FRAMSTEG

Innovation Catalogue Issue 1

Always one standard deviation ahead of the normal distribution

Table of contents

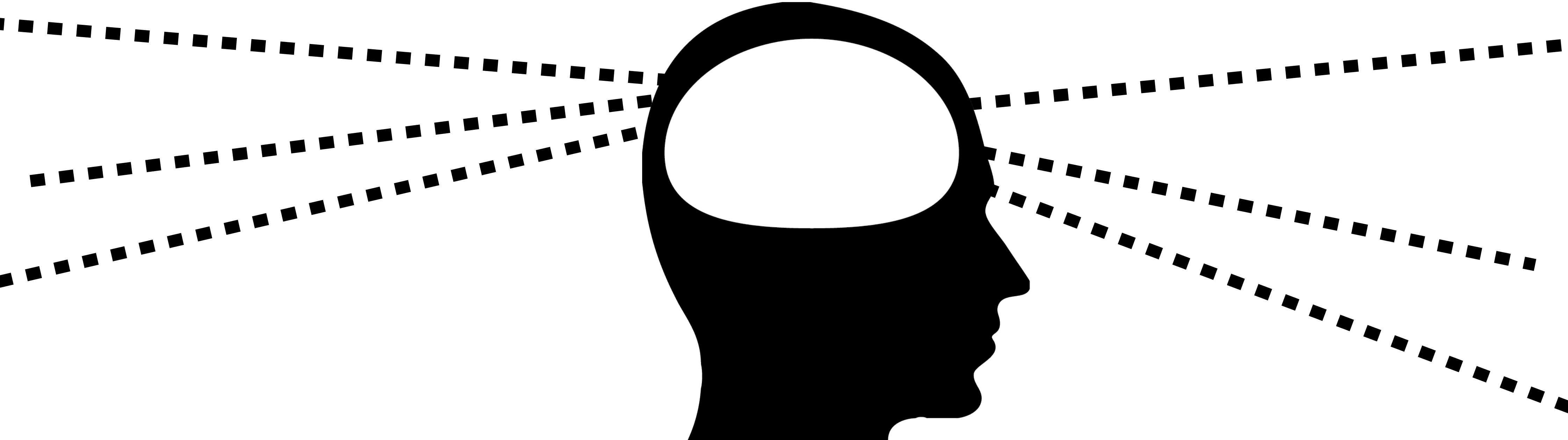
Perceived security and freedom

People today need an increasing amount of perceived security to bring optimal results at work in the freedom of the modern world. Also Millennials are essentially the same as all the other generations in the current workforce.

To derive at this conclusion, there are four necessary building bricks: Humans have a need for security, are choice-averse, want to make the right decisions and freedom is intertwined with complexity. Let me back this up:

1. Humans have a natural need for security in their life. This fact was most famously pointed out in Maslow's hierarchy of needs. Although the model itself is under harsh critic for various reasons, the existence of the need for security is commonly accepted. In this model, security was defined as the availability of resources and the absence of harms. On a more abstract level, security can be thought of as a state during which you don't have to question your current situation and can concentrate on other things.
2. secondly, humans are choice-aversive, as choice makes decision-making too complicated and unfluent. With increasing complexity of choice, we tend to fall back to simple cognitive heuristics as Daniel Kahnemann pointed out. Barry Schwartz even argues in his book "The Paradox of Choice" that too much choice results in no choice at all.
3. Thirdly, the terms freedom and complexity need be put in a relationship. Complexity is the non-linear reaction of two or more system agents together. Freedom in this context is to be understood as liberty, the power to do as one likes. As the modern social systems engage in increasing liberty through globalization, democratization and the internet, a greater amount of non-linear reactions of system agents is possible. This increases complexity.
4. Lastly, people want to make the right choices. Wrong choices generally are unpleasant because they have two results: You need to deal with the consequences (aka thinking about the situation), which is unpleasant, and you have to justify why you made the mistake in the first place (Yes cognitive dissonance, I meant you!), which is also unpleasant.

To the point: The more freedom people have, the more they crave for security to back up their choices and avoid wrong choices. Otherwise they



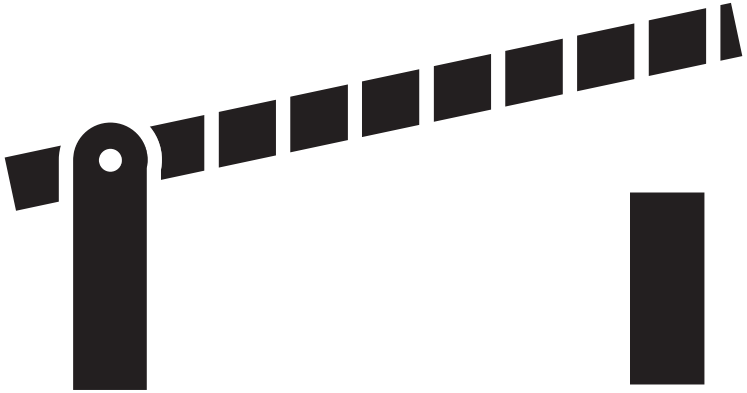
would be constantly challenged to evaluate all possible choices and non-linear reactions. This is to be avoided at all costs. This need for security in decisions can be expressed through several ways:

1. Feedback: A different opinion, preferably by someone more experienced in this sector, gives a reinforcement of the own choice.
2. Values and norms: A clear set of rules on how to evaluate things makes decision-making easier and therefore is emphasized. This can be done through the very central search for sense in life and work.
3. Structure: Routines and rituals try to avoid change under the assumption that essentially nothing has changed and therefore there is no need for constant re-evaluation.

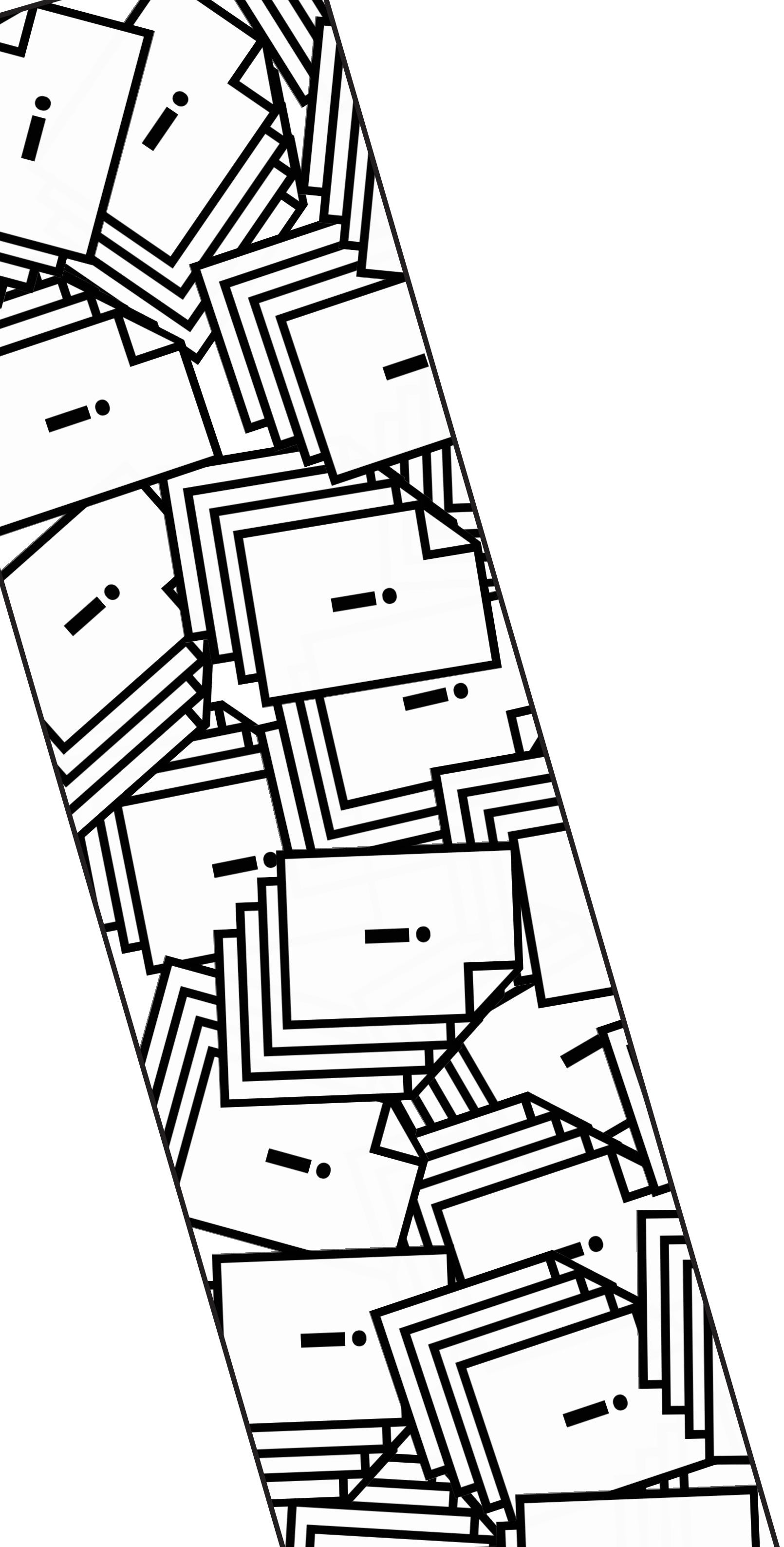
The first two aspects are perceived as traits of the Generation Y, but as I pointed out, they are rather expressions of fundamental human traits. They are just more expressed in the modern generation as it has the greatest amount of freedom possible until now. Interestingly, the third way is a more common choice for older people, independently of the generation. This backs up the cumulating proof that there is in fact no big difference between generations.

To get some implications from the theoretical construct up until now, there is just one more analogy for the connection of perceived security and active exploration in the current world. The attachment theory, famously proposed by Bowlby and later refined by Ainsworth: A healthy lifelong development through exploration first needs an equivalent of a safe haven. This safe haven provides us with security that everything is good and we can come back at any time. In attachment theory, kids with a healthy exploration tribe and which are open to the world see the mother as a safe haven to which they can find back at any time. The existence of this safe haven is what enables them to use their freedom. In the proposed theoretical framework, this theory is applicable not only to children but also to anyone in the work context.

