Journal of Universal Computer Science, vol. 26, no. 7 (2020), 783-804
submitted: 20/1/2020, accepted: 4/5/2020, appeared: 28/7/2020

(De-)Constructing Attacker Categorisations: A Typology Iteration for the Case of Digital Banking

Caroline Moeckel
(Royal Holloway, University of London, Egham, United Kingdom
caroline.moeckel.2012@live.rhul.ac.uk)

Abstract: In this extended and updated paper, the experimental construction of a new attacker typology grounded in real-life data is proposed, using grounded theory analysis and over 300 publicly available documents containing details of digital banking related cybercrime and involved attackers. Seven attacker profiles forming the typology specific to the case of digital banking are presented. An initial light-touch evaluation approach based on peer review feedback and basic heuristics is suggested. A short excursus on circumplex models is added to address this visualisation tool used across past categorisation efforts.

Keywords: attackers, threat agents, categorisation, typology, taxonomy, threat modelling, grounded theory, circumplex models, digital banking
Categories: K.4.3, K.6.5, L.4

1 Introduction

Attacker analysis and profiling have long been part of the analytical toolkit of investigators and date back centuries [Nykodym, 05], both for planning defence strategies and to aid forensics post-attack. Researchers have been interested in finding out more about the individuals behind cybercrime since the first illegal activities were observed in the early beginnings of the cyber era, initially in the area of telecommunications. In this context, attacker typologies and taxonomies are commonly used vehicles to represent attacker types and categories applicable to either a specific system or for generic usage.

Early research in the area [Gordon, 96][Hollinger, 88][Landreth, 89][Pfleeger, 06], mostly based on relatively small numbers of interviews, documented case studies and anecdotes, indicated variations amongst attackers, for example their technical skills, motives or level of damage done to the system targeted. Such observations ultimately lead to the creation of attacker categories, e.g. like the three types of computer criminals (crackers, criminals, and vandals) identified by the FBI in 1997 [Ivoce, 97]. More recent works include the widely referenced work by [Rogers, 06] from 2006, with nine attacker types and a two-dimensional matrix visualisation aligning attacker motivations and resources. Based on a literature analysis of previous works on attacker taxonomies, individual hacker categories and subcategories, [Meyers, 09] in 2009 then consolidated research efforts to date into eight common categories of attackers. In 2012, [Hald, 12] carefully updated known attacker categories, using current terminology and threat properties. More recently, in 2015, [Seebruck, 15] proposed an updated attacker typology. While closely built on the mentioned earlier works, it has been adapted with the intent to capture “recent increases in ideologically and socially motivated hacking”.

A comprehensive and critical assessment of the state of attacker typologies and taxonomies can also be found in the 2017 work by [De Bruijne, 17].

And while they certainly provide interesting and accessible visualisations of human threat actor landscapes, attacker typologies and taxonomies suffer from a range of limitations and shortcomings at this point in time, with [De Bruijne, 17] concluding on "a disheartening picture of state-of-the-art thinking on threat actor typologies" after their initial literature review. For them, problems are mostly methodological, with used data sources, classification and construction methods including evaluation and validation efforts not adequately accounted for. In our opinion, many taxonomies seem to be built on each other, reference previous literature rather than using independent real-life datasets (e.g. [Hald, 12][Meyers, 09]), with one of the key references in the area [Rogers, 99] not meeting certain standards (clear publication date and route, named data sources, methodology). As a further methodological consideration, the introduction [Rogers, 06] and continuous use [Hald, 12][Seebruck, 15] of circumplex models as a highly theorised concept from clinical psychology and sociology shows limited theoretical grounding.

As a direct extension of [Moeckel, 19] as presented at the First International Workshop on Information Security Methodology and Replication Studies (IWSMR) at the 14th International Conference on Availability, Reliability and Security (ARES '19), this paper proposes the experimental construction and initial evaluation of a new attacker typology grounded in data, using grounded theory analysis of over 300 publicly available documents containing details of digital banking related cybercrime and involved attackers. New elements to this paper over the short paper version presented at ARES'19 include an extension of the background section as well as further detailing of the methodology. Additionally, a brief validation exercise incorporating both peer review feedback and heuristic evaluation elements is now also included. A new compact excursus and critique of circumplex models as used in previous categorisations is also provided before moving onto a concluding reflection.

2 Background

[Moeckel, 19] has prepared a number of theoretical aspects surrounding attacker categorisations that can be referred to in this work. Firstly, the definition of common categorisation terminology has been undertaken, including the distinction between the two terms taxonomy and typology as categorisation types (typology is preferred in this work as the categorisation is likely to be non-exhaustive, presenting ideal summaries of attacker groups rather than truly empirical, in-depth and formally measurable attacker characteristics from a complete, finite dataset; which would constitute a taxonomy). Secondly, previously used categorisation strategies and criteria found in previous literature have been discussed in [Moeckel, 19]. Lastly, "a heavily consolidated, non-exhaustive view of common attacker types as found in literature" has been provided in [Moeckel, 19], including labels such as novices, browsers & cyber punks, ethical hackers, insiders, hacktivists, crackers & coders, professional criminals, government agents and other attacker types, e.g. the 'crowdsourcer' newly proposed in [Seebruck, 15], which describes large scale human collaboration to obtain confidential information, potentially using illegal means.

For this extended paper, two gaps in this literature review are addressed: firstly, by adding further information on the perceived value and usefulness of attacker categorisations, and secondly by considering the case study context of digital banking.

2.1 Purpose and value of attacker categorisations in previous literature

When it comes to the underlying reasoning behind the creation and maintenance of attacker categorisations, key works in the area of attacker categorisations are largely in agreement over their purpose and value.

At a strictly formal level, [DeBruijne, 17] view the format of a typology as "appealing because it promises to yield a concise yet parsimonious framework to describe and classify observed patterns". Simply put, attacker categorisations such as typologies and taxonomies are seen to help identify, structure and classify information gathered on attackers [Seebruck, 15]. Hence, at the most basic level, authors generally agree on categorisations supporting a better understanding of adversaries and helping with the aim of "knowing your enemy" [Rogers, 99/06][Hald, 12][Long, 12][Seebruck, 15]. [Rogers, 99] adds the appreciation of the heterogeneous nature of attackers as a further benefit gained from attacker categorisations, with [Gordon, 96] supporting this in her study on virus writers as a unique attacker group.

