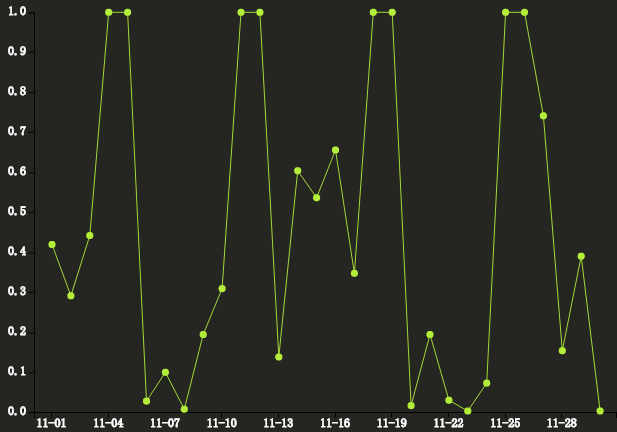
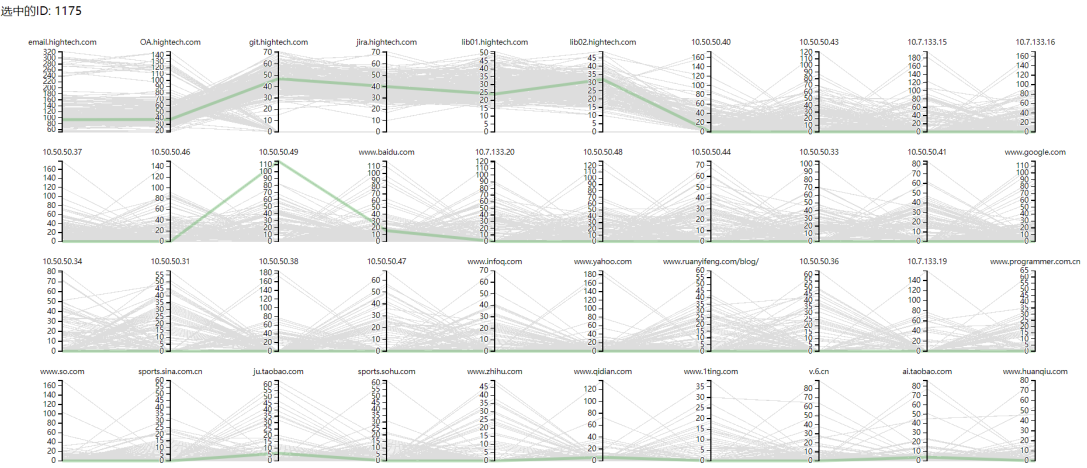
我们用多方法预检测并交叉比对的方式查找公司内部可能存在的异常事件，基本思路是独立地使用不同的方法发现疑似异常事件/人员，然后将不同方法的结果汇总。若一个员工同时出现在两种及以上异常检测的名单中，则该员工可能确实存在异常动机和行为；出现次数越多，则可能性越大。通过数据可视化，我们为公司管理者提供了四种异常预检测方法：（1）排查异常时间登录公司服务器的现象；（2）排查大量tcp流量；（3）排查对公司服务器过于频繁的页面访问；（4）排查异常的上下班时间。另外，也可以关注员工对敏感内容网站的访问，比如对个别黑客技术网站的访问：hackdig, web2hack。

基于如上方法，我们发现了如下异常员工：

1175号: 该员工是一名来自研发部门的普通员工，其对应的ip地址为10.64.105.10,。他在11月30日较晚时间（21点43分）登录了ip为10.50.50.49的服务器且在当日的tcp记录中为下载量第一（如图10所示），而在网页访问记录中，该员工在11月29日是所有员工中对该服务器访问次数最多的（如图11所示）。对于这样的行为，可以怀疑该员工在11月30日对ip为10.50.50.49的服务器中的数据进行窃取。