To summarize and transfer this theoretical framework, there is one conclusion: If we want people to engage in the freedom and complexity of todays world, we must provide a sense of a safe haven. This will enable the current and the following workforces to work more broad-minded and think more in a connected and holistic way about the business problems of tomorrows world. This safe haven can be achieved through a clear valuation norm (aka tell people what is good and what is bad) or more frequent feedback loops (which is essentially the same, but in a more implicit way).



Clear values and norms will improve perceived security



Data Retention

Lately we have repeatedly reported on safety aspects of IT since the topic of data retention has come to public awareness internationally. We consider the currently proposed solution of data retention in Germany as wrong but will provide a better solution.

First of all, data retention is illegal according to law and violates the fundamental rights of the European Union. Secondly, the main argument for data retention is the ability to facilitate crime detection, but in fact there is not a single example for the contribution of data retention to crime detection. Thirdly, data retention involves a great risk: Since being able to grant the enduring security of the collected data in today's rapidly moving IT-world is a fairy tale, skilled hackers will be able to get hold of the data. This is not an acceptable option. Additionally, the german law enforcement has a history of damp squibs like the "Bundestrojaner" which doesn't quite encourage reliability.
Lastly, many perpetrators are aware of effective counters to data retention like VPN, Proxy and encryption. Therefore the proposed solutions in data retention would not apply to the target group, making the solution not only useless but dangerous for the illustrated reasons.

We propose a dual solution for the desired purpose: a first instance analyzes data in real time, making storage of random data dispensable. Based on algorithms like predictive behavioral targeting (as discussed in a previous article) the program is able to deduct further actions. If the acquired information is relevant, further information could be used. For example the program would now work with anonymous data such as hashed or alienated IPs or with the fingerprinting process. If the program would find concrete evidence of criminal activities, personal data could be acquired via court order.

As a result, the balance between security and privacy is ensured. Big companies like Google, Zanox and plista are already using these techniques. Why shouldn't we? 🟩

The impact of big data on privacy

There is much discussion lately about the advantages and disadvantages of Big Data. Prof. Dirk Helbing - for example - reflects fundamental social aspects in his analysis “Big data, big impact“ of the subject. But let us approach this issue properly without spewing just buzzwords.

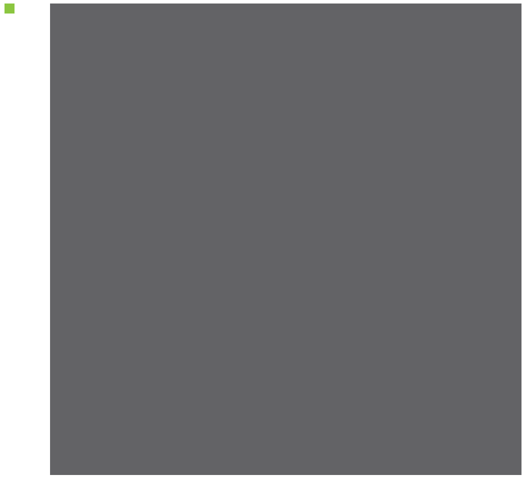
What is big data?

The term “big data” describes the collection of data records in such a large and complex scale that it becomes difficult to process them using on-hand database management tools or traditional data processing applications. The challenges of this topic includes the process of capturing, analysis, transferring as well as storing, searching and visualizing.

In a dynamic, global economy, organizations have begun to more heavily rely on insights from their customers, internal processes and business operations in order to uncover new opportunities for growth. In the process of discovering and determining these insights, large complex sets of data are generated that then must be managed, analyzed and manipulated by skilled professionals. The compilation of this large collection of data is collectively known as “big data.”

So, how big is big data?

Most professionals in the industry consider multiple terabytes or petabytes to be the current big data benchmark. Others, however, are hesitant to commit to a specific quantity, as the rapid pace of technological development may render today’s concept of “big” as tomorrow’s “normal.” Still others will define big data relative to its context. In other words, big data is a subjective label attached to situations in which human and technical infrastructures are unable to keep pace with a company’s data needs.



One terabyte (green), the size of a regular consumer drive, in relation to the size of a petabyte (grey)

What would be an example of big data?

The affiliate industry is the best example of this theme.There unimaginable amounts of information are often needed in this case to generate efficient advertising proposals. For a better understanding I will differentiate this topic:

A part of the affiliate industry works with so-called „subtle advertising“ like the zanox AG and the Affilinet GmbH. They provide fixed banners, which are often perceived by users as a nuisance.

On the other side there are companies like plista which provide rather recommendations. These recommendations are tailored to the interests and desires of the consumer. In order to realize such tailoring, they resort to scientific methods like collaborative filtering which are based on the Predictive Behavioral Targeting Method. Now, if such algorithms are to be used in real time it is necessary to resort to a variety of information. The more information corresponds to this technique, the more precise the failure behavior analysis will be.

In order to get an impression of this topic, one is able to look into the Open Recommendation Platform of plista. The ORP is a pretty cool feature developed by Torben Brodts engineering team where one is able to concept and implement its own algorithms. Furthermore these algorithms may be tested with valid recommendation traffic whereas the ORP returns the actual impressions of unique visitors. This should be considered one of the first real intel into such a deep and important Big Data topic. Researchers are able to collect huge amounts of anonymous data about visitors - and behavior data in order to analyze them.

How does this data analysis work?

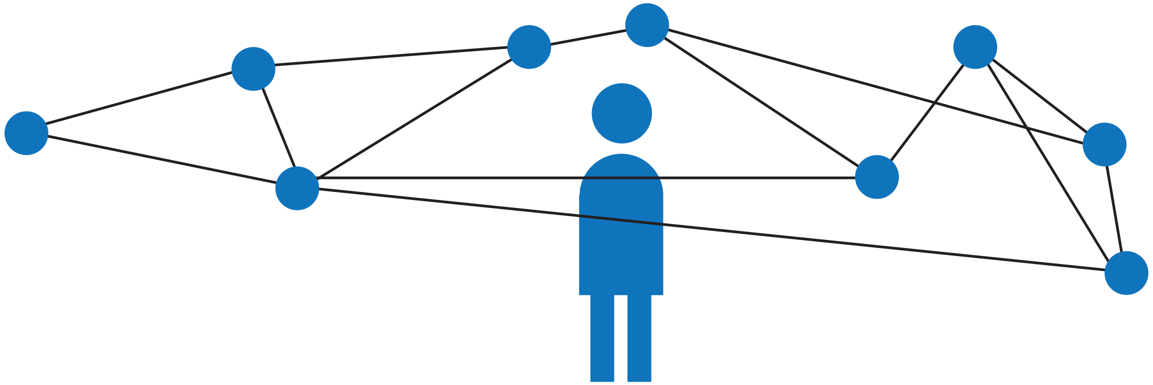
This basic concept is trivial. Imagine you have twins. They are exactly same looking, have the same interests and same character settings. At the point where you now all about one of this twins, you also know everything about the other twin. So the interesting part is to find the twin. And this stage is equivalent to finding the fitting target-group. Based on theories like collaborative filtering and predictable behavioral targeting the system establishes a matrix with so called interest tables. It also knows sample values of each target group.
For example the target group 50+, male, employee of the government. They are most likely interested in vacation trips, tips about losing fat and troubleshooting with losing hair. No offense. I probably will have the same problems. Another classical example would be a women between 25 and 35, which lastly bought sweet - sour products and searched for pregnancy topics. A typical recommendation would suggest baby products.
So basically the systems collects the connection between the targeted user and the tables of interest in the matrix and compares them to known target groups. The best fidding group calculated with a well secured algorithm is the solution after which the recommendation is chosen. Due to this procedure the user was put into one or several user groups in order to identify his interest and or behavior.

What could be a danger?

Like I already mentioned, Big Data provides a huge step for improving the work of researcher and personalized services. But we should also be aware of the possible side effects. These data and technologies could also be used in data mining & -profiling in order to create advanced behavioral model. As long as our government and authorities are not capable of ensuring the needed security- and awareness aspects as well as an applicable and thoroughgoing data protection, Big Data will remain a possible danger in my point of view. And we are already being confronted with this aspect like in this case, or that one, or this, that, that or that shows us from the last weeks.

What's new about this topic?

Security issues are not a real new danger. They always appear and will be always present. Belong this point a combination of attacks increase the potential risk exponentially. With the help of tracking pixel, its possible to collect even more information about users until the point where the user is fully transparent. The average user may know that he or she should not provide so much personal details. But one is not aware of the fact that several dots could still get connected in order to form a line.



Connecting the dots of information users reveal allows the assumption and attribution of specific target groups

A different spin on ad targeting

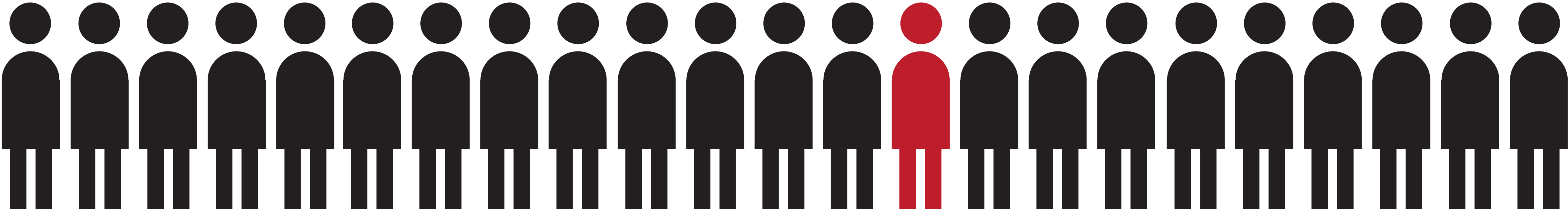
Whether geo-targeting, predictable behavioral targeting or retargeting, the intention remains the same - to provide targeted and personalized advertising. For this, the ad industry spans nearly a global network around the Internet users. But let us slow down and divide the problem into its constituent parts.