But how can this theoretical understanding obtained through such categorisation efforts then be translated into tangible benefits applicable to security practice? Here, both [Hald, 12] and [Seebruck, 15] mention the definition of common, up-to-date terminology in this area as crucial for shareability and collaboration initiatives. [De Bruijne, 17] define the goals for their typology as an update to previous typologies forming part of a large-scale security assessment exercise (Cyber Security Assessment Netherlands in their case), used and contributed to by security analysts in both public and private organisations. Similarly, [Shostack, 14] sees attacker categorisations as a useful resource and tool for security professionals, including example attacker lists and personas in his key textbook on threat modelling to supplement other structured, system- or asset-centric approaches.

In contrast, Fötinger and Ziegler ([Ziegler, 04] in collaboration with RSA Security) see their enquiry into the psychology of attackers mostly as an attempt to close a gap in literature. Overall, much work in this area seems to be theoretically and methodologically driven, with taxonomies and typologies directly extending and building on each other (e.g. [Hald, 12][Seebruck, 15]). With the distinct exception of [De Bruijne, 17], attacker categorisation works [compare to reference list in [Moeckel, 19] seem to be high-level, generic representation aiming to theoretically highlight the variation of attackers, their motives and potential modus operandi rather than very specific, ready-made models produced to be used in practice by security analysts (although they may inform and support these practical perspectives).

In alignment with previous works in the research field, the here presented work can therefore primarily be understood as a replication effort partially addressing methodological issues, with the aim of adding another reference to the research field as a basis for further research. As a secondary aim, it is hoped that this data-driven, sector specific case study may be of interest to practitioners and academics to compare to their experiences and critically respond.

2.2 Digital banking as a sector-specific case example

Based on definitions mostly in the commercial space (e.g. [Ginovsky, 15] or [Epstein, 15]), digital banking can be understood as the integration of digital technologies into the overall banking business model and organisation along the entire value chain and in all areas of financial services provision, ranging from e.g. personal or business banking products, transactional services, financing or investment offerings, but also in areas such as marketing and customer services. Real-world applications may include mobile banking apps, interactive chat bots for customer support or online trading facilities.

Digital banking case studies have been present in academic works on threat modelling [Xin, 14] [Moeckel, 10] in the past, but also in works focussing on socio-technical aspects, e.g. cross-cultural comparison efforts examining customer adoption of mobile banking in [Merhi, 19], or specific technical security problems, e.g. the detailed examination of a previously undetected vulnerability affecting the EMV protocol for card payments ('chip and PIN') in [Bond, 14]. Beyond offering a wide range of interesting research angles at the intersection of security, usability and business requirements, also involving a number of stakeholders such as users, banking and security professionals as well as attackers, examples of digital banking cybercrime case information seems widely available (e.g. as part of the data sources defined in Section 3.1). However, dedicated attacker categorisations limited to digital banking were not found by the researcher at the time of finishing this research project.

3 Methodology

3.1 Data sources

To help build the categorisation, publicly available materials containing details on digital banking attackers were analysed as described in Sections 3.2 to 3.5. To obtain these resources, four different reference data sources were consulted and reviewed to extract items suitable for analysis (based on criteria such as e.g. strict relevance to digital banking only, level of detail as well as avoiding duplication across datasets):

- British Computer Society Cybercrime Forensics Specialist Group weekly briefings (2010–2014; worldwide): 487 lists containing 7,305 web articles in total, with 127 ultimately selected for analysis [BCS, 14];
- Cambridge Computer Crime Database (2010–current; UK): 689 incidents described accompanied by linked evidence (also web articles), with 90 ultimately selected for analysis [Hutchings, 19];
- FBI Cyber Most Wanted list (current; worldwide, subject to US prosecution): 43 attacker profiles, with 32 ultimately selected for analysis [FBI, 19]; and
- Vocabulary for Event Recording and Incident Sharing (VERIS) Community Database (2012–current; worldwide, US focus): 7,833 incidents accompanied by linked evidence (also web articles), with 688 related to the finance and insurance industry and 78 ultimately selected for analysis [VERIS, 19].

For transparency reasons, a full list of the ultimately selected and subsequently analysed materials (over 300 web articles) is available under [Moeckel, 20]. It is also worth noting that this data analysis forms part of a larger research project on attacker-centric security in the context of digital banking carried out by the researcher (see also [Moeckel, 20]).

3.2 Re-coding and preparing data for categorisation

In order to build a data-driven typology, which can be understood as a classification and representation of groups with shared characteristics and behaviours, a reference dataset describing the items to be categorised is required. Using the data sources as outlined in Section 3.1, a grounded theory-based qualitative data analysis of attacker characteristics specific to digital banking had been carried out at this point and was available to the researcher [Moeckel, 20] — only, this information had not yet been grouped into clusters of common traits. So how was this heterogeneous community transitioned into relatively homogeneous categories, ultimately forming a typology?

Initially, all analytical codes describing attacker characteristics and behaviour that had emerged through the grounded theory analysis were reviewed for their relevance to categorisations. To help with this, previously employed categorisation criteria for taxonomies and typologies from literature were used (refer also to [Moeckel, 19]), yielding 12 codes of interest (see Table 1 below).

Attack: monetary damage inflicted	Attack: geographic reach and scope
Means: modus operandi (general)	Means: insider knowledge
Attacker: profit (financial motivation)	Attacker: motivations (other than profit)
Attacker: resources (funding)	Attacker: resources (equipment)
Attacker: skills	Attacker: entry/paths into criminality
Attacker group, e.g. size or cohesion	Attacker group: business model

Table 1: Analytical codes holding attacker characteristics and behaviour

At this point, the original dataset was then coded again using the qualitative data analysis software package NVivo, employing another round of initial coding and subcoding as methods, focussing specifically on the 12 identified codes. Subcoding is a coding method that assigns a "second-order tag […] after a primary code to detail or enrich the entry" [Saldana, 12] p.77. This was to ensure that all data present in the original sample and potentially relevant to the planned categorisation was captured in a structured way. In principle, this step aimed at adding an additional layer of detail and depth to certain codes to support the categorisation process.