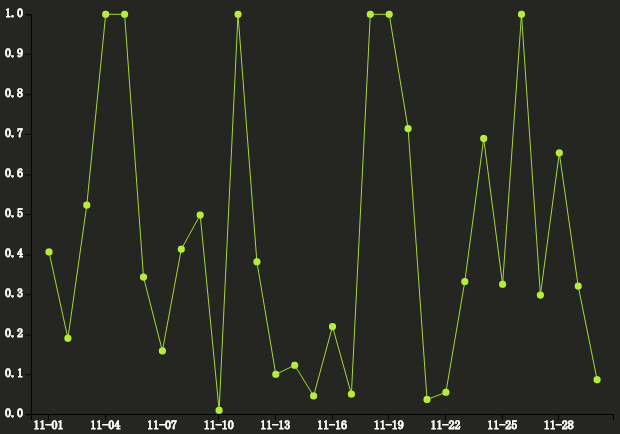


**图10：1175号员工11月下载量统计**

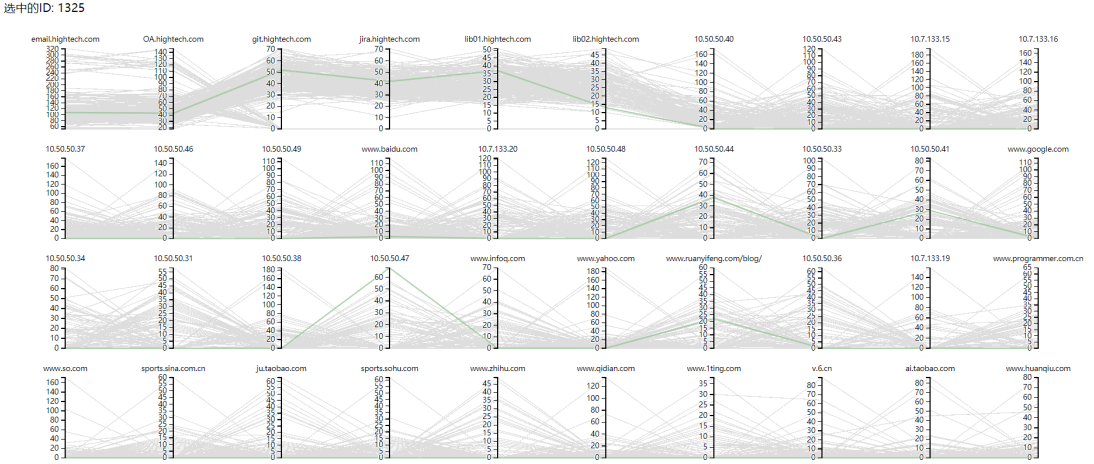


**图11：1175号员工11月网页访问统计**

1325号：该员工是一名来自研发部门的普通员工，其对应的ip地址为10.64.105.14。他在11月10日较晚时间（22点22分）登录了ip为10.50.50.44的服务器且在当日的tcp记录中为下载量第4名（如图12所示），同时他当天23点10分才下班，为当天最晚。而在11月的网页访问记录中，该员工对ip为10.50.50.47的服务器的访问次数为所有员工中的第一名（如图13所示）。对于这样的行为，可以怀疑该员工在当月对两台服务器进行过不同程度的数据采集和窃取。