Looking back, what was the biggest challenge for the developers of targeted malware like stuxnet? The exclusive infection of the target group with malware, as uncomplicated as possible.Therefore from the perspective of the attacker it should be avoided to fall for traps and Fake Clients of security agencies and security companies like Kaspersky and Norton. These could then use the information that was obtained as a warning. And there is nothing worse than a sensitive users. Hence, targeted delivery of malware is desirable in order to minimize the risk of detection. But now what do these two topics have in common?

They combine the subject of targeting. For suppose if you want to specifically attack a person or group of people, an ad network would be the ideal instrument for this purpose. More than 50 % of all websites are integrated with ad networks worldwide. Google is already integrated with AdSense and Google Analytics in more than 60 % of websites globally. The affiliate network around the zanox AG as a member of United Internet has more than one million websites that act as a publisher. These publishers can now be summarized as (advertising) network to gather information about the user and deliver personalized content (see our publication on Big Data). Using such a network, attackers can deliver malware tailored to a user group, which is hidden behind a layer of advertising.

The procedure is relatively subtle: For this purpose a simple XSS attack is executed on the network structure to integrate the malware. The specific calculation and allocation of the target group as in the case of predictable behavioral targeting is mostly taken over by the affiliate network. For transmission of the parameters JSON strings are often used, which can also be intercepted and read.

Based on those parameters one may then read out the affiliation of the target group. Since the corresponding communication is not encrypted due to performance aspects, there is no reason why the interception of information should not work. However, it requires a certain basic understanding of the attacker for the affiliate industry. Instead of sending an Ad Recommendation, a potential attacker may transmit an simple malware script like Mabezat. Furthermore he could keep tracking his victim with the interaction JSON Strings.



A potential victim could be for example be the target group: male, mid-50s , an official with computer skills who likes to shred documents . This would most likely fit to the gentlemen of the Federal Office for the Protection of the Constitution (BfV) or German Federal Office of Criminal Investigation (BKA), State Office of Criminal Investigation (LKA), or Federal Information Service (BND) - outgoing to infiltrate their networks. Through the use of extensive advertising network, malware could be executed with javascript directly by the browser. Thus, further or more extensive malicious software could be downloaded to the target machine.

In most cases, only the appropriate user or the appropriate target will be affected and the detection of an attack would be more than difficult. It is doubtful whether the distributing network can ever become aware of it. Automated checks such as Norton Safe Web crawler or automated code analysis would not be succesful if they do not imitate the target group.

It lends itself to the possibility of combining this attack scenario with tracking pixels. You can now send irrelevant request to a destination and then use the tracking pixel to find out the main stages and the end node. Most mail servers such as Postfix and Exchange reload external media for faster retrieval (see caching / proxy methods). Using this information and the knowledge that large companies, universities and government agencies have received specific IP ranges, one may now selectively filter out ads by the target group and the address range in order to manipulate them.

If the potential attacker would now like to specifically attack a person because it is a terrorist, a fraudster, citizen or other arguments of the intelligence agencies, then this would be also realized by AdTargeting. The principle of the tracking cookies can in addition to the IP be used for grabbing device identifier and software characteristics (operating system, browser etc. An attacker uses this information to address the target. In addition to the IP users can be identified (with a 89% probability) using the hash value of one's browser settings and extensions. Another approach would be equivalent to the procedure already described. Due to an XSS attack malicious software is introduced on the Client side. This conceptual approach is subtle but effective. Thanks to the advertising industry. The attacker saves huge time, minimizes his personal risk and maximizes permanence of malicious software in terms of sustainability.

So what do we learn from this? A multiple cross-cutting understanding can contribute significantly to the solution of individual topics and develop new solutions but is also revealing new risks.

All script snippets are available here or just scan this QR-Code

IT self defense

Recently the question of IT self defense came up and I found this topic quite interesting so i started digging. In situations of the natural world the use of self-defense according to German law is defined with § 32 of the Criminal Code. But can this section also be used in the IT world? Is there a way to defend yourself against IT attacks? And if so, to what extent? Let us roll up this topic from the bottom and start with the legal part.

What is self defense?

"Self-defence means any defensive action that is necessary to avert an imminent unlawful attack on oneself or another." § 32 StGB . Further this paragraph defines: "A person who commits an act in self-defence does not act unlawfully."

Does this also apply in the IT world?

That is quite difficult due to the fact that there is no explicit law written yet. I talked about this topic with Monika Menz (lawyer for Ernst & Young Law GmbH, Head of Practice Group IP / IT). They referred me to two elementary paragraphs of the Telecommunications Act (TKG) and the Federal Data Protection Act (BDSG), which regulate the fundamental practices of the IT world:

§ 100 TKG: Faults in Telecommunications Systems and Telecommunications Service Fraud

This section shall apply only to telecom operators. However it regulates the usage of personal data in the German IT business industry. Paragraph 1 allows the recording of user data for purposes of interference avoidance:

"(1) Where required, the service provider may collect and use the customer data and traffic data of subscribers and users in order to detect, locate and eliminate faults and malfunctions in telecommunications systems."

This record may take place but only temporarily and must be deleted immediately. Except when paragraph 3 is applied:

"(3) Where required, the service provider may collect and use the customer data and traffic data needed to detect and put a stop to the surreptitious use of services and other unlawful use of telecommunications networks and services when there are grounds, to be recorded in writing, to suppose such use. For the purpose referred to in sentence 1 the service provider may use collected traffic data in such a way as to identify, from the total traffic data not more than six months old, the data relating to those network connections in respect of which there are grounds to suppose that unlawful use of telecommunications networks and services has been made."

Due to this statement, not only logging personal data like the IP but also data packages are allowed, if an attack takes place. But keep in mind this applies only to telecommunications companies. For "personal" matters or matters of any other kind the regulation of the Federal Data Protection Act (BSDG) is required. The Federal Data Protection Act regulates the acquisition, processing and storage of any form of data in Germany. In particular, paragraph 28 of the BSDG is relevant.

§ 28 BDSG: Collection and recording of data for own commercial purposes

The first section authorizes the use of user data when a transaction is present. One could argue that the general good is assumed to believe by the use of server capacity and the user therefore admits only the ordinary handling. Thus, a legal transaction for the time spent on site or other resource would exist.

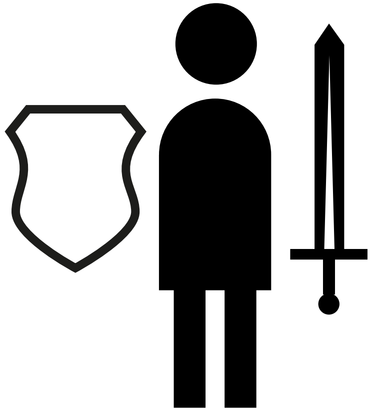
“(1) The collection, recording, alteration or transfer of personal data or their use as a means to pursue own commercial purposes shall be lawful”
“1. [...] ”
“2. as far as necessary to safeguard legitimate interests of the controller and there is no reason to assume that the data subject has an overriding legitimate interest in ruling out the possibility of processing or use, or”

However, the following section is much more interesting. This concedes the usage if the user‘s data are publicly available or responsible entity should publish them.

“3. if the data are generally accessible or the controller would be allowed to publish them, unless the data subject has a clear and overriding legitimate interest in ruling out the possibility of processing or use.”

In the legal sense the IP but also browser cookies and other personal data are understood as easily accessible data.The restriction that the legitimate interests of the person concerned obviously outweighs the exclusion of the processing or use in relation to the legitimate interests of the the responsible entity would run out here on a comparison of interest. Due to the fact, that the interest of the user is directed to its attack and therefore to harm the defender, his interest is supposed to be depreciated. Therefore, the welfare and interest of the owner and the attacked has higher priority than the interest of the attacker. This argument is also reflected in section 8.

“(6) The collection, processing and use of special categories of personal data (Section 3 (9)) for own commercial purposes shall be lawful without the data subject’s consent in accordance with Section 4a (3) if



1. necessary to protect the vital interests of the data subject or of another person where the data subject is physically or legally incapable of giving his or her consent,
2. data are involved which the data subject has manifestly made public,
3. necessary to assert, exercise or defend legal claims and there is no reason to assume that the data subject has an overriding legitimate interest in ruling out the possibility of collection, processing or use, or
4. necessary for the purposes of scientific research, where the scientific interest in carrying out the research project significantly outweighs the data subject’s interest in ruling out the possibility of collection, processing and use and the purpose of the research cannot be achieved in any other way or would require a disproportionate effort.”
Section 8 describes the handling of user data in the prosecution of offenses.

“(8) [..] Transfer or use shall be lawful also if necessary to prevent threats to state and public security or to prosecute serious crimes.”

Furthermore, when we look back to the Criminal Code (StGB) the „necessity“ in accordance with § 34 Criminal Code specifies

“A person who, faced with an imminent danger to life, limb, freedom, honour, property or another legal interestwhich cannot otherwise be averted, commits an act to avert the danger from himself or another, does not act unlawfully, if, upon weighing the conflicting interests, in particular the affected legal interests and the degree of the danger facing them, the protected interest substantially outweighs the one interfered with. This shall apply only if and to the extent that the act committed is an adequate means to avert the danger."

However, this applies only if the act is an appropriate means to avert the danger. For a calamity capable of legal protection (ife, limb, liberty, honor, property or another legal) a present danger must be present. Risk means that as soon as damage occurs due to objective circumstances, it appears probable and currently means that the risk can change at once or in a very short time in a loss. The emergency action takes place at the justifiable emergency in contrast for self-defense only against legal interests of a third party!

The danger must not be otherwise be averted, that is to say the the mildest must be selected from the resources available.

In contrast to self-defense, the interest in protecting the endangered legal interests must be the interest on impaired when emergency significantly outweigh (balance of interests). According to § 32 sentence 2 of the Criminal Code, emergency action must also be reasonable. This will ensure that the non- action against recognized values and principles of law contrary. The defender must know the circumstances justifying act and security.

Brilliant, I stopped reading at the second line. What does this mean?

Interestingly, a lot of security rights are granted by this legislation to the service owner.

You are allowed to store the IP of a user in log files in order to detect attacks on your system.