3.3 Re-checking and defining categorisation criteria

Rather than readily accepting categorisation from previous typologies and taxonomies, this study uses a combination of a literature review of categorisation criteria together with a review of the 12 codes previously identified (Table 1) to come up with a list of categorisation criteria (refer to Table 2 overleaf). A similar step is included in the typology building process used by [De Bruijne, 17] (p.16/17), who use a deductive 'first cycle analysis' to specify the "dimensions that are used in (cyber) threat actor

typologies" for their research through a comprehensive literature review (complemented by empirical interviews with security experts). At least some of these criteria are also likely to be highly specific to the individual target or victim setting: level of danger or risk posed would vary for the chosen case example of digital banking in comparison to e.g. an individual home PC user or a critical infrastructure provider (such as a power supplier) on the other end of the spectrum.

Original codes	**Proposed criteria**
Profit, other motivations (not profit)	Motives
Profit, business model, monetary damage	Criminal intent
Resources (funding and equipment), skills	Resources
Means/modus operandi, business model, insider knowledge, functions in group, group character & size	Activities
Monetary damage, geographic reach, means	Level of danger posed
Monetary damage, means	Type of risk posed
e.g. Entry/paths into criminality	Other notes

Table 2: Creating a categorisation framework: transforming codes into criteria

3.4 Transforming the data and practical categorisation process

For the categorisation process, affinity diagrams or maps as a design thinking technique[1] were used. A comprehensive definition of affinity diagrams as a synthesis method is provided in [Friis, 2020]: "affinity diagramming is a process used to externalise and meaningfully cluster observations and insights from research" through considering data-driven insights individually and by putting them on (virtual or physical) post-it notes to then be clustered around their 'affinity' (such as similar ideas, concepts or data facets). In an academic context (specifically grounded theory generation in design research), [Maher, 2018] describe traditional material approaches as a valuable support tool for data analysis in combination with software such as NVivo for data management facilities and larger coding exercises. Additionally, an affinity mapping exercise also lends itself for replication in a corporate environment: such methods may already be known to security professionals and their teams or could be introduced and adopted easily (as described for example in [Harboe, 15]).

For the practical research process surrounding the typology build, all codes and conceptual phrases were manually transferred onto post-it notes from NVivo. Following best practice guidelines for affinity diagrams, all post-it notes were reviewed for their similarity and potential connections to others and then either placed in a new group or with an existing cluster (each representing a potential attacker type). This manual process of re-grouping, merging, removing or separating as well as naming clusters and notes contained was repeated continuously by the researcher[2]. The

[1] Affinity diagrams form part of design thinking [Brown, 08] toolkits, e.g. by Stanford d.school. [Kimbell, 11] provides a comprehensive critique of the method's origin and contemporary usage. However, affinity diagramming dates further back in general project management and business practice and is also known as the K-J method after the anthropologist Kawakita Jiro [Plain, 07].

[2] Although generally described as a team-based exercise in literature [Brown, 08], affinity diagramming carried out by individuals or small groups is not unknown, as reported in [Harboe, 15] in their review of real-life practices surrounding this technique. For the purpose of this study, affinity diagramming can be seen as

categorisation process was deemed finished when all variations were accounted for and no new clues for further re-naming or -positioning presented themself — at this point, theoretical saturation (as defined in [Charmaz, 14], p.345 or [Urquhart, 13], p.194) seemed to emerge for this study and underlying dataset.

Figure 1: Example of manual clustering process in affinity diagram exercise

3.5 From categorisation to typology: presentation of results

Using the criteria list from Table 2, a list of description criteria was assembled to help build the attacker profiles in detail (refer to Table 3). With this criteria list at hand, every affinity cluster was then reviewed and where adequate details were present in the original data, details were added for every criterium, resulting in the typology as shown in Section 4. The process required continuous comparison and back-and-forth review between the manual cluster visualisation and the original NVivo file with its coded data fragments. Through this process, seven distinct groups of attackers were identified and grouped under the following descriptive terms: system challengers, insiders, supporters, ideologists, officials, professionals I: groups and gangs, professionals II: small groups and individuals.

As mandated by the nature of the data sources (incomplete and often random), not all description criteria could be filled with detail from the data. Further to that, several aspects were only indicated in a single data fragment or lacked overall detail or

a highly visual extension of the coding process in NVivo rather than a team exercise, although the same categorisation process could be carried out using a team scenario in principle.

preciseness. Where this was the case, results are marked as tentative with one asterisk* in the table view for every attacker type within the typology (see Section 4, Tables 4.1 to 4.7). For the few occurrences where an assumption without grounding in data was made, brackets with two asterisks** are used. Please also note that only a small selection of individual source materials from the original dataset analysed have been referenced in the tables and hence added to the References Section of this paper — a full list of these article links can be found under [Moeckel, 2020].

Criteria	**Description**
Group	Used as a name for the attackers in this group
Subgroups	Subgroup descriptions (optional)
Labels	Describes other descriptions found in the sample or literature that may be used for attackers in this group
Motives	Describes the primary driver behind the criminal activity engaged in for this group
Criminal intent	Describes the level of preparedness and intent for criminal and illegal actions present in this group
Resources	Describes the resources such as funds or equipment and the skills level present in this group
Activities	Describes the main criminal activities engaged in and modus operandi used by this group
Level of danger posed	Describes the overall impact and level of destruction this group may pose to its victims
Type of risk posed	Describes the type of risk posed to its victims

Table 3: Attacker categorisation framework: description criteria

4 Results

Group	**System Challengers**
Subgroups	System testers, hackers looking for fun or challenge
Labels	White hat or ethical hackers*, thrill seekers or glory hunters, young or novice hackers
Motives	Fun of hacking, bragging rights, challenge to break into system, exposing vulnerabilities (responsible disclosure**)
Criminal intent	Low to moderate
Resources	Range of skills and funds, can be limited
Activities	System intrusion, penetration testing, publication of vulnerabilities
Level of danger posed	Relatively low, but varies across the group and can be seen as an entry into serious criminality for some
Type of risk posed	Often reputational risk, may however also be of financial or operational nature
Other notes or comments: very heterogeneous group united by desire to overcome challenge posed by overcoming the system's defence — in our sample, the number of white hat/ethical hackers seems low with only limited evidence, e.g. in [Karia, 12]. Responsible disclosure cases were not present in the sample, but this option, where (non-malicious) attackers would notify banks about identified vulnerabilities to provide them with the opportunity to fix them before going public, is made available by a number of banks, e.g. The Royal Bank of Scotland, UK [RBS, 20].	