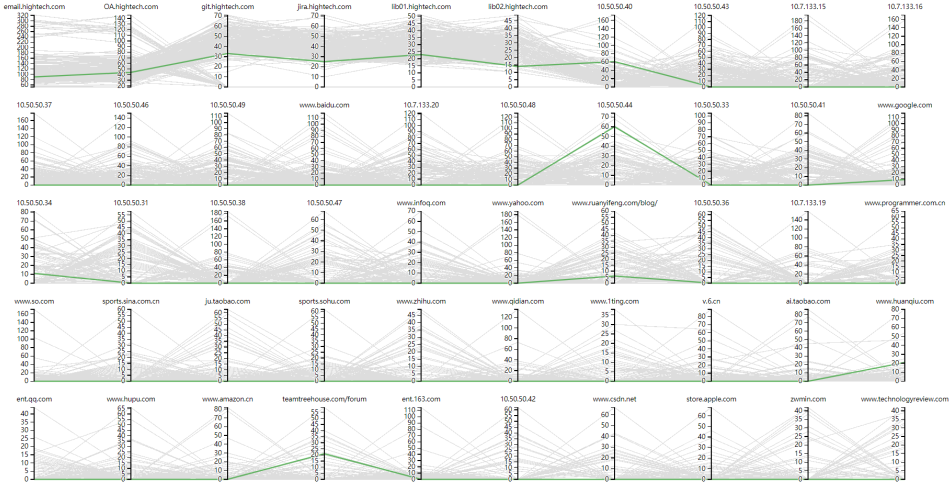


**图12：****1325号员工11月下载量统计**



**图13：1325号员工11月网页访问统计**

1173号：11月1号较晚时间（22:10）在50.47公司服务器上登录，当日下载量排名第10名；然而从weblog记录来看，该员工本月并未访问过该服务器的页面；该员工平时的网页访问量主要集中在研发组所有人都常用的几个服务器和50.44，并非50.47。总之，该员工在大多数人下班后登录并不常用的服务器并大量下载文件，我们认为这可能存在潜在的安全隐患：该员工可能利用系统漏洞私自拷贝不相关的资料，或者是电脑被安装病毒/木马程序之后自行下载资料(如图14)。

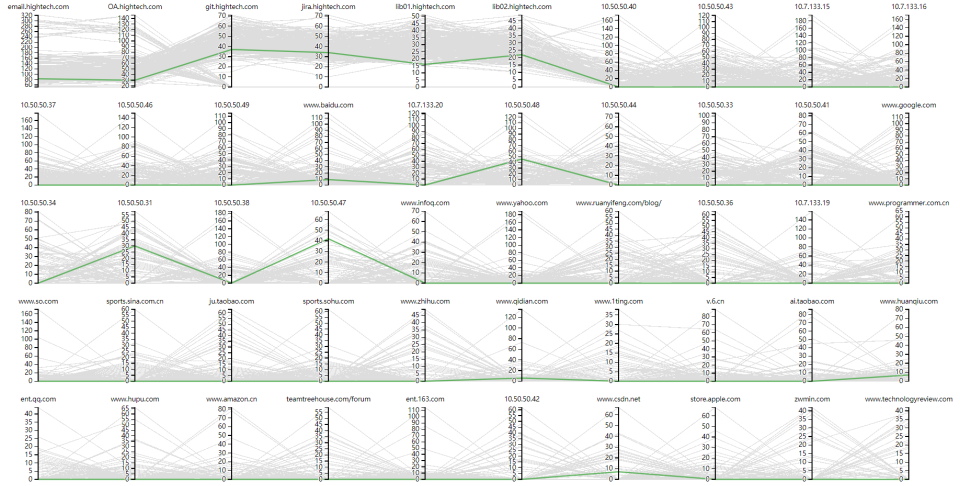


**图13：1173号员工11月网页访问统计**

1444号：该员工是一名来自研发部门的普通员工，其对应的ip为10.64.105.66。他在11月15日较晚时间（22点13分）登录了ip为10.50.50.47的服务器且在当日的tcp记录中为下载量第12名，在11月16日的记录中为第3名（如图14所示）。而在11月的网页访问记录中，该员工对ip为10.50.50.47的服务器的访问次数为所有员工中的第二名（如图15所示）。对于这样的行为，可以怀疑该员工在当月对该服务器进行过数据窃取。

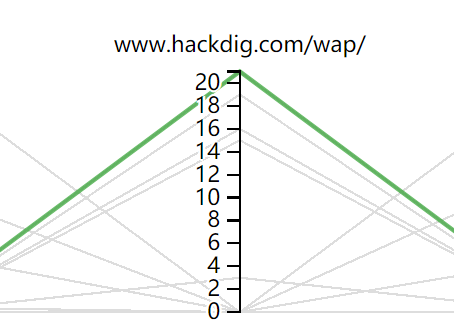


**图14：1444号员工11月下载量统计**



**图15：1444号员工11月网页访问统计**

1265号：该员工有访问黑客技术网站的记录(如图16)。11月14号，该员工比自己通常的下班时间（7-8点）晚了近两小时下班；当日，该员工下载总量在所有员工中排名第10，大部分流量来自服务器50.40；同时，该员工当日对50.40服务器的页面访问次数排名也很靠前。这几项异常让我们不得不对其动机产生怀疑，有进一步深入调查的必要。



**图16：1265号员工访问黑客网站记录**

在这5位有异常行为的员工中，1325号员工和1175号员工属于研发部门中的同一个组，因而他们两个可能存在合作偷取公司数据、窃取机密的行为；而1444号、1173号和1265号员工均属于部门中不同的组，他们的异常行为之间的关联相对较小。