You are allowed to "defend" yourself against attacks but transfer these data to a redundant storage location like your HDD.

If someone attacks you with the intention of destroying or harming your system, you are allowed to strike back with the note of proportionality. (The attack has to be present, and undeniably illegal..)

Normally, a "good" hacker uses proxies and anonymizing software in order to hide his origin. Therefore the IP doesn't help to identify the attack generally. But however, you may look into the possibilities of part 3.

And how am i able to strike back?

By now, i am able to get back to my original story. During the implementation of a game API i did a stupid mistake. I simply forgot to put a filter which checks the returned value for INTEGER.

filter_var($score, FILTER_SANITIZE_NUMBER_INT);

Due to this missing line it was possible for the attacker to execute a simple JS Injection attack. Because i was loading the value on a website in order to accomplish a leaderboard, it was a breeze for the attackers to launch a php console with the right of the web server. I'll spare you the ending of the story. However, I have thought of a potential attack and have therefore taken safety precautions:

Since the game is accessing the API with an empty user agent, other queries would always have a human origin. So I built a gateway, which recognized such a calling and executed a small script. This script does not only log the IP but also imitates different web servers in order to read the cookies of the attacker. In order to circumvent the security settings of the browser my webserver had to adjust its identification characteristics to those of the other web server. As far as i can check it, firefox and opera doesn't check the rDNS record of the webserver's IP neither the mounted ips and domains. So, i was able to access the active ID and auth of my attackers social media account.

What do we learn?

Yes, always use filters. But if you forget one or just for precaution, you can built a self defender in order to detect your enemy. You may not use these data further against him but now you have the advantage to report him. And by referring to §202c StGB your attacker has bad cards to get away with an exhortation.

A reliable protection against DDoS attacks

A Distributed Denial-of-Service Attack is an attempt to make a server or network resource unavailable to its intended users (mostly). This includes the efforts to temporarily or indefinitely interrupt services of a host server. But how can we defend ourselves?

The other day something extraordinary happened which gave me reason to think more precisely about the issue of DDoS attacks:
I participated in an IT jourfixe meeting where the security infrastructure of a successful startup was analyzed. A rather unusual subject, since this startup operates in a business area which is less prone to DDoS attacks. The question if the IT department is prepared against IT attacks came up. A number of attack scenarios were weighed. XSS and SQL Injection were double checked but the matter of DDoS was not addressed. So I questioned the topic.. The responsible employee answered short and subtle that one just had to stock up the data center if some kind of DDoS attack happens. It should not be forgotten that this jourfixe meeting concerned a free service provider who earns its surplus due to performance-based commissions. So this startup was not a financial service or product service provider - which are the preferred targets of DDoS attacks. Is it therefore always proportionate or rather appropriate to supplement the entire technical infrastructure?

Why would anyone want to reach disable a service?
Such an attack can have multiple intentions: Firstly it can harm a competitor.. For example if Amazon would be a day or more offline, that would imply the loss of profits in millions. But also the damage to their reputation would be considerable. A different intention can be by extorting businesses from the said reasons (equivalent to the ambition of data-kidnapping by using trojaner). The BKA estimates the amount of corresponding "business" volume in the millions.

How does a DDoS work?
There are countless methods how to establish a DDoS attack. A famous example is the "ping of death" method. By sending big packets (>65 kbytes), the target is overwhelmed by the size of the files and crashes. The most widely used form of attack is the Low Orbit Ion Cannon (LOIC). The LOIC floods the target server with TCP or UDP packets. Furthermore it is often used in "manual botnets".

What about the legal part?
The topic of DDoS or rather IT attacks in the legal sciences is pretty new. The German case law has just a few right-looking legal paragraphs. Law experts evaluated an DDoS attack as a kind of computer sabotage which includes up to 3 years imprisoned after § 303b Abs. 1 StGB or a fine. The so called "hackerparagraph" as part of § 202c StGB defines the creation, usage or distribution of every kind of software - which could be used for improper or dishonest intention - as a federal crime. The LOIC would be such a software.
An interesting fact is, that the reseller is subject to the contract risk - which means - in case of a DDoS the reseller has to pay for possible unavailability of infrastructure (due to case law of the district court Gelnhausen, Germany).

So, how can we prevent this kind of attacks?
Prevention is in most cases the best method, but it is not always possible or necessary. There are 3 counter opportunities:
1. "Software Settings", like limiting the amount of possible accesses to every single IP address mapped to the corresponding server or decreasing the priority of every request send by the same dynamic IP (recognizable by dissolving the rDNS entry).
2. "Infrastructural Design Patterns": using "smart" infrastructures. it is not necessary to use the biggest and most badass server to handle the entire

incoming traffic. Think smart and have an eye on scalability. A well designed and scalable infrastructure is the best way to handle DDoS. There are different design patterns which solves such "traffic overheads" like the following design pattern:
In this case there is a bigger cluster performing the task of a inverse multiplexer. All incoming traffic is randomly or ordered splitted and handed over to a different slave server which processes the package.
On the way back, the corresponding answer packet is handed back the same way it came. This process makes the infrastructure just flexible and scalable. If more resources are needed, you only have to clone a "cheap" server. only the first component which simulates the multiplexer as kind of proxy has to be powerful (e.g. dual 10 Gbit/s depending on the needs).
Actually there are two option how to solve outgoing traffic. The first one is to let each server response with its own ip. This technique decreases the risk of system failure but also increases the efficiency of DDoS attacks, because a potential attacker may collect the slave server ips in order to attack them directly. The second technique is to reroute the traffic over a proxy. This could be the same server cluster which handles the incoming traffic or a separate on which specializes on the outgoing traffic. The advantages and disadvantages would be opposite to the first option.

3. "Smart Handling of Traffic Packages": This involves the analysis of the incoming packets in order to discard unwanted, unimportant or redundant packets. A so called smart system, which is based on the predictives behavioral pattern recognition on which we are currently researching.
The predictive behavioral pattern recognition helps us to understand the intention of the client which is sending the packet, however it presupposes the possibility to read and analysis incoming messages.
In Addition, there is the option to use Round Robin for DNS. This technique is used for load balancing which already differs from the requirements for higher availability. One may use multiple DNS A Records (for IPv6 AAA Records) equivalent to the MX setup. At the moment the first destination of the DNS entry is unreachable, the host will try to connect the second one. And then the third, fourth and so one. The „Resource Record Set" method neglects the weight of the items. Accordingly, the DNS server returns all the records on request, however, this order changes during every query. Bind-level name servers support three types: cyclical, random and fixed. In more modern resource record types like SRV or NAPTR one is able to define a weight that determines which server IP addresses most commonly come first. Therefore the relevant servers are addressed more often. This concept could also be expand in order to make a DDoS more complicated. If the list of possible destination server is long enough, the first will be back online when the last entry is under attack.

And if all named approaches are not sufficient, there is still a last resort option to throw more hardware against the attacker. Loadmasters like the LM-5300 Server Load Balancer are able to handle up to 8.8 Gbps with 9,300 SSL Transactions Per/Second (TPS), 110,000 HTTP Requests per second or 400,000 L7 concurrent connections. Each of them. There also have been companies dedicated to the protection against DDoS attacks like CloudFlare, but we should always preserve the aspect of proportionality and economy in mind. In addition, this kind of service increases the danger of man-in-the-middle attacks which could be much more harmful in terms of long-term vision. Even CloudFare struggles with NTP-Reflection Attacks which easily generated 400 Gigabit/s.

Does it make sense to extract the heavy guns as in the initial example? Or isn't it smarter to go back to the smaller dedicated methods in order the solve our DDoS problem? That should everyone decide for themselves. ■

BPM against organized IT crime

During the Chaos Communication Congress 2013 (30C3) we were confronted with the subject of electronic robberies. The first attacks on ATM have already taken place. A comparatively burgeoning security issue is the ISO/IEC 14443 standard, which is also known as the paypass method provided by mastercard. You may want to purchase one of these programmable card readers and walk through a mall - then you are able to credit 10 bucks from each person you meet. Remember how many people are there?

Criminals are enjoying increasingly technological progress and the lack of sensitivity of its users. These effects become clear particularly in the areas of (IT) fraud and the social engineerings. Since criminals are always looking for the weak point in systems, we go after it as well. In most cases, these vulnerabilities are now more based due to gaps in business processes.

A Forbes.com story from Adrian Kingsley-Hughes explains that a former contributor for Gizmodo, Mat Honan, was the original victim of the attack. Hackers were able to access Honan's iCloud account, and remotely wipe his iPhone, iPad, and MacBook. The original theory was that the hackers used a brute force attack to crack Honan's iCloud password, but further investigation revealed that social engineering was used to convince Apple the attackers were Honan, and Apple gave them the keys to walk right in.

Funny. Now we are looking for no more gaps in firewalls and web servers but in business processes. A similar behavior is evident when we look at the problem of fraud in the advertising industry. By sending skillful requests to a known advertising platform one is able to receive informations about the internal audit process. This knowledge can be used in order to reconstruct the original audit process. This information can be used equivalently to the knowledge of safety barriers in conventional systems for the purpose of circumvent them.

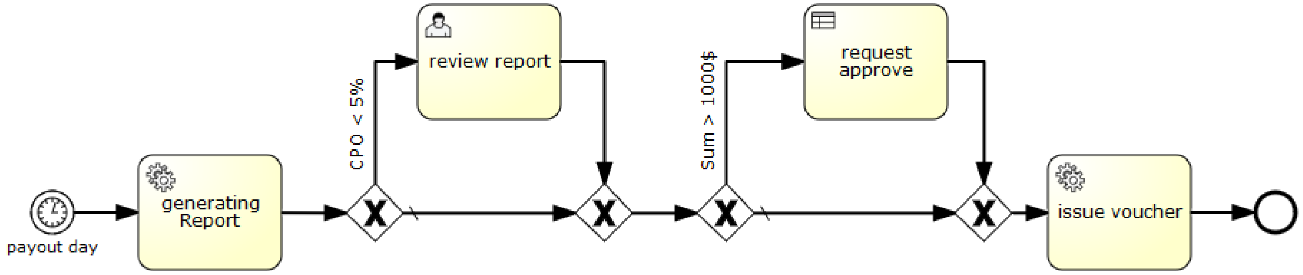
There is no way to reconstruct an entire internal audit process, it will be surely always missing information, huh?

