

Email Dynamic Contact Network

In order to prove the wide applicability of our framework, we also apply our framework to an email dataset, which contains 1.2 million lines of communication records of employees in Hacking Team company. We try to classify employees and explore the development process of Hacking Team company through our visual analysis framework.

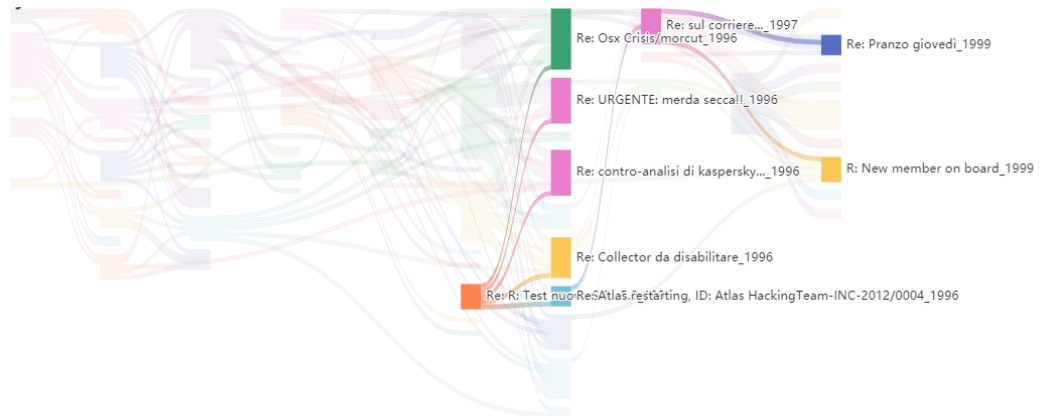


Fig. 1. The email dataset shows the distribution trend of the event involving "Atlas restarting".

In the trend view, each line represents an employee of hacking team company, and each band represents an event. We can see that each line in the graph develops over time. We search for interested users and code interested users with highlighted colors. For example, we search for employee Antonio MAZZEO. Figure 1 below shows the distribution of events that Antonio MAZZEO participated in. We can see that in 1996, Antonio MAZZEO sent an email "re: Atlas restarting, ID: Atlas hacking team-inc-2012 / 0004" to other users, which was received by many people. From 1996 to 1999, there were relevant e-mails to reply and inquire about "Atlas restarting". We can infer that atlas is a project, and the project cycle lasts for three years. After 1996, "Atlas restarting" continued to advance, and the project ended in 1999.

In addition, we can further explore the detailed structure information of a certain time slice node and community throughout our framework. As shown in Fig 2 below, the topic distribution of employee research in 2020 is shown. The nodes represent employees. We can see that the main events of hacking team company in 1996 include "Atlas restarting", "OSX crisis / morcut", "legal Trojan", etc., and the "legal Trojan" project involves the most employees. In the wrapped contour graph of Fig. 3 below, the nodes represent employees, and the connection between employees depends on whether they participate in the same event (project). It can be seen that multiple communities are formed, and there are overlaps between communities, such as the communities surrounding the legal "Trojan" project and the events surrounding the "Atlas" project.

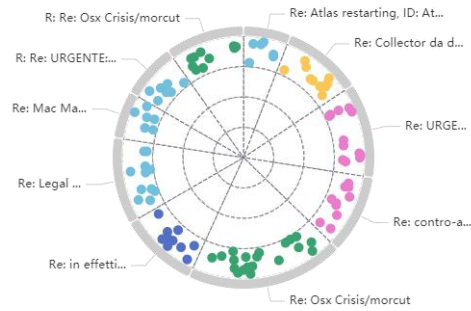


Fig. 2. The distribution of employee participation events in the data set of email dynamic contact diagram. Each node represents an employee. The closer the node is to the center of the circle, the more active the employee is in the event (project)



Fig. 3. outline of employee cooperation in the dataset of email dynamic contact network. Each node represents an author, and the mask of package represents a community