Generate new / unknown attacks from libraries of known ones
3. Machine Learning classifier to identify attacks in the honeypot and and protect in real-time at the customer’s firewall (e.g. Telekom “Speedport” Router or Avira Firewall)
4. Predict trend in attack rate, so that companies can increase their resources and load balance ahead in time