Unfortunately, there is always a way to handle missing informations. Analogously to the method used in our Fraud Detection API there are also methods allowing us to solve the same is in this case. For example the Advanced Behavioral Appropriateness Metric or this one .

Therefore, we can summarize that an one-sided view on security analysis is no longer sufficient in the current situation. Rather it also requires to involve the non-technical vulnerabilities in complex systems.

How does this non-technical vulnerability look like?

Let me illustrate this with a much simplified Business Process Model example by falling back on a modelling technique called BPMN. BPMN is not directly intended to identify such vulnerabilities but we can misuse the use case.



Example Process 1: Internal Audit of an affiliate company

Lets imagine, we were able to collect the necessary information about the internal audit process of an affiliate platform (Example Process 1). Activities with the little wheel in the top left corner represent task which will be done by a technical system. Activities with a person in the top left corner will be executed from a human being.

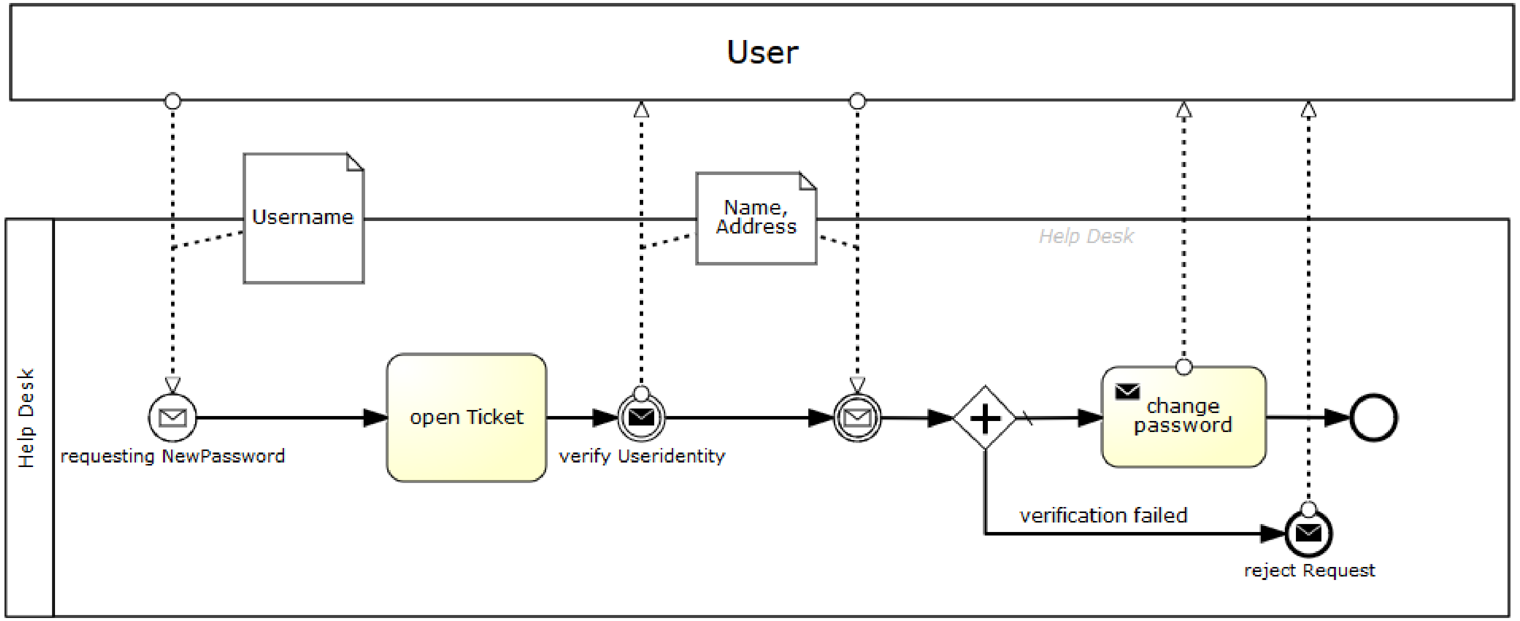
As a potential attacker or fraudster, we are now able to analyse the path of least resistance. So we notice, when our report has a lower value than 5% we can bypass the review from a accounting manager. Additionally, when the sum is less than 1000$ we are also able to bypass the approval request from a supervisor.

Do you think the example is too abstract?

Alright, lets take another one (Example Process 2). Imagine you forgot your password and therefor you request a new one from our iCompany. So you contact the help desk in order to change your password.

You provide your username. Additionally you are asked to identify yourself. With the knowledge how to help desk will verify the identity of the user, a potential hacker may gather the necessary information by checking social networks or public telephone books in order to successfully verify a wrong user identity.

This knowledge is dangerous when dealing with non-sensitized users. But it also allows companies to assess their current business processes to security gaps. 🟩



Example Process 2: Verifying a user

Increasing the efficiency of Predictable Behavioral Targeting (PBT)

In the context of developing the fraud data base and a social networking project, the idea of combining those topics came to my mind. There is a possibility to increase the performance of predictable behavioural targeting algorithms due to the analysis of social networks. That may sound a bit weird in first place, but let me explain it.

What is Predictable behavioral targeting?

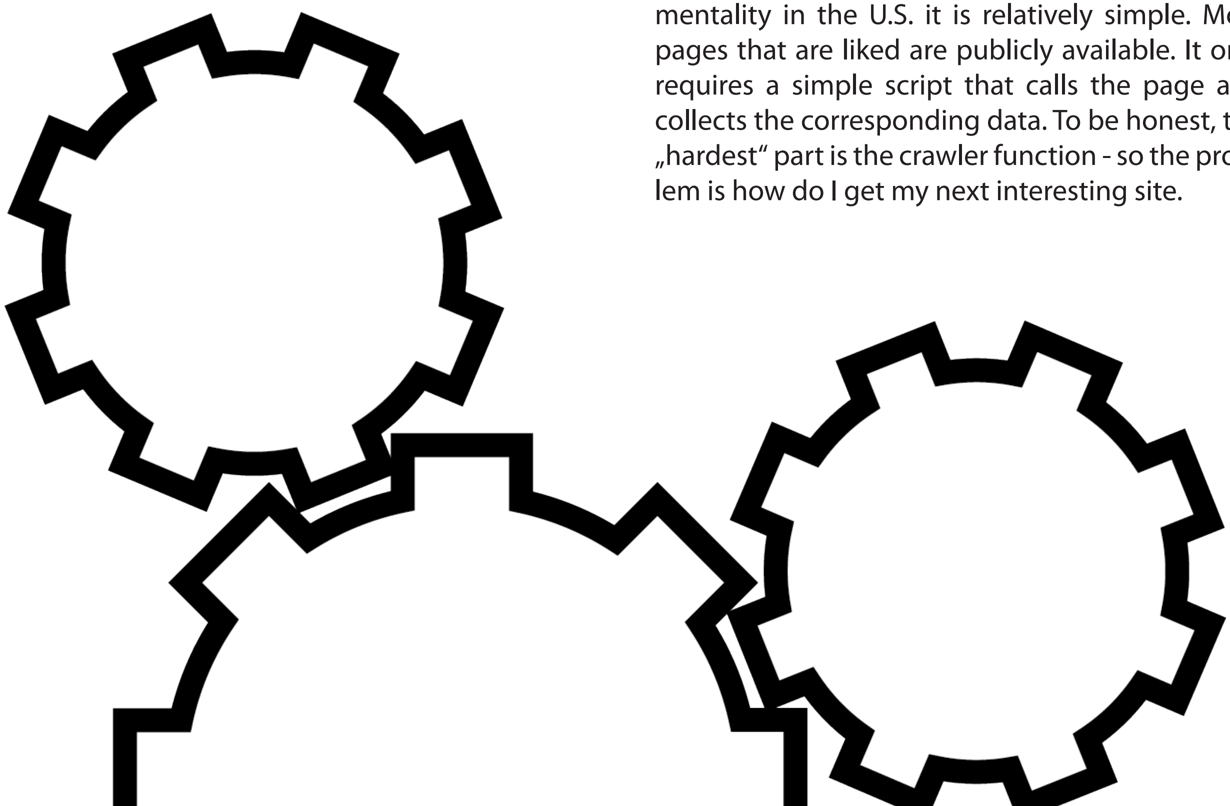
Predictable behavioral targeting (PBT) is used in the affiliate and marketing industry as well as in the security industry in order to forecast the most probable upcoming behavior of a user or target. This goes back to the topic of probabilities in the scientist field of calculus. I already described a technique called Collaborative Filtering which is implementing the topic. Collaborative Filtering works with predefined groups. The system has enough information about a group of users and if a new user appears, it tries to match parameters in order to match this user in its best fitting group.

And what about this social networking thing?

The interests of today's world are changing rapidly. 5 years ago instant messenger like ICQ and MSN were used all over the world, today everyone has a SMS Flatrate or uses WhatsApp. It seems like the interests of the people changes over time, hence the equivalent matching groups in the Predictable behavioral targeting pattern have to adjust to these changes. For many it is uncertain or they neglect this aspect. So it can occur quite possible that after 4-6 years the accuracy of a system decreases by 50%. Accordingly, the predefined groups must be adapted restive. One way to make that happen is by using a learning system, but what if our system now learns from wrong false or incorrect mappings (also considered as Type I and type II errors)? A possible starting point would be to make use of the information provided by social networks. In the American culture, the so-called liking of relevant pages is quite large. On average, each U.S. user likes 10 pages per day. If one is able to access these data, it is easy to refresh or rather update the target groups and in addition, it is also possible to refine the mapping between them. This requires a simple matrix analysis of the collected data.

And how do you want to get these data from social networks like facebook?