Table 4.1: Attacker profile for system challengers group

Group	**Supporters**
Subgroups	Money mules, non-technical support functions
Labels	Non-technical support functions: mules, cash collectors, business functions such as recruitment, marketing or customer service
Motives	Financial gain, 'making ends meet'
Criminal intent	Moderate to high (in some cases unwittingly)
Resources	Limited technical skill levels and funding
Activities	Supporting a larger group or system through all stages of money laundering and other business support functions
Level of danger posed	Low on their own, but part of a group or system
Type of risk posed	Usually financial risk, although operational and reputational risk may be indirectly posed
Other notes or comments: Supporters are not technically attackers themselves, but support others to commit their crimes. These functions are relatively well evidenced in the sample, with over 100 related references present.	

Table 4.2: Attacker profile for supporters group

Group	**Insiders**
Subgroups	Banking employees, third party supplier employees
Labels	-/-
Motives	Financial gain, retaliation
Criminal intent	Moderate to high
Resources	Range of skills and funds, enabled through insider knowledge and capabilities including elevated access rights
Activities	Supporting a larger group or system through all stages of money laundering and other business support functions
Level of danger posed	High, significant levels of damage possible
Type of risk posed	Often financial, but also significant potential for operational (IT sabotage*) and potentially reputational risk

Table 4.3: Attacker profile for insiders group

Group	**Ideologists**
Subgroups	-/-
Labels	Hacktivists, online activists or cyber terrorists
Motives	Cause, ideology, in rare cases also status and ego (secondary motives such as financial gains may be present*)
Criminal intent	Moderate to high
Resources	Moderate to high skill levels and funding rights
Activities	Social or political background to attacks
Level of danger posed	High, significant levels of damage and destruction intended
Type of risk posed	Reputational risk and linked operational risk, financial risk as a secondary motive*
Other notes or comments: Ideologists are usually motivated by cause and ideology, but examples of attackers being motivated by selfish reasons such as financial gain or simply to engage in petty vandalism can be found, e.g. in [Ward, 12]	

Table 4.4: Attacker profile for ideologists group

Group	**Officials**
Subgroups	-/-
Labels	Nation states, sovereign countries, government or its agencies, military functions*
Motives	Cause, ideology, cyber warfare*
Criminal intent	High**
Resources	Very high skill levels and funding**
Activities	Espionage, counterespionage, information monitoring and destructive attacks, cyber warfare*
Level of danger posed	High, although limited evidence and confirmed cases to date*
Type of risk posed	Operational risk as a main focus with reputational and financial risk directly linked**
Other notes or comments: Not much is known about this group and references in the data sample are sparse, e.g. in [Lee, 12] where nation state involvement in attacks affecting digital banking services is indicated, but no further detail on for example skill levels or activities are included (most likely as they are unknown). This does not necessarily mean that such attackers are not relevant to financial institutions, but more likely that they haven't found entry into the analysed sample — potentially due to these attackers being able to stay under the radar.	

Table 4.5: Attacker profile for officials group

Group	**Professionals I: groups and gangs**
Subgroups	-/-
Labels	Sophisticated large criminal groups or gangs and organised online crime syndicates with members often professionally recruited (e.g. in [Prince, 11]) or potentially acting under instructions of others as paid, service-based attackers*.
Motives	Financial gain
Criminal intent	High
Resources	High skill levels and funding: broad range of skills and resources available through group setup
Activities	Phishing, ransomware, trojans and malware attacks as well as system intrusion at large scale, physical attacks e.g. against cash machines/ATMs also possible. May also offer their services through criminal-to-criminal franchise models.
Level of danger posed	High, significant level of damage
Type of risk posed	Financial, operational and reputational risk directly linked
Other notes or comments: Primary/key category for digital banking attackers. These attackers should be viewed as highly professional criminals. Well supported in the sample, with over 200 references supporting activities and modus operandi aspects and a further 200 references on roles and functions in attacker groups.	

Table 4.6: Attacker profile for professionals I: groups and gangs

Group	Professionals II: Small Groups and Individuals
Subgroups	-/-
Labels	Lone hackers and individual attackers, small criminal groups and gangs (can be relatives or friends rather than recruited, e.g. in [Krebs, 13]). Also potentially acting under instructions of others as paid, service-based attackers*.
Motives	Financial gain
Criminal intent	High
Resources	Moderate to high skill levels and funding
Activities	Phishing, ransomware, trojans and malware attacks as well as system intrusion, physical attacks. Similar to professionals I, but usually at smaller scale.
Level of danger posed	Medium to high
Type of risk posed	Financial, operational and reputational risk directly linked
Other notes or comments: Primary/key category for digital banking attackers — small to medium group size including lone attackers based on approx. 100 references. Again, these attackers should be viewed as professional criminals.	

Table 4.7: Attacker profile for professionals II: small groups and individuals

5 Validation and Feedback

Validation efforts for this work are made up of two components: firstly, feedback from academic peers was used to improve on the initial iteration of the typology (as published in [Moeckel, 19]). Secondly, this initial typology has been compared to a number of heuristic criteria formally defining a 'good' typology or taxonomy (based on [De Bruijne, 17]). From these activities and their results, a number of changes have been made to help build the current, here presented typology iteration.

5.1 Peer review feedback and resulting amendments

For the peer review of our typology, three sets of comments where used: two sets of reviewer comments from conference submissions and direct feedback from the UK PhD examination process. The venues submitted to were the First Workshop on Attackers and Cyber-Crime Operations (WACCO) at the IEEE European Symposium on Security and Privacy 2019 in Stockholm (weak reject) and IWSMR at ARES'19 in Canterbury (accepted & published as [Moeckel, 19]). The feedback obtained throughout the PhD examination process was provided by two senior academics working in the area of information security. As an excerpt of the extensive feedback provided, Table 5 lists the highlighted aspects as well as actions taken.