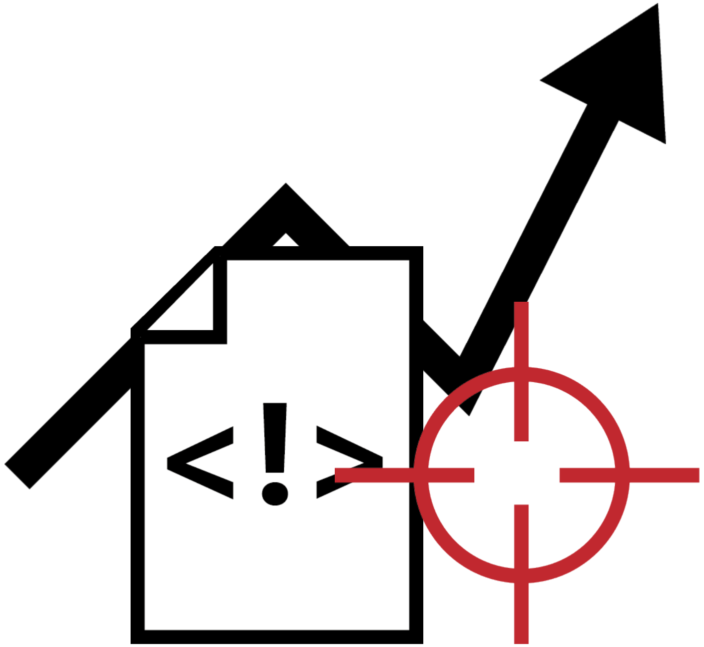
Thanks to the lax privacy practices and the open mentality in the U.S. it is relatively simple. Most pages that are liked are publicly available. It only requires a simple script that calls the page and collects the corresponding data. To be honest, the „hardest" part is the crawler function - so the problem is how do I get my next interesting site.



Where could be a possible starting point for such a crawler?

The topic of interpersonal relationships or rather the network of acquaintances and interests is best represented in so called “traffic exchange networks” or “follower exchange network”. The sense of these platforms is to generate a high amount of likes, followers and subscribers in social networks like facebook, twitter and so on. Like you can read in this article from IBN Live (Global Broadcast News) the topic of exchange networks become such a big deal during the last years, especially during the last election campaign in the US. In order to accomplish this goal, the webmaster publish their website on one of the network platforms and pays for the display of a link. In most cases huge companies or political parties, who want to push their public relations, advertise their links. Just start to crawl them and you will have a nice and long starting list for your crawler.

After you were able to collect a bigger amount of data, please be sure to anonymize them. It is not just about the privacy topic but rather that no one want to be named in that matter. Not you, not me, nobody. So please respect that. Hereafter you can use these information as predefinings for the matching groups. The rest of the procedure is followed by the well-known process of collaborative filtering. 🟩



Precise Predictable Behavioural Targeting will increase your ads performance

Browser fingerprinting

During the work for our fraud DB project we were stuck at the point of identifying users without using their IP address due to privacy issues.But there are several other reasons not to rely on the IP solely. What happens if someone uses a VPN or proxy service? What about public hotspots (which are increasing)? Therefore we concluded that we needed a different approach which focuses on the device, not the IP.

What is a (web) browser?

A browser is a software application which is used to locate, retrieve and display content. In the client/server model, the browser is the client which is run on a computer. It contacts the server and requests information. The server sends the information back to the browser, which displays the results on the computer.

We discovered that browsers are very unique. The adjusted profile of plugins (like AdBlock, Norton or Firebug), fonts, languages, Add-Ons etc. forms a highly unique fingerprint. In fact, you can be 90% sure that you've identified one exact browser.

Crazy, who would need that fancy technique?

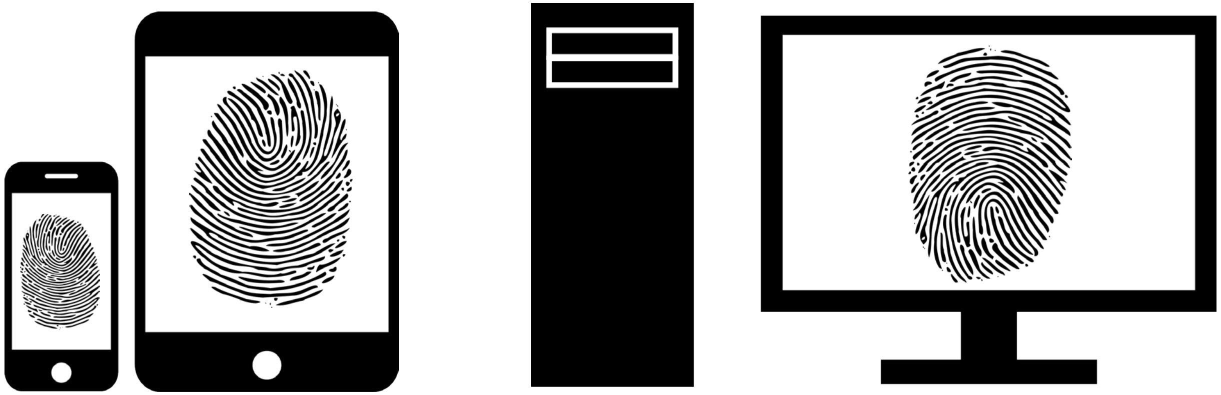
User identification is a valuable technique for many companies from the affiliate industry to market research and even for intelligence agencies. Just to make that clear, the affiliate industry, market research, IT Security companies, our dear friend the intelligence agencies and so on.
In addition, browser fingerprinting may be the first step towards overcoming cookies and the accompanied risk of session hijacking. Until now, cookies were necessary to overcome the limitations of the stateless protocol (HTTP). To conclude, browser fingerprinting would be a significant gain for IT security.

What is up with the technical implementation?

To examine the technical implementation, take a look at the thesis of Henning Tillmann. A basic understanding is provided in the following section:
The principle is very simple, grab all the information you are able to collect about the browser and use a algorithm to combine them. PHP already provides us with interesting functions like apache_request_headers, which fetches all HTTP request headers from the current request and returns an associative array. If you don't use PHP as a module of apache2, you can use a function like the one available on our website.

Does the browser fingerprint change over time?

Yes, indeed. In more than half of the cases, the fingerprint has actually changed. But through the use of predictable behavioral analysis, we could design an algorithm which calculates possible changes. Users may install or uninstall plugins, or new software. Maybe the time settings changes or the location. We predict all these changes up to a specific level of abstraction in order to improve our set of results.



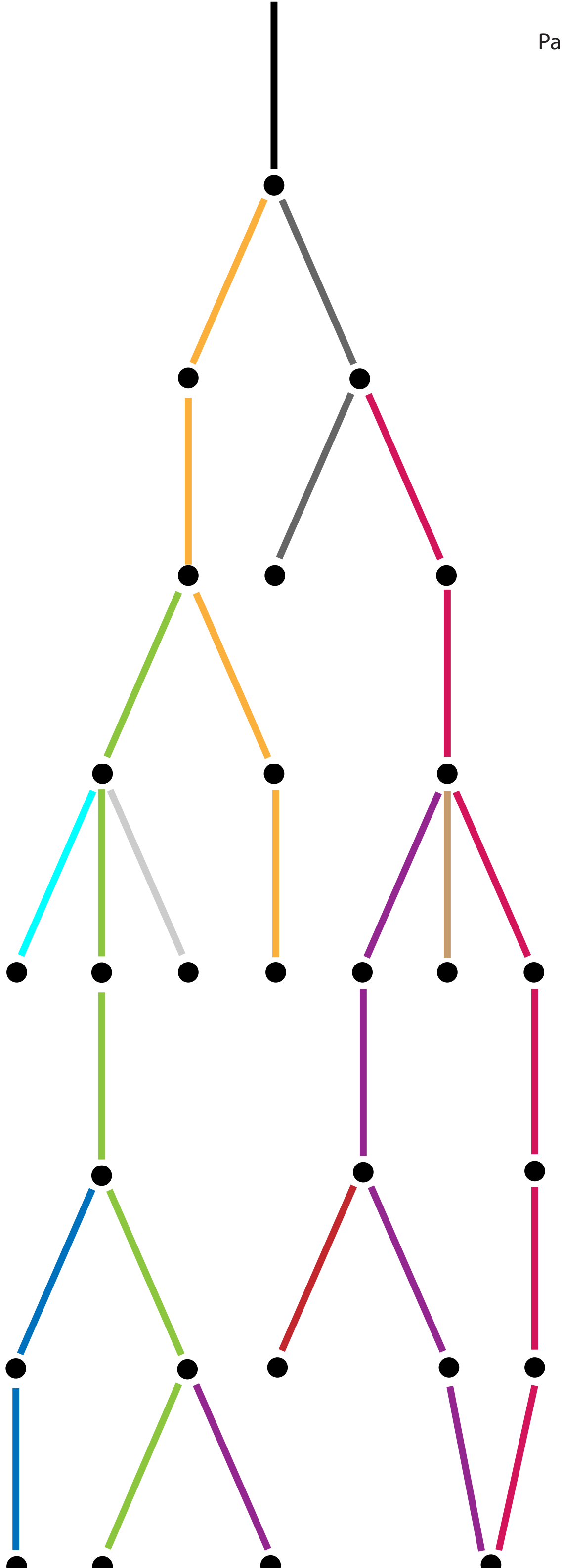
You can be 90% sure to identify an exact browser by checking its profiles and plugins

Decision graph for Predictable Behavioral Analysis

During the Chaos Communication Congress 2013 (30C3) we were confronted with the subject of electronic robberies. The first attacks on ATM have already taken place. A comparatively burgeoning security issue is the ISO/IEC 14443 standard, which is also known as the paypass method provided by mastercard. You may want to purchase one of these programmable card readers and walk through a mall - then you are able to credit 10 bucks from each person you meet. Remember how many people are there?

The evolution of the used technologies and techniques in affiliate industry grows fast and therefore changes fast too. In fact, it is one of the most increasing and changing industry currently. Some years ago affiliate platforms worked with static banners written in HTML. The first so called retargeting algorithm was used 10 years ago. We all knows this concept from amazon, who used retargeting a long time, or rather still using it. For example if you buy an external hard drive amazon is still promoting external hard drives to you even weeks after the purchase. The second step was the using of predictable behavioral analysis algorithms, which does not dumb repromote advertise to you but rather check which one is fitting to your interest and suggesting similar products. Nowadays when you purchase an external hard drive you will get ads for other hardware products like CPU or graphic chips. The latest invention was an improvement of the original collaborative filtering algorithm. It does not only recognize your interest in order to match a third fitting interest group it also detects so called forwarding behaviors. In fact, if you are a lady around 30 and purchase sweet source combinations of products the affiliate industry most likely will promote baby products to you, because there is a change of up to 80% that you are pregnant. This last algorithm works with less than 100 target groups like we earlier discussed in another article.
In addition the usage of such algorithm increased the click rate from below 1% up to ~ 4%.