Feedback/comment item	Analysis/Action item
Link to data: connection to underlying dataset, also surfacing details from the analysed materials.	Exemplary references to the original dataset have been included directly in the attacker type overview (Table 4.1 to 4.7).
Content dimensions across typology: is sufficient data available for every dimension and every attacker type from the analysed dataset?	Where evidence in data is limited or assumptions have been made, this has now been explicitly marked and commented on (asterisk usage).
Assigned weightings across typology: where possible, assign a relative importance or impact to the attacker types	Where applicable, a note has been made within the tables to indicate the most relevant categories based on number of references within the sample.
Nature of categories within the typology: are the attacker types mutually exclusive and collectively exhaustive?	Suggests a formal examination of the typology structure — while this has not been included in related works such as [Rogers, 99/06][Hald, 12][Seebruck, 15], an approach is proposed in Section 5.
Circumplex models: critical analysis and reasoning behind their inclusion required	A dedicated excursus on circumplex models has been included in this iteration in Section 6.
Validation: how can the new typology be tested/confidence be instilled?	Section 5 has been included in this iteration of the typology, with further validation efforts envisioned to help build next iterations.

Table 5: Consolidated feedback and action items for initial iteration of typology

5.2 Heuristic review

In addition to analysing the direct feedback received for the first iteration of our typology, the structure of the typology is evaluated in a second step, based on a list of criteria for evaluating the quality of a taxonomy/typology as identified from literature in [De Bruijne, 17]. Build from general (not information security or attacker specific) literature on classifications such as Bailey in [De Bruijne, 17], this list is seen to help provide confidence in our typology, but also yield recommendations for improvements. The following lists the eight abbreviated evaluation criteria adapted from [De Bruijne, 17] p.14 and a brief assessment on how these are currently met (or not met without amendments) in our typology.

1. Exhaustive— all potential attackers should be classified. As the typology introduced directly builds on the analysis of a relatively large, varied dataset of real-world digital banking cybercrime cases including information on attackers, it can be viewed as representative for this population. New and unknown attack vectors and attacker types however may only be represented in future iterations.

2. Mutually exclusive— all potential attackers should fit into just one class. This criterion was not satisfactorily met in the original, initial iteration with eight attacker

types as presented in [Moeckel, 19]: the attacker type class 'toolkit users' shows continuous overlaps with other categories, i.e. toolkit users would always also be part of small or larger criminal groups or insiders — it was therefore decided to remove this group in the current iteration (see also Section 5.3).

3. Relevant— the classification method should enable for consistent replication based on available information and lead to meaningful classification. The abbreviated overview of relevant research procedures in Section 3 enable replication using a similar dataset containing cybercrime cases.

4. Pragmatic— the typology must contain a manageable number of classes/attacker types that can be clearly distinguished, requiring observable heterogeneity between them, but also meaning a relatively high level of abstraction overall. The number of attacker type classes in our typology is limited (7), all showing relative levels of heterogeneity between each other (see also criterion 2). Also remarked on in [De Bruijne, 17], typology quality criteria may be conflicting, owed to the balancing act between creating a compact and abstract, yet complete and detail-rich typology (as required under criterion 1).

5. Efficient— the classification method must enable efficient classification efforts. The research procedures in Section 3 enable a structured classification process — however, any grounded and data-driven typology will require immersion into the source dataset and further analysis efforts.

6. Transparent— the classification method should be based on a defined list of descriptive dimensions, accessible and clearly documented. The research procedures in Section 3 present such a methodology in its abbreviated form.

7. Dynamic— the classification method should enable continuous updates to the typology to accommodate new available data. As stated throughout this chapter, the typology building exercise presented in this work is viewed as an iterative process — to accommodate new data in the typology, the research procedures as outlined in Section 3 including analytical coding of new materials would need to be conducted.

8. Iterative— new attacker types can be added as a result of criterion 7. In direct relation to the last point and as evidenced in the response to criterion 2, attacker type classes can be added or removed in a new iteration of the typology.

5.3 From initial iteration to current typology and beyond

A number of changes to the initial iteration of our typology (published and available as [Moeckel, 19]) have been made following the above described evaluation steps to arrive at the current state as presented in this thesis.

The most significant change was certainly the reduction from eight to seven attacker types, removing the 'toolkit user', owed to the reasoning that such attackers may qualify for more than one attacker category (as system challengers, insiders or professionals from Section 4 may all use toolkits for their attacks) — this logically invalidates the categorisation by violating the principle of 'mutually exclusiveness' required for a well-structured typology (Section 5.2, point 2.). For transparency, the overview table for the removed class is shown in Table 6. Furthermore, data from the original dataset the typology builds on has been surfaced via example references to individual source materials. Furthermore, where the level of confidence in findings is limited due to only few supporting data sources, this has now been marked (as * or **).

Where possible, a note on the weight of an attacker type class has been made, based on the amount of supporting references and occurrences in the sample.

Beyond this current state of the typology, further iterations should be expected to develop and improve the typology based on new developments and trends in attacks influencing the attacker landscape, but also further evaluation efforts. Viable options here could include replication of the typology building exercise using a new source dataset on digital banking related cybercrime cases, but also ongoing further input and feedback from academic peer review or industry experts.

Group	**Toolkit users**
Subgroups	-/-
Labels	Users of attack toolkits, clients of criminal-to-criminal services
Motives	Financial gain, 'making ends meet'*
Criminal intent	Limited skills and funds (relying on tool kits), although more experienced attackers may use them for convenience and scalability too*
Resources	Moderate to high skill levels and funding
Activities	Phishing, ransomware, trojans and malware attacks through usage of toolkits and services available through criminal-to-criminal franchises
Level of danger posed	Medium to high
Type of risk posed	Financial risk, operational and reputational risk directly linked

Table 6: Attacker profile for toolkit users (removed)

6 Excursus: Circumplex Models in Attacker Categorisations

The usage of circumplex models as a methodological choice and visualisation tool can be observed throughout key works in the area of attacker typologies and taxonomies such as [Rogers, 06][Hald, 12] and lastly [Seebruck, 15] (also mentioned in [De Bruijne, 17] p.23). In this compact excursus, the origin of these models is briefly discussed, including their traditional usages, why and how they have found entry into the area of typologies/taxonomies and what their realistic value is when used as part of an attacker categorisation. To illustrate the nature of these models and introduce them to the reader, a customised circumplex representation for the case of digital banking has been created in Figure 2, following the principle approach taken by [Rogers, 06][Hald, 12] in their circumplexes.

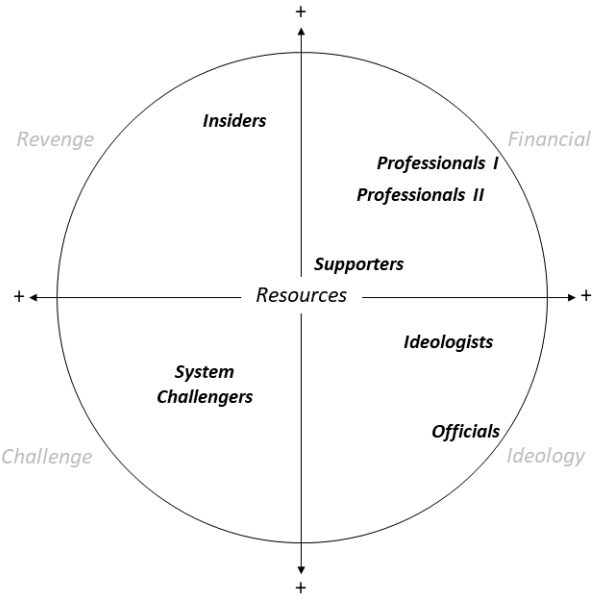

Figure 2: Circumplex visualisation for digital banking attackers

Circumplex models were initially proposed by Guttmann in 1954 as a "circular pattern of correlations in a matrix" [Tracey, 00] or a "system of variables which has a circular law of order" [Gurtman, 03], visually resulting in a two-dimensional representation of a domain (such as an attacker landscape) in which a variable set is conceptually arranged as a circle [Gurtman, 14]. These models have mostly found application in diverse clinical psychology and sociology settings such as research on e.g. interpersonal traits and interactions; personality factors and disorders; mood and affect; family and marital systems or vocational interests [Gurtman, 03][Acton, 04]. But classic circumplex models are far more than circular graphical visualisations describing a certain domain: they are statistically testable [Acton, 04] against a number of criteria to assess their circumplex properties [Gurtman, 14], and fit to underlying data [Tracey, 00]. Several criteria define circumplex models conceptually: firstly, they are best suited to accommodate two dimensions only and require interrelated variables as represented by a correlation coefficient matrix. A perfect circumplex is also signified by equal spacing of variables along the circumference of the circle and constant radius from the centre of the circle [Gurtman, 14] and can be reviewed as such using statistical testing and simulation [Acton, 04].

When it comes to adapting circumplexes as vehicles of representation for attacker categorisations, much weight is placed on the well-cited works of [Rogers 06], in line with the overall development of literature around attacker typologies and taxonomies. While [Rogers, 06] recognises that circumplexes have traditionally been used to model more complex, empirically-based behavioural concepts and personality traits as

outlined above, he sees them primarily as a representation option able to visually accommodate two interrelated variables (motivation and skill level). The circumplex as used by Rogers does not possess an attached correlation matrix defining "the exact relationship between classification variables" [Rogers, 06] or underlying statistical data to test against. At this stage, it is not a testable and empirically based circumplex aligned with previous clinical psychology or sociology works as referenced above, it seems to be a visualisation tool referring back to the circumplex shape ([Rogers, 06] indirectly acknowledges by suggesting future work using correlation coefficients).

Guided by their intention to update Rogers' work, [Hald, 12] follow his approach uncritically in their short paper, mapping their new attacker categories onto the original dimensions. [Seebruck, 15] acknowledges the origin of circumplexes as "adapted from psychology by sociologists seeking to classify groups according to attributes". While he adapts circumplexes as an intuitive way of visualisation for attacker categorisations citing Rogers' work, he regards the way circumplexes have been used previously as problematic due to their inability to depict multiple, complex motivations in attackers. In these previous models [Rogers, 06][Hald, 12] four sectors represent four distinct types of attacker motivation. As every attacker node can only be placed into one sector (or across the border between two sectors), only one (or two) type(s) of attacker motivation can be represented. An additional visualisation in form of an 'arch' to add a third dimension (secondary motivations) to the circumplex is therefore suggested by [Seebruck, 15]. In the strictest sense, this invalidates this representation as a traditional circumplex as proposed by Guttman, but falls in line with the visualisation approach taken by [Rogers, 06][Hald, 12] taken previously.

This alternative usage of the circumplex model representation in attacker categorisation literature is not necessarily a methodological issue or shortcoming. However, using a distinguishing term such as 'circumplex visualisation' or similar and an explanatory note referring back to circumplex theory may help to avoid ambiguity and strengthen the theoretical grounding for future attacker typologies and taxonomies relying on this method of visualisation.

Referring back to Figure 2, a new circumplex visualisation specific to digital banking attackers was created, based on [Rogers, 06][Hald, 12] (and to a lesser extent [Seebruck, 15]): four quadrants represent the primary four motivations for attackers (financial (gain), ideology, challenge and revenge), with the attacker's resource level mapped against this — the further out the attacker type is placed on the radius of the circumplex, the higher the resource level. In this exercise, several problems for circumplex visualisations become evident, together with some positive aspects.

Firstly, circumplex visualisations for attacker categorisations can be seen as problematic because of their relative level of ambiguity and vagueness due to the manual mapping and positioning. This may also lead to inconsistencies in practice, where different modellers may produce varying results as they assess and place individual attacker types differently within the circumplex. Furthermore, the value of circumplex approaches may be limited to practitioners due to their limitation to two classification dimensions as criticised by [Seebruck, 15] as well as the potential over-simplification of attacker landscape. Additionally, the stand-alone nature without link to existing threat modelling methods; and lastly lack of options for statistical testing and formal evaluation may hinder the practical uptake of such visualisations.

In contrast, they can be viewed as beneficial based on their highly visual nature which should make them accessible to a wide range of stakeholders. Additionally, they

enable comparisons across typologies (by overlaying circumplexes using the same dimensions) and hold the potential to visualise the full attacker landscape and consideration of all relevant attacker types. They furthermore have the capability to illustrate the relationship between skills and motivation in attackers (as mentioned in [Seebruck, 15]). Lastly, [Rogers, 06] suggests circumplex visualisations as an investigative tool, where individual, new attackers are mapped and compared to an existing circumplex showing previously identified attacker types to aid investigators.

7 Reflection

As an extension to the earlier iteration and abbreviated version of this paper presented at ARES'19 [Moeckel, 19], previous findings that have held up throughout can be highlighted, but several new aspects for the here presented typology version can now be added. From this, potential directions for future research can be enumerated.