Nice history lecture, so what's up with it?
Alright, let's not focus the current mainstream. We started a project where we develop a new algorithm within the group of predictable behavioral

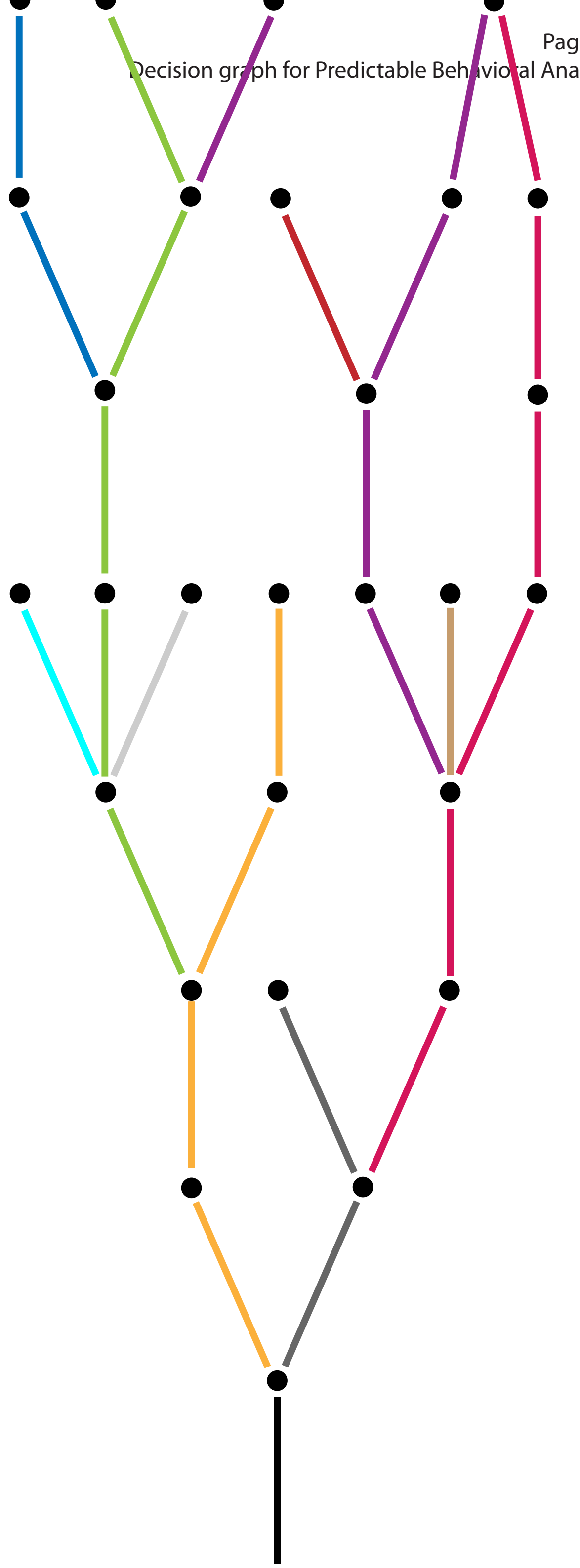


analysis. But we won't focus the topic matching of interests like in collaborative filtering but rather predict the upcoming decisions of a user. The meta level.

I don't get it. What is the difference now?
It is just a small difference in the concept but it results in a major difference of the algorithm and outcome. Collaborative filtering techniques tries to find similar interests because the factual is close that the user will like similar topics to the one it already likes. This only takes into account related topics like hardware and software but not hardware and soft drinks. We want to take all possibilities into account by predicting the upcoming decisions of a user. Collaborative Filtering (CF) bypasses this by selection similar topics. And that is basically it.

What is the origin of this idea?
The basic idea originate from the field of process mining. In process mining one uses mining algorithms in order to reproduce process models from logs. These logs were extracted from running business processes. During this stage of discovery one has also to deal with the topic of wrong or incorrect data called "noise". Therefor researchers using techniques like genetic and heuristic mining algorithms. But that's enough right here. More about the topic of process mining could be found right here. We developed a similar algorithm by adapting the target group and the different data setup and what we are doing now is mining the decision process of every user. These decision process are illustrated as workflow net or as decision node. And by comparing abstracts of these decision graphs we are able to find congruent behavioral patterns.

Results
With the first prototype we are able to generate a click rate of 5-6% but the algorithm's performance is more than bad. In order to accomplish 10 requests / second we need 32 GB RAM & 20 Ghz for our in memory database and calculations. This goes back to the point that we generate for every user request its entire decision way but only need the next possible decision gate. Further our systems currently learns from every request, because almost all decision ways are unique. Therefor the learning part was more than exponential. With the current state we have more than 100.000 unique decisions ways for as target group. In addition to increase the performance we have to optimize some of these factors. We will continue on optimizing this algorithm in order to access the mark of 7% . During the last month we were already able to increase our result set by 1% up to the benchmark of 6.5%. This goes back to the usage of linked list instead of vectors.



Building a fraud data base for the affiliate industry

Framsteg is building an anti-fraud pattern for the affiliate industry. Check out our project online!

Recently i talked to a friend of mine who is about to initialize the kickoff of his new affiliate platform with the focus on the gaming industry. We talked about the potential of this market which we were able to notice during the releases of Battlefield 4, GTA 5 and Call of Duty Ghost. The sale of these games brought together more than one and a half billion Euro already on the first day. But we also talked about the risk of affiliate marketing or rather a well known problem in this context. No, I don‘t mean the step of recruiting publishers or winning advertiser, rather we discussed the matter of click fraud and potential lawsuit outcomes.

Did you ever imagine what could happen if an advertiser cancels a contract due to click fraud or „bad traffic“? They don‘t just get their investments back, no, they also get damage compensation for non-pecuniary damage and compensation for consequential loss (claims under §§ 249, § 253 BGB German Federal Law). Lets say, someone pays you $1000 for advertising purposes and this campaign ends in smoke, the affiliate platform could have to pay twice as much back if it doesn‘t take care of click fraud. Of course this does not mean in every case of click fraud the advertiser gets his money back but this implies that the third party in this triangular relationship - the affiliate platform - has to consider the topic of inhuman traffic in their product.

Just scan this QR-Code to get to the online page of the project

Heuristic algorithm for process mining

Process mining is a technique in the field of process management which allows user to analyse business processes based on behavior and event logs. Basically the original idea is to extract knowledge from IT systems in order to visualize them on the meta level. In most cases these knowledge is represented through logs. With the usage of discovering and conformance metrics we are able to gain and verify our results.

Why is this so important?

Imagine the following situation. Your company has a high response time or the production time of a product or service takes very much unexpected time. The reason is not necessary that our employees are not good or skills, rather a typically reason is that there are weak points in the value chain within the business process. Like the situation that a department can‘t handle to amount of work due to missing or inexperienced personnel.

If you model the value chain of the german administration system in citizen center one would get an sad result to due the budgets cuts within the last years. Therefor it takes more 3 weeks to get an appointment for requesting a passport and another 3 month to receive it. Another example would be the production of a car within factory road. We all know that the most weak point will determine the speed of production. The process mining part will find the causes for that.

How to mine processes?

There are several ways and attempts in order to extract knowledge from event logs. One of the most common is the Alpha Algorithm with its improvements Alpha(+/++):

"The *Alpha(+/++) Algorithm* aims at reconstructing causality from a set of sequences of events. It was first put forward by van der Aalst, Weijters and Măruşter. Several extensions or modifications of it have since been presented, which will be listed below. Within the concept of the algorithm one takes a workflow log as input and results in a workflow net being constructed. It does so by examining causal relationships observed between tasks. For example, one specific task might always precede another specific task in every execution trace, which would be useful information."

Another attempt is the so called social miner. It aims at a different target group but is also very interesting:

"The *Social mining* or rather the analysis of social interconnected relationships is one of the most interesting topics of today‘s world. When deriving roles and other organizational entities from the event log the focus is on the relation between people or groups of people and the process. Another perspective is not to focus on the relation between the process and individuals but on relations among individuals (or groups of individuals)."

Facebook uses it, even the NSA is using social miner in order to understand the connections between people. But they have a weak point.

Whats the matter about it?

As we all know logs are not perfect. They include a lot of so called "noise". Noise could be wrong or incorrect log entries which distort the correct results with wrong data. The real issue is to detect these noise entries. There are several attempts in order to handle noise:

"The main motivation of the *genetic algorithms* (Eiben and Smith 2003) is to benefit from the global search that is performed by this kind of algorithms. Genetic algorithms are adaptive search methods that try to mimic the process of evolution. These algorithms start with an initial population of individuals. Every individual is assigned a fitness measure to indicate its quality. In our case, an individual is a possible process model and the fitness is a function that evaluates how well the individual is able to reproduce the behavior in the log. Populations evolve by selecting the fittest individuals and generating new individuals using genetic operators such as crossover (combining parts of two or more individuals) and mutation (random modification of an individual).“

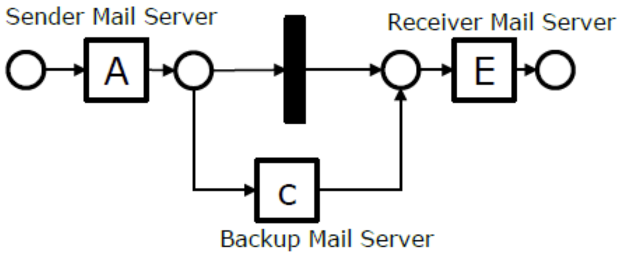
“The *Heuristic Miner* extends the alpha algorithm by consider the frequency of traces in the log. Heuristics miner can deal with noise, and can be used to express the main behavior. The Heuristics Miner Plugin mines the control flow perspective of a process model. To do so, it only considers the order of the events within a case. In other words, the order of events among cases isn‘t important. For instance for the log in the log file only the fields case id, time stamp and activity are considered during the mining. The timestamp of an activity is used to calculate these orderings.”