As acknowledged for the first iteration of this typology [Moeckel, 19], a number of new and interesting aspects specific to the case of digital banking have emerged here; e.g. the introduction of the new supporter category (including money mules) or the mentioning of reputational risk — particular relevant to financial institutions with their business model largely built on trust. As with the first iteration, this typology effort has helped to demonstrate that real-life data can be used to build viable typologies. It is also hoped that the amendments made to the first iteration and presented in this paper demonstrate the need for an ongoing review process and subsequent updates to any typology. Additionally, the inclusion of methodological issues should help to shift attention to research design for future works in this area, as this might currently be underrepresented in this area of research.

Besides providing an updated typology (reduction from eight to seven attacker types following the heuristic review of the typology), a number of new aspects have been introduced in this paper specifically and may be of value to others, such as an extended discussion of circumplex models (or visualisations) in the context of attacker categorisations. The inclusion of concepts such as affinity diagrams as a design thinking and user-centred design method provides an aspect of innovation to this research area and may be of use to fellow researchers and practitioners alike — especially in the context of digital banking, where agile ways of working are widely used already. Lastly, several issues identified as problematic in the first iteration have been further developed (as outlined in Section 5) and are awaiting the next round of validation and amendments to lead to the next typology iteration beyond this paper.

Overall, it is hoped that the here presented typology has raised awareness for this interesting research area and also presented an overview of the valuable work carried out by others in the past (as cited in this paper and [Moeckel, 19]). Directly complementing and building on this bulk of work, the presented typology has indicated that real-world data, even of secondary nature, can help to build typologies, providing a new perspective over categorisation solely based on previous literature. Several methodological areas for development have been identified and will require in-depth investigation, replication and further development, for example in the areas of evaluation and visualisation (with alternative options to be considered and tested).

Future research could benefit from replication efforts using new datasets and methodological advances, also for other industry contexts, to build new (even

experimental) typologies to help strengthen the research area. Additionally, stronger motivation and rationale needs to be provided for the usage of attacker typologies: how can academic constructs be moved into practical settings and provide real-life value on an everyday basis? In direct relation, which risk or threat modelling techniques and methods already rely on typologies or could benefit from their integration? Also, how can victimology as a research field be integrated here and how can the impact on individual targets or victims be accounted for in the best manner?

At this point, collaborating directly with digital banking practitioners is certainly called for, as it could help to strengthen the case and eradicate gaps for this particular typology. This could subsequently also enable researchers to learn more about the potential of sector-specific typologies, how they can be integrated in everyday work routines of security professionals and what is required from academia to progress this research field in a meaningful manner.

References

[Acton, 04] Acton, G.S., Revelle, W.: "Evaluation of Ten Psychometric Criteria for Circumplex Structure"; Methods of Psychological Research Online, 9, 1 (2004).

[BCS, 14] British Computer Society (BCS): "Cybercrime Forensics Specialist Group briefings"; compiled by Denis Edgar-Nevill (Canterbury Christ Church University), available via group distribution list (2010-2014).

[Bond, 14] Bond, M., Choudary, O., Murdoch, S.J., Skorobogatov, S. and Anderson, R.: "Chip and skim: Cloning EMV cards with the pre-play attack," 2014 IEEE Symposium on Security and Privacy, 49–64.

[Brown, 08] Brown, T.: "Design Thinking"; Harvard Business Review, 6 (2008).

[Charmaz, 14] Charmaz, K.: "Constructing Grounded Theory"; 2nd ed., SAGE (2014).

[Cummings, 12] Cummings, A., Lewellen, T., McIntire, D., Moore, A.P., Trzeciak, R.: "Insider Threat Study: Illicit Cyber Activity Involving Fraud in the U.S. Financial Services Sector"; Carnegie Mellon University, Software Engineering Institute (2012).

[De Bruijne, 17] De Bruijne, B., van Eeten, M., Ganan, C.H. Pieters, W.: "Towards a new cyber threat actor typology — a hybrid method for the NCSC cyber security assessment"; TU Delft (2017).

[Epstein, 15] Epstein, S.: "Understanding digital banking,"; available under https://www.finextra.com/blogposting/10390/understanding-digital-banking (2015).

[FBI, 19] Federal Bureau of Investigation (FBI): "Cyber's most wanted"; available under https://www.fbi.gov/wanted/cyber (Dec. 2019).

[FBI, 20a] FBI: "Iranian DDoS attacks"; https://www.fbi.gov/wanted/cyber/iranian-ddos-attacks (Dec. 2019).

[Friis, 2020] Friis Dam, R., Teo, Y.S.: "Affinity Diagrams – Learn How to Cluster and Bundle Ideas and Facts"; available under https://www.interaction-design.org/literature/article/affinity-diagrams-learn-how-to-cluster-and-bundle-ideas-and-facts (2020).

[Ginovsky, 15] Ginovsky, J.: "What really is 'digital banking'?"; available under http://www.bankingexchange.com/blogs-3/ making-sense-of-it-all/item/5187-what-really-is-digital-banking (Apr. 2015).

[Gordon, 96] Gordon, S.: "The generic virus writer I+II"; 6th International Virus Bulletin Conference, Brighton, UK, Sept. 1996). Also available under: https://www.virusbulletin.com/virusbulletin/2015/06/throwback -thursday-virus -writers-part-1-may-1999.

[Gurtman, 14] Gurtman, M.B.: "Circumplex Models"; in The Encyclopedia of Clinical Psychology, John Wiley & Sons (2014), 507–518.

[Gurtman, 03] Gurtman, M.B., Pincus, A.L.: "The Circumplex Model: Methods and Research Applications"; in Handbook of Psychology, Research Methods in Psychology, John Wiley & Sons (2003), 407–428.

[Hald, 12] Hald, S.L.N., Pedersen, J.M.: "An updated taxonomy for characterising hackers according to their threat properties"; 14th International Conference on Advanced Communication Technology (2012), 81–86.

[Harboe, 15] Harboe, G., Huang, E.M.: "Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap"; Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (2015), 95–104.

[Hollinger, 88] Hollinger, R.C.: "Computer hackers follow a Guttman-like progression"; Phrack Inc. 2, 22 (1988); also available under http://phrack.org/issues/22/7.html.

[Hutchings, 19] Hutchings, A.: "Cambridge Computer Crime Database"; (Dec. 2019), available under https://www.cl.cam.ac.uk/{~}ah793/cccd.html.