So what is the problem?

The biggest advantage of the heuristic algorithm is also its main problem. The threshold. By increasing the threshold we are able to remove instances with a low frequence. But we have to watch out because the threshold applies to the entire net and not single edges within it. Therefor, there is always the possibility to remove process relevant information by increasing the threshold and we try to handle this failure. Let me make that clear with a small example. We have the log containing the following entries:

$$L = [\ (a,e)^{9999}, (a,c,e)^{1}, (a,f,g,c,e)^{3}]$$

Normally with the attempt of the heuristic miner, we would increase the threshold up to 3 in order to get rid of our assumed noise. For final safety reasons we always have to interview a domain expert. That not our goal, therefore we have to think about something else.



The workflow above is the original process represented in our log. So if we would have increased our threshold up to the value of 3, our main failover plan doesn‘t work anymore. And this issue addresses all relating processes with backup and failover technology because a backup or failover should only appear in 1 of 1000000000 cases in our event log. Because its a FAIL over and not the average case. Just imagine we increase the threshold and kick out failover instances within the process. That would be the state of emergency.

So, how can we handle this disadvantage?

Basically we thought of a two step improvement of the heuristic miner.

The *Preprocessing Stage* indicates the major part in order to find the sibling model of our research instance. Here we are comparing the log against all logs in our archive with algorithms close to the predictable behavioral analysis group. But instead of comparing behaviors we take a look at the activities in order to find relative ones. If we have a match, we will mark the congruent model.

During the *Postprocessing Stage* we are able to compare the results of our heuristic miner with the results of the preprocessing stage. If we increase the threshold we can compare in time against the congruent model if process relevant activities get kicked out. Therefore we are able to increase the threshold without loosing relevant activities.

We developed this technique as SaaS. We you want to know more about it, just step over to our project site: process-mining.framsteg.de

Thanks to our partners who helped us to create a comparable archive with sample processes for the comparison part.

Paranoia for ideas

"This idea is priceless, if the chinese guys would find out about this, they would copy it this second!" I have heard this exact sentence several times during my job and it is absolutely not true.

Interestingly it always came from people who have absolutely no experience in the StartUp business. Therefore I would like to bust the myth of the evil chinese copy cats. Successful copy cats never steal ideas, they steal concepts. This means not only the idea behind the product but also the production routines, logistics, marketing, sales and most importantly: A proven business model. Successful copy cats don't steal your idea unless you have already successfully launched your product on the market. This is due to the fact that copy cats are good at two things: Optimization and marketing. These two have in common that they work best if powered by a large production and an even larger amount of money. And because copy cats are no VCs and therefore not sitting on large amounts of risk capital, they have a natural risk aversion. They won't lift a finger unless your business is up and running. If you are worried about your product idea getting stolen, try to concentrate on building a unique brand. Nobody will be able to steal your brand. Brands are what separates mediocrity from success. I would like to add that China's investment in corporate innovation will be 200% above the OECD-average until 2020. Therefore it will be faster and more efficient for a chinese corporation to develop ideas by themselves than to search and recreate existing ideas.

An exception to this rule is corporate technology. Replicating a secret and unproven patent by Audi seems like a no-brainer to print money. This is because copy cats are handling complexity like every other human: By relying on brands. Because Audi has proven it's abilities to create saleable products, it is very likely that this new technology will also be saleable. But unless you have an international brand with a reputation for solid products, this will not affect you. That being said, I would like to return to our main statement "This idea is priceless." If your going to conduct logic, it is deductable that your idea is therefore worthless. And because things of no worth can be shared without losing anything, sharing will be your best bet. Try to tell your idea to as many people as possible and get feedback. Ask the magic question: "Would you pay for this?" Most ideas fail due to the fact that no one will actually spend money on the product.

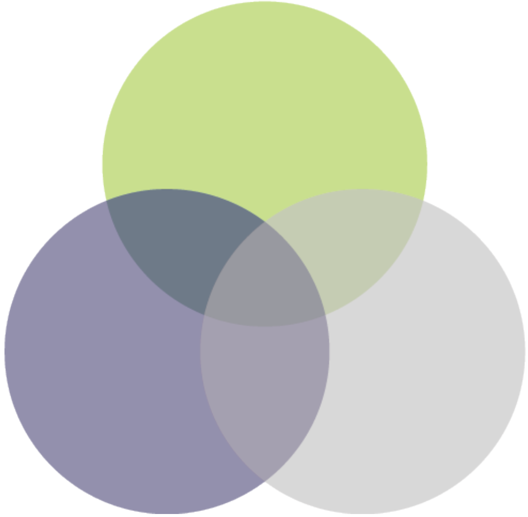
I would like to close with a common recipe: "10 Inspiration, 90% Transpiration". This is what makes you a successful business owner. And a copy cat will always be in the exact same position. The lion's share is to get from the idea to actual turnover. But copy cats lack an important and unique factor of your idea: your belief and sacrifice for the idea. This is why you will be always ahead of copy cats.

Talking about problems

Everyone got problems. Most people react to their problems on a range from slightly unpleasant to a full blown life crisis. That is mainly due to the fact that we are mostly intimidated and overstrained by them. But with a little help, every problem may be fixed. To help you with this, the terminology and the kinds of problems are discussed.

When considering problems, there are generally speaking three attributes a problem can have: complicated, complex and difficult. These attributes are not mutually exclusive, therefore a problem can have every combination of these attributes.
Let's start with complicated problems. Complicated means the context of the problem consists of a high number of individual pieces that are stringed together in a linear way. The best example is a clockwork. Many small gear-wheels work together. You can predict that the first wheel will spin the second and so on, but the sheer amount of wheels makes it hard to find the reason the clock stopped working in the first place. An overwhelming number of components results in complicacy. Whenever you deal with people, the linearity of interaction comes crashing down. It is not possible to predict a human interaction since people are not acting consistent. Therefore whenever we are confronted with a problem concerning the interaction between people, like a fight, we call it complex. Complexity results from non-linear interaction between components.
Finally, there are difficult problems. Difficult means that we are clueless about how to solve the problem. We have no starting point for a solution. The best example is the nine-dots puzzle. You have to connect all 9 points in a 3x3 Matrix by using just 4 lines. It is impossible to do so while staying inside the Matrix, you have to literally think outside the box. By the way, that's how the phrase got coined. Difficulties could arise if the wrong premises are assumed (staying inside the box). On the other hand the lack of premises could lead to difficulties as well: Ask someone without physical knowledge how magnets work and they will find it difficult to do so. To conclude, difficulties result from an absence of premises or the presence of wrong ones.

Let's move on to the scaling of these problem attributes: The question at hand is under which circumstances a problem can be labeled complicated, complex or difficult and if they have gradients.


Complex Problems
Circular interaction results in a self-evolving system which is different to it's parts.
Difficult Problems
Not having a starting point to solve a problem makes it hard.
Complicated Problems
The overwheming number of elements results in a hardly manageable problem.

Complexity is an absolute attribute. A problem can be labeled complex if there are two components interacting in a non-linear way. The degree of complexity depends on the number of components in the system, their structure and their intensity of interaction.

Complicacy in contrast is a partially relative problem attribute: It depends on the training and capacity of the individual working memory. But since humans are indeed humans, you can be fairly certain that there is a bunch of problems which are complicated for everyone. This means an absolute border of complicacy can be postulated, but this border may be lower, depending on the individual. The degree of complicacy depends on the number of components at hand, their kind of interaction and the individual ability and training to process the given problematic situation.

Difficulty is the attribute with the most individual variation. Since it depends on the individual previous experience and the resulting personal premises concerning the given situation (short version: subjectivism), difficulties can arise from a variety of factors. The easy part is to label a current problem: It can be either difficult or not difficult, depending on whether we have an idea on how to solve the problem. Retrospectively the difficulty could be graded by the amount of time used to solve the problem, but this indicator is also highly individual and of considerable methodological flaws.

To understand the usefulness of this differentiation, we have to go on and talk about methods to overcome these different problems. This is where this concept becomes practically applicable: An understanding of different kinds of problems as well as their most fitting solution methods allows for faster solving. Understand it as a toolkit to use in your everyday life.

To address complicated problems, you should conduct an analysis. Analysing things means breaking the problematic condition into it's components and describe their linear interactions one by one. An analysis always asks the question "What is happening?" Go from there, find the failing part and fix it.

Difficult problems know two solution paths: The first one is the transfer. A difficult problem can be addressed by transferring a known solution method from a different but similar context to the current problem. The solution performance of a transfer depends on the similarity of the situation it was taken from and the situation it is applied to. A more thorough solution method for difficult problems is the diagnosis. A diagnosis is an hypothesis about the cause of a problem. This hypothesis allows to fight the problem at its roots. A diagnosis therefore asks the question "Why is it happening?" Try to understand what premise causes your troubling symptoms and adjust your premises.

Solving complex problems is the hardest part: Professor Peter Kruse, a german professor of psychology and a famous systems thinker, said there were four ways to deal with complexity: Trial and error, ignoring complexity, simplifying/over-rationalizing complexity and intuition. The last one is what you should go to as the other ones are definitely not fitted to address a complex problem. Try to head on complex problems with your gut, it will make better judgments than your brain.

To wrap it up, the understanding of your specific problem will give you guidance on how to solve it. There goes an old saying: For someone with a hammer, every problem is a nail. But for someone with a toolbox, a problem can be solved with a screwdriver, a wrench or a saw.

This article is part of our upcoming project Words. Stay tuned for more content! Co-authored by Max Kinninger

FRAMSTEG

We are a Think Tank.

Our goal is to develop new, interesting concepts for tomorrows intelligent world. As a Think-Tank, we are a non-profit research institute. We want to contribute to the concepts of the future and therefore like to share our knowledge. We are not interested in financal or personal gain. Our results are published in our project archive and in our developer blog.