[Hurtado, 15] Hurtado, P., Smythe, C.: "Ex-JPMorgan Employee charged with stealing customer data"; in Bloomberg News (Apr. 2015), available under https://www.bloomberg.com/news/articles/2015-04-28/exjpmorgan-worker-arrested-by-fbi-over-theft-of-consumer-data.

[Ivoce, 97] Ivoce, D.J.: "Collaring the cybercrook: An investigator's view"; IEEE Spectrum, 34, 6 (1997), 31–36.

[Karia, 12] Karia, K.: "Hacker exposes three million Iranian bank account details"; in TechWeek Europe (April 2012), available under https://www.silicon.co.uk/workspace/hacker-three-million-iranian-bank-accounts-73161.

[Kimbell, 11] Kimbell, L.: Rethinking Design Thinking: Part I. Design and Culture: The Journal of the Design Studies Forum, 3, 3 (2011), 285–306.

[Krebs, 13] Krebs, B.: "Feds charge Calif. brothers in cyberheists"; available under https://krebsonsecurity.com/2013/11/feds-charge-calif-brothers-in-cyberheists.

[Landreth, 89] Landreth, W.: "Out of the Inner Circle: A Hacker's Guide to Computer Security"; Microsoft Press (1989).

[Lee, 12] Lee, D.: "Flame: Massive cyberattack discovered, researchers say", in BBC News (May 2012), available under https://www.bbc.co.uk/news/technology-18238326.

[Long, 12] Long, L.A. and Hadsell, E.: "Profiling hackers"; (Jan. 2012) available under http://www.sans.org/reading_room/whitepapers/hackers/profiling-hackers_33864.

[Maher, 2018] Maher, C., Hadfield, M., Hutchings, M., de Eyto, A.: "Ensuring Rigor in Qualitative Data Analysis: A Design Research Approach to Coding Combining NVivo With Traditional Material Methods"; International Journal of Qualitative Methods (July 2018).

[Merhi, 19] Merhi, M., Hone, K., and Tarhini, A.: "A cross-cultural study of the intention to use mobile banking between lebanese and british consumers: Extending UTAUT2 with security, privacy and trust," Technology in Society, 59, 1 (2019), 101–151.

[Meyers, 09] Meyers, C., Powers, S. and Faissol, D.: "Taxonomies of Cyber Adversaries and Attacks: A Survey of Incidents and Approaches"; Technical Report, U.S. Department of Energy, Lawrence Livermore National Laboratory (2009).

[Moeckel, 10] Moeckel, C. and Abdallah, A.E.: "Threat modeling approaches and tools for securing architectural designs of an e-banking application," Sixth International Conference on Information Assurance and Security, 149–154.

[Moeckel, 19] Moeckel, C.: "Examining and Constructing Attacker Categorisations : an Experimental Typology for Digital Banking"; Proc. 14th International Conference on Availability, Reliability and Security (ARES '19): 1st International Workshop on Information Security Methodology and Replication Studies (IWSMR 2019), ACM, (2019).

[Moeckel, 20] Moeckel, C.: "Public research profile at Royal Holloway University and current list of individual data sources ('Activities')"; available under https://pure.royalholloway.ac.uk/portal/en/persons/caroline-moeckel(cc765b08-a56e-4a71-8348-c32c1edb580e).html.

[Nykodym, 05] Nykodym, N., Taylor R., and Vilela, J.: "Criminal profiling and insider cyber crime"; Digital Investigation 2, 4 (2005), 261–267.

[Pfleeger, 06] Pfleeger, C.P., Pfleeger, S.L.: "Security in Computing"; Prentice Hall (2006).

[Plain, 07] Plain, C.: "Build an Affinity for K-J Method"; Quality Progress, 40, 3 (2007), 88.

[Prince, 11] Prince, B.: "Inside cybercrime money mule operations"; (Jan. 2011) available under https://www.eweek.com/security/inside-cyber-crime-money-mule-operations.

[RBS, 20] The Royal Bank of Scotland: "Security disclosures for professionals"; (2020) available under https://personal.rbs.co.uk/personal/fraud-and-security/responsible-disclosure.html.

[Rogers, 99] Rogers, M.K.: "A new hacker taxonomy"; (1999) available under http://homes.cerias.purdue.edu/~mkr/hacker.doc.

[Rogers, 06] Rogers, M.K.: "A two-dimensional circumplex approach to the development of a hacker taxonomy"; Digital Investigation 3, 2 (2006).

[Rosenquist, 09] Rosenquist, M. and Casey, T.: "Prioritizing Information Security Risks with Threat Agent Risk Assessment (TARA)"; available under https://www.researchgate.net/publication/335589639_Prioritizing_Information_Security_Risks_with_Threat_Agent_Risk_Assessment_TARA.

[Saldana, 12] Saldaña, J.: "The Coding Manual for Qualitative Researchers"; SAGE (2012).

[Schwartz, 12] Schwartz, M.J.: "Bank hacks: Iran blame game intensifies"; in Dark Reading, available under https://www.darkreading.com/attacks-and-breaches/bank-hacks-iranblame-game-intensifies/d/d-id/1106857.

[Seebruck, 15] Seebruck, R.: "A typology of hackers: Classifying cyber malfeasance using a weighted arc circumplex model"; Digital Investigation, 14 (2015), 36–45.

[Shostack, 14] Shostack, A.: "Threat Modeling: Designing for Security"; John Wiley & Sons, UK (2014).

[Tracey, 00] Tracey, T.J.G.: "Analysis of Circumplex Models. Handbook of Applied Multivariate Statistics and Mathematical Modeling"; Academic Press (2000) 641–664.

[Urquhart, 13] Urquhart, C.: "Grounded Theory for Qualitative Research"; SAGE (2013).

[VERIS, 19] VERIS Github: "Vocabulary for Event Recording and Incident Sharing (VERIS) Community Database"; in GitHub: https://github.com/vz-risk/VCDB (Dec. 2019).

[Ward, 12] Ward, M.: Anti-Sec: “Who are the world’s most wanted hackers?”; in BBC News (Mar. 2012), available under https://www.bbc.co.uk/news/technology-17548704.

[Xin, 14] Xin, T. and Xiaofang, B. “Online Banking Security Analysis based on STRIDE Threat Model,” International Journal of Security and Its Applications, 8, 2 (2014), 271–282.

[Ziegler, 04] Ziegler, W. and Fötinger, C.S.: “Understanding a hacker’s mind — a psychological insight into the hijacking of identities” (2004).